Universidade de Brasília

Instituto de Ciências Exatas

Departamento de Matemática

Finite Groups of Bounded Rank Admitting an Almost Regular Automorphism: a Lie-Theoretic Approach

por

João Pedro Papalardo Azevedo

Orientadora: Prof^{a}. Dr^{a}. Cristina Acciarri

Brasília

2020

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

Azevedo, João Pedro Papalardo
Af Finite Groups of Bounded Rank Admitting an Almost
Regular Automorphism: a Lie-Theoretic Approach / João Pedro
Papalardo Azevedo; orientador Cristina Acciarri. --
Brasília, 2020.
120 p.

Dissertação (Mestrado - Mestrado em Matemática) --
Universidade de Brasília, 2020.

1. Group Theory. 2. Lie Methods in Group Theory. 3. Rank
of a Finite Group. 4. Automorphisms. 5. Lie Rings. I.
Acciarri, Cristina, orient. II. Título.

Agradecimentos

Gostaria de agradecer primeiramente a Deus, pela oportunidade de continuar meus estudos nesses dois anos, pela saúde e pela vida.

Aos meus pais, Robertha e Ronaldo, por tudo que já fizeram e fazem por mim e pela inestimável ajuda para que esse trabalho pudesse ser concluído. Às minhas irmãs, Júlia e Ana, pela leveza que trazem à minha vida durante todos esses anos.

Um agradecimento especial para minha namorada, Regina, que me escuta e opina em todas as minhas questões, mesmo que ainda não saiba de cor os axiomas que tornam um conjunto um grupo. Muito obrigado por cada conversa, e por tudo que você me ensina sobre a vida. Só eu sei como você foi importante nesse caminho. Gratidão.

Mais um agradecimento especial para minha orientadora, a professora Cristina Acciarri. Para mim, seu jeito incansável e dedicado de trabalhar é um exemplo muito maior que os seus sempre bons conselhos, aos quais tentei seguir a todos. Não esqueço de nenhuma conversa que tivemos. Muito obrigado pela paciência e por dividir comigo seu jeito de enxergar o mundo e a matemática.

Ao professor Emerson de Melo, pelo ano em que foi meu orientador. Obrigado por me mostrar que existe intuição por trás das coisas mais abstratas.

Aos amigos e colegas que estiveram comigo em todos esses anos de universidade, aos quais não conseguiria citar sem cometer a injustiça de deixar de mencionar algum.

Ao CNPq, pelo apoio financeiro concedido durante a realização deste trabalho.

Resumo

Seja G um grupo finito admitindo um automorfismo coprimo ϕ com exatamente m pontos fixos. Ao longo dos anos, estudou-se a influência do subgrupo dos pontos fixos de ϕ sobre a estrutura de G. Por exemplo, o célebre teorema de Thompson [29] garante que, quando a ordem de ϕ é um primo p e $m = 1$, então G é nilpotente; em [11] Higman prova que existe uma cota superior para a classe de nilpotência de G que depende só de p. O objetivo do presente texto é discorrer sobre resultados similares, mas concernindo o comprimento derivado de um grupo solúvel em vez da classe de nilpotência.

As condições sobre a ordem de ϕ são essenciais em muitos resultados. Entretanto, em [27], Shalev encontra uma substituta conveniente para as hipóteses sobre a ordem de ϕ: conhecer o posto de G. Nesse contexto, Shalev prova que se G é um grupo finito de posto r que admite um automorfismo ϕ com m pontos fixos, então existe um subgrupo característico solúvel H de G tal que o índice $[G : H]$ é limitado superiormente por uma função de m e r apenas. Além disso, se ϕ é coprimo, isto é, se $(|\phi|, |G|) = 1$, o comprimento derivado $dl(H)$ também é limitado superiormente por uma função de m e r apenas. O autor pergunta se a cota para $dl(H)$ poderia depender apenas de r e se a condição de coprimalidade na ordem de ϕ poderia ser derrubada.

Em [17], Khukhro responde a primeira questão, mostrando que a cota em $dl(H)$ pode ser tomada dependendo só de r. Por fim, em [16], Jaikin-Zapirain derruba a hipótese de coprimalidade sobre a ordem de ϕ.

Palavras-chave: Grupos Finitos, Automorfismos, Posto de um Grupo Finito, Anéis de Lie.

Abstract

Let G be a finite group admitting an automorphism ϕ with exactly m fixed points. For years, the influence of the fixed point subgroup of ϕ over the structure of G has been studied. For instance, a celebrated theorem of Thompson [29] ensures that, whenever the order of ϕ is a prime p and $m = 1$, then G is nilpotent; in [11] Higman proves that there exists an upper bound for the nilpotency class of G depending only on p. The goal of the present text is to discuss similar results, but concerning the derived length of a soluble group instead of the nilpotency class.

The conditions on the order of ϕ are essential in many results. However, in [27], Shalev finds a convenient substitute to the hypotheses on the order of ϕ: to know the rank of G. In this context, Shalev proves that if G is a finite group of rank r admitting an automorphism ϕ with m fixed points, then there exists a characteristic soluble subgroup H of G such that the index $[G : H]$ is bounded from above by a function of m and r only. Moreover, when ϕ is coprime, i.e., when $(|\phi|, |G|) = 1$, it is also possible to bound the derived length $dl(H)$ from above by a function of m and r only. The author asks whether the bound on $dl(H)$ could depend only on r and the coprimality condition on $|\phi|$ could be dropped.

Later, in [17], Khukhro answers the first question, showing that the bound on $dl(H)$ can be taken depending on r only. Lastly, in [16], Jaikin-Zapirain drops the coprimality hypothesis on the order of ϕ.

Keywords: Finite Groups, Automorphisms, Rank of a Finite Group, Lie Rings.

Contents

Chapter 1

Introduction

Let G be finite group admitting an automorphism ϕ. It is a well-known fact that certain properties of ϕ and of the subgroup $C_G(\phi)$ of the elements of G fixed by ϕ reflect on properties of G itself. For instance, E. Khukhro's book contains a survey on results of the following type (see, for example, Theorems 8.1, 12.1, 13.1 and 14.1 of [18]): if the order of ϕ is p^m and the order of $C_G(\phi)$ is p^n, then G has a subgroup of index bounded in terms of p, m and n that is nilpotent (or soluble, depending on the result) of class (or derived length) bounded in terms of p, m and n.

We will say that an automorphism ϕ is *regular* provided that $C_G(\phi) = 1$. When ϕ is regular and of order 2, we can conclude that G is abelian, and in general if $|\phi| = p$ a prime, a celebrated theorem of J. Thompson ([29]) ensures that G is nilpotent. Moreover G. Higman, in [11], proves that the class of G is bounded above by a function of p only. This combination of Thompson's and Higman's results is maybe the most well-known example of a result connecting properties of an automorphism ϕ of a group G and its fixed points with consequences on the structure of G itself.

In general, it is also possible to answer questions linking the properties of an automorphism ϕ of a finite group G to the solubility of G or its subgroups. To that extent, let G be a finite group admitting an automorphism ϕ with exactly m fixed points. In case ϕ is regular, P. Rowley proves in [26] that G is soluble, using the classification of finite simple groups (CFSG), in analogy with Thompson's result but without the primality condition.

Now following Thompson and Higman's results, we want to impose restrictions on the orders of ϕ and its centralizer $C_G(\phi)$ and try to derive information concerning solubility of G. Given a finite soluble group G admitting an automorphism ϕ, we ask whether it is possible to find a function that bounds the derived length of G depending only on the order of ϕ and on the number of its fixed points. In 1962, this question started to be answered affirmatively by J. Alperin, whose advisor was Higman himself. In [1], Alperin solves the following particular case of our problem: Let p be a prime and n be an integer. Let G be a finite soluble group admitting an automorphism ϕ of order p and leaving exactly p^n elements of G fixed. Then, the derived length of G is bounded above by a function of p and n only.

The formulation of Alperin's theorem is interesting because we have almost the same conditions as in the problem solved by Thompson and Higman, but in the case where $|C_G(\phi)| = p^n$. Alperin's proof relies on the use of some Lie ring associated to the group G, following Higman's ideas.

Naturally, the next generalization of such a result consists in allowing ϕ to have order any power of p, say p^m. Shalev, in [28], proves that if G is a finite p-group admitting an automorphism of order p^m with exactly p^n fixed points, then the derived length of G is bounded by some function of p, m and n only.

Now, we want to make a comment on one of the main ideas behind the proofs in Alperin's and Shalev's articles mentioned above. It is a known fact that, if G is a p-group admitting an automorphism ϕ with order p^m with p^k fixed points, then G can be generated by kp^m elements. The proof is achieved combining for example Corollary 2.7 of [18] and Burnside's Basis Theorem 2.1.6. In both [1] and [28], the orders of ϕ and $C_G(\phi)$ are used to bound the number of generators of some characteristic subgroups of a p-group G, and that bound is essential for many arguments in both papers. We call the minimal integer r such that every subgroup of a group G can be r-generated the *rank* of G, denoted by $rk(G)$. The remark above shows that an hypothesis on the rank of G may replace the one on the order of ϕ.

In another result published in the same year, that is the main result of this thesis, Shalev modifies the statement of the theorem in the following way. This is the main theorem of [27].

Theorem (Shalev, 1993). *Let G be a finite group of rank r admitting an automorphism ϕ with m fixed points. Then G has a characteristic soluble subgroup H of (m, r)-bounded index. Moreover, if $(|\phi|, |G|) = 1$, then the derived length of H is (m, r)-bounded.*

The condition $(|\phi|, |G|) = 1$ means that the g.c.d between $|\phi|$ and $|G|$ is 1, and when this happens we call the ϕ a *coprime automorphism* . Here, it is interesting to observe that the coprimality between the orders of the group and the automorphism allows him to avoid specifying the order of ϕ. Besides that, the result appears in a more general form, because here G is any finite group and Shalev finds a soluble characteristic subgroup H of bounded index. His proof is based, among other things, on CFSG to prove the existence of H and on the use of a class of Lie rings associated to uniform p-groups (see Section 4.2 for details) to bound the derived length of H. Such Lie rings, when they are soluble, interact very well with the solubility of the associated group.

Also, in [27], Shalev says not to know yet whether the coprimality hypothesis can be dropped in his statement. The subsequent investigations focused on this coprimality dependence and on the following question: is it possible to bound the derived length of H by a function of r only?

In 1996, E. Khukhro [17] approaches the problem solved by Shalev using only group-theoretic methods, and succeeds in finding a bound which depends only on the rank of G. This is another main result of this dissertation, and is the main theorem of [17].

Theorem (Khukhro, 1996). *If a finite group of rank r admits an automorphism of coprime order with exactly m fixed points, then the group has a soluble characteristic subgroup of (m, r)-bounded index whose derived length is r-bounded.*

His proof follows the same general ideas used by Shalev with some more technical group-theoretic arguments. After Khukhro's result, one question still remains unanswered: is it really necessary to have the coprimality condition on the automorphism ϕ?

This question was submitted to the Kourovka Notebook 1995 edition as Problem 13.56, with the following formulation: "Let G be a finite p-group of rank r and ϕ an automorphism of G having exactly m fixed points. Is the derived length of G bounded by a function of m and r only?" In 2002, A. Jaikin-Zapirain enhanced, in [16], Shalev's results concerning the uniform Lie ring introduced in [28], by finding correspondences between subgroups of a uniform group G and ideals of the associated Lie ring, besides obtaining other results on uniform Lie rings. Then, using ideas from both [27] and [17], Jaikin-Zapirain managed to prove the theorem in its greatest generality, without the coprimality hypothesis between the orders of ϕ and G. He also posed new problems, related to the previous ones. We will give some highlights of the result in [16] at the end of the dissertation.

Finally, some words on the structure of this text. It is self-contained enough for a reader with basic algebraic knowledge to understand all of the results. After the introduction, which is Chapter 1, the second chapter collects some definitions and basic results used within the text. Chapter 3 is devoted to the theory of Lie rings, in particular the theory of associating a Lie ring to a group, in Section 3.3, and in Section 3.4 we prove a useful result concerning solubility of some graded Lie rings, due to Shalev. In Chapter 4, we discuss the theory of powerful p-groups, a fundamental tool in the proof of the main theorems. Chapter 5 concerns a way to construct from scratch a family of finite simple groups, the so-called finite simple groups of Lie type, and ends with a full characterization of their automorphisms. The last chapter, Chapter 6 is devoted to proving the main results, from Shalev's, Khukhro's and Jaikin's papers, namely Theorems 6.0.2, 6.5.1 and 6.6.7. It also contains the conjectures Jaikin posed at that time.

Chapter 2

Preliminaries

In this chapter, for the sake of completeness, we include various results that will be referred to along the text and are considered of more elementary content. References for the results are also given. Some theorems, as the Sylow, correspondence and the isomorphism theorems, will be assumed known by the reader.

When G is a finite group we are going to denote by $|G|$ the order of G. If ϕ is an automorphism of G and $x \in G$, we will denote both by x^{ϕ} and $\phi(x)$ the image of x through ϕ depending on the context. Along the text other notations will be introduced as necessary.

2.1 Basic Properties in Groups and Rings

In this section we give some definitions that will be used along the dissertation, in order to make the text more organized and self-contained. If H and K are subgroups of a group G, we denote by $[H, K]$ the subgroup generated by the commutators $[h, k]$ with $h \in H$ and $k \in K$.

Definition 2.1.1. *Let G be a group. We define the lower central series for G recursively, as follows. Put $\gamma_1(G) = G$ and define, for $i \geq 2$, $\gamma_n(G) = [\gamma_{n-1}(G), G]$. We say that G is nilpotent of class c when c is the smallest integer such that $\gamma_{c+1}(G) = 1$.*

Definition 2.1.2. *Let G be a group. We define the derived series for G as follows. Make $G^{(0)} = G$ and, for $n \geq 1$, define $G^{(n)} = [G^{(n-1)}, G^{(n-1)}]$. If d is the smallest integer such that $G^{(d)} = 1$, we say G is soluble of derived length d and denote by $dl(G) = d$.*

Another common definition for solubility on groups is the following. Assume that G admits a series of subgroups

$$G = G_0 \triangleright G_1 \triangleright \cdots \triangleright G_d = 1$$

where all quotients G_i/G_{i+1} are abelian. Then G is soluble of derived length at most d.

Let π be a set of prime numbers. We say that a natural number n is a π-number whenever the prime factors of n are all contained in π, and we say that n is a π'-number whenever n is coprime with every element of π. We say that a finite group G is a π-group or a π'-group if $|G|$ is a π-number or a π'-number, respectively.

Definition 2.1.3. *We denote, as usual, by $O_{p'}(G)$ the largest normal p'-subgroup of G. Also, we define $O_{p'p}(G)$ as the preimage, in G, of the largest normal p-subgroup of $G/O_{p'}(G)$.*

The subgroup $O_{p'}(G)$, and also $O_{p'p}(G)$, can be defined in a different way. The first subgroup can also be defined as the subgroup

$$\langle K \mid K \triangleleft G \,\text{and}\, K \text{ is a } p'\text{-group}\rangle,$$

while the second one may be defined as the full preimage in G of the following subgroup of $G/O_{p'}(G)$

$$\left\langle \frac{KO_{p'}(G)}{O_{p'}(G)} \;\middle|\; \frac{KO_{p'}(G)}{O_{p'}(G)} \triangleleft \frac{G}{O_{p'}(G)} \text{ is a } p\text{-group} \right\rangle.$$

These definitions make it clear that both $O_{p'}(G)$ and $O_{p'p}(G)$ are characteristic subgroups of G.

Definition 2.1.4. *Given a finite group G, we define the Fitting subgroup $Fit(G)$ as the subgroup generated by all normal nilpotent subgroups of G.*

In particular, a theorem of Fitting states that $Fit(G)$ is also nilpotent, [24, Theorem 10.22]. We denote by $\pi(G)$ the set of primes dividing the order of a finite group G. It is possible to prove that $Fit(G) = \prod_{p\in\pi(G)} O_p(G)$ and that $Fit(G) = \cap_{p\in\pi(G)} O_{p'p}(G)$.

We say that a proper subgroup M of a group G is *maximal* if, whenever there exists a subgroup H of G satisfying $M \leq H \leq G$, then $H = M$ or $H = G$.

Definition 2.1.5. *If G is any finite group, we define the Frattini subgroup $\Phi(G)$ of G as the intersection of all maximal subgroups of G.*

Again, $\Phi(G)$ is a characteristic subgroup of G. It is proved in [7, Theorem 5.1.1] that, whenever $x \in \Phi(G)$ and T is a subset of G then $G = \langle T, x\rangle$ implies $\langle T\rangle$. This is equivalent to say that the elements of $\Phi(G)$ are irrelevant for the generation of the group G. They are sometimes called non-generators for G.

We say that a group G is a p-group whenever all elements of G have order a power of p. In the special case when G is a finite p-group, the Frattini subgroup has even more properties, as stated below in the so-called Burnside's Basis Theorem.

Theorem 2.1.6. *Let G be a p-group. Then, the quotient $G/\Phi(G)$ is an elementary abelian p-group. If we assume further that the set $\{x_1\Phi(G), \dots, x_n\Phi(G)\}$ is a basis for $G/\Phi(G)$ as a vector space, then the group G cannot be generated by less that n elements, and also $G = \langle x_1, \dots, x_n\rangle$.*

Now, let G be a finite group. Let $\mathcal{C}$ be the set containing all generating sets for G. That means that for any $T \in \mathcal{C}$, $\langle T\rangle = G$. In this sense, we define $d(G)$ as the following invariant of G:

$$d(G) = \min\{|T| \,|\, T \in \mathcal{C}\}.$$

Intuitively, $d(G)$ represents the smallest cardinality of a generating set for G. There is another invariant associated to G, which we call the *rank* of G. It is defined by

$$rk(G) = \sup\{d(H) \,|\, H \leq G\}.$$

This second invariant represents the smallest cardinality such that every subgroup H of G can be generated by a set of that cardinality.

Next we give the remaining definitions.

Definition 2.1.7. *A non-empty set R is called a ring if R can be endowed with two operations, addition and multiplication, such that the following axioms hold, for all $x, y, z \in R$.*

(i) $(R, +)$ *is an abelian group.*

(ii) $x(y + z) = xy + xz$.

(iii) $(x + y)z = xz + yz$.

If, moreover, there exists an element 1_R of R such that $1_R x = x 1_R = x$ for any $x \in R$, then we say that R is a ring with unit. When the identity $(xy)z = x(yz)$ also holds in R, we say that R is an associative ring. If the identity $xy = yx$ holds in R, we say that R is a commutative ring.

As examples of rings we have $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$ with usual operations, all commutative with unit. The set of square matrices $M_n(\mathbb{R})$ is an example of a ring with unit which is non-commutative, with addition componentwise and multiplication given by the matrix product. A *subring* S of R is a subset $S \subseteq R$ for which the ring axioms hold. We say that a subring $I \subseteq R$ is a *left ideal* of R provided the product ri belongs to I for every $r \in R$ and $i \in I$. Right ideals are defined in an analogous way, and when I is both a left and right ideal of R, we call it a bilateral ideal.

An endomorphism of a group A is a homomorphism of A into itself. Let A be an abelian group and consider $G = End_{\mathbb{Z}}(A)$ to be the set of endomorphisms of A. We can endow G with the structure of a ring. Given ϕ, ψ in G, define the sum $\phi + \psi$ as $(\phi + \psi)(x) = \phi(x) + \psi(x)$ for every x in A, this operation being clearly commutative. The neutral element is the endomorphism mapping A to 0, the opposite of ϕ is the map $-\phi : x \mapsto -\phi(x)$. The multiplication is the composition of functions in G. Multiplication is right distributive with respect to the addition because every element of G is a homomorphism, and the definition of addition makes it left distributive. This is an associative ring with unit being the identity map of A.

Definition 2.1.8. *Let R be an associative ring with unit 1_R, but not necessarily commutative. A left R-module is an additive group M with an operation $R \times M \to M$ mapping (r, m) to rm, for $r \in R$ and $m \in M$, provided the following axioms hold for all $r, s \in R$ and $m, n \in M$:*

(i) $r(m + n) = rm + rn$.

(ii) $(r + s)m = rm + sn$.

(iii) $1_R m = m$.

(iv) $(rs)m = r(sm)$.

Right R-modules are defined in a similar way.

A subgroup N of the R-module M will be called a *submodule* of M provided N is itself an R-module. Every ring R is naturally a left and right R-module; R-submodules of the left-R module R are exactly the left ideals of R. Abelian groups are $\mathbb{Z}$-modules, vector spaces over the field K are K-modules.

Let R be an associative commutative ring with 1. Consider M and N to be both left R-modules, even though this definition is also possible in case the modules are not both left modules, but any combination of left and right. We next construct another left R-module from M and N. Define the *tensor product* $M \otimes_R N$ of M and N over R, as the set of all finite formal sums $\sum m_i \otimes n_i$, $m_i \in M$, $n_i \in N$, subject to the following identifications, for all $m, m' \in M$, $n, n' \in N$ and $r \in R$.

(i) $rm \otimes n = m \otimes rn = r(m \otimes n)$.

(ii) $(m + m') \otimes n = m \otimes n + m' \otimes n$.

(iii) $m \otimes (n + n') = m \otimes n + m \otimes n'$.

For more information on modules and tensor products, we refer to [6, Chapter 10].

2.2 Automorphisms

Here we collect useful results concerning automorphisms of finite groups. An automorphism ϕ of a group G is a bijection from G to itself which satisfies $\phi(xy) = \phi(x)\phi(y)$, for every x and y in G. References are also given.

Lemma 2.2.1. *Let P be a p-group admitting a p'-automorphism ϕ. Then ϕ acts trivially on P if and only if ϕ acts trivially on $P/\phi(P)$.*

The proof of this result can be found in [7, Theorem 5.1.4].

If ϕ is an automorphism of G, then we are going to call a subgroup N of G *ϕ-invariant* whenever $N^{\phi} = \{\phi(n) \,|\, n \in N\}$ equals N. We denote by $C_G(\phi)$ the subgroup $\{x \in G \,|\, x^{\phi} = x\}$. We call this subgroup the *centralizer* of ϕ in G, sometimes called the fixed point subgroup of ϕ. We also define centralizers for subsets of G. If M is any subset of G, then $C_G(M) = \{x \in G \,|\, [x, M] = 1\}$ is the subgroup consisting of the elements of G that commute with every element of M. The next result can be found in Chapter VIII, Theorem 10.14 of [14]

Lemma 2.2.2. *Let G be a finite group admitting an automorphism ϕ and a normal ϕ-invariant subgroup N. Then ϕ induces an automorphism $\bar{\phi}$ of G/N such that $|C_{G/N}(\bar{\phi})| \leq |C_G(\phi)|$.*

Now, if the order of the automorphism is coprime with the order of G, even more can be said. The following is Theorem 2.11 in [18].

Lemma 2.2.3. *Let G be a finite group admitting a coprime automorphism ϕ and a normal ϕ-invariant subgroup N. If $\bar{\phi}$ denotes the automorphism induced by ϕ on G/N, then $C_{G/N}(\bar{\phi}) = C_G(\phi)N/N$.*

Given a group G and A a group of automorphisms of G, we can define the commutator subgroup between G and A as follows: given $x \in G$ and $\phi \in A$, we denote $[x, \phi] = x^{-1}x^{\phi}$. Thus the group $[G, A]$ is the subgroup of G generated by all elements of the form $[x, \phi]$, where $x \in G$ and $\phi \in A$. The next theorem is Theorem 5.2.3 in [7].

Lemma 2.2.4. *Let A be a p'-automorphism of an abelian p-group G. Then $G = C_G(A) \times [G, A]$.*

Chapter 3

Lie Rings

Definitions and first properties of Lie rings are given in Section 3.1. In Section 3.2, we discuss a method of embedding a given Lie ring L into another Lie ring with stronger structure, and in Section 3.3 we discuss the process of, given a group G, constructing a Lie ring $L(G)$ which reflects some properties of G. Finally, in Section 3.4, we give a proof of a result related to solubility of graded Lie rings.

3.1 Definitions and First Properties

We begin this section with the definition of a Lie ring.

Definition 3.1.1. *A Lie ring L is a ring, with multiplication usually denoted by brackets $[x, y]$, for $x, y \in L$, in which the following identities hold, for all $x, y, z \in L$.*

(i) $[x, x] = 0$ *(anticommutativity)*

(ii) $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$. *(Jacobi identity)*

Let R be a commutative associative ring with 1. If, moreover, L is a left R-module subject to the relations $[rx, y] = [x, ry] = r[x, y]$ for every $x, y \in L$ and $r \in R$, then we call L a Lie R-ring and R is said to be the ground ring of L. In case that R is a field, we call L a Lie R-algebra.

Non-trivial Lie rings do not have unit. Indeed, assume L is a Lie ring with unit 1. Then we have that $[1, 1] = 1$, but anticommutativity also says that $[1, 1] = 0$. So, as $1 = 0$ and $[1, x] = x$, replacing 1 by 0 we get that $x = [1, x] = [0, x] = 0$, for all $x \in L$. So, non-trivial Lie rings do not have a unit.

A consequence of anticommutativity is that $[x, y] = -[y, x]$ for all $x, y \in L$. In fact, we may just use the distributive law in $[x + y, x + y] = 0$ to prove it.

Lie rings are not associative in general. In fact, given x, y and $z \in L$, suppose we had $[[x, y], z] = [x, [y, z]] = -[[y, z], x]$. Jacobi identity then would give that $[[z, x], y] = 0$ for all $x, y, z \in L$, and this is not true in general. For example, if we consider the Lie algebra over $\mathbb{C}$ given by $L = \langle a, b \mid [a, b] = b\rangle$, then $[a, b, a]$ is not zero.

Here, we give a few examples of Lie rings. Every Lie ring is a Lie $\mathbb{Z}$-ring and $\mathbb{R}^3$, endowed with the cross product, is an example of Lie $\mathbb{R}-$algebra. The set of all $n \times n$ matrices over a field $\mathbb{K}$ forms a Lie algebra with the usual componentwise addition and the Lie product given by $[x, y] = xy - yx$, for all matrices x, y.

In what follows, we focus on the structure of a Lie ring L. A *Lie subring* H of L is a subset of L closed under all operations, and if L is a Lie R-ring, then H must also be a R-submodule. A *homomorphism* of Lie R-rings is a homomorphism of R-modules that preserves the Lie bracket, and isomorphisms and automorphisms are defined in the usual way. An *ideal* $I \subseteq L$ is a subring of L for which $[x, y]$ belongs to I, for every x in I and y in L. As $[x, y] = -[y, x]$, the entry corresponding to the element of I may be either the first or the second. Given subsets $A, B \subseteq L$, we define $[A, B]$ to be the additive subgroup $\langle [a, b] \mid a \in A, b \in B\rangle$, generated by all commutators $[a, b]$ with $a \in A$ and $b \in B$.

Assume that L is generated, as a Lie ring, by a set of elements $X = \{x_i \mid i \in I\}$. A We will say that a Lie commutator w is *simple* if the brackets involved in the expression of w all accumulate on the left, and we will write $[x_1, \cdots, x_k] = [\cdots [[x_1, x_2], x_3], \cdots, x_k]$. For example, the commutator $[[x_1, x_2], [x_3, x_4]]$ is not simple, but $[[[x_1, x_2], x_3], x_4]$ is simple. Bilinearity of the Lie product shows that any commutator w of elements in L can be written as a linear combination of commutators in the elements of X, with the brackets occurring in the same positions they occurred in w. In this sense, we define the *weight* of a commutator w, with respect to X, as follows: commutators of weight 1 are just elements of

X, and if w and v are commutators of weights n_1 and n_2, respectively, then the commutator $[w, v]$ will have weight $n_1 + n_2$. We would like to remark that, depending on the entry set, an element can have more than one weight. For example, with respect to $X = \{a, b\}$, $[a, b]$ has weight 2, but with respect to $Y = \{[a, b]\}$, its weight is just one.

As it was remarked above, the brackets in a commutator may appear anywhere along it, and calculations may become cumbersome if we take these positions into consideration. Our first result states that we can always "ignore" the bracketing places in a Lie commutator and consider only simple ones, by paying the price of substituting an initial commutator with a $\mathbb{Z}$-linear combination of others.

Lemma 3.1.2. *Let w be a Lie commutator in the elements $x_1, \ldots, x_k$ of L.*

(i) w is a $\mathbb{Z}$-linear combination of simple commutators in the elements $x_1, \ldots, x_k$.

(ii) We can choose a fixed element $x_t \in \{x_1, \cdots, x_n\}$ present in w so that the simple commutators in item (i) all begin with x_t.

Proof. The proof of the first part will proceed by induction on the weight s of the commutator w. If s is 1 or 2, then the commutator w is itself simple. Let $s > 1$; in this case $w = [w_1, w_2]$, where w_1 and w_2 are commutators of smaller weights. By the inductive hypothesis we may assume that both w_1 and w_2 are $\mathbb{Z}$-linear combinations of simple commutators, and by distributive laws we have that w is a $\mathbb{Z}$-linear combination of commutators with the structure $[[\cdot, \cdots, \cdot], [\cdot, \cdots, \cdot]]$. Then we apply induction on the weight of w_2. If the weight is 1, then $w = [w_1, w_2]$ is simple, otherwise we may assume that $w_2 = [w_{21}, w_{22}]$ and we have

$$w = [w_1, [w_{21}, w_{22}]] = [[w_1, w_{21}], w_{22}] - [[w_1, w_{22}], w_{21}]$$

by the Jacobi identity. On the right, the initial segments $[w_1, w_{21}]$ and $[w_1, w_{22}]$ are linear combinations of simple commutators, by induction on the weight. Since the weights of w_{21} and w_{22} are less than that of w_2, induction on the weight of the second factor of the commutator finishes the proof of (i).

For the second part, fix $a \in \{x_1, \cdots, x_k\}$ and consider

$$w = [x_{i_1}, \cdots, x_{i_l}, a, x_{i_{l+1}}, \cdots, x_{i_s}].$$

We argue by induction on l, where l is the length of the initial segment preceding a in w. If $l = 1$, replace $[x_{i_1}, a]$ by $-[a, x_{i_1}]$. We may assume then that $l > 1$. We may also assume that a is the rightmost element inside the commutator, because w is simple and the Lie bracket is multilinear. Then, we may write $[x_{i_1}, \cdots, x_{i_{l-1}}, x_{i_l}, a] = [[x_{i_1}, \cdots, x_{i_{l-1}}, a], x_{i_l}] + [[x_{i_1}, \cdots, x_{i_{l-1}}], [x_{i_l}, a]]$, by Jacobi identity. On the first bracket of the right side of the equation, we apply the induction hypothesis; on the last, we regard $[a, x_{i_l}]$ as a new variable, apply the induction hypothesis and write $[[x_{i_1}, \cdots, x_{i_{l-1}}], [a, x_{i_l}]]$ as a $\mathbb{Z}$-linear combination of simple commutators starting with $[a, x_{i_l}]$. As the summands are simple commutators starting with $[a, x_{i_l}]$, the result is proved. □

We end this section defining the notions of nilpotency and solubility for Lie rings. Many of the definitions and ideas work in the same way as that they work for groups.

Definition 3.1.3. *We define recursively the terms of the lower central series of a Lie ring L by $\gamma_1(L) = L$ and $\gamma_i(L) = [L, \gamma_{i-1}(L)]$, for $i \geq 2$.*

It is straightforward to see that every $\gamma_i(L)$ is an ideal of L. By Lemma 3.1.2, $\gamma_i(L)$ is spanned by the simple commutators of weight i in the elements of L, and every commutator of length $\geq i$ is contained in $\gamma_i(L)$. Also, if L is generated by a set M, then $\gamma_i(L)$ is generated just by the simple commutators of weight i in the elements of M.

Definition 3.1.4. *A Lie ring L is said to be nilpotent of class c if $\gamma_c(L) \neq 0$ and $\gamma_{c+1}(L) = 0$, or, equivalently, if the identity $[x_1, \cdots, x_{c+1}] = 0$ holds in L, but we have $[x_1, \cdots, x_c] \neq 0$ for some choice of elements x_i.*

In view of the observations made above, if $L = \langle M \rangle$ as a Lie ring, then L is nilpotent of class c if and only if $[m_1, \cdots, m_{c+1}] = 0$ for any $m_i \in M$.

Definition 3.1.5. *The terms of the derived series of a Lie ring L are defined recursively by $L^{(0)} = L$ and $L^{(i)} = [L^{(i-1)}, L^{(i-1)}]$, for $i \geq 1$.*

We define the δ-commutator recursively as $\delta_1(x_1, x_2) = [x_1, x_2]$ and, for $n \geq 2$, $\delta_n(x_1, \ldots, x_{2^n}) = [\delta_{n-1}(x_1, \ldots, x_{2^{n-1}}), \delta_{n-1}(x_{2^{n-1}+1}, \ldots, x_{2^n})]$. Then we have the following definition.

Definition 3.1.6. *A Lie ring L is said to be soluble of derived length d if $L^{(d)} = 0$ and $L^{(d-1)} \neq 0$ or, equivalently, if the identity $\delta_d(x_1, \cdots, x_{2^d}) = 0$ holds in L and $\delta_{d-1}(x_1, \cdots, x_{2^{d-1}}) \neq 0$ for some choice of elements x_i of L.*

Again, each term of the derived series is an ideal of L. One could also define a soluble Lie ring as a ring having a series of ideals, ranging from L to 0, whose quotients are all abelian, i.e, the commutator $[x, y]$ equals 0 for any x, y in the factor. We remark here that, unlike nilpotency, if $L = \langle M \rangle$, then it may not be sufficient to verify $\delta_d(x_1, \cdots, x_{2^d}) = 0$ only on the elements for M to guarantee the solubility of L. For example, in a 2-generated Lie ring $L = \langle x, y \rangle$, the identity $\delta_2 \equiv 0$ holds for the generators: every possible δ_2 commutator has a subcommutator $[x, x]$ or $[y, y]$, and the ones that do not have will include two subcommutators of the form $[x, y]$. However, there are 2-generated Lie rings which are not soluble, for instance $L = \langle \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) \rangle$, with Lie bracket $[A, B] = AB - BA$ and over a field of characteristic different from 2. Here, $L^{(2)} = L$, so L is not soluble.

Nilpotent Lie rings are certainly soluble, since the γ-series is a series of ideals with abelian quotients that ranges from L to 0. The converse may not be true, and the simplest example is the non-abelian 2-dimensional Lie algebra over an arbitrary field $L = \langle a, b \,|\, [a, b] = b \rangle$. Indeed, $L^{(2)} = \langle b \rangle$ and $L^{(3)} = 0$, but $\gamma_i(L) = \gamma_2(L) = \langle b \rangle$ for every $i \geq 2$. Observe that many results about nilpotent and soluble Lie rings can be proved in the same way as their analogues for group theory are obtained.

3.2 Extending the Ground Ring

Let us start with the following definition.

Definition 3.2.1. *Let A be an additively written abelian group. A Lie ring L is said to be A-graded if the additive group of L is the direct sum $L = \oplus_{g\in A} L_g$ of the additive subgroups L_g, $g \in A$, such that $[L_g, L_h] \subseteq L_{g+h}$ for every g, h in A. More generally, if we consider just a subset S of an abelian group A, we can define an S-grading in the same way, but now demanding that the product $[L_g, L_h]$ be zero if $g + h$ does not belong to S.*

We now turn our attention to automorphisms. Let ϕ be an automorphism of finite order n of a Lie $\mathbb{C}$-algebra L. As ϕ is invertible, none of the eigenvalues of this automorphism are equal to zero. The Jordan normal matrix of ϕ is diagonal, since for any Jordan block of size greater than 1 we have

$$\begin{pmatrix} \alpha & 1 & \\ & \ddots & \\ & & \end{pmatrix}^n = \begin{pmatrix} \alpha^n & n\alpha^{n-1} & \\ & \ddots & \\ & & \end{pmatrix},$$

and this cannot be the identity matrix since $\alpha \neq 0$. Thus all Jordan blocks have size 1×1 and L decomposes into the direct sum of eigenspaces $L = \oplus_{i=0}^{n-1} L_i$, where $L_i = \{l \in L \mid l^\phi = \omega^i l\}$, for a fixed n-th primitive root of unity ω. The fact that this sum is direct follows from the fact that a Lie $\mathbb{C}$-algebra is a vector space over $\mathbb{C}$, and the standard arguments from Linear Algebra apply here. For any $x \in L_i$, $y \in L_j$ we have $[x, y]^\phi = [x^\phi, y^\phi] = [\omega^i x, \omega^j y] = \omega^{i+j}[x, y]$, in such a way that $[L_i, L_j] \subseteq L_{i+j}$, where $i+j$ is read modulo n. The direct sum decomposition of L above gives rise to a $\mathbb{Z}/n\mathbb{Z}$-grading.

As it has been said before, a finite order automorphism ϕ of a Lie $\mathbb{C}$-algebra L gives rise to a natural decomposition of L into eigenspaces and therefore gives rise to a natural grading of the Lie algebra too. The reason why that construction was possible is that every possible eigenvalue of ϕ lies in the field over which L is defined. It is clear that, if we replace $\mathbb{C}$ by $\mathbb{R}$, then the same construction may not be possible, depending on the eigenvalues of the automorphism ϕ. What we want to do here is to construct gradings as the one considered above, and also to avoid the problem of not being able to find every eigenvalue of the considered automorphism inside the initial ground ring of L.

With this purpose in mind, we discuss the process of extending the ground ring of a Lie ring L. If $R \subseteq K$ are commutative associative rings with 1, and L is a Lie R-ring, then the process of extending the ground ring of R consists in considering the R-module $\overline{L} = L \otimes_R K$, which is a Lie K-ring with respect to the Lie multiplication

$$[l_1 \otimes k_1, l_2 \otimes k_2] = [l_1, l_2] \otimes k_1 k_2, \quad l_1, l_2 \in L, k_1, k_2 \in K. \tag{3.2.1}$$

This Lie bracket inherits bilinearity, anticommutativity and the Jacobi identity from the one of L. Of course, $\overline{L}$ is also a Lie R-ring, and we can identify L with the R-subring $L \otimes_R R = L \otimes 1 = \{l \otimes 1 \mid l \in L\}$ of $\overline{L}$. Also, in view of equation (3.2.1), it is straightforward to see that, for any subsets A, B of L, we have

$$[A \otimes_R K, B \otimes_R K] = [A, B] \otimes_R K.$$

Also, if I is an ideal of L, then $I \otimes_R K$ will be an ideal of $\overline{L}$.

The process of extending the ground ring also behaves well when we consider automorphisms. If G is a group of automorphisms of L, then G can also be regarded as a group of automorphisms of the Lie K-ring $\overline{L}$, via the action $(l \otimes k)^\phi = l^\phi \otimes k$, for every $l \in L$, $k \in K$, $\phi \in G$.

For instance, consider a Lie $\mathbb{Z}$-ring L, admitting an automorphism ϕ of finite order n. Then, the previous construction shows that ϕ induces an automorphism of the same order on $\tilde{L} = L \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$, where ω denotes a primitive n-th root of unity. As in the discussion of the previous section, we can define additive subgroups of $\tilde{L}$ via

$${}^i\tilde{L} = \{l \in \tilde{L} \mid l^\phi = \omega^i l\}.$$

Each subgroup ${}^i\tilde{L}$ is called a ϕ-component of $\tilde{L}$.

This construction plays an important role in the theory of Lie rings: if we consider a Lie ring L admitting an automorphism ϕ of finite order, then we can extend $\mathbb{Z}$ by a n-th primitive root of unity ω and embed L into the tensorized Lie ring $\tilde{L} = L \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$. Now, every eigenvalue of ϕ belongs to the new and larger ground ring $\mathbb{Z}[\omega]$, allowing us to diagonalize ϕ as we did in the remark after the definition of grading. We will give applications of this construction in the following sections.

3.3 Associated Lie Rings

The Hall-Witt identity

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

holds in groups and resembles a lot the Jacobi identity described above, which is an identity in Lie rings. Also, the usual commutator formulas for elements in a group G, $[ab, c] = [a, c]^b [a, b]$ and $[a, bc] = [a, b][a, c]^b$, are similar to the distributive laws. It is natural to try to construct a Lie ring whose addition is based on group multiplication and whose Lie product derives from the commutator operation in the group. Since Lie rings are linear, they may be easier structures to study; an automorphism of a Lie algebra, for example, can be regarded as a linear transformation and has eigenvalues in an extension of the base field. In this sense, it would be of interest to associate to a group G a Lie ring L which inherits some properties of G and to make the association in such a way that information about G can be derived back from properties of L. A Lie ring method for studying groups consists in translating conditions about the group G into the Lie ring language, then obtaining results on Lie rings, and finally translating the conclusions back into the group language, to deduce information about the group. We introduce here the so-called method of the associated Lie ring.

Definition 3.3.1. *A series of a group G*

$$G = K_1 \geq K_2 \geq \cdots \geq K_c \geq K_{c+1} = 1$$

is said to be strongly central if $[K_i, K_j] \leq K_{i+j}$ *for all* $i, j = 1, 2, \ldots, c$. *It is* *understood that $K_s = 1$ whenever $s > c$.*

Next, we wish to give some examples of strongly central series for a group G. The most natural one to consider is the lower central series $\{\gamma_i(G)\}$. Also, we have a series defined inductively by $\lambda_1(G) = G$ and

$$\lambda_i(G) = [G, \lambda_{i-1}(G)](\lambda_{i-1}(G))^p,$$

called *lower central p-series* and associated to a finite p-group G. This is the fastest descending central series whose factors have exponent p. Some other classes of p-groups may admit other kinds of strongly central series, according to their particularities, as will be explored in Section 4.2.

Let us explore the Definition 3.3.1 a little. For $i = 1$, we see that $[G, K_j] \leq K_{j+1} \leq K_j$, thus giving normality to each term of the series. Also, a straightforward inductive argument shows that $\gamma_i(G)$ is contained in K_i for every i. In particular, if G admits a finite strongly central series of length c, then G is nilpotent of class at most c; the converse also holds since the lower central series is strongly central. Moreover, since $[K_{i-1}, G] \leq K_i$, every factor group K_{i-1}/K_i is contained in the center of G/K_i, justifying the adjective *central*. It follows that each factor of the series is an abelian group.

Now we are in position to define the Lie ring associated to a strongly central series.

Definition 3.3.2. *The additive group of the associated Lie ring $L(G)$, associated to the strongly central series $\{K_i\}$ of G, is the direct sum*

$$L(G) = \oplus_{i=1}^{\infty} K_i/K_{i+1}.$$

For each $i \in \mathbb{N}$ the direct summand K_i/K_{i+1} is called the *homogeneous component of $L(G)$ of weight i.* The elements xK_{i+1} of K_i/K_{i+1} are called *homogeneous elements of weight i.* We are going to denote homogeneous elements by $\bar{x} = xK_{i+1}$. The Lie bracket is defined for the homogeneous elements as follows

$$[\bar{x}, \bar{y}] = [xK_{i+1}, yK_{j+1}] := [x, y]K_{i+j+1},$$

where xK_{i+1} and yK_{j+1} are the images of the elements $x \in K_i$ and $y \in K_j$ in the factor groups K_i/K_{i+1} and K_j/K_{j+1}, respectively. The Lie bracket is then extended to $L(G)$ by linearity. Notice here that we may have $[xK_{i+1}, yK_{j+1}] = 0$ in $L(G)$ even if $[x, y] \neq 1$ in G: we only need $[x, y] \in K_{i+j+1}$.

The next result shows that Definition 3.3.2 in fact endows $L(G)$ with a Lie ring structure.

Proposition 3.3.3. *Definition 3.3.2 defines a Lie ring structure on $L(G)$.*

Proof. Remark that K_{i+j} centralizes G modulo K_{i+j+1}, as was observed before. Wherever we will use this remark in what follows, we signal it with an $*$. Given $x, y \in K_i \backslash K_{i+1}$, we denote by $\bar{x} = xK_{i+1}, \bar{y} = yK_{i+1}$, define $\bar{x} + \bar{y} = xyK_{i+1}$ and extend this operation by linearity to the whole $L(G)$. Consider $\bar{x}$ and $\bar{y}$ as before and $z \in K_j \backslash K_{j+1}$, write $\bar{z} = zK_{j+1}$. We have

$$\begin{aligned}
[\bar{x} + \bar{y}, \bar{z}] &= [xy, z]K_{i+j+1} \\
&= [x, z]^y [y, z]K_{i+j+1} \\
&= [x, z][y, z]K_{i+j+1} \qquad \text{(because of } *,\text{)} \\
&= [\bar{x}, \bar{z}] + [\bar{y}, \bar{z}].
\end{aligned}$$

This proves the bracket to be linear in the first entry, and linearity in the second entry also follows from a similar argument.

We are going to verify now that the Lie bracket in $L(G)$ is well-defined, and it is sufficient to verify this claim just for homogeneous elements, since the defined bracket defined is bilinear. Consider $xK_{i+1} = x'K_{i+1}$, $yK_{j+1} = y'K_{j+1}$. Then $x' = xg_1$ and $y' = yg_2$, where $g_1 \in K_{i+1}$ and $g_2 \in K_{j+1}$.

$$\begin{aligned}
[\overline{x'}, \overline{y'}] &= [x', y']K_{i+j+1} \\
&= [xg_1, yg_2]K_{i+j+1} \\
&= [x, yg_2]^{g_1}[g_1, yg_2]K_{i+j+1} \\
&= [x, yg_2][g_1, yg_2]K_{i+j+1} \qquad \text{(because of } *,\text{)} \\
&= [x, yg_2]K_{i+j+1}, \qquad \text{(as } [g_1, yg_2] \in K_{i+j+1},\text{)} \\
&= [x, g_2][x, y]^{g_2}K_{i+j+1} \\
&= [x, y]K_{i+j+1} = [\bar{x}, \bar{y}]
\end{aligned}$$

since $[x, g_2] \in K_{i+j+1}$ and using $*$ again.

As the bracket is bilinear, we will prove anticommutativity and the Jacobi identity just for homogeneous elements. As for the anticommutativity, take any $x \in K_i$ and see that $[\bar{x}, \bar{x}] = [x, x]K_{i+1} = 1K_{i+1} = \bar{0}$. Now, we prove that the Jacobi identity holds in $L(G)$. Consider here $x \in K_i \backslash K_{i+1}, y \in K_j \backslash K_{j+1}$ and $z \in K_l \backslash K_{l+1}$ and $\bar{x}, \bar{y}$ and $\bar{z}$ their usual images on the associated Lie ring. As a consequence of the Hall-Witt identity we have:

$$[x, y^{-1}, z]^y [y, z^{-1}, x]^z [z, x^{-1}, y]^x K_{i+j+l+1} = K_{i+j+l+1}.$$

Then, as K_{i+j+l} centralizes G modulo $K_{i+j+l+1}$, the conjugations in the equation above are trivial. In this sense, it follows that

$$[x, y^{-1}, z][y, z^{-1}, x][z, x^{-1}, y]K_{i+j+l+1} = K_{i+j+l+1}. \tag{3.3.1}$$

which equals $\bar{0}$ in $L(G)$.

Since we are using additive notation in $L(G)$, $\overline{x^{-1}} = -\bar{x}$. Now, translating the left side of (3.3.1) in the language of $L(G)$ we get

$$[\bar{x}, -\bar{y}, \bar{z}] + [\bar{y}, -\bar{z}, \bar{x}] + [\bar{z}, -\bar{x}, \bar{y}] = -([\bar{x}, \bar{y}, \bar{z}] + [\bar{y}, \bar{z}, \bar{x}] + [\bar{z}, \bar{x}, \bar{y}]) = \bar{0},$$

and the Jacobi identity holds, completing the proof. □

It is important to say that two non-isomorphic groups may have isomorphic associated Lie rings. For instance, let $D = \langle r, s \,|\, r^4 = s^2 = (rs)^2 \rangle$, the dihedral group of order 8, and $Q = \langle a, b \,|\, aba = b, bab = a \rangle$, the quaternion group of order 8. Consider the Lie rings associated to the lower central series of D and Q. Both Lie rings are vector spaces of dimension 3 over $\mathbb{F}_2$. Now, for the former, we have

$$L(D) = D/D' \oplus D'/D'' = \langle rD', sD' \rangle / \oplus \langle r^2 D'' \rangle \rangle,$$

with

$$[rD', sD'] = r^2 D'', [rD', r^2 D''] = 0, [rD', s^2 D''] = 0,$$

while for the latter

$$L(Q) = Q/Q' \oplus Q'/Q'' = \langle aQ', bQ' \rangle / \oplus \langle a^2 D'' \rangle \rangle,$$

with

$$[aD', bD'] = a^2D'', [aD', a^2D''] = 0, [bD', a^2D''] = 0,$$

which are clearly isomorphic as Lie algebras.

The structure of G strongly influences that of $L(G)$, as the associated Lie ring inherits properties related to nilpotency and solubility from G. In some cases, $L(G)$ will also have the same order as G, and that is the content of the next proposition.

Proposition 3.3.4. *If G is nilpotent of class c, then $L(G)$ is also nilpotent of class at most c. Also, if the derived lengt of G is d, then $L(G)$ is soluble of derived length at most d. If G is also finite, then $|L(G)| = |G|$.*

Proof. The definition of multiplication in $L(G)$ implies that if some commutator identity is satisfied by G, then this identity is also satisfied by the homogeneous elements of $L(G)$. As the nilpotency and solubility identities are multilinear, this implies that the nilpotency class and the derived length of $L(G)$ do not exceed the nilpotency class and derived length of G, respectively.

For the last part, it is is enough to notice that

$$|L(G)| = \prod_{i=1}^{s} |K_i : K_{i+1}| = |G|,$$

by Lagrange's theorem, since $G = K_1 \geq \cdots \geq K_{s+1} = 1$. □

We remark that it may be possible for $L(G)$ to have nilpotency class strictly smaller than the one of G. Consider, for instance, $G = D \times C$, where $D = \langle a, b \,|\, a^4 = b^2 = (ab)^2 = 1\rangle$ is the dihedral group of order 8 and $C = \langle c \rangle$ is a cyclic group of order 2. Take the strongly central series given by $K_1 = G$, $K_2 = \langle a^2 \rangle \times \langle c \rangle = D' \times C$, $K_3 = \langle a^2 \rangle = D'$, $K_4 = 1$. The associated Lie ring $L(G)$ is abelian, since $[K_1, K_1] \leq K_3$ and K_2 and K_3 are both central in G, while G has nilpotency class 2.

It is also important to remark that, when we consider the Lie ring associated to the γ-series of a group G, then both G and $L(G)$ have the same nilpotency class. Assume that G has class c. Since we know that the class of $L(G)$ does not exceed the class of G, it suffices to show that there exists a Lie commutator of weight $c-1$ in $L(G)$ which is nonzero. Such a commutator w certainly exists in G, as G has class c. So, if we consider the image of w in $L(G)$, $\overline{w} = w\gamma_c(G)$, this cannot be zero, because otherwise we would have $w = 1$ in G too.

Next, we want to produce a map from $Aut(G)$ to $Aut(L(G))$ by considering the action of $\phi \in Aut(G)$ in each of the factors K_i/K_{i+1}.

Proposition 3.3.5. *If in the strongly central series $G = K_1 \geq \cdots \geq K_{s+1} = 1$ every term K_i is ϕ-invariant, then the automorphism ϕ of G induces an automorphism of the Lie ring $L(G)$ by its action on the factor groups K_i/K_{i+1}. If, moreover, G is finite and the order of ϕ is coprime to the order of G, then ϕ acts faithfully on $L(G)$ and $|C_{L(G)}(\phi)| = |C_G(\phi)|$.*

Proof. Let us denote by ϕ the automorphism induced by ϕ on each factor K_i/K_{i+1} of the series. For homogeneous elements $\bar{x} = xK_{i+1}$ in K_i/K_{i+1} and $\bar{y} = yK_{j+1}$ in K_j/K_{j+1}, we have

$$[\bar{x}, \bar{y}]^{\phi} = [x, y]^{\phi} K_{i+j+1} = [x^{\phi}, y^{\phi}] K_{i+j+1} = [\bar{x}^{\phi}, \bar{x}^{\phi}].$$

Extending the action of the induced automorphism ϕ to $L(G)$ from the abelian groups K_i/K_{i+1} by linearity yields an automorphism of the associated Lie ring.

Suppose now that $(|\phi|, |G|) = 1$. In the sense of Lemma 2.2.3, we have that the number of fixed points of the induced automorphism on each quotient satisfies

$$|C_{K_i/K_{i+1}}(\phi)| = |C_{K_i}(\phi)K_{i+1}/K_{i+1}| = |C_{K_i}(\phi)/C_{K_i}(\phi) \cap K_{i+1}|.$$

As $C_{K_i}(\phi) = C_G(\phi) \cap K_i$ and $C_{K_i}(\phi) \cap K_{i+1} = C_G(\phi) \cap K_i \cap K_{i+1} = C_G(\phi) \cap K_{i+1}$, we obtain $|C_{K_i/K_{i+1}}(\phi)| = [C_G(\phi) \cap K_i : C_G(\phi) \cap K_{i+1}]$. The number of fixed points of ϕ in $L(G)$ equals the product of $|C_{K_i/K_{i+1}}(\phi)|$ for each factor group K_i/K_{i+1}, and this product clearly yields $|C_G(\phi)|$, again by the Lagrange Theorem. □

It is important to say that a non-trivial automorphism ϕ of G may induce a trivial automorphism of $L(G)$, inner automorphisms figuring as examples. In general, it may not be true that the induced automorphism will have the same order as ϕ, and the number of fixed points might even increase. For a non-trivial example, consider $Q = \langle a, b \,|\, aba = b, bab = a\rangle$, the quaternion group again. The map $a \mapsto b$, $b \mapsto a$ extends to an automorphism ϕ of Q which fixes elementwise the subgroup $\langle a^2 \rangle = Q'$. The induced automorphism on the Lie ring associated to the γ-series fixes the element abQ', because $(ab)^{\phi}Q' = baQ' = abQ'$, as Q/Q' is abelian. Hence, as ϕ acts trivially on Q'/Q'', the induced automorphism has 4 fixed points, instead of just 2, and it is non-trivial on $L(Q)$. That explains why the previous result is important: it allows $L(G)$ to inherit properties related to centralizers of automorphisms in G, under suitable conditions.

3.4 Solubility of Graded Lie Rings

Let us start with the following result on complex numbers.

Definition 3.4.1. *Let S be a subset of complex numbers. A linear ordering $<$ on S will be called good if there are no $x, y \in S$ such that $xy \in S$ and $x < xy < y$.*

Lemma 3.4.2. *The complex numbers admit a good ordering.*

Proof. We first claim that the unit circle $S^1 = \{z \in \mathbb{C}, |\, |z| = 1\}$ admits a good ordering. Indeed, given $z = e^{\theta i}$, $z' = e^{\theta' i}$, with $0 \leq \theta, \theta' < 2\pi$, define $z < z'$ if $\theta < \theta'$. Then, $zz' > z, z'$ if $\theta + \theta' < 2\pi$ and $zz' < z, z'$ otherwise, so $<$ is a good ordering.

Now, if $x, y \in \mathbb{C}$, we say that $x > y$ provided that the argument of x is greater than the argument of y, and we make 0 the smallest element in $\mathbb{C}$. The argument above shows that “$>$” is indeed a good ordering. □

In general, a group does not need to admit a good ordering. For example, the Klein 4-group $G = C_2 \times C_2$ does not admit a good ordering. Indeed, regardless of the order considered, we can always name the non-trivial elements of G a, ab

and b, in increasing order. So, no order in G is good. The reason behind this is that every group that can be faithfully represented into $\mathbb{C}^*$ admits a good ordering.

The following result appears as Proposition 2.4 in [27]. The proof is inspired by Kreknin's method as it appears in [19].

Proposition 3.4.3. *Let L be an S-graded Lie ring, where S is a finite set of complex numbers not including 1. Then L is soluble. Moreover, if $|S| = d$, then $dl(L) \leq 2^{d-1} - 1$.*

Proof. Choose a good ordering on S, and label the elements of S such that $x_1 < x_2 < \cdots < x_d$. Then $L = \oplus_{i=1}^{d} L_{x_i}$, and we are going to denote $L_{x_i} = L_i$ for the sake of simplicity. For $1 \leq k \leq d$ let H_k be the subring generated by $L_{k+1}, \cdots, L_d$. Note that $H_d = 0$. We claim that

(i) $L^{(2^{k-1})} \cap L_k \subseteq H_k$, and

(ii) $L^{(2^k-1)} \subseteq H_k$,

both for $1 \leq k \leq d$. We prove (i) and (ii) simultaneously, by induction on k. In the case $k = 1$, we have to show that $L' \subseteq H_1$, and it is sufficient to show that $[L_i, L_j] \subseteq H_1$, for every i, j. If the product $x_i x_j \neq x_1$, certainly $[L_i, L_j] \subseteq H_1$, so we must consider the case $x_i x_j = x_1$. Since $1 \notin S$, then none of the numbers x_i, x_j can equal x_1. In this case, both $i, j \geq 2$, and $[L_i, L_j] \subseteq H_1$, completing the proof in this case.

Suppose now $k > 1$. The induction hypothesis for (ii) yields

$$L^{(2^{k-1}-1)} \subseteq H_{k-1}$$

so that

$$[L^{(2^{k-1}-1)}, L^{(2^{k-1}-1)}] = L^{(2^{k-1})} \subseteq (H_{k-1})'.$$

To prove (i), it is sufficient to show that $(H_{k-1})' \cap L_k \subseteq H_k$.

Let $a, b \in H_{k-1}$ be homogeneous elements with $h = [a, b] \in L_k$. We have to show that $h \in H_k$. We may assume that $b = [b_1, \cdots, b_m]$, where $m \geq 1$ and $b_i \in$

L_{n_i} for some $n_1, \cdots, n_m \geq k$. Then, h can be written as a $\mathbb{Z}$-linear combination of simple commutators of the form $[a, b_{\sigma(1)}, \cdots, b_{\sigma(m)}]$ for permutations $\sigma \in S_m$, by Lemma 3.1.2(ii). It therefore suffices to show that all these elements lie in H_k. We may assume that σ is the identity permutation, as the type of the permutation $\sigma \in S_m$ will not interfere in the proof, and consider the element $y = [c, b_m]$, where $c = [a, b_1, \cdots, b_{m-1}]$. Suppose $c \in L_i$ and $b_m \in L_j$, so that $j = n_m$. Recall here that h belongs to L_k, and as a is a homogeneous element, in particular $y \in L_k$, and then $x_i x_j = x_k$. We know that $j \geq k$, and since $x_i \neq 1$ we have $j > k$. It is also clear that $i \neq k$, for the same reason, and we claim that $i > k$. For otherwise, if $i < k$, then we would have $x_i < x_k = x_i x_j < x_j$, contradicting the assumption that $<$ is a good ordering. Then, $y \in [L_i, L_j] \subseteq H_k$ and in particular $h \in H_k$. This proves (i).

To prove (ii), consider $M = L^{(2^{k-1})}$ as an S-graded Lie ring. Apply (ii) for M with $k - 1$ to obtain

$$M^{(2^{k-1}-1)} \subset \langle M \cap L_k, \cdots, M \cap L_d \rangle \subset \langle M \cap L_k, H_k \rangle.$$

Now, condition (i) for L and k yields $M \cap L_k \subseteq H_k$. Therefore

$$M^{(2^{k-1}-1)} \subseteq H_k.$$

Since $L^{(2^k-1)} = M^{(2^{k-1}-1)}$, the result follows. □

An automorphism of a finite rank Lie ring is called *semisimple* if it can be represented by a diagonal matrix over some suitable integral extension of the integers. In this sense, consider the following corollary.

Corollary 3.4.4. *Let L be an r-generated Lie ring of finite which admits a fixed-point-free semisimple automorphism ϕ. Then L is soluble, with derived length at most $2^{r-1} - 1$.*

Proof. We may just consider the grading induced by the set S of eigenvalues of ϕ, on a suitable extension $\overline{L} = L \otimes_{\mathbb{Z}} R$ of L, where R is an integral extension of $\mathbb{Z}$ containing all eigenvalues of ϕ. We regard S as a subset of $\mathbb{C}$. Since ϕ acts fixed-point-freely on L, $1 \notin S$. Also, the number of distinct eigenvalues of ϕ is no greater than r, because ϕ can be represented as an $r \times r$ matrix over $\mathbb{C}$. We can apply Proposition 3.4.3 to $\overline{L}$, and this ensures that $L \subseteq \overline{L}$ has the desired properties. □

This result ends our general discussion about Lie rings.

Chapter 4

Powerful p-groups

This section is devoted to the study of a class of p-groups that generalize in some sense abelian groups, the so called *powerful* p-groups. They were formally introduced in [20], but the authors of this paper say that this concept was foreseen by many other authors in previous years. The reason for this anticipation of results may be the fact that, as it will be illustrated by Theorem 4.1.15, powerful p-groups arise quite naturally in the general theory of p-groups. When we deal with certain classes of p-groups, powerful p-groups also happen often as important subgroups, and this fact illustrates their great importance and generality. We are going to see that, besides sharing various linear properties with abelian groups, the powerful p-groups also are very useful to bound the derived length of an arbitrary p-group, as showed in Theorem 4.1.15. The results in this section are all proved for the case where p is an odd prime, for the sake of simplicity, despite being true also in the case $p = 2$. Some proofs apply to both cases and some others have to be slightly modified in order to hold in the case of the even prime. The omitted proofs can all be found in Section 4 of [20]. All groups considered in this chapter are finite p-groups.

4.1 General Properties

We say that a p-group G is *powerful* when G/G^p is abelian, if p is odd, and G/G^4 is abelian, when $p = 2$. The reason for requiring the quotient over G^4 to be abelian on the latter case, and not over G^2, is that G/G^2 has exponent 2 and thus is abelian, regardless of the group G we are considering. Note that this definition is equivalent to saying that $G' \leq G^p$, in the odd case, and $G' \leq G^4$, in the even one, because G/G^p or G/G^4 abelian implies that, for every $x, y \in G$ the commutator $[x, y]$ belongs to G^p or to G^4, respectively. This means that the subgroup generated by the commutators, G', is contained in G^p or G^4, according to the case. There is a similar notion involving a subgroup H of G: we say that H is *powerfully embedded* in G when $[G, H] \leq H^p$, and we are going to write $H\,p.e\,G$ for short. Then, G is powerful if and only if it is powerfully embedded in itself, and a powerfully embedded subgroup H of G must be normal and powerful, since $[G, H] \leq H^p \leq H$ and $[H, H] \leq [H, G] \leq H^p$.

First of all, we consider some non-trivial examples of powerful p-groups, besides the abelian ones. The quaternion group $Q = \langle r, s \,|\, r^2 = s^2, r^s = r^{-1}\rangle$ is an example of a 2-group that is *not* powerful, as $Q^4 = 1$ and $Q' = \langle z\rangle$ has order 2. But if we consider the direct product $G = Q \times C_8$ of Q and $C_8 = \langle t \,|\, t^8 = 1\rangle$ and the quotient over the subgroup $K = \langle(z, t^4)\rangle$, we obtain that $\overline{G} = G/K$ is powerful, as $\overline{G}' = \langle\overline{(z, 1)}\rangle = \langle\overline{(1, t^4)} = \overline{G}^4$. As a by-product, this shows that the property of being powerful is not inherited by subgroups, since $\overline{G}$ admits as a subgroup one copy of Q. Also, for p odd we have that the group $H = \langle r, s \,|\, r^{p^2} = s^p = 1, r^s = r^{p+1}\rangle$, isomorphic to $C_{p^2} \rtimes C_p$, is powerful. Here, $H' = \langle[r, s]\rangle = \langle r^p\rangle = H^p$.

Even though this feature does not pass to subgroups, it is clear that being powerful is a property that passes to quotients and direct products. Let us start with a criterion to verify whether a normal subgroup K of G is powerfully embedded in G.

Lemma 4.1.1. *Let K be a normal subgroup of G. Then K is powerfully embedded in G whenever $[K, G] \leq K^p[K, G, G]$.*

Proof. Note that both K^p and $[K, G, G]$ are normal in G. Taking the commutator subgroup with G we get that $[K, G, G] \leq [K^p, G][K, G, G, G]$. In particular, we have

$$[K, G] \leq K^p[K, G, G] \leq K^p[K^p, G][K, G, G, G] \leq K^p[K, G, G, G],$$

since K^p is normal in G. This shows that $[K, G] \leq K^p[K, G, G, G]$. By repeating the same argument we have that $[K, G] \leq K^p[K, \underbrace{G, \cdots, G}_{s}]$ for any $s \geq 2$. We conclude that $[K, G] \leq K^p$, as G is nilpotent. □

The lemma above says that we may prove that a normal subgroup K is powerfully embedded in G by proving that the image of K modulo $[K, G, G]$ is powerfully embedded in the quotient $G/[K, G, G]$ without loss of generality.

Lemma 4.1.2. *If M and N are normal subgroups of G such that $[M, N, G, G] = 1$, then $[M^p, N] \leq [M, N]^p$.*

Proof. Applying the standard commutator identity $[ab, c] = [a, c]^b[b, c]$ several times, we have

$$\begin{aligned}[m^p, n] =&[m^{p-1}, n]^m[m, n] = ([m^{p-2}, n]^m[m, n])^m[m, n] = \cdots \\ =&(\cdots(([m, n]^m[m, n])^m[m, n])^m \cdots [m, n])^m[m, n],\end{aligned}$$

for any $m \in M$ and $n \in N$.

Since all commutators starting with $[m, n]$ and having weight greater than or equal to 4 vanish and $[M, N, G] \leq Z(G)$, in the identity above, we can collect the remaining commutators to get

$$[m^p, n] = [m, n]^p[m, n, m]^{\binom{p}{2}} \in [M, N]^p, \tag{4.1.1}$$

as p is odd. By working in the quotient $G/[M, N]^p$, we need to show that $[M^p, N] = 1$. In view of Lemma 4.1.1 we know that $[m^p, n] = 1$ for any $m \in M$ and $n \in N$. This means that $m^p \in C_G(N)$, and, in particular, that $M^p \leq C_G(N)$. Thus $[M^p, N] = 1$, as required. □

Despite the property of being powerful is not inherited by subgroups in general, the next result shows that many subgroups actually do inherit this property.

Lemma 4.1.3. *Let G be a finite p-group and M and N subgroups of G. If M p.e G and N p.e. G, then M^p p.e. G, $[M, N]$ p.e. G and MN p.e. G.*

Proof. As for M^p, we need to show that $[M^p, G] \leq (M^p)^p$. In view of Lemma 4.1.1, we may assume that $[M^p, G, G] = 1$ and prove that the image of M^p

is powerfully embedded in the quotient $G/[M^p, G, G]$. Since $[M, G] \leq M^p$, it follows that $[M, G, G, G] \leq [M^p, G, G] = 1$. It follows from Lemma 4.1.2 that $[M^p, G] \leq [M, G]^p \leq (M^p)^p$.

Next, we consider the case of $[M, N]$ and assume that $[M, N, G, G] = 1$. We know that $[M, N, G] \leq [M, [N, G]][M, G, N]$, by the Three Subgroup Lemma. As the subgroups M and N are powerfully embedded in G, we have

$$[M, [N, G]][M, G, N] \leq [M, N^p][M^p, N] \leq [M, N]^p,$$

where the last inclusion holds because of Lemma 4.1.2. This yields $[M, N, G] \leq [M, N]^p$, as desired.

In order to show that $MN\, p.e.\, G$, let us observe that $[MN, G] = [M, G][N, G]$, since both M and N are normal in G. As both subgroups M and N are powerfully embedded in G, we have $[M, G][N, G] \leq M^p N^p \leq (MN)^p$. This completes the proof. □

As an application of Lemma 4.1.3, we get the following corollary.

Corollary 4.1.4. *Let G be a powerful p-group. Then the subgroups $\gamma_i(G)$, $G^{(i)}$, G^{p^i} are powerfully embedded in G. Moreover, the Frattini subgroup $\Phi(G) = G^p$ is powerfully embedded in G too.*

Proof. The assertion about $\gamma_i(G)$, $G^{(i)}$ and G^{p^i} follows from Lemma 4.1.3. Moreover, since $[G, G] \leq G^p$ and $\Phi(G) = [G, G]G^p = G^p$, the last assertion holds in the case when p is odd (note that, when $p = 2$, $G^4 \leq G^2 = \Phi(G)$), and the proof is complete. □

One of the properties that powerful p-groups share with abelian groups is the content of the following proposition.

Proposition 4.1.5. *Let G be a powerful p-group with p odd. Then, for every $i \geq 1$, we have $G^{p^i} = \{x^{p^i} \mid x \in G\}$.*

Proof. The proof will be by induction on i. Assume $i = 1$, we argue by induction on $|G|$. First of all, note that the factor group $\overline{G} = G/(G^p)^p$ is nilpotent of class at most 2, since $[G, G, G] \leq [G^p, G]$, where we use that G is powerful, and $[G^p, G] \leq (G^p)^p$, as $G^p\, p.e.\, G$. Hence, for every $\overline{u}, \overline{v} \in \overline{G}$, $(\overline{uv})^p = \overline{u}^p \overline{v}^p [\overline{v}, \overline{u}]^{\binom{p}{2}}$. As $[\overline{G}, \overline{G}]$ is central in $\overline{G}$ and $\binom{p}{2} = \dfrac{p(p-1)}{2}$ with $p \geq 3$, then

$$\overline{u}^p \overline{v}^p = (\overline{uv})^p ([\overline{u}, \overline{v}]^{\frac{p-1}{2}})^p = (\overline{uv}[\overline{u}, \overline{v}]^{\frac{p-1}{2}})^p.$$

Thus the result holds in $\overline{G}$, which means that for every $u \in G^p$ there exists $v \in G$ such that $u = v^p w$, for some $w \in (G^p)^p$. Put $H = \langle v, G^p \rangle$. Since $H \leq \langle v, \Phi(G) \rangle$, the subgroup H must be proper, otherwise G would be cyclic and the result is trivial in this case. Then, by induction, we get $u = v^p w = t^p$ for some t in H. Since $u \in H^p$, this concludes the proof in the case $i = 1$.

Now, Corollary 4.1.4 shows that G^{p^i} is itself a powerful p-group, for every $i \geq 1$. Assuming $i > 1$, we have

$$\begin{aligned}
G^{p^i} &= \langle x^{p^i} \mid x \in G \rangle \\
&= \langle (x^{p^{i-1}})^p \mid x \in G \rangle = (G^{p^{i-1}})^p && \text{(by inductive hypothesis)} \\
&= \{ g^p \mid g \in G^{p^{i-1}} \} && \text{(by the case } i = 1) \\
&= \{ (x^{p^{i-1}})^p \mid x \in G \} && \text{(by inductive hypothesis)} \\
&= \{ x^{p^i} \mid x \in G \}
\end{aligned}$$

as desired. □

Corollary 4.1.6. *Let G be a powerful p-group, with p odd. Then, for every $i, j \geq 0$ we have $(G^{p^i})^{p^j} = G^{p^{i+j}}$.*

Proof. We have

$$\begin{aligned}
(G^{p^i})^{p^j} &= \{ h^{p^j} \mid h \in G^{p^i} \} && \text{(because } G^{p^i} \text{ is powerful, by Corollary 4.1.4,)} \\
&= \{ (g^{p^i})^{p^j} \mid g \in G \} && \text{(by Proposition 4.1.5,)} \\
&= \{ g^{p^{i+j}} \mid g \in G \} \\
&= G^{p^{i+j}} && \text{(by Proposition 4.1.5.)}
\end{aligned}$$

□

In [28], Shalev shows that commutators and p-powers have a nice interaction in powerful p-groups, as it is stated in the following lemma.

Lemma 4.1.7. *If N, M are normal subgroups of G and are powerfully embedded in G, then $[N^{p^i}, M^{p^j}] = [N,M]^{p^{i+j}}$, for all $i, j \geq 0$.*

Proof. First we show that $[N^p, M] = [N,M]^p$ and then prove the lemma by induction on $i+j$. We claim that $[x^p, y] \equiv [x,y]^p \ (mod\,[N,M]^{p^2})$ for all $x \in N, y \in M$. Let $K = \langle [x,y], x\rangle$. We know, by the Hall-Petrescu identity (appendix A in [5]), that

$$[x,y]^p = (x^{-1}x^y)^p = x^{-p}(x^y)^p c_2^{\binom{p}{2}} \cdots c_{p-1}^{\binom{p}{p-1}} c_p = [x^p, y] c_2^{\binom{p}{2}} \cdots c_{p-1}^{\binom{p}{p-1}} c_p,$$

where $c_i \in \gamma_i(\langle x^{-1}, x^y\rangle)$ and for $2 \leq i \leq p$. Since $K = \langle x^{-1}, x^y\rangle$, the equality above shows that $[x^p, y] \equiv [x,y]^p \ (mod\, H)$, where $H = (K')^p\gamma_p(K)$. Let us show that $H \leq [N,M]^{p^2}$.

Note that, as both M and N are normal subgroups of G, we can regard $[x,y]$ as an element of M or N, according to our necessity. Consider the subgroup $\gamma_p(K) = \langle [c_1, \cdots, c_s], s \geq p, c_j \in \{x, [x,y]\}\rangle$. In order to consider the non-trivial commutators that generate $\gamma_p(K)$, we may assume that $c_1 \neq c_2$ in any commutator of the form $w = [c_1, \cdots, c_s]$, and certainly $[c_1, c_2] \in [N, M, N]$. Thus any commutator of the form $[c_1, \cdots, c_s]$ belongs to $[N, M, N,_{s-2} N] \leq [N, M, N,_{p-2}, N]$, as $s-2 \geq p-2$, and we get $\gamma_p(K) \leq [N, M,_{p-1} N]$. We also have $(K')^p \leq [N, M, N]^p$ by using a similar idea. Lemma 4.1.4 tells us that $[N,M]\, p.e.\, G$, thus $[N,M]\, p.e. N$ too. It follows that

$$\begin{aligned}
[N, M,_{p-1} N] &\leq [N, M, N, N] && \text{(since p is odd and then $p-1 \geq 2$,)}\\
&\leq [[N,M]^p, N] && \text{(because $[N,M]\, p.e.\, N$,)}\\
&\leq ([N,M]^p)^p && \text{(because $[N,M]^p\, p.e.\, N$,)}\\
&= [N,M]^{p^2} && \text{(because $[N,M]$ is powerful.)}
\end{aligned}$$

We also have

$$[N, M, N]^p \leq ([N,M]^p)^p = [N,M]^{p^2}.$$

This shows that $H = (K')^p\gamma_p(K) \leq [N,M]^{p^2}$, establishing the congruence $[x^p, y] \equiv [x,y]^p \,(mod\,[N,M]^{p^2})$. Applying this congruence, we get that

$$[N^p, M][N,M]^{p^2} = [N,M]^p[N,M]^{p^2} = [N,M]^p \ \ (*).$$

In particular, we have $[N^p, M] \leq [N, M]^p$. As $[N, M]^p$ is also powerful, we know that its Frattini subgroup equals $([N, M]^p)^p = [N, M]^{p^2}$. The equality in $(*)$ guarantees that $[N^p, M]$ equals $[N, M]^p$ modulo $\Phi([N, M]^p)$, which gives us the required equality $[N^p, M] = [N, M]^p$.

The general case will follow by induction on $i+j$. Note that the case $i+j = 1$ has just been proved, since the argument with $i = 0$ and $j = 1$ is symmetric to the previous one. Suppose that $i+j > 1$ and that the result holds for $i+j-1$. We can assume that $i \geq 1$, otherwise we exchange the roles of N and M. Observe that

$$\begin{aligned} [N^{p^i}, M^{p^j}] &= [(N^{p^{i-1}})^p, M^{p^j}] & \\ &= [N^{p^{i-1}}, M^{p^j}]^p & \text{(by the case } i+j=1,) \\ &= ([N, M]^{p^{i+j-1}})^p & \text{(by the inductive hypothesis,)} \\ &= [N, M]^{p^{i+j}} & \text{(as } [N, M] \text{ is powerful.)} \end{aligned}$$

This completes the proof. □

As immediate corollaries of the previous lemma, we get the following results.

Corollary 4.1.8. *If G is powerful, then, for every $i, j \geq 0$, we have*

(i) $(G^{p^i})^{(j)} = (G^{(j)})^{p^{i2^j}}$

(ii) $\gamma_j(G^{p^i}) = \gamma_j(G)^{p^{ij}}$

(iii) $(G^{p^i})^{(j)} \leq G^{p^{i2^j+2^j-1}}$

Proof. Items (i) and (ii) follow directly from Lemma 4.1.7. The inclusion $G^{(j)} \leq G^{p^{2^j-1}}$ can be proved by induction on j, and this inclusion combined with (i) and Corollary 4.1.6 gives (iii). □

If G is a p-group of exponent p^e, we call the series $G > G^p > \cdots > G^{p^e} = 1$ the *agemo series* of G. Note that all inclusions are proper, since $G^{p^{i+1}}$ is a subgroup of $(G^{p^i})^p \leq \Phi(G^{p^i})$. This series does not need to be central in general, but when G is powerful, the following holds.

Corollary 4.1.9. *If G is a powerful group of exponent p^e, then the series*

$$G \rhd G^p \rhd \cdots \rhd G^{p^{e-1}} \rhd 1$$

is central. In particular, $dl(G) \leq e$.

Proof. The commutator $[G^{p^i}, G^{p^j}]$ equals $[G, G]^{p^{i+j}}$. Since $[G, G] \leq G^p$ we have $[G, G]^{p^{i+j}} \leq (G^p)^{p^{i+j}} = G^{p^{i+j+1}}$. □

In view of Corollary 4.1.9, the factors of the agemo series all have exponent p and are elementary abelian. The next proposition gives information about how the ranks of these elementary abelian factors behave.

Proposition 4.1.10. *If G is a powerful p-group, then the following inequalities hold*

$$p^d = [G : G^p] \geq [G^p : G^{p^2}] \geq \cdots \geq [G^{p^{e-1}} : G^{p^e}], \tag{4.1.2}$$

where $d = d(G)$ and $p^e = exp(G)$.

Proof. Consider the function $f_i : G^{p^i}/G^{p^{i+1}} \to G^{p^{i+1}}/G^{p^{i+2}}$, given by $\overline{x} \mapsto \overline{x}^p$. Here we use the same bar notation to denote elements in both quotients. First of all, it is a well defined function because if $x, y \in G^{p^i}$ and $xG^{p^{i+1}} = yG^{p^{i+1}}$, then $xy^{-1} \in G^{p^{i+1}}$ and the Hall-Petrescu formula applied to this product gives

$$(xy^{-1})^p = x^p y^{-p} c_2^{\binom{p}{2}} \cdots c_{p-1}^{\binom{p}{p-1}} c_p,$$

which implies

$$x^p y^{-p} = (xy^{-1})^p (c_2^{\binom{p}{2}} \cdots c_{p-1}^{\binom{p}{p-1}} c_p)^{-1},$$

where each c_i is an element of $\gamma_i(\langle x, y^{-1} \rangle)$. As $xy^{-1} \in G^{p^{i+1}}$, $(xy^{-1})^p \in G^{p^{i+2}}$. We also know that $[G^{p^i}, G^{p^i}] = (G')^{p^{2i}} \leq G^{p^{2i+1}}$ for every $i \geq 1$, which shows that each c_i belongs to $G^{p^{2i}}$. As a consequence, $x^p y^{-p} \in G^{p^{i+2}}$ and we have $x^p \equiv y^p \,(mod\, G^{p^{i+2}})$. The fact that f_i is a homomorphism also follows from the calculation above. The surjectivity of f_i follows from the fact that $G^{p^{i+1}}$ is exactly equal to the set of p-th powers of the elements in G^{p^i}. Moreover, as $\Phi(G) = G^p$, we get $p^d = [G : G^p]$. This completes the proof. □

We include the following important result here because of its strength and for the sake of completeness, even though it will not be needed in this dissertation. It states one of the most important features of powerful p-groups, and is Theorem 1.12 of [20].

Theorem 4.1.11. *Suppose that G is a powerful p-group generated by d elements. Then every subgroup of G can be generated by d elements.*

Our main result of this section justifies why powerful p-groups play an important role in the study of finite p-groups. We want to show that every finite p-group G with rank $rk(G) = r$ has a powerful characteristic subgroup V, such that the index of V in G is bounded by a function of p and r and the derived length of G/V is bounded by a function of r only.

In order to do so, we need some preparatory work. Given a finite p-group G and a positive integer r, let us denote by $V(G, r)$ the intersection of the kernels of all homomorphisms from G into $GL_r(\mathbb{F}_p)$. Since the image of any homomorphism of a p-group G into $GL_r(\mathbb{F}_p)$ is a p-group, and every p-subgroup of $GL_r(\mathbb{F}_p)$ is conjugate to a subgroup of the lower uni-triangular group $U_r(\mathbb{F}_p)$, that is a Sylow p-subgroup of $GL_r(\mathbb{F}_p)$, we could equally well define $V(G, r)$ as the intersection of all the kernels of homomorphisms of G into $U_r(\mathbb{F}_p)$. Note that if an element g of G belongs to $V(G, r)$, then g acts trivially on every factor of G which is elementary abelian. This happens since the rank of G equals r, and every elementary abelian quotient of G is at most r-generated.

For $r \in \mathbb{N}$, we consider here $\lambda(r) = \lceil \log_2(r) \rceil$. The proof of the following lemma can be found in [5, Lemma 2.11]. We are going to use it in the proof of Theorem 4.1.15.

Lemma 4.1.12.

(i) *The group $U_r(\mathbb{F}_p)$ has a series of normal subgroups of length $\lambda(r)$ such that every factor is elementary abelian.*

(ii) *If G is a finite p-group, then $G/V(G,r)$ has a series with these properties and length at most $\lambda(r)$.*

The following lemma is useful in proofs via contradiction, as it gives necessary conditions that hold in subgroups of a group that are not powerfully embedded.

Lemma 4.1.13. *Let G be a finite p-group, p an odd prime. If N and V are normal subgroups of G such that $N \leq V$ and N is not powerfully embedded in V, then there exists a normal subgroup J of G such that*

$$N^p \leq J < N^p[N,V], \qquad \textit{and} \qquad [N^p[N,V] : J] = p.$$

Proof. As $[N,V] \nleq N^p$, $N^p < N^p[N,V]$. Since N and V are normal in G, so are N^p and $[N,V]$. Now, G is a p-group, so induction on $|G|$ shows that there exists a subgroup J normal in G such that $N^p \leq J < N^p[N,V]$ and $[N^p[N,V] : J] = p$.□

Proposition 4.1.14. *Let G be a finite p-group, where p is odd, and let r be a positive integer. Put $V = V(G,r)$. If $N \lhd G$, $d(N) \leq r$ and $N \leq V$, then $N\,p.e.\,V$.*

Proof. The proof is by induction on $|N|$. Suppose that $[N,V] \nleq N^p$. Factoring out G by J as in Lemma 4.1.13, we may assume that $N^p = 1$ and $[N,V]$ has order p. Since N is not powerfully embedded in V, we cannot have $[N,V] = N$, otherwise $N^p \leq [N,V]$; therefore $[N,V]$ must be a proper subgroup of N. Now, G is a p-group, so there exists a normal subgroup M of G with $[N,V] \leq M < N$ and $[N : M] = p$. As $N \leq V$, we have $[N,N] \leq [N,V]$, and so $N/[N,V]$ is elementary abelian. The same observation holds for $M/[N,V]$, because $M^p \leq N^p = 1$ and $[M,M] \leq [N,V]$. Taking into account that $[N : M] = p$, we have $d(M/[N,V]) = d(N/[N,V]) - 1 \leq r - 1$. Since $[N,V]$ is cyclic, it follows that $d(M) \leq r$. Hence, M satisfies the hypothesis of the proposition and, by the inductive hypothesis, we obtain that $[M,V] \leq M^p = 1$. Thus, being M central in V, it is also central in N, and as N/M is cyclic it follows that N is abelian. Then N is an $\mathbb{F}_p$-vector space of dimension at most r, so the conjugation action of V on N must be trivial, giving $[N,V] = 1$, in contradiction to our initial hypothesis. □

Finally, we are ready to deal with the proof of the main result of this section.

Theorem 4.1.15. *Let G be a finite p-group of rank r. Then G has a powerful characteristic subgroup V of index at most $p^{r\lambda(r)}$, if p is odd. Moreover, the derived length of G/V is at most $\lambda(r)$.*

Proof. Put $V = V(G,r)$. Taking $N = V$ in Proposition 4.1.14, we get that N is powerfully embedded in itself, hence it is powerful. By Lemma 4.1.12, there exists a series of normal subgroups running from G to V, of length $\lambda(r)$, with each factor elementary abelian. This fact yields that $dl(G/V)$ is at most $\lambda(r)$. Also, since G has rank r, each of these factors has order at most p^r, so $[G:V] \leq p^{r\lambda(r)}$. □

The previous result will be very useful in an certain kind of context. Indeed, for example, if we have certain conditions on a p-group P of rank r and want to bound its derived length, then assuming that the hypothesis on P can be inherited by its characteristic subgroups, we can work with $V(G,r)$ instead of P with the additional property of being powerful, since the derived length of G/V is already bounded in terms of r only.

4.2 Uniform Groups

In this section we will introduce a particular subclass of powerful p-groups, the so-called *uniform groups*. In order to do it, we begin with the following result illustrating a specific phenomenon for powerful finite p-groups, of given exponent and rank.

Lemma 4.2.1. *Let G be a powerful p-group with exponent p^e and rank r. Then, G admits a series of characteristic subgroups*

$$G \geq G^{p^{i_1}} \geq G^{p^{i_2}} \geq \cdots \geq G^{p^{i_s}} = 1,$$

(where $i_0 = 0$) such that every factor $M_l = G^{p^{i_l}}/G^{p^{i_{l+1}}}$, $l = 0, \dots, s-1$, with exponent p^{k_l} satisfies

$$[M_l : M_l^p] = [M_l^p : M_l^{p^2}] = \cdots = [M_l^{p^{k_l-1}} : M_l^{p^{k_l}}].$$

Proof. Consider

$$G \geq G^p \geq \cdots \geq G^{p^{e-1}} \geq 1$$

the agemo series of G. As the orders of the factors are non-increasing by Proposition 4.1.10, let i_1 be the first index such that $G^{p^{i_1}}/G^{p^{i_1+1}}$ is non-isomorphic to G/G^p. Then, the claimed series will begin with G and will be followed by $G^{p^{i_1}}$, as the quotient $G/G^{p^{i_1}}$ satisfies the conditions of the statement, by construction. We find i_2 as the index such that $G^{p^{i_2}}/G^{p^{i_2+1}}$ is non-isomorphic to $G^{p^{i_1}}/G^{p^{i_1+1}}$, and $G^{p^{i_2}}$ will be the next term in the series we want to construct. Notice that the agemo series of $G/G^{p^{i_1}}$ produces quotients having rank r, while on $G^{p^{i_1}}/G^{p^{i_2}}$ the agemo quotients have rank at most $r-1$. This construction produces indices $0 < i_1 < \cdots < i_s = e$ such that the sections M_l all satisfy the conditions of the statement, for $l = 0, \dots, s-1$. Also, we have $s \leq r$. □

In view of Proposition 4.1.10 and Lemma 4.2.1, we see that an arbitrary powerful p-group can be "divided" into sections where the inequalities of equation (4.1.2) all become equalities. This property inspires the following definition.

Definition 4.2.2. *Let G be a powerful finite p-group of exponent p^e. We say that G is uniformly powerful, or uniform for short, if we have*

$$[G : G^p] = [G^p : G^{p^2}] = \cdots = [G^{p^{e-1}} : G^{p^e}].$$

Since the inequalities in Proposition 4.1.2 all become equalities, the operation of taking p^j-th powers yields an isomorphism from the factor $G^{p^i}/G^{p^{i+1}}$ to the factor $G^{p^{i+j}}/G^{p^{i+j+1}}$, as long as $i+j+1 \leq e$. Also, if $d(G) = d$ and $exp(G) = p^e$, then every quotient $G^{p^i}/G^{p^{i+1}}$ is isomorphic to $\underbrace{C_p \times \cdots \times C_p}_{d}$, where $0 \leq i \leq e-1$. In particular, the order of G equals p^{de}.

The next result is strictly about uniform p-groups.

Lemma 4.2.3. *Let G be a uniform p-group of exponent p^e. Then $x^{p^i} \in G^{p^j}$ implies $x \in G^{p^{j-i}}$ for all $0 \leq i < j \leq e$*

Proof. Choose s such that $x \in G^{p^s} \backslash G^{p^{s+1}}$. It follows then that x^{p^i} belongs to $G^{p^{s+i}} \backslash G^{p^{s+i+1}}$ as long as $s+i < e$. If $s+i < e$, we have both $x^{p^i} \in G^{p^j}$ and $x^{p^i} \notin G^{p^{s+i+1}}$, whence $G^{p^{s+i+1}} < G^{p^j}$ and thus $s+i+1 > j$, giving $s \geq j-i$. It is clear that $x \in G^{p^s} \leq G^{p^{j-i}}$, as required. If $s+i \geq e$, then $s+i \geq j$, whence again $s \geq j-i$ and $x \in G^{p^s} \leq G^{p^{j-i}}$, as required. □

A natural question that arises, concerning Lemma 4.2.3, is the following: what would be an example of a powerful p-group that does not satisfy the statement of the lemma? We can answer with the group $G = \langle x, y \,|\, x^{p^2} = y^p = 1, [x, y] = x^p \rangle$. It is an example of a powerful p-group that is not uniform and for which the previous lemma fails to hold. Indeed, in G we have $G^p = \langle x^p \rangle$ and $G/G^p \cong C_p \times C_p$, while $G^{p^2} = 1$ and $G^p/G^{p^2} \cong C_p$. Taking the element y of G, for instance, we have that $y^p \in G^{p^2}$, but y does not belong to G^p, because otherwise y would be a power of x and G would be cyclic, by Theorem 2.1.6.

In the last paragraph of the previous chapter, we discussed a method concerning problems related to solubility in a p-group P of rank r that reduces the study to the class of powerful p-groups, by dealing with the subgroup $V(r, P)$ instead of P. In view of Lemma 4.2.1, we can reduce the problem even further, to the class of uniform groups, since bounding the derived length of each uniform section of $V(P, r)$ amounts to a bound for the derived length of $V(P, r)$ itself. Our main goal in this chapter is to bound the derived length of each uniform section of a powerful p-group using a special class of Lie rings that can be associated to uniform groups.

4.3 The Associated Lie Ring

Let us describe the construction of a finite Lie ring (or rather a collection of Lie rings) associated to a uniform p-group. Following [7, §1.2], we say that a finite abelian p-group G is *homocyclic* if G is isomorphic to a direct product of one or more pairwise isomorphic cyclic groups. In this case, if $G = \underbrace{C_{p^i} \times \cdots \times C_{p^i}}_{d}$, we say that G is of type $(\underbrace{p^i, \cdots, p^i}_{d})$.

Proposition 4.3.1. *Let G be a uniform group of exponent p^e with $d(G) = d$. Denote by L the quotient $G^{p^i}/G^{p^{2i}}$, for a fixed i, $0 < i \leq e$. Then we have:*

(i) *If $i \leq e/2$, the quotient $G^{p^i}/G^{p^{2i}}$ is a homocyclic group with exponent p^i and d generators.*

(ii) *The map $x \mapsto x^{p^i}$ induces a well defined epimorphism $q : G^{p^i}/G^{p^{2i}} \to G^{p^{2i}}/G^{p^{3i}}$. If, moreover, $i \leq e/3$, then q is in fact an isomorphism.*

(iii) *The commutator operation in G, which in this proposition will be denoted by $(x, y) = x^{-1}y^{-1}xy$, induces a well defined function c, mapping a pair of elements $xG^{p^{2i}}, yG^{p^{2i}}$ of $G^{p^i}/G^{p^{2i}}$ to the element $(x, y)G^{p^{3i}}$ of $G^{p^{2i}}/G^{p^{3i}}$. Using additive notation in the respective abelian groups, this function is in fact bilinear.*

(iv) *Suppose $i \leq e/4$ and regard L as an additive group. If we equip L with the bilinear operation defined as follows:*

$$[x, y] = q^{-1}(c(x, y)),$$

for every x, y in L, then L has the structure of a Lie $\mathbb{Z}/p^i\mathbb{Z}$-ring. Moreover, the quotient ring L/pL is commutative.

Proof. By Corollary 4.1.8, we know that $[G^{p^i}, G^{p^i}] = (G')^{p^{2i}} \leq (G^p)^{p^{2i}} \leq G^{p^{2i+1}} \leq G^{p^{2i}}$, since G is powerful. Thus $G^{p^i}/G^{p^{2i}}$ is abelian. Since here we are assuming $i \leq e/2$, the construction of such a quotient gives us $exp(L) = p^i$. Moreover, since G is uniform, all quotients of the series $G \rhd G^p \rhd \cdots \rhd G^{p^{e-1}} \rhd 1$ are elementary abelian of rank d, being isomorphic to the quotient $G/\Phi(G)$. Then, the factor $G^{p^i}/G^{p^{2i}}$ has order p^{id}. As $G/G^p = \langle \bar{a}_1, \cdots, \bar{a}_d \rangle$, and as $G^p/G^{p^2} = \langle \bar{b}_1, \cdots, \bar{b}_d \rangle$, we have that $G^p = \langle b_1, \cdots, b_d \rangle G^{p^2}$ and $G = \langle a_1, \cdots, a_d \rangle G^p$, giving $G = \langle a_1, \cdots, a_d \rangle \langle b_1, \cdots, b_d \rangle G^{p^2}$. As we can certainly write the b_i in terms of the a_j, by Theorem 2.1.6, it follows that $G = \langle a_1, \cdots, a_d \rangle G^{p^2}$, and the quotient G/G^{p^2} has d generators. In particular, as G^{p^s} is uniform for any $0 \leq s \leq e$, replacing G wiht G^{p^s} in this argument shows that if $s < t$, then G^{p^s}/G^{p^t} is d-generated. In particular, $G^{p^i}/G^{p^{2i}}$ is also d-generated. Since it is abelian of exponent p^i, d-generated and of order p^{id}, it must be homocyclic, i.e., $G^{p^i}/G^{p^{2i}} \cong \underbrace{C_{p^i} \times \cdots \times C_{p^i}}_{d}$. This proves (i).

Now, item (ii) follows by the Hall-Petrescu identity [5, Appendix A], using exactly the same argument used in the proof of Proposition 4.1.10. If $i \leq e/3$ then $G^{p^i}/G^{p^{2i}}$ and $G^{p^{2i}}/G^{p^{3i}}$ have the same order, and so q is actually an isomorphism. We would like to point out that c maps the pair of elements $\bar{x} = xG^{p^{2i}}$ and $\bar{y} = yG^{p^{2i}}$ of $G^{p^i}/G^{p^{2i}}$ into the element $c(\bar{x}, \bar{y}) = (x, y)G^{p^{3i}}$, and that (x, y) actually belongs to $G^{p^{2i}}$, since (x, y) belongs to $[G^{p^i}, G^{p^i}]$, which equals $[G, G]^{p^{2i}}$ by Lemma 4.1.8, and $[G, G]^{p^{2i}}$ is a subgroup of $G^{p^{2i}}$.

The first part of (iii) also follows by Hall-Petrescu formula. Observe that the section $G^{p^i}/G^{p^{3i}}$ has class at most 2, since $\gamma_3(G^{p^i}) = \gamma_3(G)^{p^{3i}} \leq G^{p^{3i}}$ by Corollary 4.1.8. Therefore the section $G^{p^i}/G^{p^{3i}}$ satisfies $(x, yz) = (x, y)(x, z)$ and this yields bilinearity of the function c. The bilinearity of $[\ ,\ ]$ follows from the previous argument and the fact that q is an isomorphism, since $i \leq e/4$ implies $i < e/3$. To prove that $[\ ,\ ]$ satisfies the Jacobi identity, note that $G^{p^i}/G^{p^{4i}}$ has class at most 3, and the Hall-Witt identity becomes exactly $(x, -y, z) + (y, -z, x) + (z, -x, y) = 1$ in this section. Applying the isomorphism q^{-1} to both sides of this equation gives exactly the Jacobi identity. The fact that L is a Lie $\mathbb{Z}/p^i\mathbb{Z}$-ring follows from the fact that L is an abelian group of exponent p^i, and this proves the first part of (iv).

If we still assume that $i \leq e/4$, we know by item (ii) that q is an isomorphism. We are going to denote by $s^{p^{-i}}$ the unique image of $s \in G^{p^{2i}}/G^{p^{3i}}$ under q^{-1}. Thus taking $\bar{x}$ and $\bar{y}$ for some $x, y \in G^{p^i}$ in L we have

$$[\bar{x}, \bar{y}] = q^{-1}(c(\bar{x}, \bar{y})) = ((x, y)G^{p^{3i}})^{p^{-i}}.$$

Finally, take $g, h \in G^{p^i}$. Since G is powerful, then $[G^{p^i}, G^{p^i}] = [G, G]^{p^{2i}} \leq G^{p^{2i+1}}$, and so $(g, h) \in G^{p^{2i+1}}$. As $G^{p^{2i+1}} = (G^{p^{2i}})^p$, we have $(g, h) = s^p$, $s \in G^{p^{2i}}$. Now given $\bar{g}$ and $\bar{h}$, the images of g and h in L, we get that

$$[\bar{g}, \bar{h}] = q^{-1}((g, h)G^{p^{3i}}) = q^{-1}(s^p G^{p^{3i}}) = q^{-1}(p \cdot \tilde{s}) = p \cdot q^{-1}(\tilde{s}) = p\tilde{s}^{p^{-i}},$$

where $\tilde{s}$ denotes the image of s in $G^{p^{2i}}/G^{p^{3i}}$. Then, since $[L, L] \subseteq pL$, the proof is complete. □

A finitely generated free Lie $\mathbb{Z}/p^i\mathbb{Z}$-ring L with the property that L/pL is commutative is said to be *powerful*. It is important to remark that L in Proposition 4.3.1 reflects an important property of a powerful p-group G, that is having the quotient G/G^p abelian. Also, any choice of $i \leq e/4$ in the previous proposition would give a powerful Lie ring associated to G. So, Proposition 4.3.1 constructs a family of Lie rings rather than just one in case $e \geq 4$.

Our final goal in this section is to clarify the connection between the derived length of the Lie ring L and the derived length of G. Then, more than one Lie ring can be associated to G via Proposition 4.3.1. In what follows, G will denote a uniform group of exponent p^e and L_i its Lie rings. We give a series of lemmas concerning inclusions of some adequate subgroups of G that culminate in the desired result about the derived length.

Lemma 4.3.2. *If $L_i^{(j)} = 0$, then $(G^{p^i})^{(j)} \leq G^{p^{i2^j+i}}$.*

Proof. Set $L = L_i$ and, for $s \geq 0$, let $H_s \geq G^{p^{2i}}$ be the subgroup of G satisfying $L^{(s)} = H_s/G^{p^{2i}}$. The definition of the Lie product in L yields $H_s^{p^i} = H'_{s-1}G^{p^{3i}} \geq H'_{s-1}$. Let us show by induction on s that $H_s^{p^{i2^s-i}} \geq H_0^{(s)}$. Observe that for $s = 0$ the claim is obvious. Assuming the claim holds for $s - 1$ and using Corollary

4.1.8(i), we obtain

$$(H_s)^{p^{i2^s-i}} \geq (H'_{s-1})^{p^{i2^s-2i}} = (H_{s-1}^{p^{i2^{s-1}-i}})' \geq (H_0^{(s-1)})' = H_0^{(s)}$$

as required. Now, $H_0 = G^{p^i}$, and $H_j = G^{p^{2i}}$, since $L^{(j)} = 0$. Hence $G^{p^{i2^j+i}} = H_j^{p^{i2^j-i}} \geq (G^{p^i})^{(j)}$, proving the lemma. □

Lemma 4.3.3. *Suppose $L_i^{(j)} = 0$, where $i2^j + i \leq e$. Then $G^{(j)} \leq G^{p^i}$.*

Proof. By Lemma 4.3.2, $(G^{p^i})^{(j)} \leq G^{p^{i2^j+i}}$. Applying Lemma 4.1.8(i) we have $(G^{p^i})^{(j)} = (G^{(j)})^{p^{i2^j}}$ and so we obtain $(G^{(j)})^{p^{i2^j}} \leq G^{p^{i2^j+i}}$. Lemma 4.2.3 now yields $G^{(j)} \leq G^{p^i}$. □

We can now prove the main result of this section. Recall that G is a uniform group of exponent p^e, with associated Lie ring L_i.

Theorem 4.3.4. *Suppose $L_i^{(j)} = 0$, where $i = \lfloor e/(2^j + 1) \rfloor$. Then the derived length of G is at most $2j + 1$.*

Proof. The conditions of Lemma 4.3.3 are satisfied, so $G^{(j)} \leq G^{p^i}$. By Corollary 4.1.8(ii), $(G^{p^i})^{(j+1)} \leq G^{p^{i2^{j+1}+2^{j+1}-1}}$. Note that, since $2^j \geq 1$, then $i2^{j+1} + 2^{j+1} - 1 \geq (i+1)(2^j+1) - 1 \geq e$, by the choice of i. Hence, $(G^{p^i})^{(j+1)} = 1$, so $G^{(2j+1)} = (G^{(j)})^{(j+1)} \leq (G^{p^i})^{(j+1)} = 1$, and this completes the proof. □

There is no parallel result relating the derived length of the more familiar Lie ring $\bigoplus \gamma_n(G)/\gamma_{n+1}(G)$ to that of G. This is why the latter Lie ring is not an adequate tool when dealing with questions related to the solubility of uniform p-groups.

At this point, it is important to make some remarks about the method used above to construct the Lie rings L_i. First of all, in order to endow one of the L_i with a Lie ring structure, one needed to choose $1 \leq i \leq e/4$, where p^e denotes the exponent of G, otherwise L_i would be a trivial Lie ring. If no such i can be chosen, then we would have $e = 1, 2$ or 3. In these cases, we would have $dl(G) = 1, 2$ or 3, respectively, since the agemo series of G is central, by Corollary 4.1.9.

Now, suppose $e \geq 4$. Assume that, whenever a Lie ring L_i is soluble, its derived length is d_i. In this case, we need i and d_i to satisfy $i2^{d_i}+i \leq e$ in order to apply Theorem 4.3.4. Then, if the inequality holds for some pair i and d_i, the derived length of G is bounded by $2d_i+1$. But, in case $i2^{d_i}+i \leq e$ does not hold for any $i = 1, \ldots, e$ it is also possible to bound $dl(G)$ in terms of one of the d_i. Indeed, assume that for a uniform group G and its associated Lie rings the inequality $i2^{d_i}+i > e$ holds whenever L_i is soluble of length d_i. In particular, if L_1 is soluble, we have $2^{d_1}+1 > e$. By Corollary 4.1.9, we know that the derived length of G does not exceed e. The inequality $2^{d_1}+1 > e$ says that $dl(G)$ cannot exceed $2^{d_1}+1$. In any case, the derived length of G can be bounded in terms of the derived lengths of its associated Lie rings L_i, in case any of the L_i is soluble.

Chapter 5

Simple Groups of Lie Type

In this chapter we are going to discuss, in an expository way, the construction of some families of finite simple groups, the so-called simple groups of Lie type. The approach that we follow is due to Claude Chevalley, who conceived a way to construct such groups as automorphism groups of Lie algebras. Other authors made adaptations to this method and were able to construct the remaining simple groups of Lie type. So, apart from the cyclic, alternating and sporadic groups, this chapter gives the main tools and ideas which led to the discovery of the said simple groups over the years. We would like to stress that there is a more fruitful approach, by way of algebraic geometry. However, the one we give here is sufficient for deriving the basic properties of those groups and also to study their automorphisms, which is our main goal.

We begin this chapter giving a survey on the fundamental results about systems of roots and Weyl groups that are used throughout it. Then, we consider the problem of classifying finite dimensional Lie algebras over $\mathbb{C}$, which inspires the construction of some "versions" of the same Lie algebras, but this time defined over finite fields. An appropriate basis and structural constants are given. After that we define the so-called Chevalley groups, describe many of their properties and describe and find a normal form for their automorphisms. A remark on the remaining finite simple groups, the so-called "twisted" simple groups, and on their automorphism groups is made at the end of the chapter. The main reference for this chapter is the book [4].

5.1 Systems of Roots

Let $V = \mathbb{R}^l$, for some integer l, and let the canonical inner product of V be denoted by $(\ ,\ )$. Here, given a non-zero vector $r \in V$, we are going to denote by w_r the reflection of V through the hyperplane orthogonal to r. Since V has finite dimension, we can construct an orthogonal basis $\mathcal{B}$ for V containing r, in such a way that, given $s \in V$, we can write s as a linear combination of vectors in $\mathcal{B}$, where the coefficient of r is given by $(r,s)/(r,r)$. Writing s wiht respect to this basis, as w_r acts trivially on the orthogonal complement of r, it simply changes the sign of the coefficient of r in this expression. We can thus write

$$w_r(s) = s - \frac{2(r,s)}{(r,r)}r. \tag{5.1.1}$$

We now define the main object of this section.

Definition 5.1.1. *A subset Φ of V is called a system of roots, or root system, in V if the following axioms are satisfied:*

(i) *Φ is a finite set of non-zero vectors.*

(ii) *Φ spans V.*

(iii) *If $r, s \in \Phi$, then $w_r(s) \in \Phi$.*

(iv) *If $r, s \in \Phi$, then $2(r,s)/(r,r)$ is an integer.*

(v) *If $r \in \Phi$, then the only multiples of $r \in \Phi$ belonging to Φ are r itself and $-r$.*

Let Φ be a root system, and denote by $W = W(\Phi)$ the group generated by the reflections w_r for all $r \in \Phi$. Axiom (iii) of Definition 5.1.1 ensures that W leaves Φ invariant, and since Φ is finite W is a finite group of orthogonal transformations of V. Because Φ spans V, W operates faithfully on the root system. We call W the Weyl group of Φ. An example of a system of roots is the set of vectors $\{(1,0),\ (0,1),\ (-1,0),\ (0,-1,),\ (1,1),\ (-1,1),\ (-1,-1),\ (1,-1)\}$, consisting of the vertices and midpoints of a square centered in the origin of $\mathbb{R}^2$. Its corresponding Weyl group is isomorphic to D_4.

Even though Φ spans V, it is not a linearly independent set, as $-r = w_r(r)$ belongs to Φ whenever r does. It can be proved that Φ contains a linearly independent set Π that spans V and such that every root in Φ is a linear combination of vectors in Π with all coefficients either non-negative or non-positive integers, as in [4, Proposition 2.1.2]. We are going to call Π a *fundamental system of roots*. In the previous example, $a = (1, 0)$ and $b = (-1, 1)$ form Π.

We are going to endow V with a linear order relation, intending to divide Φ in two halves, and this division will be largely used. Let V^+ be a subset of V satisfying the conditions:

(i) If $v \in V^+$ and $\lambda > 0$, then $\lambda v \in V^+$.

(ii) If $v_1, v_2 \in V^+$, then $v_1 + v_2 \in V^+$.

(iii) For each $v \in V$, exactly one of the conditions holds: $v \in V^+, -v \in V^+, v = 0$.

Such a subset V^+ can be obtained, for example, if we take it to be one of the open semispaces defined by a hyperplane of V.

We now introduce the order relation $\succ$ by defining $v_1 \succ v_2$ if and only if $v_1 - v_2 \in V^+$ for some fixed choice of V^+. This is a total order in V compatible with addition and scalar multiplication by positive elements of $\mathbb{R}$. A subset of Φ is called a *positive system of roots* if it has the form $\Phi \cap V^+$ for some total ordering of V; we denote it by Φ^+. Now, as $-r$ belongs to Φ whenever r does, $\succ$ divides Φ in two parts of equal cardinality, Φ^+ and $\Phi - \Phi^+$, which we will denote by Φ^-. The roots in Φ^+ will be called positive roots and the remainder negative roots.

Now we can state the results about systems of roots and Weyl groups that we need. The proofs can be found in [4, Propositions 2.1.2-2.1.6 and 2.1.8].

Proposition 5.1.2. *Let Φ be a system of roots. The following holds.*

(i) *Every positive system of roots in Φ contains just one fundamental system, and each fundamental system is contained in some positive system of roots.*

(ii) *Let $r \in \Pi$, where Π is a fundamental system of roots. Then w_r transforms r into $-r$ but every other positive root into a positive root.*

(iii) *Each root in Φ is a linear combination of roots in Π with integer coefficients, all non-negative or all non-positive.*

(iv) *Every root in Φ is the image of some root in Π under the action of some element w of W.*

(v) *W is generated by the fundamental reflections w_r for $r \in \Pi$.*

We remark here that the roots contained in Π are called *fundamental roots*, and the reflections associated to them are called *fundamental reflections.* Item (ii) of Proposition 5.1.2 above together with item (iii) in Definition 5.1.1 have strong implications on the possible values of (r_i, r_j), $i \neq j$, $r_i, r_j \in \Pi$, which will be explored later. Items (i) and (iii) of Proposition 5.1.2 imply the following useful fact: if Π is a fundamental system, take the positive system Φ^+ associated to Π. Since every root in Φ can be written as $\sum_{i=1}^{l} \lambda_i r_i$ with integer λ_i by (iii), it is possible to choose a total ordering such that roots with $\lambda_i \geq 0$ for all i belong to V^+, and the ones such that $\lambda_i \leq 0$ for all i belong to V^-. In this sense, Φ^+ is determined completely by Π, and we can easily certify whether a root r is positive or negative just knowing the sign of one of its coefficients, when r is written in the basis Π. Item (iv) shows the importance of the fundamental system of roots, since every root is the image of some fundamental one. Finally, (v) gives a good set of generators for W, which will be explored in the next results.

Next, we want to shed some light on what can be said about the Weyl group associated to a given root system. Our intent is to give presentations of W as an abstract group; the proofs of the results can be found in [4, Theorems 2.4.1 and 2.4.3]. Let Φ be a root system and associate to it a fundamental system Π and its Weyl group W. Take $r, s \in \Pi$ and denote by m_{rs} the order of the rotation $w_r w_s$. Note that $m_{rr} = 1$ for each $r \in \Pi$.

Theorem 5.1.3. *W can be defined, as an abstract group, by the following presentations*

(i) $W = \langle w_r \mid w_r^2 = 1, w_r w_s w_r = w_{w_r(s)}, r, s \in \Pi \rangle$.

(ii) $W = \langle w_r \mid (w_r w_s)^{m_{rs}} = 1, r, s \in \Pi \rangle$.

The first presentation uses the action of W on Φ to construct the conjugation action inside the group. The second one emphasizes the role of the integers m_{rs} in the structure of W. Note that, since $w_r, w_s \in \Pi$ are reflections of V, and since W is a finite group, their product is a finite order rotation. The angle of this rotation is twice the angle between r and s, which needs to be a rational multiple of π. This kind of observation can be used to classify systems of roots. The possible systems of roots occurring in a 2-dimensional Euclidean space are described on the next proposition, up to isomorphism. Information is also given by the table. An isomorphism between systems of roots Φ and Ψ is a bijection $\alpha : \Phi \to \Psi$ such that the transformations $\alpha w_r \alpha^{-1}$ and $w_{\alpha(r)}$ are the same, for each r in Φ.

Table 5.1: Isomorphism Classes of Root Systems

Isomorphism class of Φ	$A_1 \times A_1$	A_2	B_2	G_2
Angle between generators	$\pi/2$	$2\pi/3$	$3\pi/4$	$5\pi/6$
Order of Φ	4	6	8	12
Isomorphism class of W	$C_2 \times C_2$	D_3	D_4	D_6

Proposition 5.1.4. *If Φ is a 2-dimensional root system, then it is isomorphic to one of the following systems: $A_1 \times A_1, A_2, B_2, G_2$.*

5.2 Simple Lie Algebras Over $\mathbb{C}$

The simple groups we are about to introduce appear as automorphism groups of finite dimensional Lie algebras over finite fields, constructed using finite dimensional simple Lie algebras over $\mathbb{C}$. It is natural that some material on simple Lie algebras over $\mathbb{C}$ is necessary to understand the groups. Each of the said algebras has associated to it a root system and a Weyl group, to which the results of the previous section apply and become a fruitful way to study such algebras. The discussion of this section follows mainly Humphreys' book [13]. Throughout this section, $\mathcal{L}$ will denote a finite dimensional simple Lie algebra over $\mathbb{C}$, whose dimension is greater than 1.

We begin this discussion introducing the notion of the *adjoint representation* of a Lie algebra. Consider the map ad $: \mathcal{L} \to End(\mathcal{L})$ mapping each $x \in \mathcal{L}$ to $\text{ad}(x)$, the endomorphism of $\mathcal{L}$ sending y to $[x, y]$. Since $\mathcal{L}$ is simple, it has no non-trivial ideals by definition, and as ad is a Lie homomorphism, $\mathcal{L}$ is isomorphic to its image in $End(\mathcal{L})$ under the adjoint representation. In this sense, the elements of $\mathcal{L}$ can be viewed as matrices over $\mathbb{C}$ that naturally act on $\mathcal{L}$ itself. Introduce a bilinear form (,) in $\mathcal{L}$ by way of $(x, y) = \text{tr}\,(\text{ad}x \cdot \text{ad}y)$. This map, which we call the *Killing form* of $\mathcal{L}$, is a symmetric bilinear form. Simplicity of $\mathcal{L}$ implies that the only $x \in \mathcal{L}$ such that $(x, y) = 0$ for all $y \in \mathcal{L}$ is $x = 0$. We will say that (,) is *non-degenerate*, and the proof for this fact can be found in [13, Theorem 5.1].

A *Cartan subalgebra* $\mathcal{H}$ of $\mathcal{L}$ is a nilpotent subalgebra of $\mathcal{L}$ that equals $N_{\mathcal{L}}(\mathcal{H})$ $= \{x \in \mathcal{L} \,|\, [x, h] \in \mathcal{H}, \text{ for all } h \in \mathcal{H}\}$. It can be proved that Cartan subalgebras are all conjugate under some automorphism of $\mathcal{L}$, and that they are abelian when $\mathcal{L}$ is simple [13, Corollary 16.4]. We define the *rank* of $\mathcal{L}$ as the dimension of $\mathcal{H}$ over $\mathbb{C}$.

If we identify $\mathcal{L}$ with $\text{ad}\,\mathcal{L}$, we have that $\mathcal{H}$ is an algebra of commuting endomorphisms of $\mathcal{L}$, thus it is possible to choose a basis for $\mathcal{L}$ that simultaneously diagonalizes the elements of $\mathcal{H}$, by Theorem 8 of Section 6.5 in [12]. In other words, $\mathcal{L}$ can be written as the direct sum of subspaces

$$\mathcal{L}_r = \{x \in \mathcal{L} \,|\, [x, h] = r(h)x \text{ for all } h \in \mathcal{H}\},$$

indexed by linear maps r from $\mathcal{H}$ to $\mathbb{C}$; we call such a decomposition a *Cartan decomposition*. Note that $\mathcal{L}_0$ is simply $C_{\mathcal{L}}(\mathcal{H})$, which equals $\mathcal{H}$, since $C_{\mathcal{L}}(\mathcal{H}) \subseteq N_{\mathcal{L}}(\mathcal{H}) = \mathcal{H}$. Also, the aforementioned subspaces are all one-dimensional, if non-zero, by [13, Proposition 8.4]. The set of non-zero $r \in \mathcal{H}^*$ for which $\mathcal{L}_r \neq 0$ is denoted by Φ, and the remarkable fact here is that it forms a root system inside $\mathcal{H}^*$, by [13, Propositions 8.4 and 8.5]. Non-degeneracy of the Killing form is inherited by its restriction to $\mathcal{H}$, as in [13, Corollary 8.2]. This makes it possible to identify $\mathcal{H}$ with $\mathcal{H}^*$ and to find a basis $\{t_r \,|\, r \in \Phi\}$ for $\mathcal{H}$ (remember that Φ spans $\mathcal{H}^*$) such that $(t_r, h) = r(h)$, for every $h \in \mathcal{H}$, and thus there is no harm in identifying Φ with the set of the t_r in $\mathcal{H}$ such that $\mathcal{L}_r \neq 0$, and saying that Φ is a system of roots in $\mathcal{H}$.

We want to summarize some important properties about this decomposition in the next theorem.

Theorem 5.2.1. (i) *If $r, s, r+s \in \Phi$, then $[\mathcal{L}_r, \mathcal{L}_s] = \mathcal{L}_{r+s}$. If $r+s \notin \Phi$, $\mathcal{L}_s$ and $\mathcal{L}_r$ commute.*

(ii) *$\mathcal{L}$ is generated, as a Lie algebra, by the root spaces $\mathcal{L}_r$, $r \in \Phi$.*

(iii) *Let $r, s \in \Phi$, $r \neq \pm s$. Let p, q be the largest integers for which $s - pr$ and $s + qr$ are also roots. Then, for every $-p \leq i \leq r$, $s + ir$ is also a root. Moreover, $p - q = 2(r, s)/(s, s)$, and we call this number A_{rs} for short.*

Now, let

$$\mathcal{L} = \mathcal{H} \oplus \mathcal{L}_{r_1} \oplus \cdots \oplus \mathcal{L}_{r_k}$$

be a Cartan decomposition of $\mathcal{L}$. It can be shown that, if we choose any subset of Φ which is a basis for $\mathcal{H}$, then each root in Φ is a linear combination of the roots in this subset with *rational* coefficients. Also, (r, s) is rational for every pair of roots. If we denote by $\mathcal{H}_{\mathbb{R}}$ the set of all elements of $\mathcal{H}$ that are on the $\mathbb{R}$-span of Φ, then $\mathcal{H}_{\mathbb{R}}$ is a real vector space, with the same dimension as the complex dimension of $\mathcal{H}$. Also, the Killing form induces an inner product in $\mathcal{H}_{\mathbb{R}}$ and so $\mathcal{H}_{\mathbb{R}}$ can be regarded as an Euclidean space. In particular, we define the length of an element $x \in \mathcal{H}_{\mathbb{R}}$ by $||x|| = \sqrt{(x,x)}$ and, as usual, the angle θ between $x, y \in \mathcal{H}_{\mathbb{R}}$ by $(x,y) = ||x||\,||y||cos(\theta)$, $0 \leq \theta < \pi$.

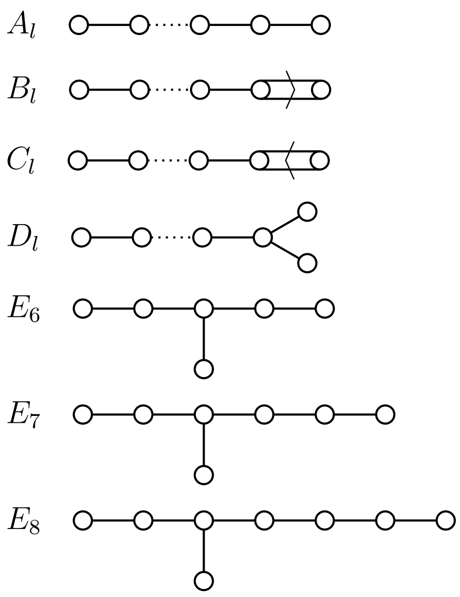
We now want to turn attention again, as at the end of the previous section, to the root systems associated to the Cartan decomposition, especially to the strong restrictions of Axiom (iv) of Definition 5.1.1. In this sense, if r and s are roots, we list on the table below the possible values for $A_{rs} = 2(r,s)/(r,r)$, the angle θ between them and the ratio between their lengths, assuming $||r|| \geq ||s||$. All values obtained in the table follow from the fact that $A_{rs}A_{sr} = 4\cos^2(\theta)$ is an integer, since A_{rs} also is, for every $r, s \in \Phi$.

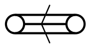
$2(r,s)/(r,r)$	$2(r,s)/(s,s)$	θ	$\|\|r\|\|^2/\|\|s\|\|^2$
0	0	$\pi/2$	undetermined
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

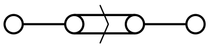
We are going to call a root system Φ *irreducible* if it cannot be partitioned into two proper subsets Φ_1 and Φ_2 such that every root in one is orthogonal to every root in the other. If a Lie algebra $\mathcal{L}$ is simple, then [13, Theorem 14.1] shows that its associated root system is irreducible, and [13, Theorem 14.2] states that, if two simple Lie algebras have isomorphic root systems, then they are isomorphic. In this way, classifying every possible root system amounts to classifying the simple Lie algebras of finite dimension over $\mathbb{C}$.

If r, s are distinct fundamental roots, we know that $A_{rs}A_{sr}$ equals 0, 1, 2 or 3, according to the previous table. We define the *Dynkin diagram* of Φ to be a graph having l vertices, where l is the cardinality of Π, the i-th joined to the j-th, with $i \neq j$, by $A_{ij}A_{ji}$ edges. If a double or triple edge occurs, we add an arrow pointing to the shorter of the two roots, and in this we call this root *short*; the longer root is called *long.* The Dynkin diagrams encode, in this way, the rank of $\mathcal{L}$, the angle between the fundamental roots, the Cartan integers and which roots are short and long. There is a classification theorem that establishes the possibilities for the Dynkin diagrams of irreducible root systems, a complete discussion can be found in [13, Theorem 11.4]. Also, if two simple Lie algebras admit the same Dynkin diagram, then their associated root systems are isomorphic and, by the previous remark, so are the considered Lie algebras. We classify all of the possibilities, then, based on the possible Dynkin diagrams below. The names given to the Lie algebras are written before the diagrams, and their Lie rank is the subscript after the name.

Theorem 5.2.2. *Let Φ be an irreducible system of roots of rank l. Then its Dynkin diagram, with l vertices in each case, is one of the following. Here, if Φ is of types A_l, B_l, C_l and D_l, then l is greater than or equal to 1, 2, 3 or 4 respectively.*



G_2 

F_4 

A question that arises naturally is if every possible Dynkin diagram presented is realizable, and the answer is yes. This is the content of [13, Theorem 12.1]. With all this theory set up, we can conclude that the only possible simple Lie algebras of finite dimension over $\mathbb{C}$ are the ones associated with the said diagrams.

Finally, we want to end this section by exhibiting a choice of basis for $\mathcal{L}$ where the structural constants are very well behaved and related to the Cartan integers. This choice of basis is the starting point for all the work on simple groups we are going to do in this chapter.

Consider

$$\mathcal{L} = \mathcal{H} \oplus \sum_{r \in \Phi} \mathcal{L}_r$$

to be a Cartan decomposition for $\mathcal{L}$, and take the elements $h_r = \dfrac{2r}{(r,r)}$ in $\mathcal{H}$, the *co-roots* associated to the roots $r \in \Phi$. By [13, Proposition 8.5(a)], for a chosen $e_r \in \mathcal{L}_r$, there is a unique $e_{-r} \in \mathcal{L}_{-r}$ such that $[e_r, e_{-r}] = h_r$, and we will suppose e_{-r} chosen in this way. The set $\{h_r, r \in \Pi;\ e_r, r \in \Phi\}$ is a basis for $\mathcal{L}$, consisting of the fundamental co-roots h_r together with the set of all root vectors e_r. The result we want to quote can be found in [4, Theorem 4.2.1]. We denote by p the greatest integer for which $s - pr$ is a root.

Theorem 5.2.3. *Let $\mathcal{L}$ be a simple algebra over $\mathbb{C}$ with Cartan decomposition as above, and let h_r and e_r be chosen as before. Then, the multiplication constants of this Lie algebra are as follows*

$$\begin{aligned}
&[h_r, h_s] = 0 \\
&[h_r, e_s] = A_{rs} e_s \\
&[e_r, e_{-r}] = h_r \\
&[e_r, e_s] = 0, && \textit{if}\ r + s \notin \Phi \\
&[e_r, e_s] = N_{rs} e_{r+s}, && \textit{if}\ r + s \in \Phi,
\end{aligned}$$

where $N_{rs} = \pm(p+1)$.

Here, $N_{rs} = (p+1)$ or $-(p+1)$, but the possible different choices of signs do not influence the isomorphism class of $\mathcal{L}$. The multiplication constants of $\mathcal{L}$ are all integers.

Such a choice of basis for $\mathcal{L}$ will be called a *Chevalley basis.* Given a choice of Cartan subalgebra of $\mathcal{L}$, the root spaces $\mathcal{L}_r$ are determined, as also is Φ. We may then choose some fundamental system Π in Φ, and in this way the fundamental co-roots h_r are determined. Then, for each $r \in \Pi$ we choose arbitrary nonzero $e_r \in \mathcal{L}_r$, and the remaining positive root vectors are determined via $[e_r, e_s] = N_{rs}e_{r+s}$ within a choice of signal, since every root in Φ is a $\mathbb{Z}$-linear combination of the fundamental ones with all coefficients non-positive or non-negative. The relation $[e_r, e_{-r}] = h_r$ determines the rest of the basis vectors, then. In this sense, one might wonder if every choice of Chevalley basis leads to structural constants that at least have the same absolute value, and the answer to this question is positive. The root system determines the absolute value of the multiplication constants, and it does not matter which Cartan subalgebra of $\mathcal{L}$ we choose in this process, as they all will have the same root system associated. In any case, the constants are all integer numbers, and that is the fact we are going to need in the future discussion.

5.2.1 Groups and Algebras over an Arbitrary Field

An endomorphism δ of $\mathcal{L}$ satisfying the following identity, for any $x, y \in \mathcal{L}$

$$\delta \cdot [x, y] = [\delta \cdot x, y] + [x, \delta \cdot y]$$

is called a *derivation* of $\mathcal{L}$. For example, the Jacobi identity turns the map $\operatorname{ad} x$ into a derivation of $\mathcal{L}$, for any $x \in \mathcal{L}$.

Lemma 5.2.4. *Let $\mathcal{L}$ be a Lie algebra over a field of characteristic 0 and δ be a derivation of $\mathcal{L}$ which is nilpotent, i.e., satisfies $\delta^n = 0$ for some n. Then*

$$\exp \delta = 1 + \delta + \frac{\delta^2}{2!} + \cdots + \frac{\delta^{n-1}}{(n-1)!}$$

is an automorphism of $\mathcal{L}$.

The proof of this result can be found in [4, Lemma 4.3.1], and this automorphism is called the *exponential* of δ. Now, let $\mathcal{L}$ be a simple Lie algebra over $\mathbb{C}$ with Cartan decomposition

$$\mathcal{L} = \mathcal{H} \oplus \sum_{r \in \Phi} \mathcal{L}_r$$

and Chevalley basis $\{h_r, r \in \Pi;\ e_r, r \in \Phi\}$. The map $\operatorname{ad} e_r$ is a derivation of $\mathcal{L}$, and this derivation is in fact nilpotent. For we have

$$\begin{aligned} \operatorname{ad} e_r \cdot \mathcal{H} &= \mathcal{L}_r, & (\operatorname{ad} e_r)^2 \cdot \mathcal{H} &= 0, \\ \operatorname{ad} e_r \cdot \mathcal{L}_r &= 0, & & \\ \operatorname{ad} e_r \cdot \mathcal{L}_{-r} &\subseteq \mathcal{H}, & (\operatorname{ad} e_r)^3 \cdot \mathcal{L}_{-r} &= 0. \end{aligned}$$

Also, $(\operatorname{ad} e_r)^{q+1} \cdot \mathcal{L}_s = 0$ if r, s are linearly independent roots, since $(q+1)r + s$ is not a root. Thus, $(\operatorname{ad} e_r \cdot \mathcal{L})^n = 0$ for sufficiently large values of n.

Let $\zeta \in \mathbb{C}$. Then $\operatorname{ad}(\zeta e_r) = \zeta \operatorname{ad} e_r$ is also a nilpotent derivation of $\mathcal{L}$, thus $\exp(\operatorname{ad}(\zeta e_r))$ is an automorphism of $\mathcal{L}$, by Lemma 5.2.4. We write

$$x_r(\zeta) = \exp(\operatorname{ad}(\zeta e_r)).$$

We now consider the effect of $x_r(\zeta)$ on the Chevalley basis. We have

$$\begin{aligned} x_r(\zeta) \cdot e_r &= e_r, \\ x_r(\zeta) \cdot e_{-r} &= e_{-r} + \zeta h_r - \zeta^2 e_r, \\ x_r(\zeta) \cdot h_r &= h_r - 2\zeta e_r. \end{aligned} \tag{5.2.2}$$

Also, if $r \neq \pm s$,

$$\begin{aligned} x_r(\zeta) \cdot h_s &= h_s - A_{sr} \zeta e_r, \\ x_r(\zeta) \cdot e_s &= e_s + N_{rs} \zeta e_{r+s} + \frac{1}{2!} N_{rs} N_{r,r+s} \zeta^2 e_{2r+s} + \cdots + \\ &\qquad + \frac{1}{q!} N_{rs} N_{r,r+s} \cdots N_{r,(q-1)r+s} \zeta^q e_{qr+s}. \end{aligned} \tag{5.2.4}$$

We write

$$M_{r,s,i} = \frac{1}{i!} N_{rs} N_{r,r+s} \cdots N_{r,(i-1)r+s},$$

for short. Then, using the fact that $N_{rs} = \pm(p+1)$ we see that

$$M_{r,s,i} = \pm\frac{(p+1)(p+2)\cdots(p+i)}{i!} = \pm\binom{p+i}{i}$$

is an integer, with $M_{r,s,0} = 1$. Thus, the automorphism $x_r(\zeta)$ transforms each element of the Chevalley basis into a linear combination of basis elements with coefficients being non-negative integral powers of ζ with integer coefficients. This property enables us to define automorphisms of this type over an arbitrary field.

Let $\mathcal{L}$ be a simple Lie algebra over $\mathbb{C}$, with the usual Chevalley basis. We denote by $\mathcal{L}_{\mathbb{Z}}$ the subalgebra of elements of $\mathcal{L}$ with integer coefficients with respect to the Chevalley basis. $\mathcal{L}_{\mathbb{Z}}$ is a subalgebra of $\mathcal{L}$ by Theorem 5.2.3. Now, let K be any field. We form the tensor product of the additive group of K with the additive group of $\mathcal{L}_{\mathbb{Z}}$ and define

$$\mathcal{L}_K = K \otimes \mathcal{L}_{\mathbb{Z}}.$$

Then $\mathcal{L}_K$ is an additive abelian group. Let 1_K be the unit element of K. Define $\bar{e}_r := 1_K \otimes e_r$ and $\bar{h}_r := 1_K \otimes h_r$. Then $\mathcal{L}_K$ is a vector space over K with basis $B = \{\bar{h}_r, r \in \Pi;\ \bar{e}_r, r \in \Phi\}$. We define a Lie multiplication on the elements of B by

$$[1_K \otimes x, 1_K \otimes y] = 1_K \otimes [x, y],$$

for all x, y in B, and extend it by linearity to the whole $\mathcal{L}_K$. Then we have the following result.

Proposition 5.2.5. *The previous construction turns $\mathcal{L}_K$ into a Lie algebra over K. The multiplication constants of $\mathcal{L}_K$ with respect to the basis $\{\bar{h}_r, r \in \Pi;\ \bar{e}_r, r \in \Phi\}$ are the same as the ones for $\mathcal{L}$ with respect to $\{h_r, r \in \Pi;\ e_r, r \in \Phi\}$ interpreted as elements of the prime subfield of K.*

We do not need to give a proof, since Theorem 5.2.3 guarantees that the structure constants are all in $\mathbb{Z}$.

5.3 Chevalley Groups

Now, we are going to translate the definition of the automorphisms $x_r(\zeta)$, considered above, for $\mathcal{L}$ into the new context of $\mathcal{L}_K$. In this sense, if $A_r(\zeta)$ is the matrix of $x_r(\zeta)$ with respect to the Chevalley basis of $\mathcal{L}$, then every entry of the matrix has the form $a\zeta^i$, where a is an integer and $i \geq 0$. In a similar way, let $t \in K$ and $\bar{A}_r(t)$ be the matrix obtained from $A_r(\zeta)$ replacing a with $\bar{a}$, its image in the prime subfield of K, i.e., its image modulo p, where p is the characteristic of K, and ζ by t. We now define $\bar{x}_r(t)$ to be the endomorphism of $\mathcal{L}_K$ given by the matrix $\bar{A}_r(t)$ with respect to the basis $\{\bar{h}_r, r \in \Pi; \bar{e}_r, r \in \Phi\}$. Then, we have the following result, whose proof can be found in [4, Proposition 4.4.2].

Proposition 5.3.1. *The $\bar{x}_r(t)$, as defined above, are automorphisms of $\mathcal{L}_K$, for each $r \in \Phi$ and $t \in K$.*

Since we are going to work with the $\bar{x}_r(t)$ from now on, we may drop the bars in the notations of both automorphisms and Chevalley basis of $\mathcal{L}_K$. We shall now define the Chevalley groups. The Chevalley group of type $\mathcal{L}$ over the field K, denoted by $\mathcal{L}(K)$, is defined to be the group of automorphisms of the Lie algebra $\mathcal{L}_K$ generated by the $x_r(t)$, for all $r \in \Phi$ and $t \in K$. The generators of $\mathcal{L}(K)$ operate on the Chevalley basis according to the formulae in equations (5.2.2) and (5.2.4). Also, a crucial result is given by [4, Proposition 4.4.3], which attests the independence of the isomorphism class of $\mathcal{L}(K)$ from the choice of the Chevalley basis. We state it here.

Proposition 5.3.2. *The group $\mathcal{L}(K)$ is determined up to isomorphism by the simple Lie algebra $\mathcal{L}$ over $\mathbb{C}$ and the field K.*

5.3.1 Chevalley's Commutator Formula

In this section, we are going to derive a formula, due to Chevalley, which expresses the commutator between two group elements of $G = \mathcal{L}(K)$ in terms of the generators $x_r(t)$ of G. The proof can be found in [4], Section 5.2, and will be omitted here. However, a sketch containing the main ideas will be included. We first need a lemma, whose proof is in [4, Lemma 5.1.1].

Lemma 5.3.3. *Let $\mathcal{L}$ be a simple Lie algebra over $\mathbb{C}$. Let y be an element of $\mathcal{L}$ such that* $\operatorname{ad} y$ *is nilpotent and let θ be an automorphism of $\mathcal{L}$. Then*

$$\theta \cdot \exp(\operatorname{ad} y) \cdot \theta^{-1} = \exp(\operatorname{ad} \theta y).$$

Theorem 5.3.4. *Let $G = \mathcal{L}(K)$ be a Chevalley group over an arbitrary field, r, s be linearly independent roots of $\mathcal{L}$ and t, u be elements of K. Then, the commutator between $x_r(t)$ and $x_s(u)$ can be written in the following way, in terms of the generators of G*

$$[x_s(u), x_r(t)] = \prod_{i,j>0} x_{ir+js}(C_{ijrs}(-t)^i u^j),$$

where the product is taken over all pairs of positive integers i, j for which $ir + js$ is a root, in order of increasing $i + j$. Each constant C_{ijrs} is one of $\pm 1, \pm 2, \pm 3$.

Sketch of the Proof. We work the proof out considering the field K to be $\mathbb{C}$ at first. It is important to note that, when we take two distinct elements, r, s from the fundamental system Π, the root system spanned by them has rank two by construction and hence is isomorphic to one of the systems listed in Proposition 5.1.4. Also, as each element of G is an endomorphism of $\mathcal{L}_K$ having finite order, we can apply Lemma 5.3.3 to the conjugation operation in the group. Analyzing each of the possible classes of root systems, together with the possible choices of linearly independent vectors in each one, it is possible to rewrite each conjugation as a product of some specific endomorphisms of $\mathcal{L}_K$, each of them indexed by a root in the span of r, s, and thus the result holds if $K = \mathbb{C}$. Now, take any e_r from the Chevalley basis and consider its images under $[x_s(u), x_r(t)]$ and $\prod_{i,j>0} x_{ir+js}(C_{ijrs}(-t)^i u^j)$. The images are linear combinations of elements of the Chevalley basis, and comparing each coefficient gives rise to some polynomials in $\mathbb{Z}[u, t]$, where the coefficients are the integers C_{ijrs}. Hence, the validity of the commutator formula if $K = \mathbb{C}$ is equivalent to the vanishing of certain polynomials, with integer entries. To pass from $\mathbb{C}$ to an arbitrary field K, we may just consider the corresponding polynomials when we read their coefficients on the prime subfield of K; the vanishing of those over $\mathbb{C}$ imply that they are identically zero over K too, and thus the formula holds in any case. □

5.3.2 The subgroups of G

Consider $G = \mathcal{L}(K)$. We are now in condition to define certain important subgroups of G. First of all, let us remark that, for fixed r and arbitrary $t, u \in K$, we have

$$x_r(t)x_r(u) = \exp\left(t\,\mathrm{ad}\,e_r\right)\exp\left(u\,\mathrm{ad}\,e_r\right) = exp\left((t+u)\,\mathrm{ad}\,e_r\right) = x_r(t+u),$$

since $\exp\left(t\,\mathrm{ad}\,e_r\right)$ and $\exp\left(u\,\mathrm{ad}\,e_r\right)$ commute.

Definition 5.3.5. *Let X_r be the subgroup of G generated by the $x_r(t), t \in K$.*

There is an epimorphism from X_r onto the additive group of K, and if we take $x_r(t)$ mapping to zero, then analyzing the action of each $x_r(t)$ on the Chevalley basis we conclude that t must be zero too. Hence, each of the subgroups X_r is isomorphic to the additive group of K, and we will call it the *root subgroup* associated to the root $r \in \Phi$. Let us define the subgroups U and V of G as follows:

Definition 5.3.6.

$$U = \langle X_r \,|\, r \in \Phi^+ \rangle,$$
$$V = \langle X_r \,|\, r \in \Phi^- \rangle.$$

If we consider $x_r(t)$ and $x_s(u)$, both indexed by positive roots, the commutator formula ensures that the product $x_r(t)x_s(u)$ equals

$$x_s(u)x_r(t) \prod_{i,j>0} x_{ir+js}(C_{ijrs}(-t)^i u^j),$$

the product taken over roots which are also positive. In this way, we see that the product of two generators of U, for example, can be written as the product of the elements $x_r(t)$ with positive r. This shows that $U \cap V = 1$, because the elements of U are products of the x_r with r positive, while the elements of V are products of the x_r with r negative.

We shall, in the next result, use the commutator formula to investigate the structure of U and V; in particular we obtain a normal form for the elements of these subgroups. The proof can be found in [4, Theorem 5.3.3].

Theorem 5.3.7. *Let $G = \mathcal{L}(K)$ be a Chevalley group, U be the subgroup of G generated by the root subgroups X_r with $r \in \Phi^+$. Then each element of U is expressible uniquely in the form*

$$\prod_{r_i \in \Phi^+} x_{r_i}(t_i),$$

where the product is taken over all positive roots in increasing order with respect to $\succ$, as defined in Section 5.1.

A similar result is valid for V, but in terms of negative roots. The above theorem allows us to evaluate the order of both subgroups. Let $|\Phi^+| = n$ and $|K| = p^b$. Since each element of U, for example, is expressible uniquely in the form mentioned above, then $|U|$ equals the order of K raised to the n-th power, giving $|U| = p^{nb}$. Actually, more can be said about U and V, as the next result attests.

Theorem 5.3.8. *Let G be a Chevalley group of type $\mathcal{L}$ over K. Then U and V are Sylow p-subgroups of G.*

A lot of theory must be constructed in order to prove this result, and the proof can be deduced from the result in [4, Theorem 9.4.10].

Now, we turn our attention to the relation between generators of the form $x_r(t)$, $x_{-r}(u)$, since Theorem 5.3.4 considers so far the case where the roots are linearly independent. In this sense, we intend to investigate on the next discussion the subgroups $\langle X_r, X_{-r} \rangle$, generated by root subgroups corresponding to opposite roots. We take r to be a fixed root of the root system Φ unless otherwise stated.

First of all, we are going to invoke a general fact about special linear groups and a good set of generators for them. The proof can be found in [4, Lemma 6.1.1].

Lemma 5.3.9. *Let K be an arbitrary field. Then the group $SL_2(K)$ is generated by the elements*

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix},$$

as t runs through K.

We can now state the main result of this section. Its proof can be found in [4, Theorem 6.3.1].

Theorem 5.3.10. *Let K be any field. Then there is a homomorphism ϕ_r from $SL_2(K)$ onto the subgroup $\langle X_r, X_{-r}\rangle$ of $G = \mathcal{L}(K)$ under which*

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mapsto x_r(t), \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \mapsto x_{-r}(t).$$

Moreover, if it is not faithful, the kernel of this homomorphism equals $\langle -I\rangle$, I being the identity matrix of $SL_2(K)$.

We consider next the images under ϕ_r of the diagonal matrices

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$$

in $SL_2(K)$ and the monomial matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The images of these matrices will play an important rôle in the sequel, but first we need to understand how they act on the Chevalley basis of $\mathcal{L}_K$.

Proposition 5.3.11. *Let*

$$h_r(\lambda) = \phi_r \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}.$$

Then $h_r(\lambda)$ operates on the Chevalley basis of $\mathcal{L}_K$ in the following manner:

$$\begin{array}{ll} h_r(\lambda)\cdot h_s = h_s, & s\in\Pi \\ h_r(\lambda)\cdot e_s = \lambda^{A_{rs}} e_s, & s\in\Phi \end{array}.$$

Moreover, if

$$n_r = \phi_r\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

then n_r operates on the Chevalley basis as follows

$$\begin{array}{l} n_r\cdot h_s = h_{w_r(s)}, \\ n_r\cdot e_s = \eta_{r,s} e_{w_r(s)}, \quad \text{where } \eta_{r,s} = \pm 1. \end{array}.$$

The constants $\eta_{r,s}$ satisfy some equalities, but they will not be essential in the sequel of this text, so they will be omitted. The above equations can be found in [4, Proposition 6.4.3]. The proof of Proposition 5.3.11 above can be found in [4, Propositions 6.4.1 and 6.4.2].

We shall also need some properties of the elements

$$n_r(t) = \phi_r\begin{pmatrix} 0 & t \\ -t^{-1} & 0 \end{pmatrix},$$

which are stated in the following lemma.

Lemma 5.3.12.

(i) $n_r(1) = n_r$,

(ii) $n_r(-1) = n_r^{-1}$,

(iii) $n_r(t) = x_r(t)x_{-r}(-t^{-1})x_r(t)$,

(iv) $h_r(t) = n_r(t)n_r(-1)$.

Proof. These results follow from the corresponding relations in $SL_2(K)$ using the homomorphism ϕ_r. □

Now, we are in position to define two new subgroups of G that play important rôles on the structure of the group. For that we allow r to be an arbitrary root of Φ and make the following definition.

Definition 5.3.13. *Let H be the subgroup of G generated by the elements $h_r(\lambda)$ for all $r \in \Phi, 0 \neq \lambda \in K$.*

Now, $h_r(\lambda)$ operates on the Chevalley basis of $\mathcal{L}_K$ by the rules given in Proposition 5.3.11. This shows that each element $h_r(\lambda)$ of H is an automorphism of $\mathcal{L}_K$ which operates trivially on $\mathcal{H}_K$, the corresponding of the Cartan subalgebra of $\mathcal{L}$ in $\mathcal{L}_K$, and transforms each root vector e_s into $\lambda^{A_{rs}} e_s$. The coefficients arising in this way define naturally a map from H to K^*, which we shall explore now.

Let $P = \mathbb{Z}\Phi$, the set of all linear combinations of elements of Φ with integer coefficients. P is the additive group generated by the roots of $\mathcal{L}$. It is a free abelian group of rank l and has a basis consisting of the set of fundamental roots $\Pi = \{p_1, \ldots, p_l\}$. A homomorphism from the additive group of P into the multiplicative group K^* of non-zero elements of K is called a K-character of P.

Now, a K-character of P is uniquely determined by its values on the fundamental roots. If we map each $r \in \Pi$ to $t_r \in K^*$, then the map taking $\sum\limits_{r \in \Pi} \lambda_r r$ to $\prod\limits_{r \in \Pi} t_r^{\lambda_r}$ is a K-character of P. Thus, if a map from Π into K^* can be extended to a K-character of P, this extension is uniquely determined.

Take χ_1, χ_2 to be K-characters of P. The K-characters of P form a multiplicative group, with multiplication given by $\chi_1\chi_2(a) = \chi_1(a)\chi_2(a), \quad a \in P$. Moreover, each K-character χ of P gives rise to an automorphism $h(\chi)$ of $\mathcal{L}_K$ defined by $h(\chi) \cdot h_s = h_s, h(\chi) \cdot e_s = \chi(s) e_s$.

The automorphisms of $\mathcal{L}_K$ of the form $h(\chi)$ form a subgroup of the full automorphism group of $\mathcal{L}_K$, which we call $\hat{H}$. Note that multiplication in $\hat{H}$ is given by $h(\chi_1) \cdot h(\chi_2) = h(\chi_1\chi_2)$, and thus the map $\chi \mapsto h(\chi)$ is an isomorphism between the group of K-characters of P and $\hat{H}$. This isomorphism allows us to compute the order of $\hat{H}$ by computing the number of different K-characters of P, which is equal to the number of maps from Π to K^*, since P is free abelian with free basis Π. The order of $\hat{H}$ is precisely $(q-1)^l$, and $\hat{H}$ is the direct product of l copies of the cyclic group C_{p^b-1}.

Now, as H is the subgroup of G generated by the $h_r(\lambda)$, for all $r \in \Phi, \lambda \in K^*$, it is a subgroup of $\hat{H}$, and every element of H therefore has the form $h(\chi)$ for some character χ. One further information that we are going to need is the exact value of the index of H in $\hat{H}$. The ideas for calculating these values can be found in [4], Sections 7.1 and 8.6. Since the characters that generate H all depend on the elements of some root system and K^*, it is expected that the values of the indices also depend somehow on the same structures. We remind the reader here that $|K| = q$ and $|\Pi| = l$. For our purposes, we will just list the values of the indices on the table below. Here, (,) denotes the gcd between two integers.

Table 5.2: Index of H in $\hat{H}$, in each case.

Class of $\mathcal{L}$	A_l	B_l	C_l	D_{2k+1}	D_{2k}
$[\hat{H} : H]$	$(l+1, q-1)$	$(2, q-1)$	$(2, q-1)$	$(4, q-1)$	$(2, q-1)^2$
Class of $\mathcal{L}$	G_2	F_4	E_6	E_7	E_8
$[\hat{H} : H]$	1	1	$(3, q-1)$	$(2, q-1)$	1

Now, we proceed to describe the relation between H and the subgroups U and V. Note first that H normalizes each root subgroup X_r. Recall here that since H is a finite group, every element of H has finite order.

Lemma 5.3.14. *H normalizes U and V in G.*

Proof. We know that

$$h(\chi)x_r(t)h(\chi)^{-1} = h(\chi)\exp \text{ad}(te_r)\, h(\chi)^{-1} = \exp \text{ad}(h(\chi)\cdot te_r) = \exp \text{ad}(\chi(r)te_r),$$

by Lemma 5.3.3. Thus $h(\chi)x_r(t)h(\chi)^{-1} = x_r(\chi(r)t)$, whence $h(\chi)X_rh(\chi)^{-1} = X_r$. It follows that $\hat{H}$, and in particular H, normalizes U and V. □

As a consequence, UH is a subgroup of G, and a natural question to ask is whether this subgroup contains or not root subgroups associated with negative roots. It is the content of the following lemma, whose proof can be found in [4, Lemma 7.1.2].

Lemma 5.3.15. *If U, V and H are as defined before, then $UH \cap V = 1$.*

As expected, a similar argument shows that $VH \cap U = 1$. In particular, we have $H \cap V = H \cap U = 1$. Observe that the same results apply when H is replaced by $\hat{H}$. Now, Dedekind modular law implies

Corollary 5.3.16. *$UH \cap VH = H$.*

Next, let us focus our attention on another subgroup of G.

Definition 5.3.17. *Let N be the subgroup of G generated by H and the elements n_r, for all $r \in \Phi$.*

In order to investigate the properties of N, we begin by understanding how the n_r act on the root subgroups.

Lemma 5.3.18. *Let $r, s \in \Phi$. Then*

$$n_r \cdot x_s(t) \cdot n_r^{-1} = x_{w_r(s)}(\eta_{r,s}t), \quad \textit{and} \quad n_r X_s n_r^{-1} = X_{w_r(s)}.$$

Proof. Using Lemma 5.3.3 and Proposition 5.3.11, we have that

$$\begin{aligned} n_r \cdot x_s(t) \cdot n_r^{-1} &= n_r \cdot \exp \operatorname{ad}(te_s) \cdot n_r^{-1} = \exp \operatorname{ad}(n_r \cdot te_s) \\ &= \exp \operatorname{ad}(\eta_{r,s} te_{w_r(s)}) = x_{w_r(s)}(\eta_{r,s}t). \end{aligned}$$ □

The group N acts on the X_r via $n_r X_s n_r^{-1} = X_{w_r(s)}$, in a way that resembles the action of the Weyl group $W(\Phi)$, and the subgroup H of N acts fixing each root subgroup, as we proved before. The most important result concerning N is the one we shall now quote, and it translates the intuition behind this idea. The proof can be found in [4, Theorem 7.2.2].

Theorem 5.3.19. *There is a homomorphism from N onto W with kernel H under which $n_r \to w_r$ for all $r \in \Phi$. Thus H is a normal subgroup of N and N/H is isomorphic to W. If $n \in N, h(\chi) \in H$, we have*

$$nh(\chi)n^{-1} = h(\chi'),$$

where $\chi'(r) = \chi(w^{-1}(r))$, w being the image of n under the above homomorphism.

We are going to call the homomorphism of Theorem 5.3.19 the *natural* homomorphism of N onto W. We end this exposition of the first properties of the subgroups U, V, H and N with a corollary of the previous theorem [4, Corollary 7.2.4].

Corollary 5.3.20. *$U \cap N = 1$ and $UH \cap N = H$.*

We are now in position to produce a decomposition of G where these subgroups will play important roles, and a lot of information about G will be derived from it.

5.3.3 The Bruhat Decomposition and (B, N)-pairs

We begin this subsection defining the general concept of a (B, N)-pair in a group. This concept was originally introduced by J. Tits in [31], and it is useful not only for deriving properties about Chevalley groups, but also for the "twisted" groups we are going to define later. We are going to work with double cosets of the form BnB.

Definition 5.3.21. *A pair of subgroups B and N of a group G is called a (B, N)-pair if the following axioms are satisfied:*

(BN1) *G is generated by B and N.*

(BN2) *$B \cap N$ is a normal subgroup of N.*

(BN3) *The group $W = N/B \cap N$ is generated by a set of elements $w_i, i \in I$, such that $w_i^2 = 1$.*

(BN4) *If $n_i \in N$ maps to w_i under the natural homomorphism of N into W and if n is any element of N, then*

$$Bn_iB \cdot BnB \subseteq Bn_inB \cup BnB.$$

(BN5) *If n_i is as above, then $n_iBn_i \neq B$.*

We would like to give examples of groups admitting (B, N)-pairs. Let $G = GL_n(K)$, where K is a finite field and $n \geq 2$ is a natural number. The construction made here is also valid if we consider, instead of $GL_n(K)$, the groups $PGL_n(K)$, $SL_n(K)$ or $PSL_n(K)$ with suitable adaptations. We claim that $G = GL_n(K)$, admits a (B, N)-pair. Let $V = \langle e_1, \ldots, e_n \rangle$ be the subjacent K-vector space on which G acts. The subgroup $B \leq G$ is the subgroup of upper triangular matrices, which is a Sylow p-subgroup of G. Let N be the subgroup consisting of matrices with exactly one nonzero entry in each row and column. Then, N acts on the set of lines $\{\langle e_1 \rangle, \ldots, \langle e_n \rangle\}$ as the symmetric group of degree n, and the kernel of this representation is precisely the diagonal subgroup of G, $B \cap N$. Then, Axioms (BN2) and (BN3) are verified.

Now, since every matrix in G can be written as a product of an upper triangular and a lower triangular matrices, to verify that B and N generate G we may just prove that $\langle B, N \rangle$ contains the lower triangular matrices. Consider the matrix $m_\sigma \in G$ which is the identity matrix with rows permuted by a permutation $\sigma \in S_n$. Then, if $a \in G$ is any matrix, the product $m_\sigma a$ is just the matrix a with rows permuted by σ, and a similar effect on the columns can be obtained by multiplying matrices of the form m_σ on the right. This shows that, by multiplying a suitable upper triangular matrix both on left and right by matrices m_σ, the result obtained can be any lower triangular matrix, thus showing that $G = \langle B, N \rangle$.

To verify Axiom (BN4) we have to show that

$$Bn_iB \cdot BnB \subseteq Bn_inB \cup BnB,$$

where n_i maps to a generator of $N/B \cap N \cong S_n$ under the homomorphism constructed above and n is any element of N. This inclusion can be rewritten as $n_iBn \subseteq Bn_inB \cup BnB$. We verify, with no loss of generality, what happens in case n_i maps to the transposition (12), in which case n_i is a matrix represented by

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

where the blank spaces are zeroes and the $*$ marks can be any element of K. Multiplying $n_iBn \subseteq Bn_inB \cup BnB$ on the right by n^{-1}, we can rewrite this inclusion as $n_iB \subseteq Bn_inBn^{-1} \cup BnBn^{-1}$. This last inclusion is equivalent to proving that any matrix in n_iB can be reduced to the identity matrix or to n_i via left multiplication by elements of B and right multiplication by elements of nBn^{-1}. It turns out that we will only need to use elementary matrices from B and nBn^{-1}.

Recall here that matrices in B are upper triangular. For example, the matrix a with 1's on the main diagonal and on position $(1, 2)$ and zeros everywhere else belongs to B, and left multiplication by a transforms any matrix m in another, where the first row equals the sum of the first and second rows of m. This process is known as *pivoting*, and this idea shows that multiplying on the left by elements of B allows us to add a multiple of a row to any other higher row. Then, since a typical element of n_iB has the form

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

by pivoting upwards we can reduce this matrix to

$$\begin{pmatrix} 0 & * & & & \\ * & * & & & \\ & & * & & \\ & & & \ddots & \\ & & & & * \end{pmatrix}.$$

In case the (2,2) entry above equals zero, we reduced the matrix to n_i. Then, we may assume that the entries (1,2), (2,1) and (2,2) are nonzero.

Now, let us use right multiplication by nBn^{-1}. To avoid confusion, we are only going to use the elements of $nBn^{-1} \cap GL_2(K)$, where $GL_2(K)$ is identified with the subgroup

$$\{g \in G \,|g\left(\langle e_1, e_2\rangle\right) = \langle e_1, e_2\rangle \; g(e_i) = e_i, i > 2\}.$$

It follows that $nBn^{-1} \cap GL_2(K)$ is the stabilizer of the line spanned by either e_1 or e_2, whichever appears first in the list $n(e_1), \ldots, n(e_n)$. This means that the matrices of $nBn^{-1} \cap GL_2(K)$ will have the following form: they consist of a 2×2 block which is either upper or lower triangular, followed by an $(n-2) \times (n-2)$ block with 1's on the main diagonal and zeroes everywhere else.

Looking at the elementary matrices of $nBn^{-1} \cap GL_2(K)$, when the 2×2 block is upper triangular, the column operation we get is to add a multiple of the first column to the second column, and when the 2×2 block is lower triangular, we can add a multiple of the second column to the first. In the first case, we use the (2,1)-entry to transform the (2,2)-entry in zero, reducing the matrix to n_i. In the second case, we use the (2,2)-entry to clear the (2,1)-entry and reduce the matrix to the identity. In each case, (BN4) is verified.

Finally, (BN5) is straightforward to verify, because any matrix of $n_i B n_i$ has the form

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

which does not belong to B, the subgroup of upper triangular matrices.

The construction above would also be true, as we remarked, if instead of $GL_n(K)$ we considered the groups $PGL_n(K)$, $SL_n(K)$ or $PSL_n(K)$.

Our first result is to show that Chevalley groups in fact admit (B, N)-pairs.

Proposition 5.3.22. *The Chevalley group $G = \mathcal{L}(K)$ has a (B, N)-pair.*

Proof. We are going to consider as B the subgroup UH and as N the subgroup with the same name from the previous section. Axioms (BN2) and (BN3) are clearly satisfied, as a consequence of Theorem 5.3.19. Now, G is generated by the root subgroups X_r. If s is not a positive root, contained in B, then there is an element $w \in W$ and an $r \in \Pi$ such that $w(r) = s$, by Proposition 5.1.2(iv). If $n \in N$ maps into w via the natural homomorphism, then $nX_rn^{-1} = X_{w(r)} = X_s$, proving that B and N in fact generate G, verifying (BN1). Also, since the n_r are mapped into the fundamental reflections in W, $n_rX_rn_r^{-1} = X_{-r}$ is not contained in B, and (BN5) holds. The proof of axiom (BN4) is more technical and can be found in [4, Corollary 8.1.6]. Therefore, G has a (B, N)-pair. □

We now derive some consequences of the axioms for a (B, N)-pair. The next results hold for arbitrary groups admitting (B, N)-pairs, not just in the context of Chevalley groups. The only result that uses the properties of Chevalley groups is item (iii) of Theorem 5.3.25. The proof of the next result can be found in [4, Proposition 8.2.2].

Theorem 5.3.23. *Let G be a group with a (B, N)-pair. Then*

(i) *$G = BNB$.*

(ii) *For each subset J of I, let W_J be the subgroup of W generated by the elements w_i for $i \in J$ and N_J be the full inverse image of W_J under the natural homomorphism. Then $P_J = BN_JB$ is a subgroup of G.*

It is evident from Theorem 5.3.23 that every double coset BgB contains an element from N. We now consider the question of whether two elements of N lie in the same double coset.

Proposition 5.3.24. *Let G be a group with a (B, N)-pair. Let n, n' be elements of N. Then $BnB = Bn'B$ if and only if n and n' map to the same element of W under the natural homomorphism of N into W. Thus, there is a natural 1-to-1 correspondence between double cosets of B in G and elements of W.*

The proof of this result can be found in [4, Proposition 8.2.3]. The axioms of (B, N)-pairs, the property that the fundamental reflections generate W and induction on the least number of fundamental reflections whose product equals n are used in the proof of the result.

Now, let us explore a little more the subgroups P_J defined in Theorem 5.3.23. We will say that a subgroup of G is *parabolic* if it contains some conjugate gBg^{-1} of B. The subgroups P_J are examples of parabolic subgroups of G, and in the next result we find exactly which subgroups of G are parabolic.

Theorem 5.3.25. *Let G be a group with a (B, N)-pair. Then the following hold*

(i) *The subgroups P_J, $J \subseteq I$ are the only subgroups of G containing B.*

(ii) *Each P_J is equal to its own normalizer; moreover distinct subgroups P_J and P_K cannot be conjugate in G, for $J, K \subseteq I$.*

(iii) *If G is a Chevalley group with associated fundamental system of roots Π, the minimal parabolic subgroups have the form $P_{\{r\}} = B \cup Bn_rB$, for any $r \in \Pi$.*

The proofs can be found in [4], Theorems 8.1.4, 8.3.2 and 8.3.3, respectively. Now, we apply the theory just described to understand better the structure of Chevalley groups, in search of a normal form for the elements of G.

Let G be, as usual, the Chevalley group $\mathcal{L}(K)$. Since G admits a (B, N)-pair, every element of G can be written in the form b_1nb_2, where $b_1, b_2 \in B$ and $n \in N$, but this expression is not unique. What we will do now is to search for an expression such that every element of G can be written uniquely in that form. For this sake, take w a fixed element of W and consider the following subsets of Φ^+

$$\Psi_1 = \{r \in \Phi^+ \mid w(r) \in \Phi^+\},$$
$$\Psi_2 = \{r \in \Phi^+ \mid w(r) \in \Phi^-\}.$$

Then Ψ_1, Ψ_2 are disjoint subsets whose union is Φ^+. We define

$$U_w^+ = \prod_{r \in \Psi_1} X_r$$
$$U_w^- = \prod_{r \in \Psi_2} X_r,$$

where the products are taken over the roots in increasing order. It is important to note that, if $r, s \in \Psi_k$, $k = 1, 2$, then the linear combinations $ir + js$ that are also roots all belong to the same Ψ_k. The commutator formula ensures that U_w^+ and U_w^- both are subgroups of G. We also have $U = U_w^+ U_w^-$ and $U_w^+ \cap U_w^- = 1$. Now we can describe the required canonical form for elements of G.

Theorem 5.3.26. *For each $w \in W$ choose a coset representative $n_w \in N$ which maps to w under the natural homomorphism. Then each element g of G is expressible in just one way in the form*

$$g = bn_w u,$$

where $b \in B$ and $u \in U_w^-$.

Proof. We are going to show that $n_w U_w^+ n_w^{-1} \subseteq U$. By Lemma 5.3.18 we have

$$n_r X_s n_r^{-1} = X_{w_r(s)},$$

for all $r, s \in \Phi$. Let $w = w_{r_1} w_{r_2} \cdots w_{r_k}$ be an expression of w as a product of reflections. Then n_w and $n_{r_1} n_{r_2} \cdots n_{r_k}$ both have image w under the natural homomorphism, hence

$$n_w = hn_{r_1} n_{r_2} \cdots n_{r_k}$$

for some $h \in H$. Thus

$$n_w X_s n_w^{-1} = hn_{r_1} n_{r_2} \cdots n_{r_k} X_s n_{r_k}^{-1} \cdots n_{r_2}^{-1} n_{r_1}^{-1} h^{-1} = hX_{w(s)} h^{-1} = X_{w(s)},$$

again by Lemma 5.3.18. It follows that

$$n_w U_w^+ n_w^{-1} = n_w \cdot \prod_{s \in \Psi_1} X_s \cdot n_w^{-1} \subseteq U. \tag{5.3.1}$$

Of course one can also prove $n_w U_w^- n_w^{-1} \subseteq V$ using the same ideas. Now, consider the double coset Bn_wB. We have

$$\begin{aligned} Bn_wB =& Bn_wHU = Bn_wHU_w^+U_w^- \\ &= BHn_wU_w^+U_w^- \subseteq BUn_wU_w^- = Bn_wU_w^-. \end{aligned}$$

But clearly $Bn_wU_w^- \subseteq Bn_wB$, so we have equality. Since every element g of G lies in some double coset, g can be expressed in the desired form.

The proof for unicity can be found in [4, Theorem 8.4.3]. □

It is important to say that this decomposition allows us to compute the order of Chevalley groups, but this computation depends on knowing the minimal length of each $w \in W$ written in terms of the fundamental reflections. Also, this canonical form can be used to prove the following useful result.

Corollary 5.3.27. *$G \cap \hat{H} = H$.*

We end this section quoting a result, Theorem 11.1.1 in [4], stating that our Chevalley groups are indeed simple. Remember here that the quotient N/H is generated by a set of involutions w_i indexed by some set I.

Theorem 5.3.28. *Let G be a group admitting a (B, N)-pair satisfying the following conditions:*

(i) $G = G'$,

(ii) B is soluble,

(iii) The intersection of all conjugates of B in G is trivial,

(iv) The set I cannot be decomposed into two non-empty complementary subsets J, K such that w_j commutes with w_k for all $j \in J$, $k \in K$.

Then G is simple.

In some sense, groups with (B, N)-pairs have some tendency to be simple, this being the case with the Chevalley groups defined so far, except for the cases $A_1(\mathbb{F}_2), A_1(\mathbb{F}_3), B_2(\mathbb{F}_2)$ and $G_2(\mathbb{F}_2)$. The proof of this result can be found in [4, Theorem 11.1.2].

5.3.4 Automorphisms of Chevalley Groups

At this point, we can finally start discussing automorphisms of finite simple groups of Lie type, which was our main goal since the beginning of the chapter. In general, we do not have a standard presentation for finite simple groups of Lie type, but the next proposition gives sufficient conditions for a bijective map of G onto itself to be an automorphism. The proof of the following result is the content of Section 12.1 in [4].

Proposition 5.3.29. *Let $G = \mathcal{L}(K)$ be a Chevalley group. A bijective map of G onto itself is an isomorphism provided it preserves the following relations*

$\mathcal{R}_1$. $x_r(t_1)x_r(t_2) = x_r(t_1 + t_2)$

$\mathcal{R}_2$. $[x_s(u), x_r(t)] = \prod\limits_{i,j>0} x_{ir+js}(C_{ijrs}(-t)^i u^j)$

$\mathcal{R}_3$. $h_r(t_1)h_r(t_2) = h_r(t_1t_2), t_1, t_2 \neq 0,$

where $h_r(t) = n_r(t)n_r(-1)$ and $n_r(t) = x_r(t)x_{-r}(-t^{-1})x_r(t)$.

We want to describe some particular kinds of automorphisms now.

Diagonal Automorphisms. It was shown in Section 5.3.2 that G is normalized by $\hat{H}$ in the group of all automorphisms of $\mathcal{L}_K$. Thus, if $h(\chi) \in \hat{H}$, conjugation by $h(\chi)$ induces an automorphism of G. If $h(\chi)$ belongs to $\hat{H}$ but not to H, then this automorphism is called a diagonal automorphism.

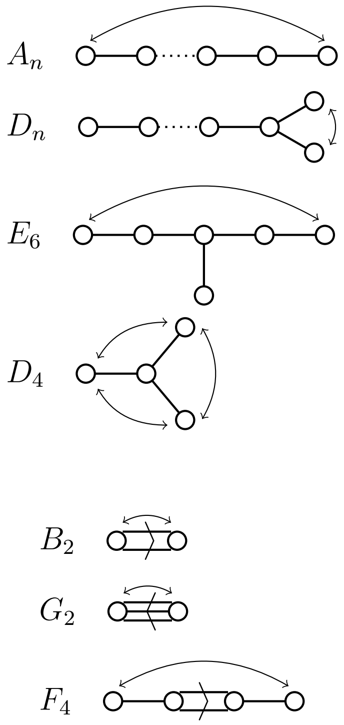
Field Automorphisms In our case, K is always assumed to be a finite field. If $|K| = p^b$, then the group of field automorphisms of K has order b, is cyclic and generated by the automorphism $x \mapsto x^p, x \in K$. If f is an automorphism of the field K, then the map

$$x_r(t) \mapsto x_r(f(t)),\ r \in \Phi, t \in K$$

can be extended to an isomorphism of G. To verify this fact, we just apply Proposition 5.3.29: relation $\mathcal{R}_1$ is preserved trivially, the fact that the C_{ijrs} all belong to the prime field of K, fixed by f, ensures $\mathcal{R}_2$ is preserved and, since $h_r(t)$

can be written in terms of $x_r(t)$, $x_{-r}(-t^{-1})$ and $x_r(-1)$, then $\mathcal{R}_3$ is also preserved. The automorphisms obtained in this way are called field automorphisms of G.

Graph Automorphisms Automorphisms of this type arise from symmetries of the Dynkin diagram. A symmetry of the Dynkin diagram of $\mathcal{L}$ is an automorphism of it as a graph, not considering the arrow present in some diagrams, as in the cases B_2, G_2 and F_4. The non-trivial symmetries of the connected Dynkin diagrams are indicated in the figure below.



However, there are some conditions under which the quoted symmetries give rise to automorphisms of G, and this is the content of the three following propositions. The proofs can be found in [4], Propositions 12.2.3, 12.3.3 and 12.4.1, respectively.

Proposition 5.3.30. *Suppose $\mathcal{L}$ is a simple Lie algebra whose roots all have the same length, and let $r \to \bar{r}$ be a map of Φ into itself arising from a symmetry of the Dynkin diagram of $\mathcal{L}$. Then there exist numbers $\gamma_r = \pm 1$ such that the map $x_r(t) \to x_{\bar{r}}(\gamma_r t)$ can be extended to an automorphism of G. The γ_r can be chosen so that $\gamma_r = 1$ if $r \in \Pi$ or $-r \in \Pi$.*

Automorphisms arising in this way have order 2 when $\mathcal{L} = A_l$ for $l \geq 2$, D_l for $l \geq 4$ and $E_6(K)$. We also obtain automorphisms of order 3 of $D_4(K)$.

Proposition 5.3.31. *Suppose G is a Chevalley group of type $B_2(K)$ or $F_4(K)$, where K is a finite field of characteristic 2. For each root $r \in \Phi$ define $\lambda(r)$ to be 1 if r is a short root and 2 if r is long. Then the map*

$$x_r(t) \to x_{\bar{r}}(t^{\lambda(\bar{r})}), \quad r \in \Phi, t \in K,$$

can be extended to an automorphism of G.

Proposition 5.3.32. *Let $G = G_2(K)$, where K is a finite field of characteristic 3. For each root $r \in \Phi$ define $\lambda(r)$ to be 1 if r is short and 3 if r is long. Then, the structure constants $N_{r,s}$ of G_2 can be chosen in such a way that*

$$x_r(t) \to x_{\bar{r}}(t^{\lambda(\bar{r})}) \quad r \in \Phi, t \in K$$

can be extended to an automorphism of G.

The proofs of Propositions 5.3.31 and 5.3.32 follow the same steps, the second one being more technical since the structure constants of $\mathcal{L}$ have to be chosen in a specific way in order to make relation $\mathcal{R}_2$ hold. We point out that the conditions in the previous two propositions may seem artificial, but in the proof of the next theorem we are going to see that perfect fields of characteristic 2 and 3 arise naturally in the construction of the automorphisms of a Chevalley group. Computation shows that automorphisms of the two last types above described have order $2b$.

Let us move finally to the main result of this chapter. The automorphisms described so far, the diagonal, field and graph automorphisms, together with the inner ones, are sufficient to generate the whole automorphism group of $G = \mathcal{L}(K)$. The main argument of the proof consists in "deconstructing" an arbitrary automorphism of G, making it fix each time more elements or specific subgroups of G, until it reaches the identity. The proof of this result is due do R. Steinberg.

Theorem 5.3.33. *Let $G = \mathcal{L}(K)$ when $\mathcal{L}$ is simple and $K = GF(q)$, where*

$q = p^b$. Let θ be an automorphism of G. Then, there exist inner, diagonal, graph and field automorphisms i, d, g, f of G such that $\theta = idgf$.

Proof. We know that U and V are Sylow p-subgroups of G, by Theorem 5.3.8, and any two Sylow p-subgroups of G are conjugate. Now $\theta(U)$ is also a Sylow p-subgroup of G, so it is G-conjugate to U. Thus there exists an inner automorphism i_1 of G such that $\theta(U) = i_1(U)$. Let $\theta_1 = i_1^{-1}\theta$. Then $\theta_1(U) = U$ and, as $\theta_1(V)$ is G-conjugate to U, being a Sylow p-subgroup of G, there exists an element $x \in G$ such that $\theta_1(V) = x^{-1}Ux$. However, x can be written in the form $x = bn_wu$, where $b \in B = UH$, $n_w \in N$, $u \in U_w^-$, as in Theorem 5.3.26. Since b normalizes U by Lemma 5.3.14 we have

$$\theta_1(V) = u^{-1}n_w^{-1}Un_wu.$$

Now $U \cap V = 1$, and so $\theta_1(U) \cap \theta_1(V) = 1$. Hence $U \cap u^{-1}n_w^{-1}Un_wu = 1$, and it follows that $n_wUn_w^{-1} \cap U$ and $n_w^{-1}Un_w \cap U$ both equal 1. By Lemma 5.3.18, this implies that w transforms every positive root into a negative one, because otherwise we would have $n_wUn_w^{-1} \cap U \neq 1$. Therefore,

$$n_wUn_w^{-1} = V,$$

and it follows that $\theta_1(V) = u^{-1}Vu$. Let i_2 be the inner automorphism of G taking $x \in G$ to $u^{-1}xu$. Thus, if $\theta_2 = i_2^{-1}\theta_1$, then $\theta_2(U) = U$ and $\theta_2(V) = V$.

We have that $N_G(U) = UH$, by [4, Theorem 8.5.2(iii)], and similarly we have $N_G(V) = VH$. Since U and V are θ_2-invariant, their normalizers also are, and also is $UH \cap VH = H$, by Corollary 5.3.16. Since θ_2 leaves $B = UH$ invariant, it permutes the minimal parabolic subgroups containing B. By Theorem 5.3.25(iii), these have the form

$$P_{\{r\}} = B \cup Bn_rB, \quad r \in \Pi.$$

We are going to show that $P_{\{r\}} \cap V = X_{-r}$. We have

$$\begin{aligned} P_{\{r\}} &= B \cup Bn_rB \\ &= (B \cup Bn_rB)n_r^{-1}, \end{aligned}$$

since $B \cup Bn_rB$ is a subgroup of G containing n_r, and

$$\begin{aligned}(B \cup Bn_rB)n_r^{-1} =& Bn_r^{-1} \cup Bn_rBn_r^{-1} \\ =& Bn_r^{-1} \cup BX_{-r}.\end{aligned}$$

Consider the subset $Bn_r^{-1} \cap V$. If we take w as before, then $n_wUn_w^{-1} = V$. Then we have

$$\begin{aligned}Bn_r^{-1} \cap V &= Bn_r^{-1} \cap n_wUn_w^{-1} \\ &\subseteq (Bn_r^{-1}n_w \cap n_wU)n_w^{-1} \\ &\subset (Bn_r^{-1}n_wB \cap Bn_wB)n_w^{-1}.\end{aligned}$$

However, $Bn_r^{-1}n_wB \cap Bn_wB = \emptyset$ by Theorem 5.3.24, since the double coset representatives clearly map into different elements of W. Thus $Bn_r^{-1} \cap V = \emptyset$ also. Hence

$$P_{\{r\}} \cap V = BX_{-r} \cap V = (B \cap V)X_{-r} = X_{-r},$$

using Dedekind modular laws and since $B \cap V = 1$, by Lemma 5.3.15.

Now θ_2 maps V into itself and permutes the subgroups $P_{\{r\}}$ for the fundamental roots $r \in \Pi$. Thus θ_2 permutes the subgroups X_{-r} for $r \in \Pi$. In a similar way, it is also possible to show that θ_2 permutes the subgroups X_r for $r \in \Pi$. However, if r, s are distinct fundamental roots, the commutator formula, Theorem 5.3.4, implies that $[X_r, X_{-s}] = 1$, since $r - s$ is not a root, whereas we also know that the commutator $[X_r, X_{-r}] \neq 1$. Thus, if $r, s \in \Pi$, $[X_r, X_{-s}] = 1$ if and only if $r \neq s$. Since this relation is preserved by θ_2 we have, for $r \in \Pi$,

$$\begin{aligned}\theta_2(X_r) &= X_{\rho(r)} \\ \theta_2(X_{-r}) &= X_{-\rho(r)},\end{aligned}$$

where ρ is a permutation of Π.

Let $\theta_2 \cdot x_r(1) = x_{\rho(r)}(t_r)$, and let χ be the K-character of P defined by $\chi(\rho(r)) = t_r$ with $r \in \Pi$. Let d be the diagonal automorphism of G mapping $x \in G$ to $h(\chi)xh(\chi)^{-1}$. Then we have

$$\theta_2 \cdot x_r(1) = d \cdot x_{\rho(r)}(1).$$

Let $\theta_3 = d^{-1}\theta_2$. Then

$$\theta_3 \cdot x_r(1) = x_{\rho(r)}(1).$$

We shall show that

$$\theta_3 \cdot x_{-r}(1) = x_{-\rho(r)}(1).$$

Let $\theta_3 \cdot x_{-r}(1) = x_{-\rho(r)}(\lambda)$. We prove that $\lambda = 1$ using the homomorphism from $SL_2(K)$ into $\langle X_r, X_{-r}\rangle$, defined in Theorem 5.3.10. We have

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

and it follows that

$$x_r(1)x_{-r}(-1)x_r(1) = x_{-r}(-1)x_r(1)x_{-r}(-1).$$

Applying θ_3, and remembering that $x_s(t)^{-1} = x_s(-t)$ we obtain

$$x_{\rho(r)}(1)x_{-\rho(r)}(-\lambda)x_{\rho(r)}(1) = x_{-\rho(r)}(-\lambda)x_{\rho(r)}(1)x_{-\rho(r)}(-\lambda).$$

However

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 2-\lambda \\ -\lambda & 1-\lambda \end{pmatrix},$$

and

$$\begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 1 \\ \lambda^2 - 2\lambda & 1-\lambda \end{pmatrix}$$

and since the kernel of the homomorphism $SL_2(K) \to \langle X_r, X_{-r}\rangle$ contains only I_2 and possibly $-I_2$, we have $\lambda = 1$ by comparing the above matrices, regardless of the field K. Thus

$$\theta_3 \cdot x_r(1) = x_{\rho(r)}(1),$$
$$\theta_3 \cdot x_{-r}(1) = x_{-\rho(r)}(1)$$

and it follows that

$$\theta_3(n_r) = \theta_3(x_r(1)x_{-r}(-1)x_r(1)) = n_{\rho(r)}.$$

We show now that the permutation ρ induces a symmetry of the Dynkin diagram. Since $\theta_3(H) = H$ and $\theta_3(n_r n_s) = n_{\rho(r)}n_{\rho(s)}$, the order of the coset

$n_r n_s H$ is the same as the order of $n_{\rho(r)} n_{\rho(s)} H$ in the group N/H. Using the isomorphism between N/H and W, the order of $w_r w_s$ is the same as the order of $w_{\rho(r)} w_{\rho(s)}$ in W. If we recall that this order, which we called m_{rs}, is related to the angle θ between r and s via $\theta = 2\pi/m_{rs}$, and that the number of bonds joining r and s is precisely $4\cos^2(\theta)$, this implies that the number of edges joining the roots r and s is the same as the number of bonds that joins $\rho(r)$ and $\rho(s)$. Hence ρ is a symmetry of the Dynkin diagram, not regarding the direction of the arrow on cases B_2, G_2 and F_4.

We prove next that there is a graph automorphism g of G such that $g(X_r) = X_{\rho(r)}$ for all $r \in \Pi$. By Theorems 5.3.30, 5.3.31 and 5.3.32, this happens in the case where every root has the same length, but it only happens provided K has characteristic 2 if $\mathcal{L} = B_2$ or F_4 or K has characteristic 3 if $\mathcal{L} = G_2$. Also, ρ cannot be the identity in any of the previous cases.

Suppose ρ is the nontrivial symmetry of the Dynkin diagram of $\mathcal{L}$, where $\mathcal{L} = B_2$ or F_4. In each case, there exist roots a and b that are interchanged by ρ such that the roots that are linear combinations of a and b form a root system of type B_2. Indeed, the case $\mathcal{L} = B_2$ is straightforward to see, and in case $\mathcal{L} = F_4$, we take a and b to be the roots connected by a double edge and consider the root system generated by them; its Dynkin diagram consists precisely of the Dynkin diagram of type F_4, removing the nodes that do not correspond to a and b and all edges inciding on those. In any case, the roots that are linear combinations of a and b are

$$\pm a, \pm b, \pm(a+b), \pm(2a+b).$$

We now use equation (5.1.1) to compute the images of some roots through elements of the Weyl group. We have that $n_b X_a n_b^{-1} = X_{w_b(a)} = X_{a+b}$. Applying θ_3 we have $n_a X_b n_a^{-1} = \theta_3(X_{a+b})$. Since $n_a X_b n_a^{-1} = X_{2a+b}$, we have $\theta_3(X_{a+b}) = X_{2a+b}$. However, $[X_{2a+b}, X_b] = 1$, and applying θ_3^{-1}, we obtain $[X_{a+b}, X_a] = 1$. But

$$[x_{a+b}(1), x_a(1)] = x_{2a+b}(-N_{a,a+b}).$$

Thus $N_{a,a+b} = 0$ in this case, and as $N_{a,a+b} = \pm 2$, K has characteristic 2.

Now suppose $\mathcal{L} = G_2$ and ρ is the nontrivial symmetry of the Dynkin diagram of $\mathcal{L}$. Then ρ interchanges the fundamental roots a and b of $\mathcal{L}$ and the other roots of Φ are

$$\pm a, \pm b, \pm(a+b), \pm(2a+b), \pm(3a+b), \pm(3a+2b).$$

Now $n_b X_a n_b^{-1} = X_{a+b}$, and so applying θ_3 we have $n_a X_b n_a^{-1} = \theta_3(X_{a+b})$. Thus $\theta_3(X_{a+b}) = X_{3a+b}$. Also, $n_a X_{a+b} n_a^{-1} = X_{2a+b}$ and, applying θ_3, we get $\theta_3(X_{2a+b}) = n_b X_{3a+b} n_b^{-1} = X_{3a+2b}$. Now $[X_{3a+2b}, X_b] = 1$. Applying θ_3^{-1} gives $[X_{2a+b}, X_a] = 1$. But

$$[x_{2a+b}(1), x_a(1)] = x_{3a+b}(-N_{a,2a+b}).$$

Again, $N_{a,2a+b}$ must be zero, but since $N_{a,2a+b} = \pm 3$ in this case, we conclude that K has characteristic 3.

Thus in each case there is a graph automorphism g of G such that $g(X_r) = X_{\rho(r)}$ for every $r \in \Pi$. Let $\theta_4 = g^{-1}\theta_3$. Then $\theta_4(X_r) = X_r$ and $\theta_4(X_{-r}) = X_{-r}$ for all r in Π. Also, if we consider the possible actions of g, as in Theorems 5.3.30, 5.3.31 and 5.3.32, it is clear that θ_4 also fixes $x_r(1)$ and $x_{-r}(1)$, fixing also n_r for all $r \in \Pi$. It follows that θ_4 fixes each X_r, $r \in \Phi$. For, if $r = w(s)$, $w \in W$ and $s \in \Pi$, via Proposition 5.1.2(v), and $w = w_{r_1} \cdots w_{r_k}$, then

$$X_r = n_{r_1} \cdots n_{r_k} X_s n_{r_k}^{-1} \cdots n_{r_1}^{-1},$$

and $\theta_4(X_r) = X_r$.

Let $r, s \in \Pi$ be fundamental roots which are joined in the Dynkin diagram. Then $r + s \in \Phi$. Here, since θ_4 fixes the root subgroups, then θ_4 induces a map from K to itself, for each root in Φ. Let

$$\theta_4 \cdot x_r(t) = x_r(f(t)),$$
$$\theta_4 \cdot x_s(t) = x_s(g(t)),$$
$$\theta_4 \cdot x_{r+s}(t) = x_{r+s}(h(t)).$$

Since

$$[x_s(u), x_r(t)] = x_{r+s}(-N_{r,s}tu) \cdots ,$$

we have, on applying θ_4,

$$[x_s(g(u)), x_r(f(t))] = x_{r+s}(h(-N_{r,s}tu)) \cdots .$$

But

$$[x_s(g(u)), x_r(f(t))] = x_{r+s}(-N_{r,s}f(t)g(u)) \cdots ,$$

and so $h(-N_{r,s}tu) = -N_{r,s}f(t)g(u)$. Now, $N_{r,s} = \pm 1$, giving $h(tu) = f(t)g(u)$ for all $t, u \in K$. As $f(1) = g(1) = 1$, we have that $h(t) = f(t)g(1) = f(t)$ and $h(u) = f(1)g(u) = g(u)$, for all $t, u \in K$. Hence, $f(t) = g(t) = h(t)$ for all $t \in K$. We also have $f(tu) = f(t)f(u)$, for all $t, u \in K$. If we consider $x_r(t+u) = x_r(t)x_r(u)$ and apply θ_4, we obtain

$$x_r(f(t+u)) = x_r(f(t))x_r(f(u)) = x_r(f(t) + f(u)),$$

and then f is additive. Since f is a bijection from K to K fixing both 0 and 1, additive and multiplicative, then f is a field automorphism of K, and it can be extended to an automorphism f of G. Let $\theta_5 = f^{-1}\theta_4$. then

$$\theta_5 \cdot x_r(t) = x_r(t)$$

for every generator $x_r(t)$ of G, and hence θ_5 must be the identity of G. Now

$$\theta_5 = f^{-1}g^{-1}d^{-1}i_2^{-1}i_1^{-1}\theta = 1.$$

Therefore $\theta = i_1 i_2 d g f$ and the theorem is proved. □

We note here that the automorphism d of the proof may or may not be in $\hat{H}\backslash H$. If the former case happens, then it is a diagonal automorphism in the sense we defined it. If the latter occurs, then the product $i_1 i_2 d$ is an inner automorphism of G, since d would also be, in this case. We also note that the proof that the function f satisfies $f(t_1)f(t_2) = f(t_1 t_2)$ breaks down in the case $L = A_1$, since we cannot choose two distinct fundamental roots. It is possible to give an alternative argument to cover this case too, and the theorem is valid anyway.

We end this section with the following remark. The CFSG claims that a finite non-abelian simple group is isomorphic to an alternating group $A_n, n \geq 5$, a sporadic group or a simple group of Lie type. We have constructed by now almost every family of simple groups of Lie type apart from the so-called twisted simple groups. The twisted groups are obtained as subgroups of certain Chevalley groups $G = \mathcal{L}(K)$, specifically for those such that the Dynkin diagram of $\mathcal{L}$ admits a non-trivial symmetry. The vast majority of the results proved for the non-twisted simple groups also hold for the twisted ones, with the appropriate modifications on the statements. They also admit a normal form similar to the one in Theorem 5.3.26, and their automorphism group also admits a normal form similar to the one in Theorem 5.3.33. The only difference is that the automorphisms of twisted simple groups decompose in the product of inner, diagonal and field automorphisms, the graph automorphism is always trivial.

Chapter 6

Main Results

The main goal of this chapter is to prove the two main resuts in Shalev's paper [27] and the main result of Khukhro in [17], which contains an improvement of Shalev's bound on the derived length. In Section 6.1, we deal with the proof of the following result.

Theorem 6.0.1 (Shalev). *Let G be a finite group of rank r admitting an automorphism with m fixed points. Then G has a characteristic soluble subgroup H whose index is (m, r)-bounded.*

The remainder of the chapter will be devoted to the proof of the following theorem.

Theorem 6.0.2 (Shalev). *Let G be a finite soluble group of rank r admitting a coprime automorphism ϕ. If $|C_\phi(G)| = m$, then the derived length of G is (m, r)-bounded.*

In Section 6.2 we begin an approach that reduces the proof of Theorem 6.0.2 to the class of p-groups. We use the ideas of Section 4.2 and some new ones in Section 6.3 to bound the derived length of some Lie rings associated to G by a function that depends only on the rank of G, in the case where ϕ is regular. In Section 6.4 we deal with the case where ϕ has few fixed points, and we bound the derived length with a function depending on the rank of G and the number of fixed points of ϕ.

6.1 First Part of the Main Theorem

The purpose of this section is to prove Theorem 6.0.1. We start with a result concerning the factor-group $Out(G) = Aut(G)/Inn(G)$ when G is a finite simple group of Lie type. Here, the Lie rank of $G = \mathcal{L}(K)$ is the rank of the root system associated to $\mathcal{L}$.

Theorem 6.1.1. *Let $G = \mathcal{L}(K)$ be a finite simple group of Lie type, with Lie rank l, over the field with p^b elements. Then the order of $Out(G)$ is (b, l)-bounded.*

Proof. Invoking Theorem 5.3.33, we know that the inner, diagonal, graph and field automorphisms of G are sufficient to generate every automorphism of G, both in the twisted and untwisted cases. Furthermore, if I, F denote the subgroups of inner and field automorphisms respectively, and if D, Γ denote the sets (Γ possibly consisting only of the identity map, as in some untwisted and all twisted cases) of diagonal and graph automorphisms of G respectively, then $Aut(G)$ equals the product $ID\Gamma F$. Since G is simple, $I \cong G$, and the group of outer automorphisms has order $|Aut(G)|/|G| = |ID\Gamma F|/|G|$. But we also know that $|ID\Gamma F| \leq |I||D||\Gamma||F|$, hence $|Out(G)| \leq |D||\Gamma||F|$.

Here, $|F| = b$. The order of Γ is 2 in case $G = A_l(G), l \geq 2$, $D_l, l \geq 5$, $E_6(K)$; it is $2b$ in case $G = B_2(K), F_4(K)$ or $G_2(K)$, and 6 in case $G = D_4(K)$, since the symmetries of the Dynkin diagram of $\mathcal{L} = D_4$ form the dihedral group of order 6. It remains then to bound the order of D in terms of l and b. Remember that D consists precisely of diagonal automorphisms induced by elements of $\hat{H}$ that are not contained in H, since $G \cap \hat{H} = H$ by Corollary 5.3.27. Also, $\hat{H}$ is an abelian group, since it is isomorphic to the group of K-characters of $P = \mathbb{Z}\Phi$. Then, if we take fixed representatives of right cosets modulo H, they will be enough to produce the elements of D, since the elements of H induce inner automorphisms, and not diagonal ones. So, we only need to consider a set of coset representatives of H in $\hat{H}$ to compute the upper bound for the order of

$Aut(G)$. We need to consider the possible values for $[\hat{H} : H]$, and these indices are listed in Table 5.2, at page 63. They are clearly bounded by $\max\{l+1, 4\}$, and then $|D|$ is l-bounded. This completes the proof of the result. □

We also need the following two results, both due to B. Hartley.

Theorem 6.1.2. *Let G be a finite non-abelian simple group admitting an automorphism of order n with at most k fixed points. Then $|G|$ is (n, k)-bounded.*

The result above is Theorem A' in [9] which gives a bound on the order of a finite simple group that depends on the order and number of fixed points of an automorphism, resembling the well-known Brauer-Fowler Theorem in [2]. In sharp contrast to this theorem, which is proved using character theory and does not depend on the CFSG, Hartley's result relies heavily on the classification, and at a certain point he even has to invoke specific calculations made for particular families of simple groups.

We remark that Hartley's approach is different from the one given in this chapter, since in [9] he studies finite simple groups as subgroups of certain algebraic groups. Namely, the finite simple groups of Lie type are subgroups of the centralizer of a Frobenius automorphism on a connected reductive algebraic group over an algebraically closed field. In [9] he does not deal with inner automorphisms, but he does so in [10]. In some sense, part of his proof is essentially a survey on similar results due to various authors, and he completes the proof with the cases not covered before. Then, using some algebraic geometry and properties of the Frobenius automorphism, he manages to prove the result.

The second result of B. Hartley is as follows.

Lemma 6.1.3. *Let $S = \mathcal{L}(K)$ be a simple group of Lie type, where $K = GF(q), q = p^b$ and let l be the Lie rank of S. Suppose S admits an automorphism ϕ with m fixed points. Then, the order of S is (b, m, l)-bounded.*

Proof. We know, by Theorem 6.1.1, that the order of $Out(S)$ is (b, l)-bounded. In particular, $\phi^e \in S$, which we identify with $Inn(S)$, for some e which is (b, l)-bounded. Then, $\langle \phi^e \rangle \subseteq C_S(\phi)$. As $|C_S(\phi)| = m$, $\phi^{em} = 1$, hence the order of ϕ is (b, m, l)-bounded. By Theorem 6.1.2, there is a bound to the order of a simple group admitting an automorphism of a given order with a given number of fixed points, depending on both values. Thus the order of S is (b, m, l)-bounded, as desired. □

We are now ready to prove Theorem 6.0.1.

Proof. We want to show that, if G is a finite group of rank r admitting an automorphism ϕ with m fixed points, then G has a characteristic soluble subgroup of (m, r)-bounded index.

When we have $H \operatorname{char} K \operatorname{char} G$, the factor group K/H is called a *characteristic section* of G. Let M be a characteristic section of G which is characteristically simple. We can suppose that M is non-abelian. Indeed, if every nontrivial characteristically simple characteristic section were abelian, since G is finite we can consider a series of characteristic subgroups

$$1 = G_0 \operatorname{char} G_1 \operatorname{char} \ldots \operatorname{char} G_k = G$$

with no characteristic subgroup of G strictly contained between G_i and G_{i+1}, for $i = 0, \ldots, k-1$. Then, the factors of this series are characteristically simple, and thus abelian, by assumption. Then this would be a soluble series for G. Also, G itself is a characteristic section, so there is always a nontrivial characteristic section of G.

Since M is characteristically simple, by [25, Theorem 8.10] it is the direct product of isomorphic simple groups S_i, $1 \leq i \leq k$, and as $rk(G) = r$, we have $k \leq r$. Now, the automorphism ϕ acts on the section M with at most m fixed points, by Lemma 2.2.2. Let $i \geq 1$ be the minimal integer such that $S_1^{\phi^i} = S_1$. Then $K := S_1 \times S_1^{\phi} \times \cdots \times S_1^{\phi^{i-1}}$ is an ϕ-invariant subgroup of M. Note that, if $s \in C_{S_1}(\phi^i)$, then $s \to ss^{\phi} \cdots s^{\phi^{i-1}}$ is an injective map from $C_{S_1}(\phi^i)$ to $C_K(\phi)$.

Hence, $|C_{S_1}(\phi^i)| \leq |C_K(\phi)| \leq m$. We conclude that S_1 is a simple group of rank at most r admitting an automorphism with at most m fixed points.

Applying the classification of finite simple groups, if S_1 is sporadic, $|S_1|$ is bounded. If $S_1 \cong A_n$ is alternating, $n \geq 5$, then we can construct an elementary abelian 3-subgroup of S_1, generated by disjoint 3-cycles. This subgroup has rank equal to the integer part of $n/3$, and as this number is bounded by above by r, n is r-bounded. Again, $|S_1|$ is bounded, but this time by a function of r only, namely $(3r+2)!/2$. It remains to tackle the case where S_1 is of Lie type. Suppose $\mathcal{L}$ has root system of rank l, $K = GF(q), q = p^b$ for some b. Since $rk(S_1) \leq r$ and the root subgroups X_r are isomorphic to the additive group of the field K, of rank b, then $b \leq r$.

Also, l is bounded in terms of r. To see this, just consider the subgroup H of G. H is a subgroup of $\hat{H}$, and $\hat{H}$ is the group of homomorphisms from $\mathbb{Z}\Phi$ into K^*. $\mathbb{Z}\Phi$ is the free group with basis Π, of cardinality l, and hence $\hat{H}$ is isomorphic to the direct product of l copies of the cyclic group of order $p^b - 1$. The index of H in $\hat{H}$ is presented on Table 5.2, page 63, for each isomorphism class of $\mathcal{L}$. We see that this index is 1, 2, 3 or 4 according to each case, except when $\mathcal{L}$ is of type A_l. So, in the cases apart from $\mathcal{L} = A_l$, the fact that $[\hat{H} : H]$ equals 1,2,3 or 4 forces the rank of H to be at least $l - 2$. In this case, l is bounded by a function of r only. For the case $\mathcal{L} = A_l$, the Weyl group N/H is isomorphic to S_{l+1}, and as G has rank $\leq r$, S_{l+1} also has rank $\leq r$. So, considering a subgroup of S_{l+1} generated by mutually disjoint 3-cycles, which is elementary abelian of rank equal to the integer part of $(l+1)/3$, we see that in this case l is also bounded by a function of r only.

We also know that the order of S_1 is (b, m, l)-bounded, by Lemma 6.1.3. It is possible to conclude that, in either case, $|S_1|$ is (m, r)-bounded. It follows that $|M|$ is (m, r)-bounded also, for any non-abelian characteristic section of G which is characteristically simple. Note also that this bound is uniform for every such section.

Given such a section M, consider $C_G(M)$. As $G/C_G(M)$ embeds in $Aut(M)$, the index of $C_G(M)$ is (m, r)-bounded. Since G is r-generated, its number of subgroups of index $\leq i$ is bounded in terms of r and i only. Therefore, the number of possibilities for $C_G(M)$ is (m, r)-bounded. Let $H := \cap C_G(M)$ where M ranges over all characteristic sections of G which are non-abelian and characteristically simple. The above discussion shows that $[G : H]$ is (m, r)-bounded. Finally, since every characteristic section of H which is characteristically simple is abelian, by construction, then H must be soluble. □

We would like to point out the quantity of tools and background that are involved in the proof of Theorem 6.0.1. The great generality of this theorem is what justifies the use of so many results.

6.2 Reduction Argument

Here, our purpose is to reduce Theorem 6.0.2 to the case where G is a finite p-group. First, we consider the case where the automorphism ϕ is fixed-point-free. The following result, which is a consequence of the classification of finite simple groups, ensures that under this condition G is soluble. This proof we discuss here is due to Peter Rowley [26].

Theorem 6.2.1. *Let G be a finite group admitting a fixed-point-free automorphism ϕ. Then G is soluble.*

We give an idea of the proof. Rowley begins with a minimal counterexample G to the theorem, that happens to be a finite non-abelian simple group. Here, G admits an abelian group of automorphisms, say $\langle \phi \rangle$. Since ϕ leaves a unique Sylow p-subgroup P of G invariant, if P is cyclic, then the fact that ϕ acts regularly on P together with the abelianity of $Aut(P)$ implies that G has a normal p-complement. Using this result, the author of [26] rules out the cases where G is sporadic, since there is always a prime to the first power diving the orders of such groups. Using Bertrand's Postulate [22, p. 367], that ensures the existence of a prime number between $n/2$ and n, for any natural number n, he deals with the case of the alternating groups. For the case where G is simple of Lie type, Rowley makes G and its automorphism group both act over the "building" associated to G. He previously proves that this action has only one fixed point, but constructs a pair of fixed points for such an action, hence getting a contradiction. This completes his proof. This shows that, in the case where ϕ is regular, we do not need to assume in the statement of Theorem 6.0.2 the further condition of solubility on G. We would like to end this comment on Rowley's paper by giving an example of the so-called "building" associated to a group that is mentioned in the previous paragraph. First of all, we define a *simplicial complex* with vertex set $\mathcal{V}$ as a collection Δ of finite subsets of $\mathcal{V}$, called simplices, such that each singleton $\{v\}$ is a simplex and every subset of a simplex A is again a simplex. Let L consist of exactly l linearly independent vectors on $\mathbb{R}^l$, where $l \geq 1$. Then the convex hull C of L is an analogue of a tetrahedron, but with dimension l. This geometric structure is a simplicial complex, with vertex set being the set of vectors L. Another example of a simplicial complex is the power set of a non-empty set K, where the vertices are just the singletons of K.

Let W be a group, possibly infinite, generated by a subset S consisting of elements of order 2. If W admits a presentation of the form

$$\langle S \mid (st)^{m(s,t)} = 1 \rangle,$$

where there is one relation for each pair $s, t \in S$ with $m(s,t) < \infty$, we call W a *Coxeter group*, following Tits [30]. For example, the symmetric groups S_n have such a presentation, where S is the set of involutions $\{(12), (23), \ldots, (n-1, n)\}$ and the $m((ij),(kl))$ equal 1, 2 or 3 depending on the values i, j, k, l. Also, dihedral groups, including the infinite dihedral group $D_\infty = \langle r, s \mid r^2 = s^2 = 1 \rangle$, are examples of Coxeter groups. Throughout the book [3] one can find many examples of such groups, arising from different areas of group theory; the classification of all finite Coxeter groups may be found in [3, p. 5]. We are going to call a pair (W, S) consisting of a Coxeter group and its set of generating involutions a *Coxeter system.*

Define a *special subgroup* T of W as a subgroup generated by some subset S' of S, $T = \langle S' \rangle$. A *special coset* of W will be, then, a coset $w\langle S' \rangle$ with w in W and $S' \supseteq S$. We are going to define a partially ordered set, poset for short, as follows. Consider $\Sigma(W, S) = \{\text{special cosets of } W\}$, ordered by the opposite of the inclusion relation, which we are going to denote by $\leq$. This means that $A \leq B$ if and only if $A \supseteq B$ as subsets of W. The singletons are the greatest elements inside $\Sigma(W, S)$ and W itself is the smallest one. The poset $\Sigma(W, S)$ is called a *Coxeter complex.* The use of the word complex here is due to the fact that such posets are indeed simplicial complexes, according to [3, Theorem III.1]. We give the lattice of the Coxeter complex associated to the Coxeter group $S_3 = \langle r, s \,|\, r^2 = s^2 = (rs)^3 = 1 \rangle$. Note that here, $(rs)^3 = 1$ implies $srs = rsr$. Also, the $\{r\}$ on the first layer also connects with the $\{1, r\}$ on the second one.

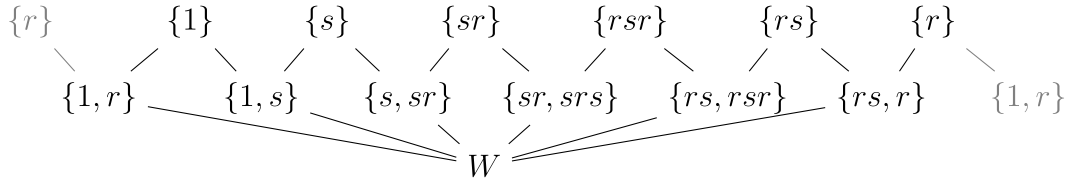


Figure 6.1: Lattice of the poset $\Sigma(S_3, \{r, s\})$

The approach of this example is totally group-theoretic, but a geometric approach is also possible. For example, if we consider inside $\mathbb{R}^3$ the set of hyperplanes $\{x_i = x_j, \,|\, i \neq j \in \{1, 2, 3\}\}$, and consider s_{ij} to be the reflection through the hyperplane $x_i = x_j$, then the reflection group generated by all the s_{ij} is isomorphic to S_3. The three hyperplanes considered above divide $\mathbb{R}^3$ in a geometric structure that can be proved to be isomorphic, in the simplicial complex sense, to $\Sigma(S_3, \{r, s\})$. Also, since S_3 acts on the just given dissecation of the euclidean space, it is also possible to prove that S_3 acts on its Coxeter complex via left translation on the cosets, in the usual way, and that these two actions are compatible with the aforementioned isomorphism. Point-stabilizers of certain subsets of Δ are the special subgroups of S_3, and one can obtain the special cosets by considering the compatible actions of S_3 on both $\Sigma(S_3, \{r, s\})$ and the geometric structure formed in $\mathbb{R}^3$. Note that this idea makes it possible to add some geometric structure to group-theoretic objects, such as the Coxeter complex $\Sigma(S_3, \{r, s\})$.

We can finally give the definition of a building. We call a simplicial complex Δ a *building* provided that Δ is the union of subcomplexes Σ, called *apartments*, satisfying the following:

(B0) Each Σ is a Coxeter complex $\Sigma(W, S)$ associated to some Coxeter group W generated by a set of involutions S;

(B1) For any two simplices A and B in Δ, there is an apartment Σ containing both of them;

(B2) If Σ and Σ' are two apartments containing A and B, then there is an isomorphsim of simplicial complexes $\Sigma \to \Sigma'$ fixing A and B pointwise.

If G is a group admitting a (B, N)-pair, according to Definition 5.3.21, a remarkable example of building is the set

$$\{gP \,|\, g \in G, P \text{ is a parabolic subgroup of } G\},$$

partialy ordered by the opposite of the inclusion relation. The proof that such a poset is a building can be found in [3, Theorem V.3]. We want to remind that when a group admits a (B, N)-pair, the subgroup $B \cap N$ is normal in N and the quotient $N/B \cap N$ is generated by a set of involutions, cf. axiom $(BN3)$ in Definition 5.3.21. So, there is a natural Coxeter group associated to groups having a (B, N)-pair, which is used to construct such buildings. In this case, G acts in a very nice way on Δ, and Rowley explores this action in his argument.

We turn back to our goal, and the next part of the discussion holds regardless of the number of fixed points of ϕ. Here on, we assume that G is always soluble.

Now, fix a prime p dividing $|G|$. If we set $P = O_{p'p}(G)/O_{p'}(G)$, then P is by construction the largest normal p-subgroup of $G/O_{p'}(G)$. Since P is normal in $G/O_{p'}(G)$, the group G acts by conjugation on P, and hence on $P/\Phi(P)$. Then, as $P/\Phi(P)$ is an $\mathbb{F}_p$-vector space of dimension at most r, G being of rank r, we get a representation of G into $GL(r, p)$. The next proposition states which subgroup of G is the kernel of this representation.

Proposition 6.2.2. *Let G and P be as above. Then the kernel of the action of G on $P/\Phi(P)$ is precisely $O_{p'p}(G)$.*

Proof. The kernel K of the considered action certainly contains $O_{p'p}(G)$, and is also a normal subgroup of G. Hence, there are two possibilities: either $K = O_{p'p}(G)$, or there exists an element x of G contained in K with order coprime to p, since $K/O_{p'}(G)$ cannot be a p-subgroup of $G/O_{p'}(G)$, as $O_{p'p}(G)/O_{p'}(G)$ is the largest normal one. Now, by Lemma 2.2.1, x must act trivially on P, because it centralizes $P/\Phi(P)$. However, [8, Lemma 1.2.3], ensures that in this situation the automorphism induced by x is non-trivial. This completes the proof. □

Proposition 6.2.2 will be used in combination with the next theorem, which is due to M.F. Newman, [21].

Theorem 6.2.3. *Let G be a soluble completely reducible subgroup of the linear group $GL(n,p)$. Then, the derived length of G is bounded above by a function ρ of n only, given by $\rho(n) = 5\log_9(n/8) + 8$.*

Since Newman's proof is very technical, we give just an idea of it. The proof uses a triple inductive argument. The author finds bounds for the derived length of soluble subgroups of permutation and linear groups. He even gives a different bound for the case when this subgroup is completely reducible, in the linear case. Newman's proof argues by induction on n, where n denotes, according to the case, either the degree of G or the dimension of the vector space over which G acts. Also, the bounds are best possible: Newman constructs groups of the given types with derived length equal to the given bound.

Let G be a soluble permutation group. In this case, we can have either that G is isomorphic to a subgroup of a wreath product $W\, wr\, K$, where both W and K are soluble permutation groups of well-behaved degrees, or that G has order p^k and admits a minimal normal subgroup A such that G/A is an irreducible linear group of degree k. In both cases, the result follows by induction and by the relations stated in [21, Lemma 1].

In the case where G is a soluble completely reducible linear group with degree n, we also have two possible situations. First, G may admit a normal subgroup N which is isomorphic to a direct product of completely reducible soluble linear groups of degree m. In this case, the factor group G/N is a soluble permutation group of degree n/m and the result follows by induction. In the second case, Newman considers the series of subgroups $G \geq F = Fit(G) \geq Z = Z(G) \geq 1$, where $Fit(G)$ denotes the Fitting subgroup of G. In this case, F/Z is an abelian group of order dividing n^2 and G/F is a subgroup of a linear group. The result also follows by induction.

Finally, let G be a soluble linear group acting on a vector space V. Newman's strategy is to consider a certain series of subspaces of V and let G act on each of the quotient spaces, which have lower degrees in each case. A careful analysis of every possible decomposition shows that in all cases G has n-bounded derived length.

One of the difficulties of this proof is illustrated by the fact that, when G is a completely reducible linear group, the factor-group G/N is permutational, for example. This fact explains why Newman needs to consider a triple inductive argument, since the reductions made on his work for a class of groups lead to factor-groups which do not belong necessarily to the same class. Another important fact to point out is that the functions that he considers as bounds for the derived length all behave well with respect to the addition, so the technique consists in bounding $dl(N) + dl(G/N)$ in each case.

Now, combining Proposition 6.2.2 with Theorem 6.2.3, we get that the factor group $G/O_{p'p}(G)$ has derived length at most $5log_9(r/8) + 8$. It is a remarkable fact that this bound does not depend on p, since we are considering a factor group whose construction is based on some prime p. The following sections will be devoted to prove that the factor group $O_{p'p}(G)/O_{p'}(G)$ has (m, r)-bounded derived length, and we assume for now that this result is true. Let $\pi(G) = \{p_1, \ldots, p_s\}$. The map taking $x \in G$ into $(xO_{p_1'}, \ldots, xO_{p_s'})$ is an injective homomorphism, with kernel $\cap_{p \in \pi(G)} O_{p'}(G) = 1$. Then, we have that G is isomorphic to a subgroup of the direct product $G/O_{p_1'} \times \cdots \times G/O_{p_s'}$, which is soluble of (m, r)-bounded derived length. This completes the proof of the fact that G has (m, r)-bounded derived length, and so in Theorem 6.0.2 we may assume that G is a p-group.

6.3 The Regular Case

Next, we are going to bound the derived length of $P = O_{p'p}(G)/O_{p'}(G)$ by a function not depending on p, but only on m and r. This work will amount to a final bound on the derived length of G itself.

In view of the reduction argument considered above, we want to prove the following result.

Theorem 6.3.1. *Let P be a finite p-group of rank r admitting a regular p'-automorphism ϕ. Then, the derived length of P is bounded above by a function depending on r only.*

In what follows, we are going to give the steps of the proof of Theorem 6.3.1. From Theorem 4.1.15, it follows that P admits a powerful characteristic subgroup $V = V(P, r)$ such that the derived length of P/V is at most $\lceil \log_2 r \rceil$. Since ϕ induces a regular automorphism of V and $dl(P/V)$ is r-bounded, we can reduce Theorem 6.3.1 to the case where P is powerful.

Now, in view of Lemma 4.2.1, we may find a series of characteristic subgroups $1 = P_t \leq \cdots \leq P_0 = P$ such that each factor group $M_i = P_{i-1}/P_i$ is uniform, for $i = 1, \ldots, t$. The fact that P_i and P_{i+1} are both characteristic ensures that ϕ induces an automorphism of the quotient M_i, which we are going to denote by ϕ_i. Lemma 2.2.2 tells us that each ϕ_i is again fixed-point-free.

Since each section M_i is uniform, we may consider one member of the family of uniform Lie rings associated to M_i, following Proposition 4.3.1, and call it L_{j_i}. In case it is not possible to construct a Lie ring as in that proposition, we know, by the remarks at the end of Chapter 4, that the exponent of M_i does not exceed p^3 and $dl(M_i) \leq 3$. At this point, we wish to bound the derived length of each uniform section of P with a function of r only. In particular we can consider just the cases where $exp(M_i) \geq 4$ and it is possible to associate a Lie ring to M_i.

Given the associated Lie ring L_{j_i}, the automorphism ϕ_i induces a fixed-point-free automorphism of L_{j_i}, which we also call ϕ_i. We know that the ranks of the sections M_i decrease, starting with $rk(M_1) = r$. So, we may assume that M_i, the i-th uniform section of P, has rank $\leq r - (i-1)$, thus giving that the rank of L_{j_i} is not greater than $r - i + 1$. We want to apply Corollary 3.4.4 to the current situation, so in order to fulfill the hypothesis of that corollary we still need to show that ϕ_i is semisimple. This is the content of the following lemma.

Lemma 6.3.2. *Let L be a Lie ring of order a power of p. If ϕ_i is a p'-automorphism which acts nontrivially on L, then ϕ_i is semisimple.*

Proof. Let p^i be the additive exponent of L. Proposition 4.3.1 shows that L is a Lie $\mathbb{Z}/p^i\mathbb{Z}$-ring. Regarding $\mathbb{Z}/p^i\mathbb{Z}$ as the subgroup generated by a p^i-th primitive root of unit of $\mathbb{C}$, let R be an extension of $\mathbb{Z}/p^i\mathbb{Z}$ containing every eigenvalue of ϕ_i. If N is the matricial representation of ϕ_i with entries in R, we can consider its Jordan block representation. Finally, let $|\phi_i| = n$, coprime to p and J be the Jordan block of N associated with each eigenvalue β. As J^n is the identity matrix, we obtain

$$\begin{pmatrix} \beta & 1 & \cdots \\ & \ddots & \cdots \\ & & \end{pmatrix}^n = \begin{pmatrix} \beta^n & n\beta & \cdots \\ & \ddots & \cdots \\ & & \end{pmatrix}.$$

Also n is invertible in $\mathbb{Z}/p^i\mathbb{Z}$, so this matrix cannot be the identity matrix if β is zero. Being n a unit in $\mathbb{Z}/p^i\mathbb{Z}$ and β different from zero, the entry $n\beta$ on the matrix cannot be zero, unless the Jordan block J has dimension 1. This proves that ϕ_i is semisimple. □

Now, since ϕ_i is semisimple, it is possible to apply Corollary 3.4.4 to this situation and we obtain that the derived length of L_{j_i} is not greater than $2^{r-i}-1$. Applying Theorem 4.3.4 to L_{j_i} and M_i we obtain that $dl(M_i) \leq 2^{r-i+1} - 1$.

In the series $1 = P_t \leq \cdots \leq P_0 = P$, we know that $t \leq r$. Denote by r_i the rank of M_i. Thus $r = r_1 \geq r_2 \geq \cdots \geq r_t$. Then, we have

$dl(P) \leq 2^{r_1}-1+2^{r_2}-1+\cdots+2^{r_t}-1 \leq 2^r-1+\cdots+2^2-1+2-1+1-1 = 2^{r+1}-r-2.$

We get the following result.

Proposition 6.3.3. *A powerful p-group P of rank r admitting a regular coprime automorphism ϕ has derived length given at most by* $2^{r+1} - r - 2$.

Applying the result above in combination with Theorem 4.1.15 we obtain that, for an arbitrary p-group P of rank r admitting a regular p'-automorphism ϕ, the derived length of P is at most

$$\lceil \log_2 r \rceil + 2^{r+1} - r - 2. \tag{6.3.1}$$

which is a function depending only on r. This completes the proof of Theorem 6.3.1.

We are now able to deal with the proof of Theorem 6.0.2 when ϕ is a regular automorphism. Remember that we are trying to bound the derived length of an arbitrary group G of rank r admitting a regular coprime automorphism. We fixed a prime p dividing $|G|$, and showed that the soluble length of the factor group $G/O_{p'p}(G)$ is bounded by a function of r only, namely $5\log_9(r/8) + 8$. Now, since $O_{p'p}(G)/O_{p'}(G)$ is a p-group of rank at most r, the derived length of this factor group is at most $\lceil \log_2 r \rceil + 2^{r+1} - r - 2$, thus yielding that the derived length of $G/O_{p'}(G)$ is at most

$$f(r) = 5\log_9(r/8) + 8 + \lceil \log_2 r \rceil + 2^{r+1} - r - 2,$$

and this expression does not depend on p. This means that the s-th term of the derived series of G, where $s = \lceil f(r) \rceil$, is a subgroup of $O_{p'}(G)$, for every

$p \in \pi(G)$. Hence $G^{(s)} \leq \cap_{p\in\pi(G)} O_{p'}(G) = 1$, and so $dl(G) \leq \lceil f(r) \rceil$, completing the proof of Theorem 6.0.2 in the case where ϕ is a regular automorphism.

6.4 The General Case

Let us begin with a uniform Lie ring L of additive exponent p^i, admitting an abelian p'-group of automorphisms A having exactly $m > 1$ fixed points. Then by Lemma 2.2.4, the additive group of L can be written as $L = C_L(A) \times [L, A]$. The order of $C_L(A)$ is at least p^i, otherwise every generator of L lies in $[L, A]$, since L is homocyclic of exponent p^i. This implies that $p^i \leq m$ and $i \leq \log_p(m) \leq \log_2(m)$.

Now, remember that L/pL is commutative. Then we have that $L^1 \subseteq pL$. Also,

$$L^{(2)} = [L^{(1)}, L^{(1)}] \subseteq [pL, pL] = p^2[L, L] \subseteq p^3 L.$$

Thus, by induction, we get the inclusion $L^{(n)} \subseteq p^{2^n-1}L$. Since the additive exponent of L is p^i, when $2^n - 1 \geq i$, or equivalently when $n \geq \log_2(i + 1)$, we have $L^{(n)} \subseteq p^{2^n-1}L = 0$. Combining $n \geq \log_2(i + 1)$ with the inequality $i \leq \log_2(m)$, we get

Lemma 6.4.1. *Let L be an uniform Lie ring, admitting a coprime automorphism ϕ with exactly $m > 1$ fixed points. Then $dl(L) \leq \lceil \log_2(\log_2(m) + 1) \rceil$.*

Now, let P be a uniform p-group of rank r admitting a p'-automorphism ϕ with $m > 1$ fixed points. Consider one of the associated uniform Lie rings L and denote the automorphism induced by ϕ on L by $\bar{\phi}$. Since we do not know exactly whether the induced automorphism $\bar{\phi}$ acts regularly or has nontrivial fixed points, a combination of both Corollary 3.4.4 and Lemma 6.4.1 yields the following bound for $dl(L)$:

$$dl(L) \leq \max\{2^{r-1} - 1, \lceil \log_2(\log_2(m + 1)) \rceil\}.$$

Theorem 4.3.4 implies that

$$dl(P) \leq \max\{2^r - 1, 2\lceil \log_2(\log_2(m+1)) \rceil + 1\}.$$

We are going to denote the maximum value defined above by $g(m,r)$. Again, this bound does not depend on p. Keeping track of the argument at the end of the preceding section, we can state the following result.

Theorem 6.0.2. *Let G be a finite group of rank r admitting a coprime automorphism ϕ. If G is soluble and $|C_G(\phi)| = m$, then the derived length of G is bounded by a function $f(m,r)$ depending on m and r only.*

Such a function $f(m,r)$ can be explicitly determined, as follows

$$f(m,r) = 5\lceil \log_9(r/8) \rceil + 8 + \lceil \log_2(r) \rceil + rg(m,r).$$

Since the ranks of the uniform sections of a powerful p-group P are decreasing, the last summand in the equation above can be replaced by

$$g(m,r) + \cdots + g(m,2) + g(m,1).$$

This discussion completes the proof of Theorem 6.0.2.

6.5 An Improvement on the Bound

It was conjectured that in the statement of Theorem 6.0.2 one could remove the dependence of the bound on the soluble length on parameter m. In the paper [17], Evgeny Khukhro answers this question affirmatively, making a clever use of centralizers in order to obtain the result. The result we are going to prove in this section can be stated as follows.

Theorem 6.5.1 (Khukhro). *Let G be a finite group of rank r, admitting an automorphism ϕ of order coprime to the order of the group with exactly m fixed points. Then G has a characteristic soluble subgroup H of (r,m)-bounded index and whose derived length is r-bounded.*

In the same way as we did previously, we first reduce the proof to the case of finite p-groups, then we consider the general case. The result from Section 6.3 will be useful.

We first deal with the proof of the following result.

Theorem 6.5.2. *Let G be a finite p-group of rank r, admitting an automorphism ϕ of order coprime to p, and such that $|C_G(\phi)| = m$. Then G admits a subgroup D of (m, r)-bounded index and r-bounded derived length.*

In what follows, we give the steps of the proof. By Theorem 4.1.15, the group G contains a powerful characteristic subgroup V, such that the derived length of the factor-group G/V is bounded by $\lceil \log_2(r) \rceil$. Therefore we can assume that G is powerful from the outset, as we did in Section 6.3.

Moreover, by Lemma 4.2.1, we know that G can be divided in $t \leq r$ uniform characteristic sections. Let K be one of those sections, which has the form $G^{p^a}/G^{p^{a+b}}$ and exponent p^b. On this section, ϕ induces a coprime automorphism as well. Denoting the induced automorphism also by ϕ, suppose that the exponent of $C_K(\phi)$ equals p^s, for some $s \leq b$. Then, $p^s \leq m$, since $p^s \leq |C_K(\phi)|$ and $|C_K(\phi)| \leq m$ via Lemma 2.2.2. Applying Lemma 4.2.3, we obtain that $C_K(\phi)$ is contained in the subgroup $K^{p^{b-s}}$, of exponent $p^s \leq m$. As this subgroup is also uniform and has rank at most r, we have $|K^{p^{b-s}}| \leq p^{sr} \leq m^r$. By Lemma 2.2.2, we know that ϕ acts regularly on the section $K/K^{p^{b-s}}$. Then it follows from Theorem 6.3.1 that this section has r-bounded derived length.

Note that the full inverse images of the subgroups K and $K^{p^{b-s}}$ are characteristic subgroups of G, namely G^{p^a} and $G^{p^{a+b-s}}$. We have, then, constructed a series $(*)$ of at most $2r$ characteristic subgroups of G, where factors of r-bounded derived length alternate, in odd positions, with factors having (r, m)-bounded order, in even positions. Now let D be the intersection of the centralizers of all the even factors having (r, m)-bounded order; this will be the sought subgroup of (r, m)-bounded index and r-bounded derived length.

First, we show that the index of D in G is (r, m)-bounded. For this purpose, let S be one of the even factors mentioned above. Its order is bounded in terms of r and m, and G clearly normalizes S. The quotient $G/C_G(S)$ embeds in $Aut(S)$, which has (r,m)-bounded order also, thus the index of $C_G(S)$ is (r,m)-bounded. We are intersecting all the centralizers $C_G(S)$ of even sections in order to construct D, so the index of D in G is not greater than the product of the indices $[G : C_G(S)]$, when S ranges over the even sections considered above. As we have at most r such even sections, the index $[G : D]$ is also (m,r)-bounded. An explicit bound can be written, in this case, as follows

$$[G : D] \leq ((m^r)!)^r.$$

Let us check now that D has r-bounded derived length. We intersect D with the terms of the series constructed in $(*)$, obtaining a normal series in D. Suppose that $G \geq G^{p^c} \geq G^{p^{c+d}}$ is the first segment of the quoted series; then the first segment for the corresponding series of D will be $D \geq D \cap G^{p^c} \geq D \cap G^{p^{c+d}}$. We are going to use both factors $D/D \cap G^{p^c}$ and $D \cap G^{p^c}/D \cap G^{p^{c+d}}$ to illustrate that D has r-bounded derived length. Observe that $D/D \cap G^{p^c}$ is isomorphic to DG^{p^c}/G^{p^c}, which is a subgroup of G/G^{p^c}, a factor-group having r-bounded derived length. The second factor, $D \cap G^{p^c}/D \cap G^{p^{c+d}}$, is isomorphic to $(D \cap G^{p^c})G^{p^{c+d}}/G^{p^{c+d}} \leq G^{p^c}/G^{p^{c+d}}$. Since D centralizes this section, $(D \cap G^{p^c})G^{p^{c+d}}/G^{p^{c+d}}$ is central in $G^{p^c}/G^{p^{c+d}}$, hence abelian and having derived length equal to 1. Since the rank of the uniform section $G/G^{p^{c+d}}$ is at most r and the rank of the following section decreases, we have that the soluble length of $D/D \cap G^{p^{c+d}}$ is bounded by $2^r - 1 + 1 = 2^r$. Repeating the same observation for all sections, we obtain that

$$dl(D) \leq 2^r + 2^{r-1} + \cdots + 2 = 2^{r+1} - 2.$$

This completes the proof of Theorem 6.5.2 when G is a finite p-group.

The reduction of the proof to the case of a finite p-group follows the same steps as in the previous sections. First of all, fix a prime $p \in \pi(G)$, consider the quotient $P = O_{p'p}(G)/O_{p'}(G)$, which is the largest normal p-subgroup of rank at most r of $G/O_{p'}(G)$. Moreover, recall that that the factor group

$G/O_{p'p}(G)$ is faithfully represented into $GL(r,p)$, which is the automorphism group of $P/\Phi(P)$. We know, by Theorem 6.2.3, that each $G/O_{p'p}(G)$ has derived length bounded by a function of r only, not depending on p. Now, the factor group $G/\bigcap\limits_{p\in\pi(G)} O_{p'p}(G)$ is isomorphic to a subgroup of the direct product $\Pi_{p\in\pi(G)}G/O_{p'p}(G)$. Since all the direct factors in this product have the same r-bounded derived length, by Theorem 6.2.3, so does $G/\bigcap\limits_{p\in\pi(G)} O_{p'p}(G)$.

It is clear that $F=\cap_{p\in\pi(G)}O_{p'p}(G)$ is a nilpotent characteristic subgroup of G. In fact, it coincides with the Fitting subgroup $F(G)$.

Now, let $q\in\pi(F)$ and S_q be the Sylow q-subgroup of F. The fact that $S_q \operatorname{char} F$ and $F \operatorname{char} G$ implies that ϕ induces an automorphism of S_q, with $|C_{S_q}(\phi)| = m_q \leq m$. Applying Theorem 6.5.2 to S_q, let us call T_q the subgroup of (m,r)-bounded index and r-bounded derived length that can be constructed, for each choice of $q\in\pi(G)$. Then T_q has (r,m_q)-bounded index and r-bounded derived length. Set

$$T=\prod_{q\in\pi(F)} T_q.$$

The index of T in F is (m,r)-bounded, since it equals the product of (m_q,r)-bounded indices, $m_q\leq m$ and the function that bounds those indices is $((m^r)!)^r$, which is ascending in m.

It remains now to take the centralizer $C=C_G(F/T)$ as a desired subgroup of (r,m)-bounded index and r-bounded derived length. Since the order of F/T is (m,r)-bounded, the index of C in G is also (m,r)-bounded, because G is the normalizer of F/T. Now, in order to bound the derived length of C, we begin by bounding the derived length of $C/C\cap F$, which is isomorphic to CF/F and has r-bounded derived length, since G/F also has. The group $C\cap F/C\cap T$ is isomorphic to $(C\cap F)T/T$, which lies in the center of F/T and thus is abelian. Finally, since T has r-bounded derived length, $C\cap T$ also does, and this completes the proof of the general case.

6.6 Jaikin-Zapirain's contribution

In this last section we make some comments on Jaikin-Zapirain's paper [16] which answers the question raised by Shalev whether it was really necessary to impose coprimality on the automorphism in Theorems 6.0.2 and 6.5.1. The main result in [16] can be stated as follows.

Theorem 6.6.1. *Let G be a finite group of rank r admitting an automorphism with m fixed points. Then G has a characteristic soluble subgroup H, whose index is (m, r)-bounded and whose derived length is r-bounded.*

As in [27] and [17], the proof of Theorem 6.6.1 can be reduced to the case where G is a uniform p-group, in the way we discussed in the previous section. Also, Zapirain makes use again of the Lie ring introduced by Shalev in [28], but in a different way. We first want to discuss his result on uniform Lie rings, which leads to an analogous result in uniform groups. The statement below is less general than Jaikin's result and can be derived as a corollary of Theorem 2.1 in [16].

Theorem 6.6.2. *Let L be a uniform Lie ring of rank r admitting an automorphism ϕ with m fixed points. Then there exists a function $f = f(m, r)$ such that L has a soluble subring H of index less than f and derived length at most 2^{r-1}.*

Above, the additive group of L is a $\mathbb{Z}/p^i\mathbb{Z}$-module. Theorem 6.6.2 says that there exist numbers $n = n(r, m)$ and $d = d(r)$ such that $(p^n L)^{(d)} = 0$. Jaikin's proof is quite technical. In his paper, he considers a class of Lie rings that are also $\mathbb{Z}_p[\phi, \phi^{-1}]$-modules, where $\mathbb{Z}_p$ denotes the ring of p-adic integers, [23, p. 26]. In his proof, he uses some machinery proved in a previous paper [15] by the same author.

For the group-theoretic part of the proof of Theorem 6.6.1, the idea is to improve the construction of Section 4.3. Let G be a uniform group of exponent p^n and write $n = 4e+f$, $0 \leq f \leq 3$. We can assume $n \geq 4$, otherwise G would be soluble of derived length ≤ 3. Set $G^{p^i} = G_i$ and consider $L = G_e/G_{2e}$. For each integer $k = 0, 1, 2$, let $\pi_k : G_{ke}/G_{(k+1)e} \to G_{(k+1)e}/G_{(k+2)e}$ be the map taking $tG_{(k+1)e}$ to $t^{p^e}G_{(k+2)e}$, with $t \in G_{ke}$. We know that these maps π_0, π_1 and π_2 are isomorphisms, from Proposition 4.1.2 and from the fact that G is uniform. For $a, b \in G_e$, we define the Lie bracket on L as

$$[xG_{2e}, yG_{2e}] = \pi_1^{-1}([a, b]G_{3e}),$$

and the ring L, together with this operation, is indeed a uniform Lie ring, by Proposition 4.3.1. Both the Lie bracket and the group commutator are going to be denoted by $[\ ,\ ]$. Instead of constructing a family of Lie rings, as in [28], only one Lie ring is considered here. Note that in Theorem 4.3.4, we need to take an index i such that the associated Lie ring $G^{p^i}/G^{p^{2i}}$ has derived length j and the condition $i(2^j + 1) \leq n$ is satisfied. Here, none of the restrictions are necessary.

We want to prove a result relating powerfully embedded subgroups of G with ideals of L, but the following lemma is needed first.

Lemma 6.6.3. *Let M and N be powerfully embedded subgroups of the p-group G. Then $M^pN^p = (MN)^p$.*

Proof. The inclusion $M^pN^p \leq (MN)^p$ always holds, so we only need to prove the reverse inclusion. For this sake, take $m \in M$, $n \in N$ and consider the Hall-Petrescu identity applied in this case:

$$(mn)^p = m^pn^pc_2^{\binom{p}{2}} \cdots c_{p-1}^{\binom{p}{p-1}}c_p.$$

Since N is powerfully embedded on G, each element c_i above belongs to $[N, G]$, which is a subgroup of N^p. Then, the product $n^pc_2^{\binom{p}{2}} \cdots c_{p-1}^{\binom{p}{p-1}}c_p$ belongs to N^p, completing the proof. □

The following lemma can be found in [16, Lemma 3.4].

Lemma 6.6.4. *Let T be a powerfully embedded subgroup of G. Then the image $\overline{T} = \pi_0(TG_e/G_e)$ is an ideal of L and $[\overline{T}, \overline{T}] = \pi_0([T, T]G_e/G_e)$.*

Proof. Since G is powerful, G_e is powerfully embedded in G and $(TG_e)^{p^e} = T^{p^e}G_e^{p^e}$, which equals $T^{p^e}G_{2e}$. So, $\overline{T} = (TG_e)^{p^e}/G_{2e}$. Let us prove that it is an ideal of L. Using that $(TG_e)^{p^e}$ is powerfully embedded in G, we obtain that $[(TG_e)^{p^e}, G_e] = [TG_e, G]^{p^{2e}}$, by Lemma 4.1.7. Hence,

$$[\overline{T}, L] = [TG_e, G]^{p^{2e}}G_{2e}/G_{2e} \subseteq \overline{T}.$$

Now we prove the second part of the lemma. We have that

$$\begin{aligned}
[(TG_e)^{p^e}, (TG_e)^{p^e}]G_{3e} &= [TG_e, TG_e]^{p^{2e}}G_{3e}, && \text{(by Lemma 4.1.7)}\\
&= ([T,T][G_e,G_e])^{p^{2e}}G_{3e} &&\\
&= [T,T]^{p^{2e}}[G_e,G_e]^{p^{2e}}G_{3e} && \text{(by Lemma 6.6.3)}\\
&= [T,T]^{p^{2e}}G_{3e} && \text{(because } [G_e,G_e]^{p^{2e}} \leq G_{4e})\\
&= ([T,T]G_e)^{p^{2e}} && \text{(by Lemma 6.6.3).}
\end{aligned}$$

Therefore, $[\overline{T}, \overline{T}] = ([T,T]G_e)^{p^e}/G_{2e} = \pi_0([T,T]G_e/G_e)$. □

As a consequence, we have

Corollary 6.6.5. *Let T be a powerfully embedded subgroup of G. Then, $\overline{T}^{(d)} = \pi_0(T^{(d)}G_e/G_e)$.*

For each non-negative integer i, define $L_i = \pi_0(G_i/G_e) = G_{i+e}G_{2e}/G_{2e}$. Then, by Lemma 6.6.4 we know that L_i are ideals of L. The following lemma is the main result connecting solubility in uniform Lie rings with solubility in uniform groups.

Lemma 6.6.6. *Let d be the derived length of L_i as a Lie ring. Then the derived length of G_i is at most $d+2$.*

Proof. By Corollary 6.6.5, $G_i^{(d)} \leq G_e$. Then, using the previous inclusion and Corollary 4.1.8,

$$G_i^{(d+2)} = [[G_i^{(d)}, G_i^{(d)}], [G_i^{(d)}, G_i^{(d)}]] \leq [[G_e, G_e], [G_e, G_e]] \leq (G^{(2)})^{p^{4e}} \leq G^{p^{4e+3}} = 1.$$

We are now able to prove Jaikin's main result of [16], which is a version of Theorem 6.6.1 for finite p-groups. We state it below.

Theorem 6.6.7 (Jaikin-Zapirain). *Let G be a finite p-group of rank r admitting an automorphism ϕ with m fixed points. Then G has a characteristic soluble subgroup H, whose index is (m, r)-bounded and whose derived length is r-bounded.*

Proof. As before, we can assume G to be powerful from the outset, construct a series of characteristic subgroups $G_0 = G \geq G_1 \geq \cdots \geq G_t = 1$ with $t \leq r$ and such that each section $H_i = G_i/G_{i+1}$ is uniform, $i = 0, \ldots, t-1$. Here, each H_i has rank $\leq r$ and $|C_{H_i}(\phi)| \leq m$ by Lemma 2.2.2.

By the above construction, we can associate to each H_i a uniform Lie ring $L(i)$ with rank $\leq r$ and such that the automorphism induced by ϕ on $L(i)$ has at most m fixed points. By Theorem 6.6.2, there are functions $t_i = t_i(m, r)$ and $s_i = s_i(r)$ such that $(p^{t_i}L_i)^{(s_i)} = 0$. Hence, by Lemma 6.6.6 we have $(G_i^{p^{t_i}})^{(s_i+2)} \leq G_{i+1}$. Since G is powerful, each G_i also is. Repeated application of Corollary 4.1.8 shows that we can find functions $f = f(m, r) = \sum_{i=0}^{t-1} t_i(m, r)$ and $g = g(r) = 2t + \sum_{i=1}^{t-1} s_i(r) \leq 2r + \sum_{i=1}^{t-1} s_i(r)$ such that $(G^{p^f})^{(g)} = 1$. Since G has rank r and is powerful, the index of G^{p^f} in G is at most p^{fr}. This completes the proof. □

The end of the argument of the proof of Theorem 6.6.1 follows exactly the same line as used in Section 6.5, so it will be omitted. The only difference is that, in that section, we applied Theorem 6.5.2 to find subgroups of (m, r)-bounded index of each Sylow p-subgroup of the Fitting subgroup, and here we apply Jaikin's Theorem 6.6.7 for that sake. The rest of the proof remains identical.

We end this text posing some conjectures raised by Jaikin in [16]. As a particular case of the result of Section 6.4, we have that a p-group G of rank r admitting a p'-automorphism ϕ with p^m fixed points has (m, r)-bounded derived length. This result suggests the following conjecture, which follows the same idea of relaxing the hypothesis on the order of the automorphism.

Conjecture 1: Let G be a finite p-group of rank r admitting an automorphism ϕ with p^m fixed points. Then the derived length of G is (m, r)-bounded. Moreover, there are functions $f = f(m, r)$ and $d = d(r)$ such that G has a subgroup of index at most p^f and derived length at most d.

In [15], Jaikin proves that there are functions $f = f(p, m, n)$ and $h(m)$ such that any finite p-group G with an automorphism ϕ of order p^n, whose centralizer has p^m fixed points, has a subgroup of derived length $\leq h(m)$ and index $\leq f(p, m, n)$. In this situation, the rank of G is also (p, m, n)-bounded. Jaikin poses, then, the following conjecture

Conjecture 2: Let G be a finite p-group of rank r admitting a p-automorphism ϕ with p^m fixed points. Then there are functions $f = f(p, m, r)$ and $d = d(m)$ such that G has a subgroup of index at most f and derived length at most d.

Bibliography

[1] J. Alperin. Automorphisms of solvable groups. *Proc. Amer. Math. Soc.*, 13:175 – 180, 1962.

[2] R. Brauer and K.A. Fowler. On groups of even order. *Ann. of Math.*, 62(2):565 – 583, 1955.

[3] K. Brown. *Buildings*. Springer-Verlag, New York/Berlin, 1989.

[4] R. W. Carter. *Simple Groups of Lie Type.* Wiley, New York, 1972.

[5] J. Dixon, M.P.F. du Sautoy, A. Mann, and D. Segal. *Analytic pro-p groups.* Cambridge University Press, Cambridge, 2nd edition, 1999.

[6] D. S. Dummit and R. S. Foote. *Abstract Algebra.* Wiley, 3rd edition, 2004.

[7] D. Gorenstein. *Finite Groups.* Chelsea Publishing Company, New York, 2nd edition, 1980.

[8] P. Hall and G. Higman. On the p-length of p-soluble groups and reduction theorems for Burnside's problem. *Proc. Lond. Math. Soc.*, 6:1–42, 1954.

[9] B. Hartley. A general brauer-fowler theorem and centralizer in locally finite groups. *Pacific J. Math.*, 152:101–117, 1992.

[10] B. Hartley and M. Kuzucuoğlu. Centralizers of elements in locally finite simple groups. *Proc. Lond. Math. Soc.*, 62(3):301 – 324, 1991.

[11] G. Higman. Groups and rings which have automorphisms without nontrivial fixed elements. *J. London Math. Soc. (2)*, 32:321 – 334, 1957.

[12] K. Hoffman and R. Kunze. *Linear Algebra.* Prentice Hall, Inc, Englewood Cliffs, New Jersey, 1971.

[13] J. E. Humphreys. *Introduction to Lie Algebras and Representation Theory.* Springer-Verlag, New York, 1972.

[14] B. Huppert and N. Blackburn. *Finite Groups II.* Springer-Verlag, Berlin, 1982.

[15] A. Jaikin-Zapirain. On the almost regular automorphisms of finite p-groups. *Adv. Math.*, 153:391 – 402, 2000.

[16] A. Jaikin-Zapirain. Finite groups of bounded rank with an almost regular automorphism. *Israel J. Math.*, 129:209 – 220, 2002.

[17] E. I. Khukhro. Almost regular automorphisms of finite groups of bounded rank. *Sib. Math. J.*, 37:1237–1241, 1996.

[18] E.I. Khukhro. *p-automorphisms of finite p-groups.* Cambridge University Press, 1998.

[19] V.A. Kreknin. Solvability of lie algebras with a regular automorphism of finite period. *Soviet Math. Dokl.*, 4:683 – 685, 1993.

[20] A. Lubotzky and A. Mann. Powerful p-groups. I. Finite groups. *J. Algebra*, 105:484 – 505, 1987.

[21] M. Newman. The soluble length of soluble linear groups. *Math. Zeit.*, 126:59 – 70, 1972.

[22] I. Niven, H. Zuckerman, and H. Montgomery. *An introduction to the theory of numbers*. Wiley, New York, 5th edition, 1991.

[23] L. Ribes and P. Zaleskii. *Profinite Groups*. Springer-Verlag, 2nd edition, 2010.

[24] H. E. Rose. *A course on finite groups*. Springer-Verlag, 2009.

[25] J. S. Rose. *A course on group theory*. Cambridge University Press, 1978.

[26] P. Rowley. Finite groups admitting a fixed point free automorphism group. *J. Algebra*, 174:724 – 727, 1995.

[27] A. Shalev. Automorphisms of finite groups of bounded rank. *Israel J. Math.*, 82:395 – 404, 1993.

[28] A. Shalev. On almost fixed point free automorphisms. *J. Algebra*, 157:271 – 282, 1993.

[29] J. G. Thompson. Finite groups with fixed-point-free automorphisms of prime order. *Proc. Nat. Acad. Sci. U.S.A.*, 45:578 – 581, 1959.

[30] J. Tits. Groups et géométries de Coxeter. *unpublished manuscript*, 1961.

[31] J. Tits. Théorème de Bruhat et sous-groupes paraboliques. *C.R. Acad. Sci. Paris*, 254:2910–2912, 1962.

Index of Notations

Symbol	**Meaning**
$\vert\ \vert$	order of a finite group
ϕ	automorphism of a finite group
$[H,K]$	the subgroup generated by commutators $[h,k]$, $h \in H$ and $k \in K$
$\gamma_k(G)$	k-th term of lower central series of G
$G^{(k)}$	k-th term of derived series of G
π-number	a number divisible only by the numbers in the set π
π'-number	a number coprime to all the numbers in the set π
$O_{p'}(G)$	the largest normal p'-subgroup of G
$O_{p'p}(G)$	full inverse image of $O_p(G/O_{p'}(G))$
$Fit(G)$	the Fitting subgroup of G
$\Phi(G)$	the Frattini subgroup of G
$d(G)$	the minimal number of generators of G
$rk(G)$	the rank of G
$[G:H]$	the index of H in G
$exp(G)$	the exponent of G
R	a ring
$M_n(K)$	square matrices of order $n \times n$ and entries in K
$End_{\mathbb{Z}}(A)$	endomorphisms of the abelian group A
$C_G(\phi)$	the fixed point subgroup of the automorphism ϕ in G
$C_G(M)$	elements of G that centralize $M \subseteq G$ simultaneously
$N_G(H)$	the normalizer of H in G
$L(G)$	a Lie ring associated to a group G
$\gamma_k(L)$	k-th term of lower central series of L
$L^{(k)}$	k-th term of derived series of L
G^n	the subgroup of G generated by n-th powers
$V(G,r)$	powerful characteristic subgroup of the p-group G, of rank r

$GL_r(\mathbb{F}_p)$	the group of invertible matrices over $\mathbb{F}_p$
$U_r(\mathbb{F}_p)$	subgroup of $GL_r(\mathbb{F}_p)$ consisting of lower unitriangular matrices
$\lambda(r)$	the function $\lceil \log_2(r) \rceil$
Φ	a system of roots
w_r	the reflection through the hyperplane orthogonal to the vector r
W	the Weyl group associated to a system of roots
Π	a fundamental system of roots
$\succ$	total order on a vector space V
V^+	subset of V consisting of vectors v such that $v \succ 0$
V^-	subset of V consisting of vectors v such that $0 \succ v$
Φ^+	a positive system of roots
Φ^-	a negative system of roots
A_l, B_l, C_l, D_l E_6, E_7, E_8, F_4, G_2	isomorphism classes of simple $\mathbb{C}$-Lie algebras
$\mathcal{H}$	a Cartan subalgebra
A_{rs}	a Cartan integer related to the roots r and s, on this order
h_r	the co-root associated to the root r
exp	the exponential map
ad	the adjoint map
$\mathcal{L}(K)$	Chevalley group of type L over the field K
$x_r(t)$	$\exp(\operatorname{ad} t \cdot e_r)$
C_{ijrs}	the coefficients appearing in Chevalley's commutator formula
X_r	the root subgroup associated to the root r
U	the subgroup generated by X_r with r positive
V	the subgroup generated by X_r with r negative
ϕ_r	homomorphism from $\langle X_r, X_{-r} \rangle$ to $SL_2(K)$
$h_r(\lambda)$	image of $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{smallmatrix}\right)$ via ϕ_r
H	the subgroup of G generated by the $h_r(\lambda)$
$\mathbb{Z}\Phi$	free abelian group with basis Φ
$\hat{H}$	the group of K-characters of $\mathbb{Z}\Phi$

$n_r(\lambda)$	image of $\left(\begin{smallmatrix} 0 & \lambda \\ -\lambda^{-1} & 0 \end{smallmatrix}\right)$ via ϕ_r
N	the subgroup of G generated by the $h_r(\lambda)$ and n_r
P_J	the parabolic subgroup associated to the subset $J \subseteq I$
$Out(G)$	the quotient $Aut(G)/Inn(G)$
I, D, Γ, F	the set if inner, diagonal, graph and field automorphisms of $\mathcal{L}(G)$
CFSG	Classification of Finite Simple Groups
$GF(q)$	the field with $q = p^b$ elements
$\Sigma(W, S)$	the Coxeter complex associated to W and S