Quantifying the Residual Finiteness of Linear Groups

Daniel Elliot Franz
Laurinburg, NC

Bachelor of Arts, Kenyon College, 2010

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Mathematics

University of Virginia
May, 2016

Abstract

A residually finite group is a group for which the intersection of all finite index subgroups is trivial; such a group can be studied using its finite quotients. Normal residual finiteness growth measures how well a finitely generated residually finite group is approximated by its finite quotients. We show that any linear group $\Gamma \leq \mathrm{GL}_d(K)$ has normal residual finiteness growth asymptotically bounded above by $(n \log n)^{d^2-1}$; notably this bound depends only on the degree of linearity of Γ. If char $K = 0$, then this bound can be improved to n^{d^2-1}. We also give lower bounds on the normal residual finiteness growth of Γ in the case that Γ is a finitely generated subgroup of a Chevalley group G of rank at least 2. These lower bounds agree with the computed upper bounds, providing exact asymptotics on the normal residual finiteness growth. In particular, finite index subgroups of $G(\mathbb{Z})$ and $G(\mathbb{F}_p[t])$ have normal residual finiteness growth $n^{\dim(G)}$. We also compute the non-normal residual finiteness growth in the above cases; for the lower bounds the exponent $\dim(G)$ is replaced by the minimal codimension of a maximal parabolic subgroup of G.

Acknowledgments

Dr. Mikhail Ershov

For his continued guidance and support. His suggestions and advice were instrumental in completing this work.

Dr. Andrei Rapinchuk, Dr. Peter Abramenko, and Dr. Abhi Shelat

For serving on my defense committee.

My parents Jonathan and Elaine Franz

For their unwavering love and support.

My fiance Jeff Curtis

For believing in me and inspiring me to fulfill my potential.

Contents

Chapter 1

Introduction

This thesis contains two results on the residual finiteness growth of linear groups. The first result is computing an upper bound on the residual finiteness growth of finitely generated linear groups that depends only on the degree of linearity and not the field of coefficients. The second is computing a lower bound on the residual finiteness growth of Chevalley groups. These results combine to provide exact asymptotics on the residual finiteness growth of Chevalley groups. We first give some background on residual finiteness growth and then summarize the new results.

1.1 Definitions and Overview

Let Γ be a finitely generated group with finite generating set X which is symmetric, i.e. $X = X^{-1}$. If $\gamma \in \Gamma$, the word length of γ with respect to X is $||\gamma||_X = \min\{n : \gamma = x_1 \cdots x_n, x_i \in X\}$. Set $w_{\Gamma,X}(n) = |\{\gamma \in \Gamma : ||\gamma||_X \leq n\}$. In other words, $w_{\Gamma,X}(n)$ is the size of the ball of radius n in the Cayley graph of Γ with respect to X. Then the word growth of Γ, sometimes just called the growth of Γ, is the the asymptotic growth of $w_{\Gamma,X}(n)$. Using a different generating set for Γ changes $w_{\Gamma,X}(n)$ by a multiplicative constant, so the growth of Γ is independent of the choice of generating set X. For example, the growth of $\mathbb{Z}^d$ is n^d, and the growth of a nonabelian free group is exponential in n.

The central tenet of geometric group theory is that the geometry of a group, via its Cayley graph, can be used to understand algebraic properties of a group. This is illustrated by one of the main early results of geometric group theory, Gromov's theorem. It is relatively straightforward to show that if Γ is virtually nilpotent, i.e. Γ has a finite index subgroup which is nilpotent, then Γ has polynomial word growth; that is, the growth of Γ is bounded above by a polynomial in n. Gromov proved the converse: every finitely generated group Γ which has polynomial word growth is virtually nilpotent.

Prior to 1984, no groups were known to have growth that was not polynomial or exponential, but in [13] Grigorchuk constructed a group of intermediate growth, with growth function strictly between $e^{\sqrt{n}}$ and e^n. The group can be realized as a group of autmorphisms on an infinite rooted regular tree, and the study of this group has spurred research into branch groups, self-similar groups, and other areas.

Studying the word growth of groups increased understanding of groups and helped spur the development of new areas of mathematics, so it is natural to study other asymptotic invariants of finitely generated groups. One of these invariants is the subgroup growth of Γ, defined to be the asymptotic growth of $s_\Gamma(n) = |\{H \leq \Gamma :$

$[\Gamma : H] \leq n\}|$. If $R(\Gamma)$ is the intersection of all finite index subgroups of Γ, then $s_\Gamma(n) = s_{\Gamma/R(\Gamma)}(n)$, so it is enough to consider groups with $R(\Gamma) = 1$; such groups are called residually finite. In a similar spirit as Gromov's theorem, Lubotzky, Mann, and Segal proved that a finitely generated, residually finite group Γ has polynomial subgroup growth if and only if Γ is virtually solvable of finite rank, where Γ has finite rank if there is some positive integer N such that every finitely generated subgroup of Γ can be generated by at most N elements (see [20]).

Residually finite groups have many subgroups of finite index, which places restrictions on the properties of the group. For example, there exist groups of bounded exponent which are infinite, but by the solution to the restricted Burnside problem, every residually finite finitely generated group of finite exponent is finite. Since every finitely generated linear group is residually finite, we also have a large class of examples to work with.

It is natural to try to understand a residually finite group G by studying its finite index subgroups. One approach is to compute how many subgroups of a certain index G has, as mentioned above, but one can also try to quantify how quickly the intersection of finite index subgroups becomes trivial. In [3], Khalid Bou-Rabee introduced a new asymptotic invariant, the normal residual finiteness growth of G, as a way to quantify how residually finite a given group is. In this thesis we further investigate this invariant for finitely generated linear groups.

An equivalent definition of a group Γ being residually finite is that for every nontrivial $\gamma \in \Gamma$, there is a homomorphism $\varphi : \Gamma \to Q$ of Γ onto a finite group such that $\varphi(\gamma) \neq 1$; if this is the case we say that the quotient Q detects γ. If Γ is finitely generated by X, then we define $F^{\trianglelefteq}_{\Gamma,X}(n)$ to be the smallest natural number N such that every nontrivial γ in the ball of radius n is detected in some quotient of size at most N. The asymptotic growth of this function does not depend on X and is called the normal residual finiteness growth of Γ, denoted by $F^{\trianglelefteq}_{\Gamma}(n)$.

With this definition, the normal residual finiteness growth of a group Γ can be thought of as quantifying how well Γ is approximated by finite quotients. If $F^{\trianglelefteq}_{\Gamma}(n)$ grows very quickly in n, then there are many elements of Γ of short word length that vanish even in large quotients of Γ. Conversely, if $F^{\trianglelefteq}_{\Gamma}(n)$ grows slowly, then Γ is well approximated by finite quotients.

Estimates for normal residual finiteness growth have been found for virtually nilpotent groups [3], linear groups [8], arithmetic groups [6], and free groups [4] [15] [27]. In particular, the normal residual finiteness growth of a virtually nilpotent group grows slower than a power of $\log n$, and the normal residual finiteness growth of a linear group is slower than n^k for some k. It is still an open problem whether the converses of the above statements are true, or if we can infer algebraic properties of a group from knowing that its residual finiteness growth is bounded by a power of $\log n$ or n. A first step is to compute the normal residual finiteness growth of many classes of groups to provide evidence for possible conjectures.

The primary difficulty in computing normal residual finiteness growth lies in finding lower bounds; indeed for the case of the free group, this amounts to finding a group law which is satisfied by all finite groups of size at most n. In contrast, to establish an upper bound one must find a single quotient of appropriate size which detects a given element, which is in general more straightforward.

An element $\gamma \in \Gamma$ is detected by a quotient of size at most N if and only if $\gamma \notin H$ for some normal subgroup H of Γ of index at most n. By generalizing this statement to include all subgroups instead of just normal subgroups, one can define the non-normal residual finiteness growth of Γ, sometimes called the residual finiteness growth of Γ. Specifically, $F^{\leq}_{\Gamma,X}(n)$ is defined as the smallest natural number N such that for all nontrivial $\gamma \in \Gamma$ with $||\gamma||_X \leq n$, there exists $H \leq \Gamma$ with $\gamma \notin H$ and $[G : H] \leq N$, and the asymptotic growth of $F^{\leq}_{\Gamma,X}(n)$, denoted by $F^{\leq}_{\Gamma}(n)$, is the non-normal residual finiteness growth of Γ.

Non-normal residual finiteness growth has also been studied for various classes of groups, including right angled Artin groups and virtually special groups in [5] and free groups in [7] [9] [17].

It is difficult to compute the normal and non-normal residual finiteness growth of a group, even for such well understood groups as linear groups. While exact asymptotics exist for normal residual finiteness growth for some arithmetic groups in characteristic 0, a uniform upper bound on the normal and non-normal residual finiteness growth of finitely generated subgroups of $\mathrm{GL}_d(\mathbb{C})$ had not been established, and there were very few results in positive characteristic, e.g. for subgroups of $\mathrm{GL}_d(K)$

where K is a field of characteristic p. The new results presented in this thesis add to our understanding of the normal and non-normal residual finiteness growth of linear groups and present a unified strategy for proving statements in both characteristic 0 and positive characteristic.

1.2 Summary of New Results

It was shown in [8] that if Γ is a finitely generated linear group over an infinite field, then $F_\Gamma^{\trianglelefteq}(n) \preceq n^k$ for some k depending on the field and the degree of linearity. A natural question is whether the degree of polynomial growth actually depends on the field of coefficients. Our first result is that in fact there is a uniform bound on the normal and non-normal residual finiteness growth of finitely generated linear groups with a fixed degree of linearity. We write $f(n) \preceq g(n)$ if for some C, $f(n) \leq Cg(Cn)$ for all n.

Theorem 1.2.1. *Let K be a field and let $\Gamma \leq \mathrm{GL}_d(K)$ be a finitely generated linear group with $d \geq 2$.*

(i) If char $K > 0$, then $F_\Gamma^{\trianglelefteq}(n) \preceq (n \log n)^{d^2-1}$ and $F_\Gamma^{\leq}(n) \preceq (n \log n)^{d-1}$.

(ii) If char $K = 0$ or K is a purely transcendental extension of a finite field, then $F_\Gamma^{\trianglelefteq}(n) \preceq n^{d^2-1}$ and $F_\Gamma^{\leq}(n) \preceq n^{d-1}$.

The proof is contained in chapter 6; we give a brief outline of the argument. Since Γ is finitely generated, it is contained in $\mathrm{GL}_d(R)$ for some finitely generated integral domain R. We let $A \in \Gamma$ have word length n and find a ring homomorphism $\varphi : R \to \mathbb{F}$, where $\mathbb{F}$ is a field of size approximately $n \log n$ or n, depending on if we are in case (i) or (ii) of the theorem, such that A remains nontrivial under the induced group homomorphism $\varphi^* : \mathrm{GL}_d(R) \to \mathrm{GL}_d(\mathbb{F})$. With the proper choice of φ, the image of A remains nontrivial in $\mathrm{GL}_d(\mathbb{F})/Z(\mathrm{GL}_d(\mathbb{F}))$, the size of which has order n^{d^2-1} or $(n \log n)^{d^2-1}$. This establishes the bound on normal residual finiteness growth. One then shows that the image of A is not in a maximal parabolic subgroup of index approximately n^{d-1} to provide the bound on non-normal residual finiteness growth and complete the proof.

The key step is finding the correct ring homomorphism φ. This is straightforward when K is purely transcendental, but in the general situation we must use variations of the Chebotarev density theorem, a result from number theory concerning the density of primes with certain splitting properties in Galois extensions. In characteristic 0 we are able to use a higher dimensional generalization of the Chebotarev density theorem proved by Serre in [25], while in characteristic p we use an effective version of the Chebotarev density theorem which produces slightly weaker bounds.

The proof of Theorem 1.2.1 is based on inducing a group homomorphism of general linear groups using a ring homomorphism, so the argument generalizes to linear algebraic groups, yielding the following more specific result. We write $\dim(G)$ for the dimension of a linear algebraic group. If G is a simple Chevalley group, i.e. a Chevalley group whose root system is irreducible, then we let $a(G)$ be the minimal codimension of a proper parabolic subgroup. The values of $\dim(G)$ and $a(G)$ when G is a simple Chevalley group are given in Table 1.1 and justified in Lemma 2.5.5.

Theorem 1.2.2. *Let G be a linear algebraic group defined over $\mathbb{Z}$, let K be a field, and let $\Gamma \leq G(K)$ be finitely generated.*

(i) If char $K > 0$, then $F_{\Gamma}^{\trianglelefteq}(n) \preceq (n \log n)^{\dim(G)}$ and, if G is a simple Chevalley group, $F_{\Gamma}^{\leq}(n) \preceq (n \log n)^{a(G)}$.

(ii) If char $K = 0$ or K is a purely transcendental extension of a finite field, then $F_{\Gamma}^{\trianglelefteq}(n) \preceq n^{\dim(G)}$ and, if G is a simple Chevalley group, $F_{\Gamma}^{\leq}(n) \preceq n^{a(G)}$.

The second result concerns finding lower bounds on normal and non-normal residual finiteness growth. In [6], Bou Rabee and Kaletha proved that if G is a simple Chevalley group of rank at least 2 and Γ is a finite index subgroup of $G(\mathbb{Z})$, then $F_{\Gamma}^{\trianglelefteq}(n) \succeq n^{\dim(G)}$. In addition, Bou-Rabee, Hagen, and Patel showed in [5] that $F_{\mathrm{SL}_d(\mathbb{Z})}^{\leq}(n) \succeq n^{d-1}$ if $d > 2$. Both results were proved using techniques specific to characteristic 0. We generalize the normal residual finiteness growth result to characteristic p and provide lower bounds on non-normal residual finiteness growth in both characteristic 0 and p. The restriction on the rank of G is because the congruence subgroup property plays a pivotal role in the proof.

Theorem 1.2.3. *If G is a simple Chevalley group of rank at least 2, $\mathcal{O} = \mathbb{Z}$ or $\mathbb{F}_p[t]$, and $\Gamma \leq G(\mathcal{O})$ has finite index, then $F_{\Gamma}^{\trianglelefteq}(n) \succeq n^{\dim(G)}$ and $F_{\Gamma}^{\leq}(n) \succeq n^{a(G)}$, where* $\dim(G)$ *and* $a(G)$ *are given in Table* 1.1.

Φ	$\dim(G)$	$a(G)$
$A_l, l \geq 2$	$l^2 + 2l$	l
$B_l, l \geq 2$	$2l^2 + l$	$2l - 1$
$C_l, l \geq 3$	$2l^2 + l$	$2l - 1$
$D_l, l \geq 4$	$2l^2 - l$	$2l - 2$
G_2	14	5
F_4	52	15
E_6	78	16
E_7	133	27
E_8	248	57

Table 1.1: This table gives the dimension $\dim(G)$ of a simple Chevalley group G and the minimal codimension $a(G)$ of a proper parabolic subgroup.

The proof of this theorem, contained in chapter 7, has the advantage of using the same techniques for both $G(\mathbb{Z})$ and $G(\mathbb{F}_p[t])$. We choose a specific element $A \in \Gamma$ and show that if a subgroup H of Γ does not contain A, then $[\Gamma : H]$ must be appropriately large in terms of the word length of A. Instead of dealing with subgroups of Γ directly, we use the congruence subgroup property to work with subgroups of $G(\mathcal{O}/\mathfrak{m}^k)$ for some maximal ideal $\mathfrak{m}$ of $\mathcal{O}$ and $k > 1$. Once in this setting, we define G_i to be the kernel of the natural projection $G(\mathcal{O}/\mathfrak{m}^k) \to G(\mathcal{O}/\mathfrak{m}^i)$ and give $L(G) = \bigoplus G_i/G_{i+1}$ the structure of a graded Lie algebra, where the quotients G_i/G_{i+1} are identified with the Lie algebra of G over the field $\mathcal{O}/\mathfrak{m}$. The details of this construction are given in section 7.1.

We then associate to each subgroup H of $G(\mathcal{O}/\mathfrak{m}^k)$ a graded subalgebra $L(H)$ of $L(G)$. The index of H is related to the codimension of $L(H)$ in $L(G)$, so it is enough to show that if the image of A is not in $L(H)$, then the codimension of $L(H)$ is large. Computing a bound on the codimension of $L(H)$ based on the word length of A involves fairly technical arguments, which have been collected in chapter 4 along with more general results about codimensions of certain subspaces of Lie algebras.

Normal and non-normal residual finiteness growth can only decrease when passing to a subgroup, so Theorem 1.2.3 also gives lower bounds for all finitely generated subgroups of $G(K)$, where G is a simple Chevalley group of rank at least 2 and K is a field. Combining this lower bound with the upper bound from Theorem 1.2.2 then gives exact asymptotics for normal and non-normal residual finiteness growth.

Corollary 1.2.4. *Let G be a simple Chevalley group of rank at least 2, let K be a field of characteristic 0 or a purely transcendental extension of a finite field, and let $\Gamma \leq G(K)$ be finitely generated. Put $\mathcal{O} = \mathbb{Z}$ if char $K = 0$ and $\mathcal{O} = \mathbb{F}_p[t]$ if char $K = p > 0$.*

If $\Gamma \cap G(\mathcal{O}) \leq G(\mathcal{O})$ has finite index, then $F_\Gamma^{\trianglelefteq}(n) \approx n^{\dim(G)}$ and $F_\Gamma^{\leq}(n) \approx n^{a(G)}$.

Chapter 2

Lie Algebras and Chevalley Groups

Lie algebras are an important tool in the study of the residually finite groups, and Chevalley groups are an important class of linear groups; in this chapter we review their constructions and basic properties, and set some notation. The reader is referred to [10], [14], and [26] for more details.

2.1 Root Systems

We begin by defining the notion of a root system. Let E be a Euclidean space with the usual inner product $(\cdot, \cdot)$. Any vector $\alpha \in E$ defines a reflection $\sigma_\alpha : E \to E$ by the formula

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha,$$

where $\langle \beta, \alpha \rangle = \dfrac{2(\beta, \alpha)}{(\alpha, \alpha)}$.

Definition. A **root system** in E is a subset of E satisfying the following axioms:

(R1) Φ is finite, spans E, and does not contain 0.

(R2) For all $\alpha \in \Phi$, $\mathbb{R}\alpha \cap \Phi = \{\pm\alpha\}$.

(R3) For all $\alpha, \beta \in \Phi$, $\sigma_\alpha(\beta) \in \Phi$.

(R4) For all $\alpha, \beta \in \Phi$, $\langle \beta, \alpha \rangle \in \mathbb{Z}$.

Some authors define root systems more generally. In those cases, what we are defining as a root system is both reduced (because of (R2)) and crystallographic (because of (R4)). Such a root system is sometimes called a classical root system.

The **Weyl group** of Φ is $W = \langle \sigma_\alpha : \alpha \in \Phi \rangle$, which embeds into $\mathrm{Sym}(\Phi)$ and is thus finite. An important property of W is that $\langle w(\beta), w(\alpha) \rangle = \langle \beta, \alpha \rangle$ for all roots α, β and all $w \in W$.

A **base** of a root system Φ is a subset Π which is a basis for E such that every element of Φ is either a positive or negative linear combination of elements of Π. Elements of Π are called simple roots, and $W = \langle \sigma_\alpha : \alpha \in \Pi \rangle$.

A root system Φ is **irreducible** if it is not the union of two proper orthogonal subsets. If Φ is not irreducible, then it decomposes uniquely as the union of irreducible root systems Φ_i in E_i such that $E = E_1 \oplus \cdots \oplus E_k$.

In any irreducible root system Φ, there are at most two lengths of roots; we will call the roots of smaller length **short** and the roots of greater length **long**. Any two roots of the same length are conjugate under the action of the Weyl group. If Φ has roots of only one length, then by convention we will say these are long roots.

If Φ is a root system and $\alpha, \beta \in \Phi$, then the **α-string through β** is the set $\{\beta + n\alpha : n \in \mathbb{Z}\} \cap \Phi$. Root strings are unbroken, i.e. the α-string through β can be written as $\beta - r\alpha, \cdots, \beta + q\alpha$ for some positive integers r, q. In addition, $\langle \beta, \alpha \rangle = r - q$, and every root string has length at most 4.

2.2 Lie Algebras

Definition. A **Lie algebra** over a field F is a vector space L over F with a bracket operation $[\cdot,\cdot]$ satisfying the following axioms:

(L1) The bracket operation is bilinear.

(L2) $[x,x]=0$ for all $x \in L$.

(L3) $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$ for all $x,y,z \in L$.

The third axiom is called the Jacobi identity. The dimension of L is just its dimension over F as a vector space. We will only be considering finite dimensional Lie algebras. If K is a subfield of F, then L can also be considered as a Lie algebra over K. When we need to consider L as a Lie algebra over both K and F, we will use the following notation to indicate over which field we are working(note that $K=F$ is a possibility): we write $U \leq_K L$ to mean U is a K-subspace of L, and given $U,V \leq_K L$, $[U,V]_K$ is the K-span of $\{[u,v] : u \in U, v \in V\}$. If we are considering L only over one field, no subscripts will be used.

Example.

- If F is a field, then the associative matrix algebra $\mathrm{Mat}_d(F)$ of $d \times d$ matrices over F is a Lie algebra with bracket operation defined by $[A,B]=AB-BA$.

When viewing this algebra as a Lie algebra, we denote it by $\mathfrak{gl}_d(F)$.

- The set of $d \times d$ matrices over a field F with trace 0, denoted by $\mathfrak{sl}_d(F)$, is a sub-Lie algebra of $\mathfrak{gl}_d(F)$. This set is closed under the bracket operation because if $A, B \in \mathfrak{gl}_d(F)$, then AB and BA have the same trace, so $[A, B] = AB - BA$ has trace 0. This also shows that $[\mathfrak{gl}_d(F), \mathfrak{gl}_d(F)] \subseteq \mathfrak{sl}_d(F)$.

An **ideal** I of L is a subspace satisfying $[L, I] \subseteq I$. A Lie algebra L is **simple** if it has no nonzero, proper ideals, and L is **semisimple** if it is a direct sum of ideals which are each simple Lie algebras. We say L is **abelian** if $[L, L] = 0$, and L is **perfect** if $[L, L] = L$.

On any Lie algebra L, one can define a symmetric, bilinear form κ, defined by $\kappa(x, y) = \operatorname{trace}(\operatorname{ad} x \operatorname{ad} y)$. This is called the **Killing form**. The Killing form is nondegenerate if and only if L is semisimple.

Now let L be a complex Lie algebra, i.e. a Lie algebra over $\mathbb{C}$, and assume L is semisimple for the remainder of this section.

Definition. An element $x \in L$ is called **semisimple** if $\operatorname{ad} x \in \operatorname{End}(L)$ is semisimple, i.e. $\operatorname{ad} x$ is a diagnolizable linear transformation. A subalgebra of L consisting entirely of semisimple elements is called a **toral** subalgebra.

Lemma 2.2.1. *Any toral subalgebra H of L is abelian. If H is a maximal toral subalgebra of L, then H is its own centralizer in L, and the restriction of the Killing form to H is nondegenerate.*

Let H be a maximal toral subalgebra of L. Since the restriction of κ to H is nondegenerate, we can use this form to identify H with H^* as follows: for $\varphi \in H^*$, define $t_\varphi \in H$ by $\varphi(h) = \kappa(t_\varphi, h)$ for all $h \in H$. We then define an inner product on H^* by $(\alpha, \beta) = \kappa(t_\alpha, t_\beta)$.

Since H is abelian, it consists of commuting endomorphisms $\operatorname{ad} h, h \in H$, of L. Thus L is the direct sum of the subspaces $L_\alpha = \{x \in L : [h, x] = \alpha(h)x \text{ for all } h \in H\}$ as α ranges over H^*, the dual space of H. Let $\Phi = \{\alpha \in H^* \setminus 0 : L_\alpha \neq 0\}$. It is the case that $L_0 = H$, so L has a **root space decomposition**

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

Each L_α is one-dimensional and Φ is a root system; this root system is independent of the choice of H and depends only on L.

A semisimple complex Lie algebra is simple if and only if its root system Φ is irreducible. The complex simple Lie algebras are classified by their irreducible root systems Φ, which fall into four infinite families and five exceptional cases. We will need to explicitly reference the roots in these root systems, so we present descriptions of each.

In what follows, $\{\epsilon_1, \cdots, \epsilon_n\}$ will be an orthonormal basis of $\mathbb{R}^n$ with the usual inner product.

- The root system of type $A_l, l \geq 1$, has roots $\Phi = \{\epsilon_i - \epsilon_j : 1 \leq i \neq j \leq l + 1\}$.

- The root system of type $B_l, l \geq 2$, has roots

  $$\Phi = \{\pm\epsilon_i : 1 \leq i \leq l\} \cup \{\pm(\epsilon_i \pm \epsilon_j) : 1 \leq i \neq j \leq l\},$$

  where the first set consists of short roots and the latter set of long roots.

- The root system of type $C_l, l \geq 3$, has roots

  $$\Phi = \{\pm(\epsilon_i \pm \epsilon_j) : 1 \leq i \neq j \leq l\} \cup \{\pm 2\epsilon_i : 1 \leq i \leq l\},$$

  where the first and second set are the short and long roots, respectively.

- The root system of type $D_l, l \geq 4$, has roots $\Phi = \{\pm(\epsilon_i \pm \epsilon_j) : 1 \leq i \neq j \leq l\}$.

- The root system of type E_8 has roots $\Phi = \Phi_1 \cup \Phi_2$, with

  $$\Phi_1 = \{\pm\epsilon_i \pm \epsilon_j : 1 \leq i \neq j \leq 8\},$$
  $$\Phi_2 = \{\pm\frac{1}{2}\sum_{i=1}^{8} c_i\epsilon_i : c_i = \pm 1, \prod_{i=1}^{8} c_i = 1\}.$$

  The root systems of type E_6 and E_7 can naturally be viewed as subsystems of E_8; in E_6 we restrict to $3 \leq i, j \leq 7$ in Φ_1 and require $c_1 = c_2 = c_8$ in Φ_2. In E_7, we restrict to $2 \leq i, j \leq 7$ in Φ_1, require $c_1 = c_8$ in Φ_2, and add in $\pm(\epsilon_1 + \epsilon_8)$.

- The root system of type F_4 has roots

  $$\Phi = \{\pm\epsilon_i, \pm(\epsilon_i \pm \epsilon_j) : 1 \leq i \neq j \leq 4\} \cup \{\pm\frac{1}{2}(\epsilon_1 \pm \epsilon_2 \pm \epsilon_3 \pm \epsilon_4)\}.$$

  The roots of the form $\pm\epsilon_i \pm \epsilon_j$ are long, while the remaining roots are short.

- For the root system of type G_2, we fix a base $\{\alpha_S, \alpha_L\}$, where α_S and α_L are short and long, respectively. Then the short roots are $\{\pm\alpha_S, \pm(\alpha_S+\alpha_L), \pm(2\alpha_S+\alpha_L)\}$ and the long roots are $\{\pm\alpha_L, \pm(3\alpha_S+\alpha_L), \pm(3\alpha_S+2\alpha_L)\}$.

Example. The Lie algebra of type A_l is $\mathfrak{sl}_{l+1}(\mathbb{C}) = \{A \in \mathrm{Mat}_{l+1}(\mathbb{C}) : \mathrm{trace}(A) = 0\}$.

The root systems A_l, D_l, and E_l only have one root length; we call these root systems simply laced. In the remaining cases, the ratio of root lengths is $\sqrt{2}$, except in G_2, where the ratio is $\sqrt{3}$.

We will need the following result about irreducible root systems.

Lemma 2.2.2. *Let Φ be an irreducible root system.*

(1) If $\alpha, \gamma \in \Phi$ and α is a long root, then $\gamma - 2\alpha \in \Phi$ if and only if $\gamma = \alpha$.

(2) If Φ is not of type $C_l, l \geq 2$, then there exist long roots $\alpha, \beta \in \Phi$ such that $\alpha + \beta \in \Phi$ and $\alpha - \beta \notin \Phi$.

Proof. We first prove (1). Let $\alpha, \gamma \in \Phi$ with α long. If Φ is simply laced, assume $\gamma - 2\alpha \in \Phi$ with $\gamma \neq \alpha$. Then the α-root string through γ is at least $\gamma - 2\alpha, \gamma - \alpha, \gamma$. This can never occur in the simply laced root systems, where root strings have length at most 2.

If Φ is of type B_l or F_4, then $\alpha = \pm\epsilon_i \pm \epsilon_j$ for some i, j, and by examining the descriptions of the root systems, $\gamma - 2\alpha \notin \Phi$ if $\gamma \neq \alpha$. If Φ is of type C_l, then $\alpha = \pm 2\epsilon_i$ for some i, and $\pm\epsilon_j \pm \epsilon_k \pm (4\epsilon_i) \notin \Phi$ for any j, k.

Finally, if Φ is of type G_2, then inspection of the roots verifies the claim.

We now prove (2) case by case, using the descriptions of the root systems given above. If Φ is a simply laced root system, then there are no root strings of length greater than 2, so any choice of $\alpha, \beta \in \Phi$ with $\alpha + \beta \in \Phi$ will suffice.

If Φ is of type B_l, $l \geq 3$, or F_4, set $\alpha = \epsilon_1 - \epsilon_2$ and $\beta = \epsilon_2 - \epsilon_3$. Then $\alpha + \beta = \epsilon_1 - \epsilon_3 \in \Phi$ and $\alpha - \beta \notin \Phi$.

If Φ is of type G_2, then put $\alpha = \alpha_L$ and $\beta = 3\alpha_S + \alpha_L$. Then $\alpha + \beta = 3\alpha_S + 2\alpha_L \in \Phi$ and $\alpha - \beta = -3\alpha_S \notin \Phi$. □

2.3 Chevalley Algebras

Let $\mathfrak{g}$ be a complex semisimple Lie algebra with root system Φ of rank l and Cartan subalgebra H, with root space decomposition $\mathfrak{g} = H \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$. Then $\mathfrak{g}$ has a **Chevalley basis**

$$B = \{e_\alpha : \alpha \in \Phi\} \cup \{h_i : 1 \leq i \leq l\},$$

where $e_\alpha \in \mathfrak{g}_\alpha$ for all α and $h_i \in H$ for $1 \leq i \leq l$, with the following properties:

1. $[h_i, h_j] = 0$ for all $1 \leq i, j \leq l$.

2. $[h_i, e_\alpha] = \langle \alpha, \alpha_i \rangle x_\alpha$ for all $1 \leq i \leq l$, $\alpha \in \Phi$.

3. $[e_\alpha, e_{-\alpha}] = h_\alpha \in \bigoplus_{i=1}^{l} \mathbb{Z}h_i$ for all $\alpha \in \Phi$.

4. If α and β are independent roots and $\beta - r\alpha, \cdots, \beta + q\alpha$ is the α-string through β, then $[e_\alpha, e_\beta] = \begin{cases} 0 & \text{if } \alpha + \beta \notin \Phi \\ \\ \pm(r+1)e_{\alpha+\beta} & \text{if } \alpha+\beta \in \Phi \end{cases}$.

Example. Let $\mathfrak{g} = \mathfrak{sl}_l(\mathbb{C})$, which has rank $l-1$, and let e_{ij} be the l by l matrix with a 1 in entry (i,j) and 0s everywhere else. Then the Chevalley basis for $\mathfrak{g}$ consists of $e_{\epsilon_i - \epsilon_j} = e_{ij}$ for $i \neq j$ and $h_i = e_{ii} - e_{i+1,i+1}$ for $1 \leq i \leq l-1$.

We can use a Chevalley basis to construct Lie algebras over other fields that have the same structure constants as $\mathfrak{g}$. Let $\mathfrak{g}(\mathbb{Z})$ be the $\mathbb{Z}$ span of the Chevalley basis B and, for a field K, define $\mathfrak{g}(K) = \mathfrak{g}(\mathbb{Z}) \otimes_{\mathbb{Z}} K$, the Chevalley algebra of type Φ over K. We will in particular be focusing on the case when K is a finite field.

We note that if $\mathfrak{g} = \mathfrak{g}(\mathbb{C})$ is simple, $\mathfrak{g}(K)$ may fail to be simple. For example, if $\mathfrak{g}$ is type A_l, then $\mathfrak{g}(K) = \mathfrak{sl}_{l+1}(K)$ has a one dimensional center whenever char K divides $l+1$.

2.4 Elementary Chevalley Groups

If L is a Lie algebra, then each $x \in L$ induces a linear map $\operatorname{ad} x : L \to L$ given by $\operatorname{ad} x(y) = [x,y]$. Now let $L = \mathfrak{g}$ be a complex semisimple Lie algebra with root system Φ and Chevalley basis $B = \{e_\alpha, h_i : \alpha \in \Phi, 1 \leq i \leq l\}$. Then $\operatorname{ad} e_\alpha$ is nilpotent for all $\alpha \in \Phi$, so we can define

$$\exp(\operatorname{ad} e_\alpha) = 1 + \operatorname{ad} e_\alpha + \frac{(\operatorname{ad} e_\alpha)^2}{2!} + \cdots + \frac{(\operatorname{ad} e_\alpha)^N}{N!},$$

where $(\operatorname{ad} e_\alpha)^{N+1} = 0$. In fact we can always take $N \leq 3$, as will be seen later. The maps $\exp \operatorname{ad} e_\alpha$ are automorphisms of $\mathfrak{g}$ and have the following important property.

Proposition 2.4.1. *If $\alpha \in \Phi$, then $\dfrac{(\operatorname{ad} e_\alpha)^n}{n!}$ leaves $\mathfrak{g}(\mathbb{Z})$ invariant for any $n \in \mathbb{Z}_{\geq 0}$. As a consequence,* $\exp(\operatorname{ad} e_\alpha)$ *leaves $\mathfrak{g}(\mathbb{Z})$ invariant.*

As a result, if K is a field then $\exp(\operatorname{ad} te_\alpha)$, $t \in K$, can be viewed as an automorphism of the Chevalley algebra $\mathfrak{g}(K)$. Set $x_\alpha(t) = \exp(\operatorname{ad} te_\alpha)$.

Definition. The **adjoint elementary Chevalley group $\mathbb{E}_\Phi^{\mathrm{ad}}(K)$ of type Φ over** K is

$$\mathbb{E}_\Phi^{\mathrm{ad}}(K) = \langle x_\alpha(t) : \alpha \in \Phi, t \in K \rangle \leq \operatorname{Aut}(\mathfrak{g}(K)).$$

The definition makes it clear that $\mathbb{E}_\Phi^{\mathrm{ad}}(K)$ is in fact a linear group. Also, one can replace the field K by any commutative ring R in the defintion, so that one obtains a functor $\mathbb{E}_\Phi^{\mathrm{ad}}$ from commutative rings to groups; this is the perspective we will take from now on.

While we will be working over commutative rings in general, we will also need the fact that adjoint elementary Chevalley groups over fields whose root systems are irreducible are usually simple. The proof of the following proposition can be found in Chapter 4 of [26].

Proposition 2.4.2. *Let Φ be an irreducible root system and let K be a field with at least 4 elements. Then $\mathbb{E}_\Phi^{\mathrm{ad}}(K)$ is a simple group.*

We now let $G = \mathbb{E}_\Phi^{\mathrm{ad}}(R)$, with the understanding that we are working with a fixed root system Φ and a fixed a commutative ring R, and turn our attention to describing the action of G on the Chevalley basis B. These actions can be found immediately from the definition of $x_\alpha(t)$ and the properties of a Chevalley basis. If $\alpha \in \Phi, t \in R$, then

$$\begin{aligned} x_\alpha(t) \cdot e_\alpha &= e_\alpha, \\ x_\alpha(t) \cdot e_{-\alpha} &= e_{-\alpha} + th_\alpha - t^2 e_\alpha, \\ x_\alpha(t) \cdot h_\alpha &= h_\alpha - 2te_\alpha. \end{aligned}$$

If $\alpha, \beta \in \Phi$ are linearly independent, i.e. $\beta \neq \pm\alpha$, then

$$\begin{aligned} x_\alpha(t) \cdot h_\beta &= h_\beta - \langle \alpha, \beta \rangle e_\alpha, \\ x_\alpha(t) \cdot e_\beta &= e_\beta + \sum_{i=1}^{q} M_{\alpha,\beta,i} t^i e_{i\alpha+\beta}, \end{aligned}$$

where $M_{\alpha,\beta,i} \in \{\pm 1, \pm 2, \pm 3\}$.

We now focus on the structure of G. For each $\alpha \in \Phi$, let $X_\alpha = \{x_\alpha(t) : t \in R\}$. We call X_α a **root subgroup** of G. Each X_α is isomorphic to the additive group of R, so that $x_\alpha(t)x_\alpha(s) = x_\alpha(t+s)$. To understand how the root subgroups interact, we use the Chevalley commutator formula: if $\alpha, \beta \in \Phi$ are linearly independent, then

$$[x_\alpha(t), x_\beta(s)] = \prod_{i,j>0} x_{i\alpha+j\beta}(N_{\alpha\beta ij} t^i, s^j),$$

where the product is taken over roots $i\alpha + j\beta \in \Phi$ in order of increasing $i + j$ and the $N_{\alpha\beta ij} \in \{\pm 1, \pm 2, \pm 3\}$ are independent of s and t, with $[e_\alpha, e_\beta] = N_{\alpha\beta 11} e_{\alpha+\beta}$. The following specific cases occur in all root systems, and are the only formulas needed when Φ is simply laced:

$$[x_\alpha(t), x_\beta(s)] = 1 \text{ if } \alpha+\beta \notin \Phi;$$

$$[x_\alpha(t), x_\beta(s)] = x_{\alpha+\beta}(\pm ts) \text{ if } \alpha+\beta \in \Phi, 2\alpha+\beta, \alpha+2\beta, \alpha-\beta \notin \Phi.$$

For all other pairs of roots, excepting some in G_2, one of the following formulas from B_2 applies.

$$[x_{\epsilon_1}(t), x_{\epsilon_2}(s)] = x_{\epsilon_1+\epsilon_2}(\pm 2ts);$$

$$[x_{\epsilon_1-\epsilon_2}(t), x_{\epsilon_2}(s)] = x_{\epsilon_1}(\pm ts)x_{\epsilon_1+\epsilon_2}(\pm ts^2).$$

Remark 2.4.3. We have given formulas where the signs of the coefficients are undetermined. One can choose a consistent set of signs for each root system, which depends on the choice of Chevalley basis, but the specific choice will not be relevant for what follows, so we will continue using ±.

We have defined elementary adjoint Chevalley groups as groups of automorphisms of Lie algebras, but they can also be described abstractly by a group presentation with generators and relations, allowing us to define elementary Chevalley groups more generally. Before we proceed, we define some additional elements of G.

For $\alpha \in \Phi$ and $t \in R^*$, define

$$w_\alpha(t) = x_\alpha(t)x_{-\alpha}(-t^{-1})x_\alpha(t);$$

$$h_\alpha(t) = w_\alpha(t)w_\alpha(1)^{-1}.$$

We will denote $w_\alpha(1)$ by w_α. The elements w_α act via the Weyl group action on roots, and the $h_\alpha(t)$ act diagonally. More precisely, if $\alpha, \beta \in \Phi$, $t \in R$, $s \in R^*$, then

$$w_\alpha x_\beta(t) w_\alpha^{-1} = x_{\sigma_\alpha(\beta)}(\pm t);$$
$$h_\alpha(s) x_\beta(t) h_\alpha(s)^{-1} = x_\beta(s^{\langle \beta, \alpha \rangle} t).$$

The elements w_α and $h_\alpha(t)$ act on the Chevalley basis as follows, where $\alpha, \beta \in \Phi$, $t \in R^*$.

$$h_\alpha(t) \cdot h_\beta = h_\beta,$$
$$h_\alpha(t) \cdot e_\beta = t^{\langle \beta, \alpha \rangle} e_\beta,$$
$$w_\alpha \cdot h_\beta = h_{\sigma_\alpha(\beta)},$$
$$w_\alpha \cdot e_\beta = \pm e_{\sigma_\alpha(\beta)}.$$

The following proposition also serves as a definition of elementary Chevalley groups other than the adjoint elementary Chevalley group. Traditionally these elementary Chevalley groups are defined by actions on admissible lattices similar to the action of the adjoint elementary Chevalley group on $\mathfrak{g}(\mathbb{Z})$, but for us the definition using generators and relations is more convenient and useful.

Proposition 2.4.4. *Let Φ be a root system and let R be a commutative ring. If no irreducible component of Φ has rank 1, let G be the abstract group generated by* *elements* $x_\alpha(r)$, $r \in R$, $\alpha \in \Phi$, *subject to the relations*

$$\begin{aligned} x_\alpha(t)x_\alpha(s) &= x_\alpha(ts), \\ [x_\alpha(t), x_\beta(s)] &= \prod_{i,j>0} x_{i\alpha+j\beta}(N_{\alpha\beta ij}t^i, s^j) \text{ if } \alpha+\beta \neq 0, \\ h_\alpha(t)h_\alpha(s) &= h_\alpha(st) \text{ if } s,t \in R^*. \end{aligned}$$

where $h_\alpha(t) = w_\alpha(t)w_\alpha(1)^{-1}$ *and* $w_\alpha(t) = x_\alpha(t)x_{-\alpha}(-t^{-1})x_\alpha(t)$ *for* $t \in R^*$.

If Φ *has irreducible components* Φ_i *of rank 1, let* G *be as described above, except for* $\alpha \in \Phi_i$ *we replace the commutator relation with* $w_\alpha(t)x_\alpha(s)w_\alpha(-t) = x_{-\alpha}(-t^2 s)$ *for* $s \in R, t \in R^*$. *Then*

$$G/Z(G) \cong \mathbb{E}_\Phi^{\text{ad}}(R).$$

We call G *the* ***universal, or simply connected, elementary Chevalley group*** $\mathbb{E}_\Phi^{\text{sc}}(R)$ ***of type*** Φ ***over*** R. *For any* $N \trianglelefteq Z(G)$, *we say* G/N *is an elementary Chevalley group* $\mathbb{E}_\Phi(R)$ *of type* Φ.

Example. If $\Phi = A_l$ and K is a field, then $\mathbb{E}_\Phi^{\text{sc}}(K) \cong \text{SL}_{l+1}(K)$ and $\mathbb{E}_\Phi^{\text{ad}}(K) \cong \text{PSL}_{l+1}(K)$. If R is a commutative ring, then $\mathbb{E}_\Phi^{\text{sc}}(R) \cong \text{EL}_{l+1}(R)$, the matrix group generated by elementary matrices. Whether or not this group is equal to $\text{SL}_{l+1}(R)$ depends on the ring R.

Remark 2.4.5. By Proposition 2.4.2, if Φ is an irreducible root system and $\mathbb{E}_\Phi$ is an elementary Chevalley group of type Φ, then $\mathbb{E}_\Phi(K)/Z(\mathbb{E}_\Phi(K))$ is a simple group when K is a field with at least 4 elements. Thus we will say that $\mathbb{E}_\Phi$ is a **simple** elementary Chevalley group of type Φ if Φ is irreducible.

One of the nice properties of elementary Chevalley groups is that they are usually perfect.

Proposition 2.4.6. *Let $\mathbb{E}_\Phi$ be a simple elementary Chevalley group. Then $\mathbb{E}_\Phi(\mathbb{F}_p[t])$ is perfect unless $p = 2$ and Φ is of type B_2 or G_2.*

Proof. The statement is proved in chapter 11 of [10] for elementary Chevalley groups over fields, but the same arguments apply to the polynomial ring $\mathbb{F}_p[t]$. □

2.5 Chevalley Groups As Algebraic Groups

We now let K be an algebraically closed field and discuss the connection between elementary Chevalley groups over K and linear algebraic groups, for which we now give some brief background.

A subset V of K^n is called **algebraic** if there exists a subset of polynomials $S \subseteq K[x_1, \cdots, x_n]$ such that

$$V = \{(a_1, \cdots, a_n) \in K^n : f(a_1, \cdots, a_n) = 0 \text{ for all } f \in S\}.$$

That is, V is the zero set of a collection of polynomials. Defining algebraic sets to be closed puts a topology on K^n, called the Zariski topology.

If $n = d^2 + 1$, then $\mathrm{GL}_d(K)$ can naturally be identified with an algebraic subset of K^n. To see this, number the indeterminates as x_0, x_{ij} for $1 \leq i \leq n$. Then $\mathrm{GL}_d(K)$ is the zero set of the polynomial $1 - x_0 \det(x_{ij})$.

Definition. A **linear algebraic group** G is a subgroup of $\mathrm{GL}_d(K)$ which is also an algebraic subset of K^{d^2+1}, for some algebraically closed field K and some natural number d.

If V is an algebraic subset of K^n, put

$$I(V) = \{f \in K[x_1, \cdots, x_n] : f(a_1, \cdots, a_n) = 0 \text{ for all } (a_1, \cdots, a_n) \in V\}.$$

The set $I(V)$ contains all the polynomials which vanish on V. The **coordinate ring** of G is

$$K[G] = K[x_1, \cdots, x_n]/\sqrt{I(V)},$$

where $\sqrt{I(V)} = \{f : f^m = 0 \text{ for some } m \in \mathbb{N}\}$ is the radical of $I(V)$. The linear algebraic group G is connected if and only if its coordinate ring $K[G]$ is an integral domain.

If G is a linear algebraic group and R is a subdomain of K, we say that G is **defined over** R if $I_K(V)$ has a basis of polynomials with coefficients in R.

Example. The special linear group $\mathrm{SL}_d(\mathbb{C})$ is defined over $\mathbb{Z}$ since it has defining equation $1 - \det(x_{ij}) = 0$, which has integer coefficients.

Definition. Let G be a linear algebraic group. The **radical** of G, denoted $\mathrm{rad}(G)$, is the maximal connected, solvable, normal subgroup of G. The group G is **semisimple** if $\mathrm{rad}(G) = 1$ and G is connected.

The following theorem allows us to connect elementary Chevalley groups and linear algebraic groups.

Theorem 2.5.1. *Let K be an algebraically closed field. Every elementary Chevalley group over K is a semisimple linear algebraic group defined over K, and in fact is defined over $\mathbb{Z}$.*

Proof. See Theorem 6 in chapter 5 of [26]. □

Suppose $G \leq \mathrm{GL}_d(K)$ is a linear algebraic group defined over $\mathbb{Z}$. Then we can view the set of defining polynomials S of G as having coefficients in any commutative ring R using the natural homomorphism $\mathbb{Z} \to R$ that maps 1 to 1. We define

$$G(R) = \{(r_{ij}) \in \mathrm{GL}_d(R) : f(r_{ij}) = 0 \ \forall f \in S\}.$$

Definition. Let $\mathbb{E}_\Phi(K)$ be an elementary Chevalley group of type Φ over an algebraically closed field K. Then by Theorem 2.5.1, $\mathbb{E}_\Phi(K)$ is a linear algebraic group G defined over $\mathbb{Z}$. If R is a commutative ring, then we call $G(R)$, as defined above, a **Chevalley group of type Φ over** R.

Remark 2.5.2. If $E_\Phi(K)$ and $G(K)$ are as in the above definition, then $E_\Phi(K) = G(K)$. However, it is not the case that $E_\Phi(R) = G(R)$ for an arbitrary ring R. For example, if Φ is of type A_l, then $\mathbb{E}_\Phi^{sc}(R) = \mathrm{EL}_{l+1}(R)$, while $G(R) = \mathrm{SL}_{l+1}(R)$, which contains $\mathrm{EL}_{l+1}(R)$ but in general can be larger.

We note that every Chevalley group G embeds into SL_d for some d.

The Chevalley group $G(\mathbb{C})$ corresponding to $\mathbb{E}_\Phi^{\mathrm{sc}}(\mathbb{C})$ is simply connected, topologically, so when referring to any Chevalley or elementary Chevalley group arising from E_Φ^{sc}, we say it is simply connected or of simply conected type.

Both Chevalley groups and elementary Chevalley groups have certain nice properties with regards to changing rings of coefficients. If G is a Chevalley group and $R \subseteq S$ are commutative rings, then it is clear that $G(R) = G(S) \cap \mathrm{GL}_d(R)$. In contrast, $\mathbb{E}_\Phi(R) \neq \mathbb{E}_\Phi(S) \cap \mathrm{GL}_d(R)$ in general.

However, elementary Chevalley groups are better behaved under homomorphisms. Any ring homomorphism $\varphi : R \to S$ induces a group homomorphism $\varphi^* : \mathrm{GL}_d(R) \to \mathrm{GL}_d(S)$, where $\varphi^*(a_{ij}) = (\varphi(a_{ij}))$. If φ is surjective, then the image of $\mathbb{E}_\Phi(R)$ is $\mathbb{E}_\Phi(S)$, since the generators $x_\alpha(t)$ of $\mathbb{E}_\Phi(R)$ are mapped to $x_\alpha(\varphi(t))$, which generate $\mathbb{E}_\Phi(S)$. However, in the case of Chevalley groups, the image of $G(R)$ is not necessarily $G(S)$.

When an elementary Chevalley group agrees with a Chevalley group, we can take advantage of both sets of nice properties. The following result, proved in [26] for the case R is a Euclidean domain and in [1] when R is semi-local, i.e. has only finitely many maximal ideals, gives a set of conditions for which this is the case.

Lemma 2.5.3. *If R is a Euclidean domain or a semi-local ring and G is a simply connected Chevalley group of type Φ, then $G(R) = \mathbb{E}^{\mathrm{sc}}_\Phi(R)$.*

We will mostly be concerned with Chevalley groups G associated with irreducible root systems. Following Remark 2.4.5, in such cases we will say G is a **simple Chevalley group**.

Let G be a simple Chevalley group of type Φ, considered as an algebraic group. If F is a field, one can recover the Chevalley algebra $\mathfrak{g}(F)$ of type Φ from $G(F)$ as follows. Details can be bound in chapter 2 of [21] Let $F[\epsilon] = F[x]/(x^2)$, so $\epsilon^2 = 0$, and define

$$\mathrm{Lie}(G(F)) = \ker(G(F[\epsilon]) \to G(F)),$$

where the map is induced by sending ϵ to 0. If we fix an embedding $G \hookrightarrow \mathrm{SL}_d(F)$, then $\mathrm{Lie}(G(F)) = \{I_d + \epsilon A \in \mathrm{SL}_d(F) : I_d + \epsilon A \in G(F)\}$ can be viewed as an F-vector space because

$$(I_d + \epsilon A)(I_d + \epsilon B) = I_d + \epsilon(A + B).$$

Thus $\mathrm{Lie}(G(F))$ naturally embeds into $\mathfrak{sl}_d(F)$ under the map $I_d + \epsilon A \to A$, and defining the Lie bracket to be $[A, B] = AB - BA$ turns $\mathrm{Lie}(G(F))$ into a Lie algebra over $\mathbb{F}$ which is isomorphic to the Chevalley algebra $\mathfrak{g}(F)$. This embedding of $\mathfrak{g}(F)$ into $\mathfrak{sl}_d(F)$ is particularly nice in that the action of $G(F)$ on $\mathfrak{g}(F)$ by conjugation, using matrix multiplication, is the same as the adjoint action of $G(F)$ on $\mathfrak{g}(F)$ given in section 2.4.

If $\alpha \in \Phi$ and $t \in F$, then keeping in mind that $\epsilon^2 = 0$,

$$x_\alpha(\epsilon t) = \exp(\epsilon \,\mathrm{ad}\, te_\alpha) = 1 + \epsilon \,\mathrm{ad}\, te_\alpha \in \mathrm{Lie}(G(F)).$$

Then under the map $\mathrm{Lie}(G(F)) \to \mathfrak{sl}_d(F)$, $x_\alpha(\epsilon t)$ is sent to $\mathrm{ad}\, te_\alpha$. In particular, $x_\alpha(\epsilon)$ is mapped to $\mathrm{ad}\, e_\alpha$. Since ad is a Lie algebra homomorphism which is faithful on $\bigoplus_{\alpha \in \Phi} Fe_\alpha$, we may identify the image $\mathrm{ad}\, e_\alpha$ with e_α when working in the Lie algebra.

We finish this section with some size estimates of Chevalley groups and linear algebraic groups over finite fields. To state the results we need the notion of the dimension of a linear algebraic group. Dimension is more naturally viewed as a geometric property, but the equivalent algebraic definition is more applicable to our setting.

Definition. Let G be a linear algebraic group. The **dimension** of G, denoted by $\dim(G)$, is the transcendence degree of the field of quotients of $K[G]$.

Remark 2.5.4. The construction of $\mathrm{Lie}(G(F))$ done above can be carried out for any linear algebraic group. The dimension of this Lie algebra is equal to the dimension of G as defined above. In particular, if G is a Chevalley group with corresponding Chevalley algebra $\mathfrak{g}$, then $\dim(G) = \dim(\mathfrak{g})$.

The values of $a(G)$ and $\dim(G)$ for simple Chevalley groups G are given in Table 1.1; $a(G)$ is the minimal codimension of a proper parabolic subgroup of G, but the following lemma also acts as a definition for $a(G)$. We note that $a(G)$ and $\dim(G)$ depend only on the root system Φ of G (and not, for example, on whether G is simply connected or adjoint).

Lemma 2.5.5. *Let G be a simple Chevalley group with an embedding into SL_d, q be a prime power, and $H \leq G(\mathbb{F}_q)$ be a proper subgroup of minimal index. Then $|G(\mathbb{F}_q)/Z(G(\mathbb{F}_q))| \geq \frac{1}{2d}q^{\dim(G)}$ and $\frac{1}{2}q^{a(G)} \leq [G(\mathbb{F}_q) : H] \leq 2q^{a(G)}$.*

Proof. The size bound of $|G(\mathbb{F}_q)/Z(\mathbb{F}_q)|$ follows from Theorem 25, §9, in [26]. The index of the largest maximal subgroup of $G(\mathbb{F}_q)$ can be found in [16] (Theorem 5.2.2) if G is of type A_l, B_l, C_l, or D_l, and in [28], [29] for the remaining cases. □

Lemma 2.5.6. *Let G be a linear algebraic group defined over $\mathbb{Z}$ and q be a prime power. There exists a constant C independent of q such that $|G(\mathbb{F}_q)| \leq Cq^{\dim(G)}$.*

Proof. The connected component of G containing the identity is a normal subgroup of finite index, so it suffices to prove the lemma in the case G is connected, so that $K[G]$ is an integral domain.

There is a natural bijection between $G(\mathbb{F}_q)$ and $\mathrm{Hom}_K(K[G], \mathbb{F}_q)$, so we bound the size of the latter. By Noether normalization, $K[G]$ is a finitely generated module over a polynomial ring $K[x_1, \cdots, x_d]$, where $d = \dim(G)$. If $K[G]$ is generated as a module by $y_1, \cdots, y_m$, then each y_i is integral over $K[x_1, \cdots, x_d]$, so for each $1 \leq i \leq m$ we can find a polynomial

$$f_i(x_1, \cdots, x_d, Y) \in k[x_1, \cdots, x_d][Y]$$

such that $f_i(x_1, \cdots, x_d, y_i) = 0$. Let $c = \max\limits_{1 \leq i \leq m} \deg f_i$. An element $\varphi \in \mathrm{Hom}_K(K[G], \mathbb{F}_q)$ is determined by the images of the x_i and y_j. Given choices of $\varphi(x_i)$, which can be made arbitrarily, for each $1 \leq j \leq m$ there are at most c choices of $\varphi(y_j)$ that will satisfy $f_j(\varphi(x_1), \cdots, \varphi(x_d), \varphi(y_j)) = 0$. Thus $|\mathrm{Hom}_K(K[G], \mathbb{F}_q)| \leq c^m q^d$, so $C = c^m$ is a constant satisfying the lemma. □

2.6 Invariant Ideals Under The Adjoint Action Of Chevalley Groups

We now examine the action of a Chevalley group on its Lie algebra over a finite field in more detail.

Proposition 2.6.1. *Let $\mathbb{F}$ be a finite field of characteristic p and G a simple simply connected Chevalley group. For all but finitely many p, the adjoint action of $G(\mathbb{F})$ on $\mathfrak{g}(\mathbb{F})$ is irreducible. The exceptions are given in Table* 2.1*, along with the largest possible dimension of a proper ideal $I \subseteq \mathfrak{g}(\mathbb{F})$ invariant under the action of $G(\mathbb{F})$ in those cases. If G is of type B_2 and $p = 2$, then any invariant ideal I is either the center or contains $\mathbb{F}e_\alpha$ for all short roots α.*

Proof. See Theorem 2.1 in [11]. □

We will be concerned with $\mathbb{F}_p$-subspaces of $\mathfrak{g}(\mathbb{F})$ which are invariant under the action of a simple simply connected Chevalley group $G(\mathbb{F})$. The following lemma allows us to apply Proposition 2.6.1 to this situation.

Lemma 2.6.2. *Let $\mathbb{F}$ be a finite field of characteristic p such that $|\mathbb{F}| \geq 4$, and let G be a simple simply connected Chevalley group of type Φ. Let V be a proper $\mathbb{F}_p$-subspace of $\mathfrak{g}(\mathbb{F})$. If V is $G(\mathbb{F})$-invariant, then $\mathbb{F}V$, the $\mathbb{F}$-subspace spanned by V, is a proper ideal of $\mathfrak{g}(\mathbb{F})$ which is invariant under the action of $G(\mathbb{F})$.*

Proof. Let Φ have rank l and fix a Chevalley basis $B = \{e_\alpha : \alpha \in \Phi\} \cup \{h_1, \cdots, h_l\}$

Φ	p	$\max\dim(I)$	$\min\operatorname{codim}(I)$
$A_l, l \geq 2$	$p\|(l+1)$	1	l^2+2l-1
$B_l, l \geq 3$	2	$2l+2$	$2l^2-l-2$
$C_l, l \geq 2$	2	$2l^2-l$	$2l$
$D_l, l \geq 4$	2	2	$2l^2-l-2$
G_2	3	7	7
F_4	2	26	26
E_6	3	1	77
E_7	2	1	132

Table 2.1:

of $\mathfrak{g}(\mathbb{F})$. The $\mathbb{F}$-subspace $\mathbb{F}V$ is an ideal of $\mathfrak{g}(\mathbb{F})$ if $[\mathfrak{g}(\mathbb{F}), \mathbb{F}V] \subseteq \mathbb{F}V$, but it is sufficent to check that $[e_\alpha, \mathbb{F}V] \subseteq \mathbb{F}V$ for all $\alpha \in \Phi$, as we now show.

Recall that Φ has a base $\Pi = \{\alpha_1, \cdots, \alpha_l\}$ and $h_i = [e_{\alpha_i}, e_{-\alpha_i}]$ for $1 \leq i \leq l$. Then using the Jacobi identity, for $v \in \mathbb{F}V$ and $1 \leq i \leq l$,

$$[h_i, v] = [[e_{\alpha_i}, e_{-\alpha_i}], v] = [e_{\alpha_i}, [e_{-\alpha_i}, v]] - [e_{-\alpha_i}, [e_{\alpha_i}, v]].$$

Thus if $[e_\alpha, \mathbb{F}V] \subseteq \mathbb{F}V$ for all $\alpha \in \Phi$, then $[h_i, \mathbb{F}V] \subseteq \mathbb{F}V$ as well, so $[\mathfrak{g}(\mathbb{F}), \mathbb{F}V] \subseteq \mathbb{F}V$ and $\mathbb{F}V$ is an ideal. We now proceed to the proof of the lemma.

First assume that V is actually an $\mathbb{F}$-subspace of $\mathfrak{g}(\mathbb{F})$, so $V = \mathbb{F}V$. If $\alpha \in \Phi$, $\lambda \in \mathbb{F}$, and $v \in V$ then

$$x_\alpha(\lambda) \cdot v - v = \lambda[e_\alpha, v] + \lambda^2 \frac{1}{2}[e_\alpha, [e_\alpha, v]] + \lambda^3 \frac{1}{6}[e_\alpha, [e_\alpha, [e_\alpha, v]]] \in V. \qquad (2.6.1)$$

As noted in the construction of elementary Chevalley groups, the last two terms in (2.6.1) can be considered as elements in the $\mathbb{Z}$-span of B. With this interpretation the equation is valid for any characteristic.

Using (2.6.1) for three distinct nonzero elements $s, t, u \in \mathbb{F}$, one can use linear combinations to obtain $[e_\alpha, v] \in V$ for all $v \in V$ and $\alpha \in \Phi$. To see this, fix $v \in V$ and $\alpha \in \Phi$ and write the right hand side of (2.6.1) as $\lambda z_1 + \lambda^2 z_2 + \lambda^3 z_3 \in V$. Since V is an $\mathbb{F}$-subspace, this implies $z_1 + \lambda z_2 + \lambda^2 z_3 \in V$. Using s, t, u in place of λ, we have

$$\begin{aligned} v_1 &= z_1 + s z_2 + s^2 z_3 \in V, \\ v_2 &= z_1 + t z_2 + t^2 z_3 \in V, \\ v_3 &= z_1 + u z_2 + u^2 z_3 \in V. \end{aligned}$$

Then

$$\begin{aligned} v_4 &= t^2 v_1 - s^2 v_2 = (t^2 - s^2) z_1 + st(t-s) z_2 \in V, \\ v_5 &= u^2 v_1 - s^2 v_3 = (u^2 - s^2) z_1 + su(u-s) z_2 \in V. \end{aligned}$$

Finally,

$$\begin{aligned} u(u-s)v_4 - t(t-s)v_5 &= (u(u-s)(t^2 - s^2) - t(t-s)(u^2 - s^2)) z_1 \\ &= s(u-s)(t-s)(u-t) z_1 \in V. \end{aligned}$$

Since s, t, and u are all nonzero and distinct, we conclude that $z_1 = [e_\alpha, v] \in V$.

Thus V is an ideal of $\mathfrak{g}(\mathbb{F})$. Since V is assumed to be proper, $V = \mathbb{F}V$ is a proper ideal of $\mathfrak{g}(\mathbb{F})$, so the lemma is proved in this case.

Now assume V is not an $\mathbb{F}$-subspace of $\mathfrak{g}(\mathbb{F})$. Then $\mathbb{F}V$ is a $G(\mathbb{F})$-invariant $\mathbb{F}$-subspace of $\mathfrak{g}(\mathbb{F})$ and thus an ideal by the above argument. It remains to show that $\mathbb{F}V \neq \mathfrak{g}(\mathbb{F})$.

When $p \neq 2$ and Φ is not of type G_2, this is straightforward. For any $s \in \mathbb{F}$, $\alpha \in \Phi$, and $v \in V$,

$$x_\alpha(s) \cdot v + x_\alpha(-s) \cdot v = 2s[e_\alpha, v] \in V,$$

so $s[e_\alpha, v] \in V$, and thus also $s[h_i, v] \in V$ for all $s \in \mathbb{F}$, $1 \leq i \leq l$, $v \in V$ by the argument at the beginning of this proof.

Therefore the $\mathbb{F}$-span of $\{[x, v] : x \in \mathfrak{g}(\mathbb{F}), v \in V\}$ is contained in V and is not all of $\mathfrak{g}(\mathbb{F})$. But this set is just $[\mathfrak{g}(\mathbb{F}), \mathbb{F}V]$. Since char $\mathbb{F} \neq 2$, $[\mathfrak{g}(\mathbb{F}), \mathfrak{g}(\mathbb{F})] = \mathfrak{g}(\mathbb{F})$, so we must have $\mathbb{F}V \neq \mathfrak{g}(\mathbb{F})$.

Treating the general case requires using the structure of each root system. We now assume p is any prime, until we reach the case of Φ being of type C_l.

Assume $\mathbb{F}V = \mathfrak{g}(\mathbb{F})$; we will show that this implies $V = \mathfrak{g}(\mathbb{F})$, a contradiction. Let E_L and E_S be the $\mathbb{F}$-subspaces of $\mathfrak{g}(\mathbb{F})$ spanned by $\{e_\alpha : \alpha \text{ long}\}$ and $\{e_\alpha : \alpha \text{ short}\}$, respectively, so $\mathfrak{g}(\mathbb{F}) = H \oplus E_S \oplus E_L$, with the convention that $E_S = 0$ if Φ is simply laced. Since V is $G(\mathbb{F})$-invariant and $x_\alpha(t) \cdot e_{-\alpha} = e_{-\alpha} + th_\alpha - t^2 e_\alpha$ for $\alpha \in \Phi$, $t \in \mathbb{F}$, to show that $V = \mathfrak{g}(\mathbb{F})$ it suffices to show $E_S \oplus E_L \subseteq V$.

Fix $v \in V$, which we write as

$$v = h + \sum_{\beta \in \Phi} s_\beta e_\beta \in V, \tag{2.6.2}$$

where $h \in H$. If $\gamma \in \Phi$ is a long root, then $2\gamma + \delta \in \Phi$ if and only if $\delta = -\gamma$ by Lemma 2.2.2, so for any $t \in \mathbb{F}$,

$$x_\gamma(t) \cdot v - v = t[e_\gamma, v] - t^2 s_{-\gamma} e_\gamma \in V. \tag{2.6.3}$$

Assume Φ is not of type C_l, $l \geq 2$. By Lemma 2.2.2, we can find long roots α andβ such that $\alpha + \beta \in \Phi$ and $\alpha - \beta \notin \Phi$. Also by Lemma 2.2.2, if $\gamma \in \Phi$, then $\gamma - 2\alpha \in \Phi$ if and only if $\gamma = \alpha$.

Now put $\gamma_1 = \beta$, $\gamma_2 = -\alpha$, and $\gamma_3 = -(\alpha + \beta)$. We show that for any $t \in \mathbb{F}$ and any $v \in V$ written as in (2.6.2),

$$t[e_{\gamma_3}, [e_{\gamma_2}, [e_{\gamma_1}, v] - s_{-\gamma_1} e_{\gamma_1}]] = \pm t s_\alpha e_{-\alpha} \in V.$$

So fix $t \in \mathbb{F}$, $v \in V$. Set $v_1 = [e_{\gamma_1}, v] - s_{-\gamma_1} e_{\gamma_1}$, which is in V by (2.6.3). Since

$$-(\gamma_1 + \gamma_2) = \alpha - \beta \notin \Phi,$$

the coefficient of $e_{-\gamma_2}$ in v_1 is 0, and thus $v_2 = [e_{\gamma_2}, v_1] = [e_{\gamma_2}, [e_{\gamma_1}, v]] \in V$ by (2.6.3). We also have $v_2 \in E_S \oplus E_L$. Similarly, the coefficient of $e_{-\gamma_3}$ in v_2 is 0 since $-(\gamma_1 + \gamma_2 + \gamma_3) = 2\alpha \notin \Phi$, so $v_3 = t[e_{\gamma_3}, v_2] \in V$ by (2.6.3) and $v_3 \in E_S \oplus E_L$.

For any $\gamma \in \Phi$,

$$\gamma + \gamma_1 + \gamma_2 + \gamma_3 = \gamma - 2\alpha,$$

and $\gamma - 2\alpha \in \Phi$ if and only if $\gamma = \alpha$. We also have $\gamma_2 + \gamma_3 = \beta - 2\alpha \notin \Phi$, so in fact $v_3 = \pm t s_\alpha e_{-\alpha}$ as claimed.

By assumption, $\mathbb{F}V = \mathfrak{g}(\mathbb{F})$, so there exists $v \in V$ with $s_\alpha \neq 0$, when v is written as in (2.6.2). Using the above computation, we conclude that $\mathbb{F}e_{-\alpha} \subseteq V$.

Since $\langle w_\gamma : \gamma \in \Phi \rangle \leq G(\mathbb{F})$ acts transitively on $\{e_\alpha : \alpha \text{ long}\}$, we conclude that $E_L \subseteq V$. If Φ is simply laced, this immediately implies $V = \mathfrak{g}(\mathbb{F})$. We treat the remaining root systems case by case, using the fact that $E_L \subseteq V$. We use the descriptions of the root systems given in section 2.2.

If Φ is of type $B_l, l \geq 3$ or F_4, then for $t \in \mathbb{F}$,

$$x_{\epsilon_1}(t) \cdot e_{\epsilon_2-\epsilon_1} - e_{\epsilon_2-\epsilon_1} = \pm t e_{\epsilon_2} \pm t^2 e_{\epsilon_1+\epsilon_2} \in V,$$

so $\mathbb{F}e_{\epsilon_2} \subseteq V$. By the transitive action of $G(\mathbb{F})$ on $\{e_\alpha : \alpha \text{ short}\}$, $E_S \subseteq V$ and hence $V = \mathfrak{g}(\mathbb{F})$.

If Φ is of type G_2, then for $t \in \mathbb{F}$,

$$x_{-\alpha_S-\alpha_L}(t) \cdot e_{\alpha_L} - e_{\alpha_L} = \pm t e_{-\alpha_S},$$

so $\mathbb{F}e_{-\alpha_S} \subseteq V$. Hence $E_L \subseteq V$ and $V = \mathfrak{g}(\mathbb{F})$.

In every case we contradict the assumption that V is proper, so we must have $\mathbb{F}V \neq \mathfrak{g}(\mathbb{F})$.

We now consider the remaining case. Assume Φ is of type C_l, $l \geq 2$ and $p = 2$. Let $\gamma_1 = 2\epsilon_2$ and $\gamma_2 = -2\epsilon_1$. These are long roots with $\gamma_1 + \gamma_2 \notin \Phi$ and $\gamma + \gamma_1 + \gamma_2 \in \Phi$ if and only if $\gamma = \epsilon_1 - \epsilon_2$. Then by the same reasoning as in the argument for root systems not of type C_l, for $t \in \mathbb{F}$ and $v \in V$ written as in (2.6.2) we have

$$t[e_{\gamma_2}, [e_{\gamma_1}, v] - s_{-\gamma_1} e_{\gamma_1}] = \pm t s_{\epsilon_1-\epsilon_2} e_{\epsilon_2-\epsilon_1} \in V.$$

Therefore $\mathbb{F}e_{\epsilon_2-\epsilon_1} \subseteq V$ and hence $E_S \subseteq V$.

To show $E_L \subseteq V$, let $v \in V$ with $s_{2\epsilon_2} \neq 0$. Since $E_S \subseteq V$, we can write v as $v = h + \sum_{\alpha \text{ long}} s_\alpha e_\alpha$. The only long roots α satisfying $\epsilon_1 - \epsilon_2 + \alpha \in \Phi$ are $\alpha = -2\epsilon_1$ and $\alpha = 2\epsilon_2$, and

$$x_{\epsilon_1-\epsilon_2}(1) \cdot e_{-2\epsilon_1} - e_{-2\epsilon_1} = \pm e_{-\epsilon_1-\epsilon_2} \pm e_{2\epsilon_1},$$

$$x_{\epsilon_1-\epsilon_2}(1) \cdot e_{2\epsilon_2} - e_{2\epsilon_2} = \pm e_{\epsilon_1+\epsilon_2} \pm e_{-2\epsilon_2}.$$

Therefore

$$x_{\epsilon_1-\epsilon_2}(1) \cdot v - v = se_{\epsilon_1-\epsilon_2} \pm s_{-2\epsilon_1} e_{-\epsilon_1-\epsilon_2} \pm s_{-2\epsilon_1} e_{2\epsilon_1} \pm s_{2\epsilon_2} e_{\epsilon_1+\epsilon_2} \pm s_{2\epsilon_2} e_{-2\epsilon_2} \in V.$$

Again using the fact that $E_S \subseteq V$, we conclude that

$$v_1 = \pm s_{2\epsilon_2} e_{2\epsilon_1} \pm s_{-2\epsilon_1} e_{-2\epsilon_2} \in V.$$

Then if $t \in \mathbb{F}$,

$$x_{e_{-\epsilon_1-\epsilon_2}}(t) \cdot v_1 - v_1 = \pm t s_{2\epsilon_2} e_{\epsilon_1-\epsilon_2} \pm t^2 s_{2\epsilon_2} e_{-2\epsilon_2} \in V,$$

so $t^2 s_{2\epsilon_2} e_{-2\epsilon_2} \in V$. But $\mathbb{F}$ is a finite field with characteristic 2, so $\mathbb{F}^2 = \mathbb{F}$, and thus we conclude that $E_L \subseteq V$ and hence $V = \mathfrak{g}(\mathbb{F})$, the desired contradiction. □

Chapter 3

Number Theory Background

3.1 The Prime Number Theorem

In this section we recall estimates on the number of primes up to a certain size in $\mathbb{Z}$ and $\mathbb{F}_q[t]$, where q is a prime power. Asymptotics will be measured as follows. If f, g are two real valued functions on $\mathbb{R}$ or $\mathbb{N}$, we will write $f \sim g$ if $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = 1$.

We start with primes in $\mathbb{Z}$. Let $\pi(x)$ be the number of primes $p \in \mathbb{Z}$ with $p \leq x$. Then the classical prime number theorem states that

$$\pi(x) \sim \frac{x}{\log x},$$

where we write $\log x$ to indicate the natural logarithm of x. In addition to this result, we will need the following two equivalent statements of the prime number theorem:

$$\operatorname{lcm}(1, \cdots, n) \sim e^n,$$

$$\prod_{\substack{p \leq n \\ p \text{ prime}}} p \sim e^n.$$

Primes in $\mathbb{F}_q[t]$ are irreducible polynomials, which we will always assume to be monic, so counting primes up to a certain size is the same as counting the number of irreducible polynomials in $\mathbb{F}_q[t]$ of a given degree. This was computed by Gauss and can be found, for example, in [24]. Before stating the formula, recall that the Mobius funciton μ is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n = p_1 \cdots p_k \text{ for distinct primes } p_i \\ 0 & \text{otherwise} \end{cases} .$$

Proposition 3.1.1. *If q is a prime power, then the number of irreducible polynomials of degree k in $\mathbb{F}_q[t]$ is*

$$I_q(k) = \frac{1}{k} \sum_{d|k} \mu(d) q^{k/d}.$$

3.2 Integral Extensions

The reference for this section is Chapter 1 of [18]. All rings in this chapter will be integral domains, i.e. commutative rings with unity which do not contain zero divisors. We begin with some definitions.

Definition. Let A be a ring contained in a field L. An element $x \in L$ is **integral** over A if it satisfies an equation

$$x^n + a_{n-1}^n + \cdots + a_1 x + a_0 = 0,$$

where n is a positive integer and each $a_i \in A$. Such an equation is called an integral equation.

Definition. If $A \subseteq B$ are rings and every element of B is integral over A, then we say B is integral over A, or that B is an integral extension of A.

A subset $S \subseteq A$ containing 1 is called multiplicatively closed if $s_1 s_2 \in S$ for all $s_1, s_2 \in S$. If $S \subseteq A$ is multiplicatively closed, then $S^{-1}A$, the set of quotients a/s for $a \in A, s \in S$, is a ring, and there is a canonical inclusion of A into $S^{-1}A$. If $\mathfrak{p}$ is a prime ideal of A, then $S = A \setminus \mathfrak{p}$ is a multiplicatively closed set. Then $S^{-1}A$ is the localization of A at $\mathfrak{p}$ and denoted by $A_{\mathfrak{p}}$. If B is a ring containing A, we will denote $S^{-1}B$ by $B_{\mathfrak{p}}$.

We now collect some facts about integral extensions.

Proposition 3.2.1. *Let A be a ring with field of fractions K.*

1. *If A is contained in a field L, then the set of elements in L which are integral over A is a ring, which is called the **integral closure** of A in L. The ring A is said to be **integrally closed** if it is equal to its integral closure in its field of fractions K.*

2. *If A is a unique factorization domain, then A is integrally closed.*

3. *If L/K is a field extension, B is the integral closure of A in L, and $\mathfrak{p}$ is a prime ideal of A, then $A_{\mathfrak{p}}$ is integrally closed and $B_{\mathfrak{p}}$ is the integral closure of $A_{\mathfrak{p}}$ in L.*

Our main interest in integral extensions concerns the properties of prime ideals. Let B be an integral extension of A and let $\mathfrak{p}$ be a prime ideal of A. If $\mathfrak{P}$ is a prime ideal of B such that $\mathfrak{P} \cap A = \mathfrak{p}$, then we say $\mathfrak{P}$ lies above $\mathfrak{p}$. In this case, there is a natural injection $A/\mathfrak{p} \to B/\mathfrak{P}$. By the following proposition, if $A/\mathfrak{p}$ is a field, then $B/\mathfrak{P}$ is a field extension of $A/\mathfrak{p}$.

Proposition 3.2.2. *Let B be an integral extension of A, and let $\mathfrak{p}$ be a prime ideal of A. Then there exists a prime ideal $\mathfrak{P}$ of B lying above $\mathfrak{p}$. If $\mathfrak{P}$ lies above $\mathfrak{p}$, then $\mathfrak{P}$ is maximal if and only if $\mathfrak{p}$ is maximal.*

We close this section by showing the connection between the maximal ideals lying above $\mathfrak{p}$ and the factorization of an irreducible polynomial $f(x) \in A[x]$ when taken modulo $\mathfrak{p}$.

Let A be integrally closed, with field of fractions K. Let $f(x) \in A[x]$ be an irreducible polynomial, and let $\mathfrak{p}$ be a maximal ideal of A. Define the discriminant of f to be

$$\Delta(f) = \prod_{i \neq j} (\alpha_i - \alpha_j),$$

where the α_i are the roots of f in some algebraic closure of K. The image of $f(x)$ in $(A/\mathfrak{p})[x]$, which we denote by $\overline{f}(x)$, may fail to be irreducible. The factorization of $\overline{f}(x)$ is controlled by the maximal ideals lying above $\mathfrak{p}$ in a certain integral extension. The following lemma shows when $\overline{f}(x)$ factors into distinct linear factors.

Lemma 3.2.3. *Let A be an integrally closed ring with field of fractions K, and let $f(x) \in A[x]$ be an irreducible, separable polynomial. Set $L = K[x]/f(x) \cong K[\alpha]$ for some root α of $f(x)$ and let B be the integral closure of A in L. Let $\mathfrak{p}$ be a* *maximal ideal of A with $\Delta(f) \notin \mathfrak{p}$. If $B/\mathfrak{P} = A/\mathfrak{p}$ for every $\mathfrak{P}$ lying above $\mathfrak{p}$, then $\overline{f}(x) \in (A/\mathfrak{p})[x]$ is a product of distinct linear factors.*

Proof. We prove the statement by localizing at $\mathfrak{p}$. By Proposition 3.2.1, $B_{\mathfrak{p}}$ is the integral closure of $A_{\mathfrak{p}}$. If $\mathfrak{P}$ lies above $\mathfrak{p}$, then $\mathfrak{P} \cap A = \mathfrak{p}$, so no element of $\mathfrak{P}$ becomes a unit in $B_{\mathfrak{p}}$, and hence $\mathfrak{P}B_{\mathfrak{p}} \neq B_{\mathfrak{p}}$. We observe that $B_{\mathfrak{p}}/\mathfrak{P}B_{\mathfrak{p}} = B/\mathfrak{P}$.

Denote $A/\mathfrak{p}$ by $\mathbb{F}$ and let z be a root of an irreducible factor $\overline{P}(x)$ of $\overline{f}(x) \in \mathbb{F}[x]$. Since $\Delta(f)$ is a unit in $A_{\mathfrak{p}}$, $B_{\mathfrak{p}} = A_{\mathfrak{p}}[\alpha]$ by Lemma 5.3 in [12]. Then the map $B_{\mathfrak{p}} = A_{\mathfrak{p}}[\alpha] \to \mathbb{F}[z]$ given by $g(\alpha) \to g(z) \bmod \mathfrak{p}$ for $g(X) \in A_{\mathfrak{p}}[X]$ is a ring homomorphism. Its kernel is a maximal ideal $\mathfrak{P}B_{\mathfrak{p}}$ for some maximal ideal $\mathfrak{P}$ of B. But

$$B_{\mathfrak{p}}/\mathfrak{P}B_{\mathfrak{p}} = B/\mathfrak{P} = A/\mathfrak{p} \cong \mathbb{F},$$

so $z \in \mathbb{F}$. Hence $\overline{P}(x)$ is linear. We also have $\Delta(\overline{f}) \neq 0$ since $\Delta(f) \notin \mathfrak{p}$, so $\overline{f}(x)$ is separable. Hence $\overline{f}(x)$ is the product of distinct linear factors. □

3.3 The Chebotarev Density Theorem

Let K be a global field, i.e. a finite extensions of $\mathbb{Q}$ or the function field $\mathbb{F}_p(t)$. The ring of integers of K, denoted by $\mathcal{O}_K$, is the integral closure of $\mathbb{Z}$ or $\mathbb{F}_p[t]$ in K, where char K is 0 or p, respectively. Let L/K be a finite Galois extension with Galois group G. Then $\mathcal{O}_L$ is the integral closure of $\mathcal{O}_K$ in L. Let $\mathfrak{p}$ be a maximal ideal of $\mathcal{O}_K$ and let $\mathfrak{P}$ be a maximal ideal of $\mathcal{O}_L$ lying above $\mathfrak{p}$. We have $\sigma\mathcal{O}_L = \mathcal{O}_L$ for all $\sigma \in G$, so if

$\sigma\mathfrak{P} = \mathfrak{P}$, then there is a natural action of σ on $\mathcal{O}_L/\mathfrak{P}$ which leaves $\mathcal{O}_K/\mathfrak{p}$ invariant. We call the group $G_\mathfrak{P} = \{\sigma \in G : \sigma\mathfrak{P} = \mathfrak{P}\}$ the **decomposition group** of $\mathfrak{P}$.

The field extension $(\mathcal{O}_L/\mathfrak{P})/(\mathcal{O}_K/\mathfrak{p})$ is a finite extension of a finite field, so it is a Galois extension; the natural map $G_\mathfrak{P} \to \mathrm{Gal}(\overline{B}/\overline{A})$ is surjective. The kernel of this map is called the **inertia group** $I_\mathfrak{P}$ of $\mathfrak{P}$.

The Galois group G acts transitively on the maximal ideals lying above $\mathfrak{p}$, so the decomposition groups are all conjugate, as are the inertia groups. If the inertia group of $\mathfrak{P}$ is trivial, then the inertia group of every maximal ideal lying above $\mathfrak{p}$ is trivial, and we say $\mathfrak{p}$ is **unramified** in L.

We define the norm of $\mathfrak{p}$ to be $N\mathfrak{p} = |\mathcal{O}_K/\mathfrak{p}|$. Then if $\mathfrak{p}$ is unramified in L and $\mathfrak{P}$ lies above $\mathfrak{p}$, then $G_\mathfrak{P}$ is isomorphic to the Galois group of a finite extension of a finite field of size $N\mathfrak{p}$. This group is cyclic with a canonical generator φ, called the Frobenius automorphism, which acts by $\varphi x = x^{N\mathfrak{p}}$ on $\mathcal{O}_L/\mathfrak{P}$. Since $G_\mathfrak{P} \leq G$, this Frobenius automorphism can be realized as an element of G. The conjugacy class of this element in G depends only on $\mathfrak{p}$; we will denote this conjugacy class or an element of it by the Artin symbol $\left(\dfrac{L/K}{\mathfrak{p}}\right)$. We note that by writing this symbol, it is implied that $\mathfrak{p}$ is unramified.

Let $P(K)$ be the set of maximal ideals of $\mathcal{O}_K$. If $x \in \mathbb{R}$, define

$$\pi(x) = |\{\mathfrak{p} \in P(K) : N\mathfrak{p} \leq x\}|.$$

Definition. Let $S \subseteq P(K)$ and set $\pi_S(x) = |\{p \in S : p \leq x\}|$. Then S has **natural**

density λ in $P(K)$ if $\lim\limits_{x\to\infty} \dfrac{\pi_S(x)}{\pi(x)} = \lambda$, i.e. if $\pi_S(x) \sim \lambda\pi(x)$.

If $\mathcal{C}$ is a conjugacy class of G, define

$$P(K)_{\mathcal{C}} = \left\{ \mathfrak{p} \in P(K) : \left(\frac{L/K}{\mathfrak{p}}\right) = \mathcal{C} \right\},$$

$$\pi^{\mathcal{C}}(x) = |\{\mathfrak{p} \in P(K)_{\mathcal{C}} : N\mathfrak{p} \leq x\}|.$$

We are now ready to state the Chebotarev density theorem.

Theorem 3.3.1 (Chebotarev density theorem)**.** *Let K be a global field and let L/K be a finite Galois extension with Galois group G. If $\mathcal{C}$ is a conjugacy class of G, then $P(K)_{\mathcal{C}}$ has natural density $|\mathcal{C}|/|G|$ in $P(K)$, i.e. $\pi^{\mathcal{C}}(x) \sim \dfrac{|\mathcal{C}|}{|G|}\pi(x)$.*

Our main application of the Chebotarev density theorem is to the factorization of a polynomial modulo a prime ideal $\mathfrak{p}$. In particular we are interested in the case when $\left(\dfrac{L/K}{\mathfrak{p}}\right) = \{1\}$, which implies that $\mathcal{O}_L/\mathfrak{P} = \mathcal{O}_K/\mathfrak{p}$ for any $\mathfrak{P}$ lying above $\mathfrak{p}$.

Lemma 3.3.2. *Let K be a global field with ring of integers $\mathcal{O}_K$. Let $f(x) \in \mathcal{O}_K[x]$ be a separable polynomial with splitting field L. If $\Delta(f) \notin \mathfrak{p}$ and $\left(\dfrac{L/K}{\mathfrak{p}}\right) = \{1\}$, then $f(x)$ mod $\mathfrak{p}$ is a product of distinct linear factors.*

Proof. Let $f(x)$, L, and $\mathfrak{p}$ be as in the statement of the lemma. It is enough to show that if $g(x)$ is an irreducible factor of $f(x)$, then $g(x)$ mod $\mathfrak{p}$ is a product of distinct linear factors. Let α be a root of $g(x)$ and put $F = K[\alpha]$, with ring of integers $\mathcal{O}_F$.

Consider a maximal ideal $\mathfrak{q}$ of $\mathcal{O}_F$ lying above $\mathfrak{p}$. Then there is a maximal ideal

$\mathfrak{P}$ of $\mathcal{O}_L$ lying above $\mathfrak{q}$, and thus also $\mathfrak{p}$. Since $\left(\dfrac{L/K}{\mathfrak{p}}\right) = \{1\}$, we have

$$\mathcal{O}_K/\mathfrak{p} \hookrightarrow \mathcal{O}_F/\mathfrak{q} \hookrightarrow \mathcal{O}_L/\mathfrak{P} = \mathcal{O}_K/\mathfrak{p},$$

so $\mathcal{O}_F/\mathfrak{q} = \mathcal{O}_K/\mathfrak{p}$ for every $\mathfrak{q}$ lying above $\mathfrak{p}$. We also have $\Delta(g) \notin \mathfrak{p}$ because $\Delta(f) \notin \mathfrak{p}$, so by Lemma 3.2.3, $g(x)$ mod $\mathfrak{p}$ is a product of distinct linear factors. □

Chapter 4

Codimension Bounds

4.1 Subspaces

Let $\mathfrak{g}(F)$ be a Chevalley algebra with root system Φ and let F/K be a finite, separable field extension. In this section we investigate how large K-subspaces U and V of $\mathfrak{g}(F)$ can be and still satisfy $[U,V]_K \neq \mathfrak{g}(F)$. In particular, we find upper bounds on the sum of the dimensons of U and V which satisfy $Fe_\alpha \not\subseteq [U,V]_K$ for some $\alpha \in L$. These bounds will depend on α and char F.

The methods used to compute the upper bounds do not use the full Chevalley algebra but rather a subalgebra with certain properties. We prove results in the setting of an abstract subalgebra with the desired properties and then apply them to the Chevalley algebra setting.

Lemma 4.1.1. *Let F/K be a finite, separable field extension and let L be a Lie algebra over F. Fix $x \in L$, and suppose $J = \bigoplus_{i=1}^{n} J_i \leq_K L$ such that*

1. $\dim_F(J_i) = 2$ *for all i.*

2. $[J_i, J_j]_F = 0$ *for* $i \neq j$.

3. $[J_i, J_i] = Fx$ *for all* i.

If $U, V \leq_K L$ *and* $Fx \not\subseteq [U, V]_K$, *then there exist* $W_U, W_V \leq_K J$ *such that*

1. $W_U \cap U = 0$ *and* $W_V \cap V = 0$.

2. $\dim_K(W_U) + \dim_K(W_V) = \dim_K(J) = [F : K] \dim_F(J)$.

Proof. Let $\mathrm{Tr} : F \to K$ be the nondegenerate trace on F and write $[U, V] \cap Fx = Tx$, $T \leq F$. Since T is proper, there exists some nonzero $a \in F$ such that $\mathrm{Tr}(at) = 0$ for all $t \in T$. Replacing U by aU, we may assume $\mathrm{Tr}(t) = 0$ for all $t \in T$.

Let $B = \{b_1, \cdots, b_m\}$ be a K-basis of F, and let $B' = \{b'_1, \cdots, b'_m\}$ be the dual basis of B with respect to the trace, so that

$$\mathrm{Tr}(b_i b'_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Recalling the properties of J from the lemma statement, let J_i be generated over F by $\{y_{2i-1}, y_{2i}\}$ with $[y_{2i-1}, y_{2i}] = x$, and set

$$X = \{by_{2i-1} : b \in B, 1 \leq i \leq n\} \cup \{b' y_{2i} : b' \in B', 1 \leq i \leq n\}.$$

Define an involution on X by $\overline{b_i y_{2j-1}} = b'_i y_{2j}$. Then for all $w_1, w_2 \in X$,

$$[w_1, \overline{w_2}] = tx \text{ with } \mathrm{Tr}(t) = \begin{cases} \pm 1 & \text{if } w_1 = w_2 \\ 0 & \text{if } w_1 \neq w_2. \end{cases} \tag{4.1.1}$$

Let $X_U \subseteq X$ be maximal with respect to the property $\langle X_U \rangle_K \cap U = 0$, and set

$$X_V = \{\overline{w} : w \in X \setminus X_U\}.$$

Put $W_U = \langle X_U \rangle_K$ and $W_V = \langle X_V \rangle_K$. Since X_U and X_V are each K-linearly independent and $|X_U| + |X_V| = |X| = \dim_K(J)$, all that remains to show is $W_V \cap V = 0$.

Assume not. Then there is some nonzero

$$v = \sum_{w \in X_V} s_w w \in V,$$

where each $s_w \in K$. We now construct $u \in U$ such that the coefficient of $[u, v] \in Fx$ has nonzero trace, a contradiction.

Some coefficient s_{w_0} is nonzero, and we may assume $s_{w_0} = 1$. Then $\overline{w_0} \notin X_U$, so by the maximality of X_U,

$$u = \overline{w_0} + z \in U$$

for some $z \in \langle X_U \rangle_K$. By the definition of X_V and (4.1.1), $[z, v], [\overline{w_0}, v - w_0] \in Fx$ each have coefficients with trace 0, so for some $t \in F$ with $\mathrm{Tr}(t) = 0$,

$$[u, v] = [\overline{w_0}, w_0] + tx = (\pm bb' + t)x \in [U, V]$$

for some $b \in B$. But $\mathrm{Tr}(bb' + t) = \pm 1 \neq 0$, giving the desired contradiction. □

Proposition 4.1.2. *Let F/K be a finite, separable field extension and let L be a Lie algebra over F. Suppose I is an ideal of F and $[I, I]_F = Z(I) = Fx$ for some $x \in L$.*

Then $I = Fx \oplus J$, and J satisfies the assumptions of Lemma 4.1.1*. If $[L, Fx]_F = Fx$ and $U, V \leq_K L$ such that $Fx \not\subseteq [U, V]_K$, then there exist $W_U, W_V \leq L$ such that*

1. $W_U \cap U = 0$ *and* $W_V \cap V = 0$.

2. $\dim_K(W_U) + \dim_K(W_V) = [F : K](\dim_F(J) + 2)$.

3. $\dim_K(W_U \cap J) + \dim_K(W_V \cap J) = \dim_K(J) = [F : K] \dim_F(J)$.

Proof. Write I as $I = Z(I) \oplus J$ and let $\{y_1, \cdots, y_m\}$ be an F-basis for J. We show $J = J_1 \oplus \cdots \oplus J_{m/2}$ with each J_i 2 dimensional, $[J_i, J_j] = 0$ if $i \neq j$, and $[J_i, J_i] = Fx$.

By assumption, $[I, I]_F = Fx$ for some $x \in Z(I)$, so since y_1 is not central in I, $[y_1, y_i] = ax$ for some $i > 1, a \neq 0$; after reordering and rescaling we may assume $[y_1, y_2] = x$. If $[y_1, y_j] = a_j x$ and $[y_2, y_j] = b_j x$ for $j > 2$, replace y_j by $y_j - a_j y_2 + b_j y_1$. Then for $j > 2$, $[y_1, y_j] = [y_2, y_j] = 0$. After repeating this process inductively on $\{y_3, \cdots, y_m\}$, the subspaces $J_i = \langle y_{2i-1}, y_{2i} \rangle_F$ are seen to satisfy the desired conditions.

Set $n = 2m$. We proceed as in the beginning of the proof of Lemma 4.1.1, constructing the same basis and dual basis of F/K and the same set X with an involution. In addition, we may assume $[U, V] \cap Fx = Tx$ with $\mathrm{Tr}(t) = 0$ for all $t \in T$, so equation (4.1.1) is still true. We now deviate from that proof. For the remainder of this proof, all subspaces will be considered as K-subspaces, and $\langle W \rangle$ will mean $\langle W \rangle_K$.

Let $B_U \subseteq \{bx : b \in B\}$, respectively $B_V \subseteq \{b'x : b' \in B'\}$, be maximal with respect to the property $\langle B_U \rangle \cap U = 0$, respectively $\langle B_V \rangle \cap V = 0$. Let $X_U \subseteq X$ be maximal with respect to the property $\langle B_U \cup X_U \rangle \cap U = 0$, and let $X_V = \{\overline{w} : w \in X \setminus X_U\}$. Note that $X \subseteq I$ and $B_U, B_V \subseteq Z(I)$. In what follows, all linear combinations are $K-$linear combinations, and all instances of s are in K.

We first show $\langle B_V \cup X_V \rangle \cap V = 0$. If not, then for some $ax \in \langle B_V \rangle$, there is a nonzero element

$$v = ax + \sum_{w \in X_V} s_w w \in V.$$

Since $ax \notin V$, $s_{w_0} \neq 0$ for some $w_0 \in X_V$; we may assume $s_{w_0} = 1$. Since $\overline{w_0} \notin X_U$, by the maximality of X_U there is a nonzero

$$u = \overline{w_0} + z \in U$$

for some $z \in \langle B_U \cup X_U \rangle$. Using (4.1.1) and the fact that $[B_U, X] = [B_V, X] = [B_U, B_V] = 0$, we see that

$$[u, v] = [z, v] + [\overline{w_0}, v - \overline{w_0}] + [\overline{w_0}, w_0] = (t \pm bb')x \in [U, V]$$

for some $b \in B$, with $\mathrm{Tr}(t) = 0$. Then $\mathrm{Tr}(t \pm bb') = \pm 1$, giving a contradiction.

Since $[L, Fx]_F = Fx$, there exists $h \in L$ such that $[h, x] = x$. Define

$$H_U = \{bh : b'x \notin B_V\}$$

$$H_V = \{b'h : bx \notin B_U\},$$

where b' is the element of B' corresponding to $b \in B$. Now set $W_U = \langle B_U \cup X_U \cup H_U \rangle$. We claim $W_U \cap U = 0$. If not, then there is a nonzero element

$$u = z_1 + \sum_{bh \in H_U} s_b bh \in U$$

for some $z_1 \in \langle B_U \cup X_U \rangle$. Since $\langle B_U \cup X_U \rangle \cap U = 0$, $s_c \neq 0$ for some $c \in B$; we may assume $s_c = 1$. By definition of H_U, $c'x \notin B_V$, so $\langle B_V, c'x \rangle \cap V \neq 0$ by the maximality of B_V. Hence we can find some nonzero element

$$v = z_2 + c'x \in V$$

with $z_2 \in \langle B_V \rangle$.

Using (4.1.1), the choice of h, and the properties of the dual basis, we find that

$$[u, v] = [u, z_2] + [u - ch, c'x] + [ch, c'x] = (t + cc')x \in [U, V]$$

with $\mathrm{Tr}(t) = 0$. Then $\mathrm{Tr}(t + cc') = 1 \neq 0$, yielding a contradiction.

Set $W_V = \langle B_V \cup X_V \cup H_V \rangle$. Then the same argument as above with U and V switched shows that $W_V \cap V = 0$. To compute the sum of the dimensions of $W_U = \langle B_U \cup X_U \cup H_U \rangle$ and $W_V = \langle B_V \cup X_V \cup H_V \rangle$, observe that the sets $B_U \cup X_U \cup H_U$ and $B_V \cup X_V \cup H_V$ are each linearly independent over K, so we just need to calculate their cardinalities. By construction,

$$|X_U| + |X_V| = |X|$$

$$|B_U| + |H_V| = |B|$$

$$|B_V| + |H_U| = |B|.$$

Since $|X| = \dim_K(J) = [F : K] \dim_F(J)$ and $|B| = [F : K]$,

$$\dim_K(W_U) + \dim(W_V) = [F : K](\dim_F(J) + 2).$$

Also, $W_U \cap J = \langle X_U \rangle$ and $W_V \cap J = \langle X_V \rangle$, so

$$\dim_K(W_U \cap J) + \dim_K(W_V \cap J) = [F : K] \dim(J). \qquad \square$$

We now use this proposition to find results for subspaces of Chevalley algebras. Let Φ be an irreducible root system of rank $l \geq 2$, F/K a finite, separable field extension, and $\mathfrak{g}(F)$ the corresponding Chevalley algebra with Chevalley basis $\{e_\alpha : \alpha \in \Phi\} \cup \{h_1, \cdots, h_l\}$. Set

$$E = \bigoplus_{\alpha \in \Phi} F e_\alpha, \qquad H = \bigoplus_{i=1}^{l} F h_i.$$

We seek to apply Proposition 4.1.2 to a Chevalley algebra $\mathfrak{g}(F)$ in the case $x = e_\alpha$ for some $\alpha \in \Phi$. To accomplish this, we first construct a subspace J of E which satisfies the conditions of Lemma 4.1.1; in particular, $[J, J]_F = Fe_\alpha$. This fails in the case when Φ is of type C_l, $l \geq 2$, char $F = 2$, and α is a long root, because then $\mathfrak{g}(F)$ is not perfect and $e_\alpha \notin [\mathfrak{g}(F), \mathfrak{g}(F)]_F$, but otherwise such a J can be constructed.

One then hopes that $I = Fe_\alpha \oplus J$ and $L = H \oplus I$ satisfy the assumptions of Proposition 4.1.2. Fortunately this is the case except for certain instances when Φ is of type $B_l, l \geq 2$, or G_2. The details are provided in the proof of the following proposition and in Tables 4.1 and 4.2.

Proposition 4.1.3. *Let $\mathfrak{g}(F)$ be a Chevalley algebra with irreducible root system Φ and fix $\alpha \in \Phi$. Then there exists $J \leq_F E$, with dimension given in Tables* 4.1 *and* 4.2 *and depending on the length of α and char F, such that if $I = Fe_\alpha \oplus J$ and $L = H \oplus I$, then either*

1. *I is an ideal of L, and I, L, and $x = e_\alpha$ satisfy the assumptions of Proposition* 4.1.2*, or*

2. I is not an ideal of L, and J and $x = e_\alpha$ satisfy the assumptions of Lemma 4.1.1*.*

Proof. Each J we construct is of the form $J = \bigoplus_{\gamma \in \Phi_J} Fe_\gamma$ for some $\Phi_J \subseteq \Phi$, so we will define J by defining the appropriate set of roots Φ_J. Because the Weyl group acts transitively on the set of roots of a given length, for each length it suffices to consider a specific root α of that length. The choices of Φ_J and α are given in Table 4.3 for every root system except E_l, for which the corresponding Φ_J is slightly more complicated.

In each case Φ_J was found by first considering the set $\{\gamma, \alpha - \gamma : \gamma, \alpha - \gamma \in \Phi\}$ and then removing roots that resulted in $[J, J] \neq Fe_\alpha$. In most cases, this was enough for $I = Fe_\alpha \oplus J$ to be an ideal of $L = H \oplus I$ with $Z(I) = [I, I] = Fe_\alpha$; when this was not the case, I is still an ideal in small characteristic.

That J and I satisfy the desired conditions is a straightforward computation. Except in the case of $\Phi = C_l$, $l \geq 2$, and α long, one can always find $\beta \in \Phi$ with $[h_\beta, e_\alpha] = \pm e_\alpha$, so that $[L, Fe_\alpha] = Fe_\alpha$ for every characteristic. We are excluding the aforementioned case already when char $F = 2$, so we always have $[L, Fe_\alpha] = Fe_\alpha$. We will point out why I sometimes fails to be an ideal and explain some of the differences in Φ_J when the characteristic changes.

When Φ is type $B_l, l \geq 2$, and $\alpha = \epsilon_1$, we have $e_{\epsilon_k}, e_{\epsilon_1 - \epsilon_k} \in J$ for $2 \leq k \leq l$. Then $[e_\alpha, e_{\epsilon_k}] = \pm 2 e_{\epsilon_1 + \epsilon_k}$, so I is an ideal when char $F = 2$ but not an ideal otherwise. Also, $[e_{\epsilon_1}, e_{-\epsilon_2}] = \pm 2 e_{\epsilon_1 - \epsilon_2}$, so when $\alpha = \epsilon_1 - \epsilon_2$, ϵ_1 and $-\epsilon_2$ are only added to Φ_J

when char $F \neq 2$.

When Φ is type $C_l, l \geq 3$, and char $F = 2$, then $[\mathfrak{g}(F), \mathfrak{g}(F)] \cap Fe_\alpha = 0$ for any long root α, so no appropriate J can be constructed in this case.

When Φ is type E_l, $l = 6, 7, 8$, let $\alpha = \sum_{i=1}^{8} \epsilon_i$, and set $\Phi'_J = \{\epsilon_i + \epsilon_j : 1 \leq i, j \leq 8\} \cap \Phi$. Then define

$$\Phi_J = \Phi'_J \cup \{\alpha - \beta : \beta \in \Phi'_J\}.$$

We have $\epsilon_3+\epsilon_4 \in \Phi$ for each l, and $[h_{\epsilon_3+\epsilon_4}, e_\alpha] = \pm e_\alpha$. The values of $|\Phi_J|$ are computed by examining the root systems as described in the chapter on Lie algebras.

When Φ is type F_4 and $\alpha = \epsilon_1$, Φ_J can be enlarged when char $F = 2$ because for $\beta, \gamma \in \Phi_J$ with $\beta + \gamma \in \Phi \setminus \{\alpha\}$, $[e_\beta, e_\gamma] = \pm 2e_{\beta+\gamma}$. The reverse is true when α is long; now Φ_J is larger when char $F \neq 2$ because there are pairs $\gamma, \beta \in \Phi$ with $[e_\gamma, e_\beta] = \pm 2e_\alpha$.

Finally we comment on the case when Φ is type G_2. There are three pairs of roots $\{\beta, \gamma\}$ in Φ which sum to α_S. Examing the coefficients of $[e_\gamma, e_\beta]$, $[e_{\alpha_S}, e_\beta]$, and $[e_{\alpha_S}, e_\gamma]$ leads one to the choices of Φ_J given in Table 4.3, as well as the conclusion that I is an ideal only when F has characteristic 2 or 3. In contrast, when $\alpha = \alpha_L$, characteristic 3 is the only exception to the general rule, since $[e_{-\alpha_S}, e_{\alpha_S+\alpha_L}] = \pm 3e_{\alpha_L}$. $\square$

Φ	char F	$I \trianglelefteq L$?	$\dim_F J$
$A_l, l \geq 2$	any	yes	$2(l-1)$
$D_l, l \geq 4$	any	yes	$4(l-1)$
E_6	any	yes	20
E_7	any	yes	32
E_8	any	yes	56

Table 4.1: This table gives the dimensions of the subspaces J given in Proposition 4.1.3 and indicates when I is an ideal when Φ is simply laced.

We can now state the result we will need when computing lower bounds for normal and non-normal residual finiteness growth.

Corollary 4.1.4. *Let F/K be a finite, separable field extension. Fix $\alpha \in \Phi$ and assume that α is a short root if Φ is of type $C_l, l \geq 2$. Suppose $U, V \leq_K \mathfrak{g}(F)$ satisfy $Fe_\alpha \not\subseteq [U,V]_K$. Then*

$$\operatorname{codim}(U) + \operatorname{codim}(V) \geq 2[F:K].$$

Proof. By Proposition 4.1.3 and Lemma 4.1.1, there exist $W_U, W_V \leq_K \mathfrak{g}(F)$ such that $W_U \cap U = W_V \cap V = 0$ and $\dim_K(W_U) + \dim_K(W_V) \geq 2[F:K]$. This immediately gives the desired inequality. □

Φ	length of α	char F	$I \trianglelefteq L$?	$\dim_F J$
B_2	short	2	yes	2
		$\neq 2$	no	2
	long	$\neq 2$	yes	2
$B_l, l \geq 3$	short	2	yes	2
		$\neq 2$	no	2
	long	2	yes	$4(l-2)$
		$\neq 2$	yes	$4(l-2)+2$
$C_l, l \geq 3$	short	any	yes	$4(l-2)$
	long	$\neq 2$	yes	$2(l-1)$
F_4	short	2	yes	8
		$\neq 2$	yes	2
	long	2	yes	8
		$\neq 2$	yes	14
G_2	short	2	yes	4
		3	yes	2
		$\neq 2, 3$	no	2
	long	3	yes	2
		$\neq 3$	yes	4

Table 4.2: This table gives the dimensions of the subspaces J given in Proposition 4.1.3 and indicates when I is an ideal in the case Φ is not simply laced.

Φ	α	char F	Φ_J
$A_l, l \geq 2$	$\epsilon_1 - \epsilon_2$	any	$\{\epsilon_1 - \epsilon_k, \epsilon_k - \epsilon_2 : 3 \leq k \leq l+1\}$
B_2	ϵ_1	any	$\{\epsilon_2, \epsilon_1 - \epsilon_2\}$
	$\epsilon_1 - \epsilon_2$	$\neq 2$	$\{\epsilon_1, -\epsilon_2\}$
$B_l, l \geq 3$	ϵ_1	any	$\{\epsilon_1 - \epsilon_k, \epsilon_k : 2 \leq k \leq l\}$
	$\epsilon_1 - \epsilon_2$	2	$\{\epsilon_1 \pm \epsilon_k, \pm\epsilon_k - \epsilon_2 : 3 \leq k \leq l\}$
		$\neq 2$	$\{\epsilon_1 \pm \epsilon_k, \pm\epsilon_k - \epsilon_2 : 3 \leq k \leq l\} \cup \{\epsilon_1, -\epsilon_2\}$
$C_l, l \geq 3$	$\epsilon_1 - \epsilon_2$	any	$\{\epsilon_1 \pm \epsilon_k, \pm\epsilon_k - \epsilon_2 : 3 \leq k \leq l\}$
	$2\epsilon_1$	$\neq 2$	$\{\epsilon_1 \pm \epsilon_k : 2 \leq k \leq l\}$
$D_l, l \geq 4$	$\epsilon_1 - \epsilon_2$	any	$\{\epsilon_1 \pm \epsilon_k, \pm\epsilon_k - \epsilon_2 : 3 \leq k \leq l\}$
F_4	ϵ_1	2	$\{\frac{1}{2}(\epsilon_1 \pm \epsilon_2 \pm \epsilon_3 \pm \epsilon_4)\}$
		$\neq 2$	$\{\frac{1}{2}(\epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4), \frac{1}{2}(\epsilon_1 - \epsilon_2 - \epsilon_3 - \epsilon_4)\}$
	$\epsilon_1 - \epsilon_2$	2	$\{\epsilon_1 \pm \epsilon_k, \pm\epsilon_k - \epsilon_2 : k = 3, 4\}$
		$\neq 2$	above and $\{\epsilon_1, -\epsilon_2, \frac{1}{2}(\epsilon_1 - \epsilon_2 \pm \epsilon_3 \pm \epsilon_4\}$
G_2	α_S	2	$\{-\alpha_S, \alpha_S + \alpha_L, -2\alpha_S - \alpha_L, 3\alpha_S + \alpha_L\}$
		$\neq 2$	$\{-\alpha_S - \alpha_L, 2\alpha_S + \alpha_L\}$
	α_L	3	$\{3\alpha_S + 2\alpha_L, -3\alpha_S - \alpha_L\}$
		$\neq 3$	$\{3\alpha_S + 2\alpha_L, -3\alpha_S - \alpha_L, \alpha_S + \alpha_L, -\alpha_S\}$

Table 4.3: This table gives the set of defining roots Φ_J of $J = \bigoplus_{\gamma\in\Phi_J} Fe_\gamma$ for each root system and a root α of each length. The details for E_l are given in the proof of Proposition 4.1.3.

4.2 Group action

We now restrict ourselves to the case that $F = \mathbb{F}_q$ is a finite field and consider the action of the Chevalley group $G(\mathbb{F}_q)$ on $\mathbb{F}_p$-subspaces of $\mathfrak{g}(\mathbb{F}_q)$. The following result was initially found while working on lower bounds of non-normal residual finiteness growth but ultimately was not needed there.

We note that certain choices of α are excluded in the proposition. The proof uses the sets Φ_J constructed in the proof of Proposition 4.1.2 and given in Table 4.3. For this proposition to be valid we require $\alpha + \beta \notin \Phi$ for all $\beta \in \Phi_J$, so we exclude short roots in $B_l, l \geq 3$, and G_2. In addition, the proof requires the existence of $\delta \in \Phi$ such that $\langle \alpha, \delta \rangle = 1$. This fails only when α is a long root in $C_l, l \geq 2$.

Proposition 4.2.1. *Let $q = p^m$ for some prime p. Let $K \leq G(\mathbb{F}_q)$, $V \leq_{\mathbb{F}_p} \mathfrak{g}(\mathbb{F}_q)$ such that K acts on V. Let $\alpha \in \Phi$ and assume that α is a long root if Φ is type $B_l, l \geq 3$, or G_2, and that α is a short root if Φ is type $C_l, l \geq 2$. If $\mathbb{F}_p e_\alpha \not\subseteq V$, then*

$$\log_p([G(\mathbb{F}_q) : K]) + \operatorname{codim}(V) \geq m(\dim_{\mathbb{F}_q}(J) + 2),$$

where $\dim_{\mathbb{F}_q}(J)$ is given in Table 4.2, with $F = \mathbb{F}_q$.

Proof. Let $\operatorname{Tr} : \mathbb{F}_q \to \mathbb{F}_p$ be the trace, and write $V \cap \mathbb{F}_q e_\alpha = T e_\alpha$, $T \leq_{\mathbb{F}_p} \mathbb{F}_q$. As in the proof of Lemma 4.1.1, we may assume $\operatorname{Tr}(T) = 0$.

Let $\{b_1, \cdots, b_n\}$ be an $\mathbb{F}_p$-basis for T, and extend it to a basis $B = \{b_1, \cdots, b_m\}$

of $\mathbb{F}_q$. Let $B' = \{b'_1, \cdots b'_m\}$ be the dual basis with respect to the trace, so that

$$\mathrm{Tr}(b_i b'_j) = \begin{cases} 1 & \text{if } i = j \\ \\ 0 & \text{if } i \neq j. \end{cases}$$

For each i, let $\overline{b_i} = b'_i$ and $\overline{b'_i} = b_i$.

Let Φ_J be as given in Table 4.3, depending on Φ, p, and the length of α. Note that $|\Phi_J| = \dim_{\mathbb{F}_q}(J)$ is even, so we can write $\Phi_J = \Phi_{J,1} \sqcup \Phi_{J,2}$, where $\Phi_{J,2} = \{\alpha - \beta : \beta \in \Phi_{J,1}\}$. Let

$$E_V = \{be_\beta : b \in B, \beta \in \Phi_{J,1}\} \cup \{b' e_\beta : b' \in B', \beta \in \Phi_{J,2}\}$$

$$E_K = \{x_\beta(b) : b \in B, \beta \in \Phi_{J,1}\} \cup \{x_\beta(b') : b' \in B', \beta \in \Phi_{J,2}\}.$$

Define a map from E_K to E_V by $\overline{x_\beta(b)} = \overline{b} e_{\alpha-\beta}$. This map will be useful when we bound the codimension of V. Before we get to V, we investigate the index of K in $G(\mathbb{F}_q)$.

Let $\delta \in \Phi$ such that $\langle \alpha, \delta \rangle = 1$; this is possible because of our restrictions on α. Set $N = \langle X_\alpha, \{X_\beta\}_{\beta \in \Phi_J} \rangle$ and $H_\delta = \{h_\delta(s) : s \in \mathbb{F}_q^*\}$. Note that X_α commutes with N since $\alpha + \beta \notin \Phi$ for all $\beta \in \Phi_J$, and $[N, N] = X_\alpha$ since $[J, J] = \mathbb{F}_q e_\alpha$. Define $G_0 = \langle H_\delta, N \rangle$ and $K_0 = K \cap G_0$, and observe that $N \trianglelefteq G_0$. Since $[G(\mathbb{F}_q) : K] \geq [G_0 : K_0]$, it suffices to find a lower bound for $[G_0 : K_0]$.

Let $\varphi : G_0 \to G_0/X_\alpha$ be the natural projection map. Let $B_K \subseteq \{x_\alpha(b) : b \in B\}$ be maximal with respect to $\langle B_K \rangle \cap K_0 = 1$, and let $X_K \subseteq E_K$ be maximal with respect to the property $\langle \varphi(X_K) \rangle \cap \varphi(K_0) = 1$.

We first compute $|K_0 \cap N|$. The kernel of φ restricted to $K_0 \cap N$ is $K_0 \cap X_\alpha$, so

$$|K_0 \cap N| = |\varphi(K_0 \cap N)||K_0 \cap X_\alpha| \leq |\varphi(K_0) \cap \varphi(N)||K_0 \cap X_\alpha|.$$

The groups X_α and $\varphi(N)$ are elementary abelian p-groups, so we can view them as vector spaces over $\mathbb{F}_p$. With this perspective and the fact that $\langle B_K \rangle \cap K_0 = 1$, we see that

$$\dim(K_0 \cap X_\alpha) \leq m - \dim(\langle B_K \rangle) = m - |B_K|.$$

Similarly, since $\langle \varphi(X_K) \rangle \cap \varphi(K_0) = 1$,

$$\dim(\varphi(K_0) \cap \varphi(N)) \leq \dim(\varphi(N)) - \dim(\langle \varphi(X_K) \rangle) = m|\Phi_J| - |X_K|.$$

Thus

$$|K_0 \cap N| \leq p^{m|\Phi_J| + m - |X_K| - |B_K|}.$$

Of course, what we really want is $|K_0|$. Let $\psi : G_0 \to H_\delta$ be reduction mod N. Then $\psi(K_0) \cong K_0/(K_0 \cap N)$, so $|K_0| = |\psi(K_0)||K_0 \cap N|$. Thus it remains to find $|\psi(K_0)|$.

Since $[J, e_\alpha] = 0$, N acts trivially on e_α, so the action of K_0 on e_α descends to the action of $\psi(K_0) \leq H$. Suppose $h_\delta(s) \in \psi(K_0)$, where $s = \sum_{i=1}^m s_i b_i'$. If $s_j \neq 0$ for some $1 \leq j \leq n$, then since $b_j e_\alpha \in V$ and $\langle \alpha, \delta \rangle = 1$,

$$h_\delta(s) \cdot b_j e_\alpha = s b_j e_\alpha \in V$$

with $\mathrm{Tr}(s b_j) = \mathrm{Tr}(s_j b_j' b_j) = s_j \neq 0$, a contradiction. Therefore

$$\psi(K_0) \subseteq \{h_\delta(s) : s \in \langle b_{n+1}', \cdots, b_m' \rangle \setminus \{0\}\},$$

so $|\psi(K_0)| \leq p^{m-n} - 1$. Thus

$$[G_0 : K_0] \geq \frac{(p^m - 1)p^{m|\Phi_J|+m}}{(p^{m-n} - 1)p^{m|\Phi_J|+m-|X_K|-|B_K|}} \geq p^{n+|X_K|+|B_K|}.$$

We now turn our attention to V. Set $B_V = \{b_{n+1}e_\alpha, \cdots, b_m e_\alpha\}$ and $X_V = \{\overline{w} : w \in E_K \setminus X_K\} \subseteq E_V$, and recall that $\mathbb{F}_q e_\alpha = T e_\alpha \oplus B_V$.

Because $[e_\alpha, J] = 0$, $x_\alpha(s) \cdot \overline{w} = \overline{w}$ for all $w \in E_K$, $s \in \mathbb{F}_q$. Also, by the properties of Φ_J, for all $w_1, w_2 \in E_K$ we have

$$w_1 \cdot \overline{w_2} = \overline{w_2} + te_\alpha \text{ with } \operatorname{Tr}(t) = \begin{cases} \pm 1 & \text{if } w_1 = w_2 \\ \\ 0 & \text{if } w_1 \neq w_2. \end{cases} \tag{4.2.1}$$

We begin by showing $\langle B_V \cup X_V \rangle \cap V = 0$. If not, then X_V is nonempty and

$$v = z + \sum_{i=1}^{k} s_i \overline{w_i} \in V$$

for some $z \in \langle B_V \rangle, w_1, \cdots, w_k \in E_K \setminus X_K$. We may assume $s_1 \neq 0$. Since $w_1 \notin X_K$, the maximality of X_K implies $g = w_1 w_0 \in K$ for some $w_0 \in \langle X_K, X_\alpha \rangle$. By (4.2.1) we have $g \cdot v - v = te_\alpha \in V$ for some $t \in \mathbb{F}_q$ with $\operatorname{Tr}(t) = \pm s_1 \neq 0$, a contradiction.

Now define $H_V = \{b' h_\delta : x_\alpha(b) \notin B_K\}$ and suppose $\langle B_V \cup X_V \cup H_V \rangle \cap V \neq 0$, so that

$$v = z + \sum_{b' h_\delta \in H_V} s_{b'} b' h_\delta \in V$$

for some $z \in \langle B_V \cup X_V \rangle$, some $s_{c'} \neq 0$. By the maximality of B_K, $g = x_\alpha(c) w_0 \in K_0$ for some $w_0 \in \langle B_K \rangle$. Then $g \cdot v - v = te_\alpha \in V$ with $\operatorname{Tr}(t) = \pm s_{c'} \neq 0$, a contradiction.

Recalling that

$$\log_p([G_0 : K_0]) \geq n + |X_K| + |B_K|,$$

we now have

$$\begin{aligned}\log_p([G:K]) + \operatorname{codim}(V) &\geq n + |X_K| + |B_K| + |B_V| + |X_V| + |H_V| \\ &= n + |X_K| + |B_K| + m - n + |X_V| + m - |B_K| \\ &= m(|\Phi_J| + 2) = m(\dim_{\mathbb{F}_q}(J) + 2),\end{aligned}$$

since $|X_K| + |X_V| = |E_K| = m|\Phi_J|$. □

Chapter 5

Residual Finiteness Growth Background

5.1 Residually Finite Groups

We recall the definition of a residually finite group.

Definition. A group Γ is residually finite if for every $1 \neq g \in \Gamma$, there is a finite group Q and a group homomorphism $\varphi : \Gamma \to Q$ such that $\varphi(g) \neq 1$.

Sometimes it is useful to work with an equivalent definition, two of which are given below.

Proposition 5.1.1. *Let Γ be a group. The following are equivalent.*

(1) Γ is residually finite.

(2) The intersection of all finite index normal subgroups of Γ is trivial.

(3) The intersection of all finite index subgroups of Γ is trivial.

Proof. We first show (1) and (2) are equivalent. Let $1 \neq g \in \Gamma$. If $\varphi : \Gamma \to Q$ is a homomorphism to a finite group with $\varphi(g) \neq 1$, then $g \in \ker \varphi$, a normal subgroup of Γ of finite index, so (1) implies (2). If (2) is true and $1 \neq g \in \Gamma$, then $g \notin N$ for some $N \trianglelefteq \Gamma$ of finite index. Then g is nontrivial in the finite group G/N, giving (1).

We clearly have that (2) implies (3) since the intersection is being taken over a larger set. The other direction follows from the fact that every finite index subgroup H contains a finite index normal subgroup of G, namely $\bigcap\limits_{g \in G} gHg^{-1}$. □

We will freely switch among these equivalent definitions whenever appropriate.

Example.

- Any finite group is trivially residually finite.

- Finitely generated free groups are residually finite.

- Subgroups and direct products of residually finite groups are residually finite.

- The Baumslag-Solitar group $B(m, n) = \langle a, b : ba^m b^{-1} = a^n \rangle$ is not residually finite if $|m| \neq |n|$ and $|n|, |m| \neq 1$.

A large class of examples of residually finite groups are linear groups.

Definition. We call a group Γ a **linear group**, or say Γ is **linear**, if $\Gamma \leq \mathrm{GL}_d(K)$ for some positive integer d and some field K.

It is a classical result by Mal'cev that all finitely generated linear groups are residually finite. As the goal of this thesis is to provide a more detailed description of this fact, we give a proof of Mal'cev's theorem. The proof requires the following standard number theory fact.

Lemma 5.1.2. *Let A be a finitely generated integral domain. Then the intersection of all maximal ideals of A is trivial and if A is a field, then A is finite.*

Theorem 5.1.3 (Mal'cev's Theorem)**.** *Every finitely generated linear group Γ is residually finite.*

Proof. We first prove the statement in the case $\Gamma = \mathrm{SL}_d(\mathbb{Z})$ to illustrate the idea. For a prime p, let $\Gamma(p) = \ker(\Gamma \to \mathrm{SL}_d(\mathbb{Z}/p\mathbb{Z})$. We observe that each $\Gamma(p)$ is finite index in Γ since $\mathrm{SL}_d(\mathbb{Z}/p\mathbb{Z})$ is finite. It is easy to see that $\bigcap_{p \text{ prime}} \Gamma(p)$ is trivial, so Γ is residually finite by Proposition 5.1.1.

We now address the general case. Assume $\Gamma \leq \mathrm{GL}_d(K)$ for some field K and that Γ is generated by a finite set X. The entries of the matrices in X and their inverses generate a finitely generated subdomain A of K with $\Gamma \leq \mathrm{GL}_d(A)$. For a maximal ideal $\mathfrak{m}$ of A, let $\Gamma(\mathfrak{m}) = \ker(\mathrm{GL}_d(A) \to \mathrm{GL}_d(A/\mathfrak{m}))$. Since $A/\mathfrak{m}$ is a finitely generated domain which is a field, it is finite by Lemma 5.1.2. Hence $\Gamma(\mathfrak{m})$ has finite index. Also by Lemma 5.1.2, the intersection of all $\Gamma(\mathfrak{m})$ as $\mathfrak{m}$ ranges over the maximal ideals of A is trivial. Therefore $\mathrm{GL}_d(A)$ is residually finite, so Γ is as well. □

5.2 Residual Finiteness Growth Definition

Let Γ be a residually finite, finitely generated group, generated by a finite symmetric set X, by which we mean $x \in X$ if and only if $x^{-1} \in X$. If $\gamma \in \Gamma$ is nontrivial, define

$$D_{\Gamma}^{\trianglelefteq}(\gamma) = \min\{[\Gamma : N] : N \trianglelefteq \Gamma, \gamma \notin N\},$$

$$D_{\Gamma}^{\leq}(\gamma) = \min\{[\Gamma : H] : H \leq \Gamma, \gamma \notin H\}.$$

Since Γ is residually finite, and using Proposition 5.1.1, both $D_{\Gamma}^{\trianglelefteq}(\gamma)$ and $D_{\Gamma}^{\leq}(\gamma)$ are finite integers for nontrivial $\gamma \in \Gamma$.

Remark 5.2.1. While $D_{\Gamma}^{\trianglelefteq}(\gamma)$ and $D_{\Gamma}^{\leq}(\gamma)$ are defined similarly, in practice we will use the following equivalent definition of $D_{\Gamma}^{\trianglelefteq}(\gamma)$ to take advantage of the restriction to normal subgroups.

$$D_{\Gamma}^{\trianglelefteq}(\gamma) = \min\{|Q| : \varphi : \Gamma \to Q, \varphi(\gamma) \neq 1\}.$$

That is, the strategy to compute $D_{\Gamma}^{\trianglelefteq}(\gamma)$ will be to find a small finite quotient of Γ in which the image of γ is nontrivial.

Definition. A set $X \subseteq \Gamma$ is **symmetric** if $X = X^{-1}$. Let Γ be finitely generated by a symmetric set X. If $g \in G$, we define the word length of g with respect to X to be

$$||g||_X = \min\{n : g = x_1 \cdots x_n, x_i \in X\}.$$

The normal and non-normal residual finiteness growth of Γ are determined, respectively, by the residual finiteness growth functions

$$F^{\trianglelefteq}_{\Gamma,X}(n) = \max\{D^{\trianglelefteq}_{\Gamma}(\gamma) : ||\gamma||_X \leq n, \gamma \neq 1\},$$

$$F^{\leq}_{\Gamma,X}(n) = \max\{D^{\leq}_{\Gamma}(\gamma) : ||\gamma||_X \leq n, \gamma \neq 1\}.$$

We are interested not in the exact values of these functions for a given n but in how they grow as n goes to infinity. We will compare the asymptotic growth of two functions $f, g : \mathbb{N} \to \mathbb{N}$ by writing $f \preceq g$ if there exists C such that $f(n) \leq Cg(Cn)$ for all $n \in \mathbb{N}$. If $f \preceq g$ and $g \preceq f$ we will write $f \approx g$.

The first advantage of considering asymptotic growth is that it is independent of the choice of generating set. The following lemma is Lemma 1 from [3]; we give its proof for completeness.

Lemma 5.2.2. *Let $H \leq \Gamma$ be residually finite groups finitely generated by S and X respectively. Then $F^{\trianglelefteq}_{H,S}(n) \preceq F^{\trianglelefteq}_{\Gamma,X}(n)$ and $F^{\leq}_{H,S}(n) \preceq F^{\leq}_{\Gamma,X}(n)$.*

Proof. If $K \leq \Gamma$, then $[K : K \cap H] \leq [\Gamma : K]$, and $K \cap H \trianglelefteq K$ if $K \trianglelefteq \Gamma$. It follows that $D^{\trianglelefteq}_{H}(h) \leq D^{\trianglelefteq}_{\Gamma}(h)$ and $D^{\leq}_{H}(h) \leq D^{\leq}_{\Gamma}(h)$ for all $h \in H$.

Since S and X are finite, there exists some $C > 0$ such that $||s||_X \leq C$ for all $s \in S$. Hence

$$\{h \in H : ||h||_S \leq n\} \subseteq \{g \in \Gamma : ||g||_X \leq Cn\}.$$

Therefore

$$\begin{aligned} F^{\trianglelefteq}_{H,S}(n) &= \max\{D^{\trianglelefteq}_{H}(h) : 1 \neq h \in H, ||h||_S \leq n\} \\ &\leq \max\{D^{\trianglelefteq}_{\Gamma}(h) : 1 \neq 1 \in H, ||h||_S \leq n\} \\ &\leq \max\{D^{\trianglelefteq}_{\gamma}(g) : 1 \neq g \in G, ||g||_X \leq Cn\} \\ &= F^{\trianglelefteq}_{\Gamma,X}(Cn). \end{aligned}$$

The same argument works replacing $F^{\trianglelefteq}$ by $F^{\leq}$. □

If X_1 and X_2 are two finite generating sets of Γ, then applying Lemma 5.2.2 twice with $H = \Gamma$ shows that the choice of generating set does not affect the asymptotic growth of either residual finiteness growth function. We thus drop the reference to the generating set.

Another important consequence of Lemma 5.2.2 is that when computing upper bounds for $F^{\trianglelefteq}_{\Gamma}(n)$ or $F^{\leq}_{\Gamma}(n)$, we may pass to a larger group, and when computing lower bounds we may pass to a subgroup.

5.3 Basic Results

We will need the following result when proving lower bounds; it is contained in Lemma 2.4 in [6]. In particular it will allow us to pass from a Chevalley group to its simply connected cover.

Lemma 5.3.1. *Assume Γ and Δ are finitely generated, residually finite groups. If $f : \Gamma \to \Delta$ is surjective with finite kernel, then $F^{\trianglelefteq}_{\Gamma}(n) \preceq F^{\trianglelefteq}_{\Delta}(n)$ and $F^{\leq}_{\Gamma}(n) \preceq F^{\leq}_{\Delta}(n)$.*

Proof. Since $f(\Gamma) \leq \Delta$, $F^{\trianglelefteq}_{f(\Gamma)}(n) \preceq F^{\trianglelefteq}_{\Delta}(n)$ by Lemma 5.2.2. Hence it suffices to show $F^{\trianglelefteq}_{\Gamma}(n) \preceq F^{\trianglelefteq}_{f(\Gamma)}(n)$.

Assume $\Gamma = \langle X \rangle$, $|X| \leq \infty$. Then $f(\Gamma)$ is generated by $f(X) = \{f(x) : x \in X\}$. Since the kernel of f is finite, if n sufficiently large then $f(\gamma) \neq 1$ for all $\gamma \in \Gamma$ with $||\gamma||_X = n$. Let n be large enough to ensure this and let $\gamma \in \Gamma$ with $||\gamma||_X = n$. We have $||f(\gamma)||_{f(X)} \leq n$ and $f(\gamma) \neq 1$, so there exists a normal subgroup $N \trianglelefteq f(\Gamma)$ such that $f(\gamma) \notin N$ and $[f(\Gamma) : N] \leq F^{\trianglelefteq}_{f(\Gamma),f(X)}(n)$. Hence $N' = N \ker(f) \trianglelefteq \Gamma$ satisfies

$$\gamma \notin N' \text{ and } [\Gamma : N'] \leq F^{\trianglelefteq}_{f(\Gamma),f(X)}(n),$$

so $F^{\trianglelefteq}_{\Gamma,X}(n) \leq F^{\trianglelefteq}_{f(\Gamma),f(X)}(n)$ and thus $F^{\trianglelefteq}_{\Gamma}(n) \preceq F^{\trianglelefteq}_{f(\Gamma)}(n)$.

The same argument with N replaced by an arbitrary subroup H shows that $F^{\leq}_{\Gamma}(n) \preceq F^{\leq}_{\Delta}(n)$. □

The following proposition was proved as Theorem 2.2 in [3]. We present a slightly different proof more aligned with the strategies we will use in the next section.

Proposition 5.3.2. *We have $F^{\trianglelefteq}_{\mathbb{Z}}(n) \approx \log n$. In particular, if n is sufficiently large then there is a prime p not dividing n with $p \leq 2 \log n$.*

Proof. We first show $F^{\trianglelefteq}_{\mathbb{Z}}(n) \preceq \log n$. Let $X = \{-1, 1\}$ and fix n a sufficiently large positive integer. If we set $m = \lceil \log 2n \rceil$, the smallest integer which is at least $\log 2n$, then by the prime number theorem,

$$\prod_{\substack{p \leq m \\ p \text{ prime}}} p > \frac{1}{2}e^m \geq \frac{1}{2}(2n) = n.$$

Thus there is a prime $p \leq m$ such that p does not divide n. Since $m \leq 1 + \log 2n \leq 2\log n$, we have a homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ with $\varphi(n) \neq 0$ and $|\mathbb{Z}/p\mathbb{Z}| \leq 2\log n$. Hence $F^{\trianglelefteq}_{\mathbb{Z},X}(n) \leq 2\log n$, so $F^{\trianglelefteq}_{\mathbb{Z}}(n) \preceq \log n$.

To show the lower bound is also $\log n$, let $k > 0$ be sufficiently large and set $n = \mathrm{lcm}(1, \cdots, k)$. Then the prime number theorem implies $n \leq 2e^k$, so $k \geq \log(n/2)$. Thus if m does not divide n, then $m > k \geq \log(n/2)$. Since every finite quotient of $\mathbb{Z}$ is of the form $\mathbb{Z}/m\mathbb{Z}$, this shows that $F^{\trianglelefteq}_{\mathbb{Z},X}(n) \geq \log(n/2)$, so $F^{\trianglelefteq}_{\mathbb{Z}}(n) \succeq \log n$. □

This proof illustrates the basic strategy we will apply to linear groups, and it already indicates the added difficulty in finding a lower bound in general. One needs to account for every subgroup up to a certain index. While this is easy in the case of $\mathbb{Z}$, we will need to use the congruence subgroup property to approach linear groups.

Chapter 6

Residual Finiteness of Linear Groups: Upper Bounds

In this chapter we provide upper bounds on the residual finiteness of finitely generated linear groups. The essential ideas are conveyed by treating the case of $\mathrm{SL}_d(\mathbb{Z})$. We then prove a characteristic p specific result before proving the general upper bound.

6.1 Special Linear Group Over $\mathbb{Z}$

We begin by providing proofs of the upper bounds on the normal and non-normal residual finiteness growth of $\mathrm{SL}_d(\mathbb{Z})$ that indicate the strategy for the general case. These bounds were first proved in [3] and [5], respectively.

Proposition 6.1.1. *Let* $\Gamma = \mathrm{SL}_d(\mathbb{Z})$. *Then* $F_\Gamma^{\trianglelefteq}(n) \preceq n^{d^2-1}$ *and* $F_\Gamma^{\leq}(n) \preceq n^{d-1}$.

Proof. Let X be a finite symmetric generating set for Γ, let $n > 0$ be sufficiently large, and let $A \in \Gamma$ with $||A||_X = n$. For $B \in \Gamma$, let $||B||_1$ be the maximum absolute value of an entry of B. If $||B||_1 \leq c$ for all $B \in X$, then the properties of matrix multiplication imply $||A||_1 \leq d^{n-1}c^n$.

Since A is nontrivial and $\mathrm{SL}_d(\mathbb{Z})$ has only finitely many diagonal matrices, we may assume A has a nonzero entry a off the diagonal. Then $|a| \leq (dc)^n$, so by Proposition 5.3.2, there is a prime $p \leq 2\log|a| \leq 2n\log(cd)$ such that p does not divide a. The ring homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ naturally induces a group homomorphism $\varphi^* : \mathrm{SL}_d(\mathbb{Z}) \to \mathrm{SL}_d(\mathbb{Z}/p\mathbb{Z})$, given by applying φ to each entry of a matrix. Since p does not divide a, $\varphi^*(A)$ is nontrivial, so $F^{\trianglelefteq}_{\Gamma,X}(n) \leq |\mathrm{SL}_d(\mathbb{Z}/p\mathbb{Z})|$. By Lemma 2.5.6, $|\mathrm{SL}_d(\mathbb{F}_p)| \leq Cp^{d^2-1}$ for some constant C independent of p, so

$$F^{\trianglelefteq}_{\Gamma,X}(n) \leq C(2n\log(cd))^{d^2-1} = C(2\log(cd))^{d^2-1}n^{d^2-1}.$$

Therefore $F^{\trianglelefteq}_{\Gamma}(n) \preceq n^{d^2-1}$.

Let $\overline{A}$ be the image of $\varphi^*(A)$ in $\mathrm{PSL}_d(\mathbb{F}_p) = \mathrm{SL}_d(\mathbb{F}_p)/Z(\mathrm{SL}_d(\mathbb{F}_p))$. Since $\varphi^*(A)$ has a nonzero entry off the diagonal, $\overline{A}$ is nontrivial. By Lemma 2.5.5, $\mathrm{PSL}_d(\mathbb{F}_p)$ has a proper subgroup H_0 of index at most $2p^{d-1}$. The intersection of all the conjugates of H_0 is normal in $\mathrm{PSL}_d(\mathbb{F}_p)$ and hence trivial since $\mathrm{PSL}_d(\mathbb{F})$ is simple (it is safe to assume $p > 3$). Hence $\overline{A}$ is not in some conjugate of H_0, so if H is the preimage of H_0 under the map $\Gamma \to \mathrm{PSL}_d(\mathbb{F}_p)$, then $A \notin H$ and

$$[\Gamma : H] \leq 2p^{d-1} \leq 2(2\log(cd)^{d-1})n^{d-1}.$$

Therefore $F^{\leq}_{\Gamma}(n) \preceq n^{d-1}$. □

Remark 6.1.2. In [3], Bou-Rabee proved the normal residual finiteness upper bound of $\mathrm{SL}_d(\mathbb{Z})$ as part of showing the bound held for $\mathrm{SL}_d(\mathcal{O}_K)$, where $\mathcal{O}_K$ is the ring of integers in a number field K. The proof uses the Chebotarev density theorem (see section 3.3). We will use variations of the Chebotarev density theorem for the case where $\mathcal{O}_K$ is replaced by a more general domain.

The above proof provides an outline in the general case where $\Gamma \leq G(R)$ is finitely generated, for some linear algebraic group G defined over $\mathbb{Z}$ and some ring R. We let $A \in \Gamma$ have word length n and find a bound on the size of its entries, which we use to create a homomorphism $\varphi^* : \Gamma \to G(\mathbb{F})$ for some finite field $\mathbb{F}$ with $|\mathbb{F}| \leq Cn^{\dim(G)}$, thus providing an upper bound on $F_\Gamma^{\trianglelefteq}(n)$. If G is a simple Chevalley group, we ensure that $\varphi^*(A)$ does not vanish in the simple group $G(\mathbb{F})/Z(G(\mathbb{F}))$ and use Lemma 2.5.5 to find a subgroup H of index approximately $n^{a(G)}$ not containing the image of A, which establishes the desired upper bound on $F_\Gamma^{\leq}(n)$.

The main difficulty in the above argument is in finding a finite field $\mathbb{F}$ of the correct size so that A does not become trivial in $G(\mathbb{F})$. The strategy will differ between characteristic 0 and p, but each will use variations of the Chebotarev density theorem.

6.2 Purely Transcendental Extensions

In this section we prove Theorem 1.2.2 in the case K is a purely transcendental extension of a finite field. We first need the following lemma.

Lemma 6.2.1. *Let $f(t) \in \mathbb{F}_q[t]$ be nonzero with degree at most n. Then there exists a finite field $\mathbb{F}$ with $2n < |\mathbb{F}| \leq 2nq$ and a homomorphism $\phi : \mathbb{F}_q[t] \to \mathbb{F}$ such that*

$\phi(f(t)) \neq 0$.

Proof. Recall from section 3.1 that $I_q(k)$ is the number of irreducible polynomials in $\mathbb{F}_q[t]$ of degree k; the inequality $kI_q(k) \geq \frac{1}{2}q^k$ for $k \geq 2$ follows immediately from Proposition 3.1.1.

Given $f(t) \in \mathbb{F}_q[t]$, we wish to find an irreducible polynomial of appropriate degree that does not divide $f(t)$. To that end, note that if $f(t)$ is divisible by all irreducible polynomials of degree k, then

$$\deg f(t) \geq kI_q(k) \geq \frac{1}{2}q^k.$$

So now let $f(t) \in \mathbb{F}_q[t]$ have degree at most n. Choose $M \in \mathbb{N}$ with $\frac{1}{2}q^{M-1} \leq n < \frac{1}{2}q^M$. Then by the above observation, there is some irreducible polynomial $h(t)$ with degree M such that $h(t)$ does not divide $f(t)$. From the choice of M we have

$$2n < q^M \leq 2nq,$$

so $f(t)$ is not zero in the field $\mathbb{F} = \mathbb{F}_q[t]/(h(t))$, which satisfies $2n < |\mathbb{F}| \leq 2nq$. □

We now set some notation to make the proof of the following proposition clearer. Let $\mathbb{F}_q(t)(x_1, \cdots, x_s)$ be a purely transcendental extension of $\mathbb{F}_q$ of degree at least 1. If $f(t, x_1, \cdots, x_s) \in \mathbb{F}_q[t][x_1, \cdots, x_s]$, we will view it as a polynomial in the indeterminates $x_1, \cdots, x_s$ with coefficients in $\mathbb{F}_q[t]$. Then the degree of f is the largest degree of a monomial term of f, where the degree of $x_1^{n_1} \cdots x_s^{n_s}$ is $n_1 + \cdots + n_s$. The height of f is the maximum degree in t of a coefficient of f.

Proposition 6.2.2. *Let G be a linear algebraic group defined over $\mathbb{Z}$ and let K be a purely transcendental extension of $\mathbb{F}_q(t)$ for some prime power q. If $\Gamma \leq G(K)$ is finitely generated, then $F_\Gamma^{\trianglelefteq}(n) \preceq n^{\dim(G)}$ and, if G is a simple Chevalley group, $F_\Gamma^{\leq}(n) \preceq n^{a(G)}$. If $G = \mathrm{GL}_d$, then $F_\Gamma^{\trianglelefteq}(n) \preceq n^{d^2-1}$ and $F_\Gamma^{\leq}(n) \preceq n^{d-1}$.*

Proof. Fix an embedding $G \hookrightarrow \mathrm{GL}_d$, allowing us to treat elements of Γ as invertible matrices with entries in K. Because Γ is finitely generated, we may assume the transcendence basis of K is finite, so write $K = \mathbb{F}_q(t)(x_1, \cdots, x_s)$ for some indeterminates x_i. For notational convenience write $R = \mathbb{F}_q[t][x_1, \cdots, x_s]$. Again using the fact that Γ is finitely generated, $\Gamma \leq G(S)$ for some $S = R[g^{-1}]$, $g \in R$.

Let X be a symmetric finite generating set of Γ. Let $m > 0$ such that $g^m\gamma \in \mathrm{Mat}_d(R)$ for all $\gamma \in X$. Fix $A \in \Gamma$ with $||A||_X = n$ and put $B = g^{mn}A \in \mathrm{Mat}_d(R)$.

Since A is a word of length n in the elements of X, we may view B as a word of length n in the elements of $g^mX = \{g^m\gamma : \gamma \in X\}$. Let N be larger than the degree or height of any entry of an element of g^mX.

If A is not a scalar matrix, then B has a nonzero off-diagonal entry or two diagonal entries with nonzero difference; in this case put f equal to one of these nonzero values. We can ignore the finitely many instances where A is a scalar matrix of determinant 1. If $A = aI_d$ is scalar with determinant not equal to 1, put $f = g^{mnd}(a^d - 1)$. Our general strategy is to map $R[x_1, \cdots, x_s]$ to an appropriately sized finite field $\mathbb{F}$ so that fg is not mapped to 0. This map will then extend to a homomorphism $\varphi : S \to \mathbb{F}$

with $\varphi(f) \neq 0$, so that under the induced homomorphism

$$\varphi^* : G(S) \to G(\mathbb{F}),$$

the image of A is not a scalar matrix or has determinant not equal to 1.

We must first bound the degrees of the entries of B. Recall that B can be represented as a word of length n in $g^m X$, and each entry of an element of $g^m X$ has degree less than or equal to N. Thus each entry of B has degree bounded above by nN; in particular, $\deg f \leq ndN$, so if we set $h = fg$, then $\deg h \leq 2ndN$ for sufficiently large n. Similar reasoning shows $\mathrm{ht}(f) \leq 2ndN$.

Since h is nonzero, it has some nonzero coefficient $h_0(t) \in \mathbb{F}_q[t]$ with $\deg h_0(t) \leq 2ndN$. By Lemma 6.2.1, there exists a field $\mathbb{F}$ and homomorphism $\tau : \mathbb{F}_q[t] \to \mathbb{F}$ such that

$$2n(2dN) \leq |\mathbb{F}| \leq 2qn(2dN)$$

and $\tau(h_0) \neq 0$. Extending τ in the natural way to

$$\tau : \mathbb{F}_q[t][x_1, \cdots, x_s] \to \mathbb{F}[x_1, \cdots, x_s],$$

note that $\tau(h) \neq 0$ and $\deg \tau(h) \leq 2ndN < |\mathbb{F}|$. Hence there exist $\alpha_1, \cdots, \alpha_s \in \mathbb{F}$ so that $\tau(f)(\alpha_1, \cdots, \alpha_s) \in \mathbb{F}^\times$, as is easily shown by induction on s.

Composing this evaluation map with τ yields a homomorphism $\theta : R \to \mathbb{F}$ such that $\theta(h) \neq 0$. Since the image of θ is a field and $h = fg$, g is mapped to a unit by θ, so θ extends to a ring homomorphism $\varphi : S \to \mathbb{F}$ satisfying $\varphi(f) \neq 0$. Finally, φ

induces a group homomorphism

$$\varphi^* : G(S) \to G(\mathbb{F})$$

with $\varphi^*(A)$ nontrivial. By Lemma 2.5.6, $|G(\mathbb{F})| \leq C|\mathbb{F}|^{\dim(G)}$ for some constant C depending only on G, so $|\mathbb{F}| \leq 4qdNn$ and $F_\Gamma^{\trianglelefteq}(n) \preceq n^{\dim(G)}$.

Now assume G is a simple Chevalley group. If A is a scalar matrix, then by the choice of f we have $\det \varphi^*(A) \neq 1$, so the image of $\varphi^*(A)$ in $\mathbb{F}$ under the determinant map is nontrivial. So suppose A was not scalar. Then $\varphi^*(A)$ is not scalar by the choice of f, so its image $\overline{A}$ is nontrivial in the simple group $G(\mathbb{F})/Z(G(\mathbb{F}))$. Let P be a maximal subgroup of minimal index in $G(\mathbb{F})/Z(G(\mathbb{F}))$, so $[G(\mathbb{F})/Z(G(\mathbb{F})) : P] \leq 2|\mathbb{F}|^{a(G)}$ by Lemma 2.5.5. The intersection of all conjugates of P is normal, so since $G(\mathbb{F})/Z(G(\mathbb{F}))$ is simple, this intersection is trivial. Thus $\overline{A}$ is not in a subgroup of $G(\mathbb{F})/Z(G(\mathbb{F}))$ of index at most $2|\mathbb{F}|^{a(G)}$. Hence

$$A \notin H \leq \Gamma \text{ with } [\Gamma : H] \leq 2|\mathbb{F}|^{a(G)},$$

so $F_\Gamma^{\leq}(n) \preceq n^{a(G)}$.

Now suppose $G = \mathrm{GL}_d$. If A is a scalar matrix, then $\det(\varphi^*(A)) \neq 1$ and the image of $\varphi^*(A)$ in $\mathbb{F}^*$ under the determinant map is nontrivial. Otherwise $\varphi^*(A) \in \mathrm{GL}_d(\mathbb{F})$ is not a scalar matrix, so the image of $\varphi^*(A)$ is nontrivial in $\mathrm{GL}_d(\mathbb{F})/Z(\mathrm{GL}_d(\mathbb{F}))$, the size of which is bounded by a constant multiple of $|\mathbb{F}|^{d^2-1}$. Hence $F_\Gamma^{\trianglelefteq}(n) \preceq n^{d^2-1}$.

In addition, the image of $\varphi^*(A)$ in $\mathrm{GL}_d(\mathbb{F})/Z(\mathrm{GL}_d(\mathbb{F}))$ is in the image of $\mathrm{SL}_d(\mathbb{F})$, which is isomorphic to $\mathrm{PSL}_d(\mathbb{F})$. Applying the Chevalley group argument from above to $G = \mathrm{PSL}_d$ and using the fact that $[\mathrm{PGL}_d(\mathbb{F}) : \mathrm{PSL}_d(\mathbb{F})] \leq d$, we find that $F_\Gamma^{\leq}(n) \preceq n^{d-1}$. □

6.3 Chebotarev Density Theorems

We now prepare to prove Theorem 1.2.1 and finish proving Theorem 1.2.2. To work with coefficients in arbitrary fields, we need two variations of the Chebotarev density theorem, which was discussed in section 3.3.

The standard Chebotarev density theorem applies to global fields; in particular it applies to finite Galois extensions of $\mathbb{Q}$. We will need a version of the Chebotarev density theorem which applies to finite Galois extensions of $K = \mathbb{Q}(x_1, \cdots, x_s)$, a purely transcendental extension of $\mathbb{Q}$. Its ring of integers, i.e. the integral closure of $\mathbb{Z}$ in K, is the polynomial ring $\mathcal{O}_K = \mathbb{Z}[x_1, \cdots, x_s]$.

We recall the notation of section 3.3. We denote by $P(K)$ the set of maximal ideals of $\mathcal{O}_K$ and put

$$\pi(x) = |\{\mathfrak{p} \in P(K) : N\mathfrak{p} \leq x\}|,$$

where $N\mathfrak{p} = |\mathcal{O}_K/\mathfrak{p}| < \infty$.

Let L/K be a finite Galois extension with Galois group G and ring of integers $\mathcal{O}_L$. If $\mathfrak{p}$ is a maximal ideal of $\mathcal{O}_K$ which is unramified in L and $\mathfrak{P}$ is a maximal ideal of $\mathcal{O}_L$ lying above $\mathfrak{p}$, then G contains a unique element which acts on $\mathcal{O}_L/\mathfrak{P}$ as the Frobenius automorphism $x \mapsto x^{N\mathfrak{p}}$. We denote the conjugacy class of this element by

$\left(\frac{L/K}{\mathfrak{p}}\right)$.

If $\mathcal{C}$ is a conjugacy class of G, define

$$P(K)_{\mathcal{C}} = \left\{\mathfrak{p} \in P(K) : \left(\frac{L/K}{\mathfrak{p}}\right) = \mathcal{C}\right\},$$

$$\pi^{\mathcal{C}}(x) = |\{\mathfrak{p} \in P(K)_{\mathcal{C}} : N\mathfrak{p} \leq x\}|.$$

The following result is obtained by applying Theorem 9.11 of [25], which is a broader generalization of the Chebotarev density theorem, to the setting $K = \mathbb{Q}(x_1, \cdots, x_s)$.

Theorem 6.3.1. *Let $K = \mathbb{Q}(x_1, \cdots, x_s)$ be a purely transcendental extension of $\mathbb{Q}$ and let L/K be a finite Galois extension with Galois group G. If $\mathcal{C}$ is a conjugacy class of G, then $P(K)_{\mathcal{C}}$ has natural density $|\mathcal{C}|/|G|$ in $P(K)$.*

To apply this result to residual finiteness growth, we need to know how $\pi(x)$ grows. We would also like to work with maximal ideals of the form $(p, x_1 - a_1, \cdots x_s - a_s)$ instead of arbitrary maximal ideals. If $\mathfrak{p}$ is a maximal ideal of $\mathcal{O}_K$, then $\mathcal{O}_K/\mathfrak{p}$ is a finite field, so it is isomorphic to $\mathbb{F}_{p^d}$ for some d. Define the degree of $\mathfrak{p}$ to be d, the degree of the field extension $\mathcal{O}_K/\mathfrak{p}$ over $\mathbb{F}_p$. Set

$$\pi_1(x) = |\{\mathfrak{p} \in P(K) : N\mathfrak{p} \leq x, \mathfrak{p} \text{ is degree } 1\}.$$

If $\mathfrak{p}$ has degree 1, then $\mathcal{O}_K/\mathfrak{p} \cong \mathbb{F}_p$ for some prime p. Thus the map $\mathcal{O}_K \to \mathbb{F}_p$ sends p to 0 and each x_i to an element a_i of $\mathbb{F}_p$, so $p, x_i - a_i \in \mathfrak{p}$ for each $1 \leq i \leq s$. Since $(p, x_1 - a_1, \cdots, x_s - a_s)$ is maximal and contained in $\mathfrak{p}$, it must equal $\mathfrak{p}$. So we want to work with degree 1 maximal ideals. Fortunately, in terms of density most maximal ideals are degree 1, by the following lemma.

Lemma 6.3.2. *Let $K = \mathbb{Q}(x_1, \cdots, x_s)$. Then $\pi_1(x) \sim \pi(x) \sim \dfrac{x^{s+1}}{\log(x^{s+1})}$.*

Proof. The statement follows from Corollary 9.2 and Lemma 9.3 in [25]. □

We are most interested in applying the Chebotarev density theorem with the trivial conjugacy class $\{1\}$. If $\left(\dfrac{L/K}{\mathfrak{p}}\right) = \{1\}$, then $\mathcal{O}_L/\mathfrak{P} = \mathcal{O}_K/\mathfrak{p}$ for any $\mathfrak{P}$ lying above $\mathfrak{p}$. In particular, if $\mathfrak{p}$ is degree one, then both these fields are just $\mathbb{F}_p$ for some prime p. We define $\pi_1^{\mathcal{C}}(x)$ in the natural way, counting the maximal ideals counted by $\pi^{\mathcal{C}}(x)$ which have degree 1. Applying Theorem 6.3.1 and Lemma 6.3.2 to this case gives the following result.

Corollary 6.3.3. *Let $K = \mathbb{Q}(x_1, \cdots, x_n)$ be a purely transcendental extension of $\mathbb{Q}$ and let L/K be a finite Galois extension with Galois group G. Then*

$$\pi_1^{\{1\}}(x) \sim \frac{1}{|G|}\pi_1(x) \sim \frac{1}{|G|}\frac{x^{n+1}}{\log(x^{n+1})}.$$

Remark 6.3.4. We will apply this result to find ideals over which a polynomial factors into a product of distinct linear polynomials. We note that the analogue of Lemma 3.3.2 is true in this setting, with the same proof.

Returning to the standard setting of global fields, we will need an effective version of the Chebotarev density theorem for $\mathbb{F}_p(t)$. We note that each maximal ideal $\mathfrak{p}$ of $\mathbb{F}_p[t]$ is a principal ideal generated by an irreducible polynomial $q(t)$ satisfying

$N\mathfrak{p} = p^{\deg(q(t))}$; we will conflate these two notions when convenient. The following theorem is a specific case of Theorem 1 in [22]. Recall that $I_p(x)$ is the number of irreducible polynomials in $\mathbb{F}_p[t]$ with degree x, and define

$$I_p^{\{1\}}(x) = |\{\mathfrak{p} \in P(K)_{\{1\}} : N\mathfrak{p} = p^x\}|.$$

That is, we count ideals with norm equal to p^x. This is more natural to consider in characteristic p because the norm of each maximal ideal is a power of p.

Theorem 6.3.5. *Let L be a finite Galois extension of $\mathbb{F}_p(t)$ with Galois group G, let P be the set of irreducible polynomials in $\mathbb{F}_p[t]$ which ramify over L, and set $D = \deg(\prod_{q(t)\in P} q(t))$. Let $\mathbb{F}_{p^m}$ be the algebraic closure of $\mathbb{F}_p$ in L. If m divides x and $x > \max\{\deg q(t) : q(t) \in P\}$, then*

$$\left| I_p^{\{1\}}(x) - \frac{m}{|G|} I_p(x) \right| \leq \frac{p^{x/2}(2+D)}{x|G|} + D\left(1 + \frac{1}{x}\right)$$

Remark 6.3.6. In the statement of Theorem 1 in [22], there is no requirement on the size of x, and $I_p(x)$ is replaced by the number of unramified irreducible polynomials of degree x. Since there are only finitely many ramified irreducible polynomials, these numbers are equal for x sufficiently large, which we have taken x to be in our statement of the result.

The following technical lemma will enable us to apply Theorem 6.3.5 to finding upper bounds on residual finiteness growth of linear groups over fields of positive characteristic.

Lemma 6.3.7. *Fix $c_1, c_2 > 0$. Let $f(y) \in \mathbb{F}_p[t][y]$ be separable with degree k. If n is sufficiently large, $h(t) \in \mathbb{F}_p[t]$ has degree at most $c_1 n \log n$, and $f(y)$ has discriminant $\Delta(f) \in \mathbb{F}_p[t]$ of degree less than $c_2 \log n$, then there exists $c \leq 2c_1(k!)p^{k!}$, dependent on n, so that there exists an irreducible polynomial $g(t) \in \mathbb{F}_p[t]$ of degree at most $\log_p(cn \log n)$ not dividing $h(t)$ such that $f(y)$ factors into distinct linear factors mod $g(t)$.*

Proof. Assume $h(t)$ and $f(y)$ satisfy the assumptions of the lemma. Let L be the splitting field of $f(y)$, so that $L/\mathbb{F}_p(t)$ is a finite Galois extension. Let

$$Q_x = |\{\mathfrak{p} \in P(K)_{\{1\}} : N\mathfrak{p} = p^x\}|,$$

so that $I_p^{\{1\}}(x) = |Q_x|$. If $x > \deg(\Delta(f))$, then $\Delta(f) \notin \mathfrak{p}$ for all $\mathfrak{p} \in Q_x$. Lemma 3.3.2 then immediately implies that $f(y)$ factors into distinct linear factors mod $\mathfrak{p}$ for all $\mathfrak{p} \in Q_x$.

We now treat the elements of Q_x as irreducible polynomials of degree x. We want to find an x of appropriate size so that some $g(t) \in Q_x$ does not divide $h(t)$. To that end, observe that

$$\deg\left(\prod_{q(t) \in Q_x} q(t)\right) = x I_p^{\{1\}}(x).$$

Let $\mathbb{F}_{p^m}$ be the algebraic closure of $\mathbb{F}_p$ in L. Since $[L : \mathbb{F}_p(t)] \leq k!$, we have $|G| \leq k!$ and $m|k!$, so if $m|x$ and x is sufficiently large, then Theorem 6.3.5 yields

$$\begin{aligned} I_p^{\{1\}}(x) &\geq \frac{m}{k!} I_p(x) - \frac{p^{x/2}(2+D)}{x|G|} - D\left(1 + \frac{1}{x}\right) \\ &\geq \frac{1}{k!} I_p(x) - \frac{p^{x/2}(2+D)}{x} - 2D. \end{aligned}$$

Since $\deg(h(t)) \leq c_1 n \log n$, if $xI_p^{\{1\}}(x) > c_1 n \log n$, then some $g(t) \in Q_x$ will not divide $h(t)$. Using the estimates $xI_p(x) \geq \dfrac{p^x}{2}$ and $D \leq \deg(\Delta(f)) \leq c_2 \log n$, we have

$$\begin{aligned} xI_p^{\{1\}}(x) &\geq \frac{xI_p(x)}{k!} - p^{x/2}(2+D) - 2Dx \\ &\geq \frac{p^x}{2(k!)} - p^{x/2}(2 + c_2 \log n) - 2c_2 x \log n. \end{aligned}$$

Then

$$xI_p^{\{1\}}(x) - c_1 n \log n \geq \frac{p^x}{2(k!)} - p^{x/2}(2 + c_2 \log n) - 2c_2 x \log n - c_1 n \log n. \qquad (6.3.1)$$

If we set $x = \log_p(c' n \log n)$ for some $c' > 0$, then the right hand side of (6.3.1) becomes

$$\frac{c' n \log n}{2(k!)} - \sqrt{c' n \log n}(2 + c_2 \log n) - 2c_2 \log_p(n \log n) \log n - c_1 n \log n.$$

The highest order terms in n are $\dfrac{c' n \log n}{2(k!)}$ and $c_1 n \log n$. Hence if $c' > 2c_1(k!)$ and n is sufficiently large, then the above expression is positive.

However, we also need x to be an integer divisible by m, while $\log_p(c' n \log n)$ may not even be an integer. Since m divides $k!$, it is enough to have $k!$ divide x. For any n, the interval $(\log_p(2c_1(k!)n \log n), \log_p(2c_1(k!)p^{k!} n \log n)]$ has length $k!$, so it contains an integer multiple of $k!$. Thus there exists $c > 0$ satisfying

$$2c_1(k!) < c \leq 2c_1(k!)p^{k!}$$

such that $x = \log_p(cn \log n) \in k!\mathbb{Z}$. Note that while the choice of c depends on n, its absolute value is bounded independent of n.

For n sufficiently large, the above choice of c yields $xI_p^{\{1\}}(x) > \deg(h(t))$, so we conclude that there is some irreducible $g(t) \in \mathbb{F}_p[t]$ such that $g(t)$ does not divide $h(t)$ and $f(y) \bmod g(t)$ factors into distinct linear factors. □

6.4 The Main Result

We prove two straightforward lemmas to aid the argument in characteristic p.

Lemma 6.4.1. *Let $f \in \mathbb{F}_p[t][x_1, \cdots, x_s]$ be nonzero with $\deg f \leq 2^m$. Then there exist $g_1(t), \cdots, g_s(t) \in \mathbb{F}_p[t]$ with $\deg g_i(t) \leq m$ for each i such that $f(g_1(t), \cdots, g_s(t)) \neq 0$.*

Proof. We induct on s. Suppose $s = 1$. Since $\mathbb{F}_p[t]$ is an integral domain and $\deg f \leq 2^m$, $f(x)$ has at most 2^m roots. There are at least 2^{m+1} elements of $\mathbb{F}_p[t]$ with degree at most m, so $f(g(t)) \neq 0$ for some $g(t)$ with $\deg g(t) \leq m$.

Now assume the lemma is true for $s = n-1$ and suppose $s = n$. When considered as a polynomial over x_s with coefficients in $\mathbb{F}_p[t][x_1, \cdots, x_{s-1}]$, f has at most 2^m roots, so there is some $g_s(t) \in \mathbb{F}_p[t]$ with $\deg g(t) \leq m$ such that

$$f(x_1, \cdots, x_{s-1}, g(t)) \neq 0.$$

Applying the inductive hypothesis finishes the proof. □

The primitive element theorem will play an important role in the proof of the main theorem of this chapter. This theorem is not guaranteed to hold in positive characteristic because finite extensions are no longer necessarily separable. The following lemma allows us to reduce to the case of a separable extension.

Lemma 6.4.2. *Let p be a prime and put $E_0 = \mathbb{F}_p(x_1, \cdots, x_s)$ for some $x_1, \cdots, x_s$ algebraically independent over $\mathbb{F}_p$. If L_0/E_0 is a finite extension, then there is some positive integer m such that if $\tilde{x}_j = x_j^{1/p^m}$ for $1 \leq j \leq s$, then $L = L_0(\tilde{x}_1, \cdots, \tilde{x}_s)$ is a separable extension of $E = \mathbb{F}_q(\tilde{x}_1, \cdots, \tilde{x}_s)$.*

Proof. First note that since L_0/E_0 is finite, there are some $\alpha_1, \cdots, \alpha_k \in L_0$ such that $L_0 = E_0(\alpha_1, \cdots, \alpha_k)$. Each α_i is the root of an irreducible polynomial $f_i(y) \in E_0[y]$. In turn, each $f_i(y) = g_i(y^{p^{m_i}})$ for some irreducible, separable $g_i(y) \in E_0[y]$ and some positive integer m_i. Set $m = \max\{m_i\}$, put $\tilde{x}_j = x_j^{1/p^m}$ for $1 \leq j \leq s$, and let $E = \mathbb{F}_q(\tilde{x}_1, \cdots, \tilde{x}_s)$. For each i, form $\tilde{g}_i(y) \in E[y]$ by replacing each x_j in $g_i(y)$ by $\tilde{x}_j^{p^{m-m_i}} = x_j^{1/p^{m_i}}$.

Since we are in characteristic p, we then have

$$f_i(y) = g_i(y^{p^{m_i}}) = \tilde{g}_i(y)^{p^{m_i}},$$

so $\tilde{g}_i(\alpha_i) = 0$. Each $\tilde{g}_i(y)$ is separable, so $L = L_0(\tilde{x}_1, \cdots, \tilde{x}_s)$ is separable over E. □

We now prove Theorems 1.2.1 and 1.2.2 together.

Theorem 6.4.3. *Let G be a linear algebraic group defined over $\mathbb{Z}$, K a field, and $\Gamma \leq G(K)$ a finitely generated subgroup. Put $g(n) = n$ if char $K = 0$ and $g(n) = n \log n$ if char $K > 0$.*

Then $F_\Gamma^{\trianglelefteq}(n) \preceq g(n)^{\dim(G)}$ and, if G is a simple Chevalley group, $F_\Gamma^{\leq}(n) \preceq g(n)^{a(G)}$. If $G = \mathrm{GL}_d$, then $F_\Gamma^{\trianglelefteq}(n) \preceq g(n)^{d^2-1}$ and $F_\Gamma^{\leq}(n) \preceq g(n)^{d-1}$.

Proof. Fix an embedding $G \hookrightarrow \mathrm{GL}_d$ ($G \hookrightarrow \mathrm{SL}_d$ if G is a Chevalley group) and let $\Gamma = \langle X \rangle \leq G(K)$, where X is finite and symmetric. We may assume K is the field generated by the entries of the elements of X. We first consider the case $\operatorname{char} K = 0$, though we will see later that most of the arguments leading up to the use of the Chebotarev density theorem apply when $\operatorname{char} K > 0$.

Since K is a finitely generated field, K is a finite extension of $F = \mathbb{Q}(x_1, \cdots, x_s)$ for some algebraically independent elements $x_1, \cdots, x_s$. Replacing K by its Galois closure if necessary, we may assume K/F is Galois. By the primitive element theorem, $K = \mathbb{Q}(x_1, \cdots, x_s)[\alpha]$ for some $\alpha \in K$, which we can choose to be integral over $\mathbb{Z}[x_1, \cdots, x_s] = \mathcal{O}_F$. Let $f(y) \in \mathcal{O}_F[y]$ be the minimal polynomial for α over $\mathbb{Q}(x_1, \cdots, x_s)$ and set $k = \deg f(y)$.

The entries of the elements of X generate a ring contained in $\mathcal{O}_F[g^{-1}][\alpha]$ for some $g \in \mathcal{O}_F$. Set $R = \mathcal{O}_F[g^{-1}]$ and J be the ideal of $R[y]$ generated by $f(y)$. If the ring homomorphism

$$\epsilon_\alpha : F[y] \to F[\alpha] = K$$

is evaluation of y to α, then clearly $\ker \epsilon_\alpha$ is the ideal generated by $f(y)$. We claim the kernel of $\epsilon_\alpha|_{R[y]}$ is J, so that $R[\alpha] \cong R[y]/J$. This follows from the fact that $f(y)$ is monic; if some element of $R[y]$ is a multiple of $f(y)$ in $K[y]$, then it must in fact be a multiple in $R[y]$, as is seen by an easy computation of coefficients.

Before proceeding, let us set up some convenient notation. If $h \in R[y]$, set $\widetilde{h}$ to be the element of $R[y]$ with $\widetilde{h} \equiv h \bmod J$ and $\deg_y \widetilde{h} < k$. If $b = h + J \in R[\alpha]$, set

$\widetilde{b} = \widetilde{h}$.

We now present an outline of the proof. Let $A \in \Gamma$ with $||A||_X = n$. Using the above argument, we consider Γ as being embedded in $G(R[y]/J)$. By using appropriate coset representatives and multiplication to eliminate inverses, we examine the entries of A as elements of $\mathcal{O}_F[y]$, i.e. as polynomials with integer coefficients. We then use the variation of the Chebotarev density theorem given in Corollary 6.3.3 to produce a homomorphism $\mathcal{O}_F[y] \to \mathbb{F}_p[y]$ under which an element related to the entries of A remains nontrivial and the image of $f(y)$ factors as a product of distinct linear factors. The end result is a homomorphism $R[y] \to \mathbb{F}_p$ which factors through J, inducing a homomorphism $G(R[y]/J) \to G(\mathbb{F}_p)$ in which A remains nontrivial. This suffices to prove the normal residual finiteness growth bound; for the non-normal residual finiteness growth upper bound, we use similar techniques as in the proof of Proposition 6.2.2.

Accomplishing this with no regard for the size of $G(\mathbb{F}_p)$ is fairly straightforward, but to achieve the desired bound, we must keep track of certain details. This is the reason we prefer to work in $\mathcal{O}_F[y]$; its elements are just polynomials with integer coefficients, with easily tracked “size” properties.

So let $A = (A_{ij}) \in \Gamma$ be nontrivial with $||A||_X = n$. For each $\gamma = (\gamma_{ij}) \in X$, let $\widetilde{\gamma} = (\widetilde{\gamma}_{ij})$ be the element of $\text{Mat}_d(R[y])$ with $\widetilde{\gamma}_{ij} = \widetilde{\gamma_{ij}}$. Put $\widetilde{X} = \{\widetilde{\gamma} | \gamma \in X\}$ and let $m > 0$ such that $g^m\widetilde{\gamma} \in \text{Mat}_d(\mathcal{O}_F[y])$ for all $\gamma \in X$. Let N_0 be the maximum degree in $x_1, \cdots, x_s$ of the entries of all the $g^m\widetilde{\gamma}$.

If $A = \gamma_1 \cdots \gamma_n$, $\gamma_i \in X$, set $\widetilde{A} = \widetilde{\gamma_1} \cdots \widetilde{\gamma_n}$. Then $(g^m)^n \widetilde{A} = B = (B_{ij})$ is a product of n elements chosen from $g^m \widetilde{X}$, so $B \in \mathrm{Mat}_d(\mathcal{O}_F[y])$. For convenience, suppose B has a nonzero off-diagonal entry $h(x_1, \cdots, x_s, y) = h$ which is not divisible by $f(y)$; the case when B is diagonal can be treated as in the proof of Proposition 6.2.2. Then for some constant α_0 depending on g^m, X, and s, we have

$$\deg_y h \leq (k-1)n, \;\; \deg_{x_1, \cdots, x_s} h \leq N_0 n, \;\text{ and } \mathrm{ht}(h) \leq \alpha_0^n,$$

where $\mathrm{ht}(h)$, called the height of h, is the largest absolute value of a coefficient of h.

We want to ensure that h continues to not be divisible by $f(y)$ when we evaluate the x_i; the easiest way to accomplish this is by degree considerations, so we now replace h by $\widetilde{h}$. Since our goal is to map this element to something nonzero, it will then suffice to clear denominators and map the resulting polynomial to something nonzero. We need to do this carefully to keep track of how the x degrees and coefficient sizes change.

Write

$$f(y) = y^k + \sum_{j=0}^{k-1} a_j(x_1, \cdots, x_s) y^j,$$

where $a_j \in \mathcal{O}_F$, and put M to be the maximum degree of the a_j. Then for $r > k$, the coefficients of $\widetilde{y^r}$ will be sums of products of the a_j. For example,

$$\begin{aligned} y^{k+1} = y \cdot y^k &\equiv y\left(-\sum_{j=0}^{k-1} a_j y^j\right) \mod J \\ &\equiv -a_{k-1}y^k - \sum_{j=0}^{k-2} a_j y^{j+1} \mod J \\ &\equiv a_{k-1}\sum_{j=0}^{k-1} a_j y^j - \sum_{j=0}^{k-2} a_j y^{j+1} \mod J \\ &\equiv a_0 a_{k-1} + \sum_{j=1}^{k-1}(a_j a_{k-1} - a_{j-1})y^j \mod J \\ &= \widetilde{y^{k+1}}. \end{aligned}$$

As the above example helps illustrate, each $\widetilde{y^r} \in \mathcal{O}_F[y]$, and the coefficients of $\widetilde{y^r}$ will include products of at most $r-(k-1)$ coefficients of $f(y)$, so $\widetilde{y^{(k-1)n}}$ includes products of at most $(k-1)(n-1)$ terms. Hence if $a(x_1, \cdots, x_s)$ is a coefficient of $\widetilde{y^r}$ with $r \le (k-1)n$, then

$$\deg a \le M(n-1)(k-1) \text{ and } \mathrm{ht}(a) \le \beta^{(n-1)(k-1)}$$

for some β independent of n.

We obtain $\widetilde{h} \in \mathcal{O}_F[y]$ by replacing each y^r by $\widetilde{y^r}$. Using the size and degree estimates on $\widetilde{y^r}$, we have

$$\deg_y \widetilde{h} < k, \ \deg_{x_1,\cdots,x_s} \widetilde{h} \le N_1 n, \text{ and } \mathrm{ht}(\widetilde{h}) \le \alpha_1^n,$$

where N_1 and α_1 are independent of n.

Viewing $\widetilde{h}$ as a polynomial with coefficients in $\mathcal{O}_F = \mathbb{Z}[x_1, \cdots, x_s]$, some coefficient $b(x_1, \cdots, x_s) = b$ of $\widetilde{h}$ is nonzero. Let $\Delta(f(y))$ be the discriminant of $f(y)$, an element of $\mathcal{O}_F$. Consider the polynomial

$$b'(x_1, \cdots, x_s) = g(x_1, \cdots, x_s)\Delta(f(y))b(x_1, \cdots, x_s) \in \mathcal{O}_F. \tag{6.4.1}$$

The only term in this product that depends on n is $b(x_1, \cdots, x_s)$, so b' retains the properties from $\widetilde{h}$ that its degree is linear in n and its height is exponential in n, so bounded above by Nn and α^n, respectively, for some N, α independent of n.

Our goal is to find a homomorphism from $R[y]$ to an appropriately small finite field such that the image of b' is nontrivial and the image of $f(y)$ is a product of distinct linear factors. We will use Corollary 6.3.3 to accomplish this.

Let $\mathcal{O}_K$ be the integral closure of $\mathcal{O}_F$ in K, and recall the definitions of $\pi_1(x)$ and $\pi_1^{\{1\}}(x)$ from Corollary 6.3.3. We wish to find $\mathfrak{p} \trianglelefteq \mathcal{O}_F$ of degree 1 with $|\mathcal{O}_F/\mathfrak{p}| \leq Cn$ for some constant C independent of n such that $b'(x_1, \cdots, x_s) \bmod \mathfrak{p} \neq 0$ and $f(y) \bmod \mathfrak{p}$ factors into distinct linear factors.

There are $\pi_1^{\{1\}}(Cn)$ degree 1 maximal ideals $\mathfrak{p}$ with $|\mathcal{O}_K/\mathfrak{p}| \leq Cn$ and $\left(\dfrac{K/F}{\mathfrak{p}}\right) = \{1\}$; if $b' \notin \mathfrak{p}$, then $\Delta(f) \notin \mathfrak{p}$, so by Lemma 3.3.2 and Remark 6.3.4 $f(y) \bmod \mathfrak{p}$ is a product of distinct linear factors. Thus we wish to show that the number of degree 1 maximal ideals of norm at most Cn that contain b' is less than $\pi_1^{\{1\}}(Cn)$.

Each degree one maximal ideal of $\mathcal{O}_F = \mathbb{Z}[x_1, \cdots, x_s]$ is of the form $(p, x_1 - a_1, \cdots, x_s - a_s)$ for some $0 \leq a_i \leq p-1$. For such an ideal $\mathfrak{p}$, $b'(x_1, \cdots, x_s) \in \mathfrak{p}$ if and only if $b'(a_1, \cdots, a_s) \equiv 0 \bmod p$. Set

$$X_p(b') = \{\mathbf{a} \in \mathbb{F}_p^s : b'(\mathbf{a}) \equiv 0 \bmod p\}.$$

Then $|X_p(b')| = p^s$ if $p|b'$. If p does not divide b', we claim $|X_p(b')| \leq s\deg(b')p^{s-1}$.

To show this, we induct on s. When $s = 1$, $X_p(b') = \{a \in \mathbb{F}_p : b'(a) \equiv 0 \bmod p\}$, so $|X_p(b')| \leq \deg(b')$.

Now assume $s > 1$. There are at most $\deg(b')$ values of a_s in $\mathbb{F}_p$ such that $b'(x_1, \cdots, x_{s-1}, a_s) \equiv 0 \bmod p$, yielding at most $p^{s-1}\deg(b')$ elements of $X_p(b')$. For the remaining at most p values of a_s, $b'(x_1, \cdots, x_{s-1}, a_s)$ is a nonzero polynomial mod p, so by induction there are at most $(s-1)\deg(b')p^{s-2}$ tuples $(a_1, \cdots, a_{s-1})$ such that $b'(a_1, \cdots, a_s) \equiv 0 \bmod p$. This gives at most $(s-1)\deg(b')p^{s-1}$ elements of $X_p(b')$, so

$$|X_p(b')| \leq p^{s-1}\deg(b') + (s-1)\deg(b')p^{s-1} = s\deg(b')p^{s-1},$$

proving the claim. Hence if $M > 0$,

$$\sum_{p \leq M} |X_p(b')| \leq s\deg(b') \sum_{p \leq M} p^{s-1} + \sum_{p \leq M, p|b'} p^s.$$

We can split the second sum into two as

$$\sum_{p \leq M, p|b'} p^s = \sum_{p \leq \sqrt{n}, p|b'} p^s + \sum_{\sqrt{n} < p \leq M, p|b'} p^s. \tag{6.4.2}$$

Each prime in the second sum of (6.4.2) is greater than $\sqrt{n}$, so if there are l terms in the sum, the product of the involved primes is at least $(\sqrt{n})^l$. Since b' has height at most α^n, $(\sqrt{n})^l \leq \alpha^n$, so $l \leq 2n\log\alpha/\log n$. Hence

$$\sum_{\sqrt{n} < p \leq M, p|b'} p^s \leq \frac{2n\log\alpha}{\log n} M^s.$$

The sum $\sum_{p \leq \sqrt{n}} p^s$ is the number of ideals in $\mathbb{Z}[x_1, \cdots, x_s]$ of the form $(p, x_1 -$

$a_1, \cdots, x_s - a_s)$ with $p \leq \sqrt{n}$, so if n is sufficiently large, Lemma 6.3.2 gives

$$\sum_{p \leq \sqrt{n}} p^s \leq 2 \frac{(\sqrt{n})^{s+1}}{\log((\sqrt{n})^{s+1})}.$$

If we put $M = Cn$ for some $C > 1$ to be determined, we conclude that

$$\sum_{p \leq M, p | b'} p^s \leq \frac{4n \log \alpha}{\log n} M^s.$$

Using Lemma 6.3.2 again with $s-1$ in place of s, we have

$$\sum_{p \leq M} p^{s-1} \leq 2 \frac{M^s}{\log(M^s)},$$

so recalling that $\deg(b') \leq cn$, we can conclude that

$$\sum_{p \leq M} |X_p(b')| \leq 2scn \frac{M^s}{\log(M^s)} + \frac{4n \log \alpha}{\log n} M^s = 2nM^s \left(\frac{c}{\log M} + \frac{2 \log \alpha}{\log n} \right). \quad (6.4.3)$$

Thus we can find $\mathfrak{p}$ with the desired properties if $\pi_1^{\{1\}}(M)$ is greater than the right hand side of (6.4.3). If we let $m_0 = |\mathrm{Gal}(K/F)| \leq k!$, then by Corollary 6.3.3, $\pi_1(M) \geq \dfrac{1}{2m_0} \dfrac{M^{s+1}}{\log(M^{s+1})}$, so we want

$$\frac{1}{2m_0} \frac{M^{s+1}}{\log(M^{s+1})} > 2nM^s \left(\frac{c}{\log M} + \frac{2 \log \alpha}{\log n} \right) \Leftrightarrow M > 4m_0(s+1)n \left(c + \frac{2 \log \alpha \log M}{\log n} \right).$$

Recalling that $M = Cn$, the above inequality implies we want C to satisfy

$$C > 4m_0(s+1) \left(c + 2 \log \alpha \left(1 + \frac{\log C}{\log n} \right) \right).$$

Since m_0, c, s, and α are all independent of n, such a C exists independent of n for n sufficiently large. Then with $M = Cn$, $\pi_1^{\{1\}}(M) > \sum_{p \leq M} |X_p(b')|$. The number on the right side of this inequality is the number of degree one maximal ideals containing $b'(x_1, \cdots, x_s)$, so we can in fact choose a maximal ideal $\mathfrak{p}$ of $\mathcal{O}_F$ such that $b'(x_1, \cdots, x_s) \neq 0 \bmod \mathfrak{p}$, $f(y) \bmod \mathfrak{p}$ is a product of distinct linear factors, and $\mathcal{O}_F/\mathfrak{p} \cong \mathbb{F}_p$ with $p \leq M = Cn$.

Now consider the homomorphism

$$\psi : \mathcal{O}_F[y] \to (\mathcal{O}_F/\mathfrak{p})[y] \cong \mathbb{F}_p[y].$$

Since $\psi(b') \neq 0$ and $g|b'$, we have $\psi(g) \neq 0$, so ψ extends to

$$\pi : R[y] = \mathcal{O}_F[g^{-1}][y] \to \mathbb{F}_p[y]$$

with $\pi(b') \neq 0$ and $\pi(f(y))$ a product of distinct linear polynomials.

Recalling that b' is a coefficient of $\widetilde{h}$, $\pi(b') \neq 0$ implies $\pi(\widetilde{h}) \neq 0$. By our choice of $\widetilde{h}$, $\deg \pi(\widetilde{h}) < \deg \pi(f)$, so $\pi(f)$ does not divide $\pi(\widetilde{h})$. In particular, $\pi(f)$ has some linear factor $y - \lambda \in \mathbb{F}_p[y]$ that does not divide $\pi(\widetilde{h})$. Hence under the evaluation map $\epsilon_\lambda : \mathbb{F}_p[y] \to \mathbb{F}_p$ that sends y to λ, $\pi(f)$ is sent to 0 and $\pi(\widetilde{h})$ remains nontrivial. Thus we have a homomorphism

$$\epsilon_\lambda \circ \pi : R[y] \to \mathbb{F}_p$$

which maps $f(y)$ to 0 and maps $\widetilde{h}(y)$ to a nonzero element of $\mathbb{F}_p$. This map thus factors through $J = (f(y)) \trianglelefteq R[y]$, yielding the commutative diagram below.

$$\begin{tikzcd}
R[y] \arrow[r, "\pi"] \arrow[dr] & \mathbb{F}_p[y] \arrow[r, "\epsilon_a"] & \mathbb{F}_p \\
& R[y]/J \arrow[ur, "\varphi"'] &
\end{tikzcd}$$

Recall from the beginning of the proof that $h = B_{ij} \equiv g^{mn} A_{ij} \bmod J$ for some $i \neq j$, and that $\widetilde{h} \equiv h \bmod J$. Then by the above diagram, the homomorphism $\varphi : R[y]/J \to \mathbb{F}_p$ satisfies

$$0 \neq \varphi(\widetilde{h} + J) = \varphi(h + J) = \varphi(g^{mn} A_{ij} + J).$$

Since $g \in R^{\times}$, $\varphi(g^{mn} + J) \neq 0$, so we conclude that $\varphi(A_{ij} + J) \neq 0$. Thus the ring homomorphism φ induces a group homomorphism

$$\varphi^* : G(R[y]/J) \to G(\mathbb{F}_p)$$

with $\varphi^*(A)$ a nontrivial, non-diagonal matrix. Restricting φ^* to Γ yields the desired homomorphism.

By the choice of $\mathfrak{p}$ we have $|\mathbb{F}_p| \leq Cn$, so by Lemma 2.5.6, we conclude $F_{\Gamma}^{\trianglelefteq}(n) \preceq n^{\dim(G)}$. The upper bounds on non-normal residual finiteness growth G is a simple Chevalley group can now be proved using the same arguments in the proof of Proposition 6.2.2, as can bounds on both growth functions in the case $G = \mathrm{GL}_d$.

Now consider the case char $K = p > 0$. Then K is a finite extension of $F = \mathbb{F}_p(t, x_1, \cdots, x_s)$ for some algebraically independent elements $t, x_1, \cdots, x_s$. By Lemma 6.4.2 we can assume K is a separable extension of F by adding appropriate roots of the indeterminates. As in the characteristic 0 case we may then replace K by its Galois closure and assume K/F is Galois. By the primitive element theorem, $K = F[\alpha]$ for some $\alpha \in K$. We again can assume α is integral, and we let $f(y) \in \mathbb{F}_p[t][x_1, \cdots, x_s][y]$ be the minimal polynomial for α over $\mathbb{F}_p(t)(x_1, \cdots, x_s)$, with $\deg_y f(y) = k$. In this context $\mathcal{O}_F = \mathbb{F}_p[t][x_1, \cdots, x_s]$ and $R = \mathcal{O}_F[g^{-1}]$, $g \in \mathcal{O}_F$. As noted before Proposition 6.2.2, the degree of an element of $\mathcal{O}_F$ is its total degree in $x_1, \cdots, x_s$ and its height is its degree in t.

One can now perform the same steps as in the characteristic 0 case, replacing $\mathbb{Z}$ by $\mathbb{F}_p[t]$ and replacing the exponential size bounds on the coefficients by linear degree bounds. Indeed, the first place where the characteristic p argument diverges is just after (6.4.1). So we pick up the argument at that point, using the same notation as before.

We have a polynomial $b'(x_1, \cdots, x_s)$ defined similarly as in (6.4.1),

$$b'(x_1, \cdots, x_s) = g(x_1, \cdots, x_s)\Delta(f(y))b(x_1, \cdots, x_s) \in \mathcal{O}_F,$$

with $\deg b' \leq c_1 n$ and the height of b' at most $c_2 n$ for some constants c_1, c_2 independent of n. Then by Lemma 6.4.1, there exist $g_1(t), \cdots, g_s(t) \in \mathbb{F}_p[t]$, each of degree at most $\log(c_1 n)$, such that $b'(g_1(t), \cdots, g_s(t)) \neq 0$. If $\epsilon : \mathcal{O}_F \to \mathbb{F}_p[t]$ is the evaluation homomorphism with $\epsilon(x_i) = g_i(t)$, we have $\epsilon(g) \neq 0$, so ϵ extends to $\epsilon : R \to \mathbb{F}_p[t]$ and thus induces a homomorphism

$$\psi : R[y] \to \mathbb{F}_p[t][y]$$

satisfying $\psi(\widetilde{h}) \neq 0$, $\psi(f) \neq 0$, and $\psi(\Delta(f)) \neq 0$.

We are now in a position to use Lemma 6.3.7. We observe that

$$\deg \psi(b') \leq c_2 n + c_1 n \log(c_1 n) \leq c' n \log n$$

for some constant c'. Also, the discriminant of $\psi(f(y))$ has degree at most $c_3 \log n$ for some constant c_3 since $f(y)$ is independent of n and each $g_i(t)$ has degree at most $\log(c_1 n)$. Thus if we let n be sufficiently large, then by Lemma 6.3.7 we can find $c > 0$ bounded above independently of n and an irreducible polynomial $F(t)$ of degree less than $\log_p(cn \log n)$ such that $F(t)$ does not divide $\psi(b')$ and $\psi(f(y)) \bmod F(t)$ is a product of distinct linear factors.

Put $\mathbb{F} = \mathbb{F}_p[t]/(F(t))$, let π_F be the homomorphism $\pi_F : \mathbb{F}_p[t][y] \to \mathbb{F}[y]$ induced by $\mathbb{F}_p[t] \to \mathbb{F}$, and define $\pi = \pi_F \circ \psi : R[y] \to \mathbb{F}[y]$. Observe that $|\mathbb{F}| \leq cn \log n$. Following the same arguments as in the characteristic 0 case, one can then show $F_\Gamma^{\trianglelefteq}(n) \preceq (n \log n)^{\dim(G)}$ using Lemma 2.5.6. The remaining upper bounds are then found using the exact same arguments as in the proof of Proposition 6.2.2. □

Chapter 7

Residual Finiteness of Linear Groups: Lower Bounds

In this chapter we establish lower bounds on normal and non-normal residual finiteness growth of simple Chevalley groups over $\mathbb{Z}$ or $\mathbb{F}_p[t]$. The methods were developed specifically to address the $\mathbb{F}_p[t]$ case but also carry over to the characteristic 0 setting, where some results had already been found. We first construct a graded Lie algebra that allows us to work with vector spaces instead of subgroups. The characteristic 0 and p settings are then treated separately.

7.1 A Graded Lie Algebra

Let G be a simple, simply connected Chevalley group with irreducible root system Φ of rank $l \geq 2$. Let $\mathfrak{g}(\mathbb{C})$ be the Lie algebra of G, with Chevalley basis $\{e_\alpha : \alpha \in \Phi\} \cup \{h_1, \cdots, h_l\}$.

Fix an embedding $G \leq SL_d$. As discussed in section 2.5, there is an embedding of $\mathfrak{g}(\mathbb{C})$ into $\mathfrak{sl}_d(\mathbb{C})$ respecting the Lie bracket so that the action of G on $\mathfrak{g}(\mathbb{C})$ by conjugation, via matrix multiplication, is the same as the adjoint action of G on $\mathfrak{g}(\mathbb{C})$. We will use this more concrete perspective for the remainder of this section.

Example. In the case Φ is type A_l and K is a field, we have $\mathfrak{g}(K) = \mathfrak{sl}_{l+1}(K) \subseteq \mathfrak{gl}_{l+1}(K)$, with $e_{\epsilon_i - \epsilon_j} = e_{ij}$ as Chevalley basis elements, as in the example in section 2.3. Then the generators of the Chevalley group $G(K)$ are the elementary matrices $x_\alpha(t) = E_{ij}(t) = 1 + te_{ij}$, which are known to generate $\mathrm{SL}_{l+1}(K)$. If $\alpha = \epsilon_i - \epsilon_j$, $\beta = \epsilon_j - \epsilon_k$ are linearly independent roots, then $e_\alpha e_\beta e_\alpha = 0$, so

$$\begin{aligned}(1 + te_\alpha)(e_\beta)(1 - te_\alpha) &= e_\beta + te_\alpha e_\beta - te_\beta e_\alpha - t^2 e_\alpha e_\beta e_\alpha \\ &= e_\beta + t[e_\alpha, e_\beta],\end{aligned}$$

which is the action described in section 2.4.

Let K be the field $\mathbb{Q}$ or $\mathbb{F}_p(t)$, with ring of integers $\mathcal{O} = \mathbb{Z}$ or $\mathbb{F}_p[t]$, respectively. We have $G(\mathcal{O}) = G(K) \cap \mathrm{SL}_d(\mathcal{O})$. Fix a maximal ideal $\mathfrak{m} \trianglelefteq \mathcal{O}$ and $k \in \mathbb{N}$. Set $R = \mathcal{O}/\mathfrak{m}^k$ and $\mathbb{F} = \mathcal{O}/\mathfrak{m}$; if $K = \mathbb{F}_p(t)$ then $\operatorname{char} \mathbb{F} = p$, and if $K = \mathbb{Q}$, put $p = \operatorname{char} \mathbb{F}$.

Let G_i be the kernel of the projection $G(R) \to G(\mathcal{O}/\mathfrak{m}^i)$ for $i \geq 1$ (note that $G_i = \{1\}$ for $i \geq k$). Since G is simply connected,

$$\begin{gathered} G(R) = \langle x_\alpha(r) : r \in R, \alpha \in \Phi \rangle, \\ G(\mathcal{O}/\mathfrak{m}^i) = \langle x_\alpha(r \bmod \mathfrak{m}^i) : r \in R, \alpha \in \Phi \rangle. \end{gathered}$$

by Lemma 2.5.3. Then the generators of $G(\mathcal{O}/\mathfrak{m}^i)$ are all in the image of the projection $G(R) \to G(\mathcal{O}/\mathfrak{m}^i)$, so the projection is surjective.

We now use the groups G_i to construct a graded Lie algebra (see [20], Chapter 7 for more details and [2] for a similar construction). In the case $K = \mathbb{Q}$, these kernels were used in [6] to study the normal residual finiteness growth of Chevalley groups.

Lemma 7.1.1. *Let $G(R)$, G_i be as above. Then $(G_n, G_m) \subseteq G_{n+m}$ and $G_n^p \leq G_{n+1}$. As a consequence, G_i/G_{i+1} is an elementary abelian p-group for $1 \leq i \leq k-1$.*

Proof. Let $n, m \geq 1$ with $n + m \leq k - 1$, and let $x \in G_n, y \in G_m$. Since $\mathcal{O}$ is a PID, we can write $\mathfrak{m} = (\pi)$ for some irreducible element π. Then $x = I_d + \pi^n A$, $y = I_d + \pi^m B$ for some $A, B \in \mathrm{Mat}_d(R)$, where I_d is the $d \times d$ identity matrix. We show that $(x, y) \bmod \mathfrak{m}^{n+m}$ is the identity. We have $(x, y) = xyx^{-1}y^{-1} = xy(yx)^{-1}$, so

$$\begin{aligned}(x,y) &= (I_d + \pi^n A + \pi^m B + \pi^{n+m}AB)(I_d + \pi^n A + \pi^m B + \pi^{n+m}BA)^{-1}\\ &= (I_d + \pi^n A + \pi^m B)(I_d + \pi^n A + \pi^m B)^{-1} \bmod \mathfrak{m}^{n+m}\\ &= I_d \bmod \mathfrak{m}^{n+m}.\end{aligned}$$

Hence $(x, y) \in G_{n+m}$.

Still letting $x = I_d + \pi^n A \in G_n$, by the binomial theorem we have

$$x^p = (I_d + \pi^n A)^p = I + p\pi^n A \bmod \mathfrak{m}^{n+1}.$$

If $\mathcal{O} = \mathbb{Z}$, then $\pi = p$, and if $\mathcal{O} = \mathbb{F}_p[t]$, then $p = 0$, so in either case $x^p = I_d \bmod \mathfrak{m}^{n+1}$ and hence $x^p \in G_{n+1}$. □

By Lemma 7.1.1, each G_i/G_{i+1} can be viewed as a vector space over $\mathbb{F}_p$, so we can define an $\mathbb{F}_p$-vector space

$$L(G_1) = \bigoplus_{i=1}^{k-1} G_i/G_{i+1}.$$

In fact, we can make $L(G_1)$ a Lie algebra over $\mathbb{F}_p$. An element of $L(G_1)$ is called homogeneous if it belongs to one of the direct summands, i.e. it is of the form xG_{i+1} for some $1 \leq i \leq k-1$, $x \in G_i$. Using Lemma 7.1.1, define the Lie bracket on homogeneous elements to be

$$[xG_{i+1}, yG_{j+1}] = \begin{cases} [x,y]G_{i+j+1} \in G_{i+j}/G_{i+j+1} & \text{if } i+j \leq k-1 \\ \\ 0 & \text{otherwise} \end{cases},$$

where $[x,y] = xyx^{-1}y^{-1}$ is the group commutator, and extend the bracket to all of $L(G_1)$ by linearity.

We note that $G(\mathbb{F}) \cong G(R)/G_1$ acts on G_i/G_{i+1} by conjugation. Indeed, $G(R)$ acts on G_i/G_{i+1} since $G_i, G_{i+1} \trianglelefteq G(R)$, and if $x \in G_1$, $y \in G_i$, then $xyx^{-1} \in yG_{i+1}$ by Lemma 7.1.1, so G_1 acts as the identity on G_i/G_{i+1}.

Since $G(\mathbb{F})$ also acts on $\mathfrak{g}(\mathbb{F})$ by conjugation, this suggests these objects are related. Indeed, the following proposition shows that G_i/G_{i+1} can be treated as a copy of $\mathfrak{g}(\mathbb{F})$. We use without proof the fact that $|G_i/G_{i+1}| \leq |\mathbb{F}|^{\dim(G)}$.

Proposition 7.1.2. *For $1 \leq i \leq k-1$, the map $\varphi : G_i/G_{i+1} \to \mathfrak{gl}_d(\mathbb{F})$ given by*

$$I_d + \pi^i A \mapsto A \bmod \pi$$

induces an isomorphism of elementary abelian p-groups $G_i/G_{i+1} \cong \mathfrak{g}(\mathbb{F})$*, and this isomorphism is equivariant with respect to the action of* $G(\mathbb{F})$ *on both sides by conjugation.*

Proof. Begin with $i = 1$. Then

$$G_1/G_2 = \ker(G(\mathbb{F}[\pi]/\pi^2) \to G(\mathbb{F})).$$

The right hand side of the above equation is how $\mathrm{Lie}(G(\mathbb{F}))$ was defined in section 2.5, with ϵ in place of π. As discussed in that section, the map $I_d + \pi A \to A \bmod \pi$ produces an isomorphism with $\mathfrak{g}(\mathbb{F}) \subseteq \mathfrak{sl}_d(\mathbb{F})$.

Now assume $i > 1$. Since $\mathcal{O}/\pi^i \cong \mathbb{F}[\pi]/\pi^i$, we have

$$G_i/G_{i+1} = \ker(G(\mathbb{F}[\pi]/\pi^{i+1}) \to G(\mathbb{F}[\pi]/\pi^i)).$$

If $A \in G_i/G_{i+1}$, then then entries of A are in

$$(\mathbb{F} \oplus \pi^i\mathbb{F})/\pi^{i+1} = (\mathbb{F} \oplus \pi^i\mathbb{F})/\pi^{2i} = \mathbb{F}[\pi^i]/\pi^{2i},$$

so we have

$$\begin{aligned} G_i/G_{i+1} &= \ker(G(\mathbb{F}[\pi^i]/\pi^{2i}) \to G(\mathbb{F}[\pi]/\pi^i)) \\ &= \ker(G(\mathbb{F}[\pi^i]/\pi^{2i}) \to G(\mathbb{F})), \end{aligned}$$

since the image of $\mathbb{F}[\pi^i]/\pi^{2i}$ modulo π^i is just $\mathbb{F}$. Using $\epsilon = \pi^i$, the same reasoning as when $i = 1$ now applies.

Finally, for any $1 \leq i \leq k - 1$, $I_d + \pi A \in G_i$, and $g \in G(\mathbb{F})$, we have

$$g(I_d + \pi^i A)g^{-1} = I_d + \pi^i gAg^{-1},$$

so the isomorphism is equivariant under conjugation by $G(\mathbb{F})$. □

For any $\alpha \in \Phi$, $x_\alpha(\pi^i) \bmod \pi^{i+1} \in G_i/G_{i+1}$ and its image in $\mathfrak{g}(\mathbb{F})$ can be identified with the Chevalley basis element e_α, as noted in section 2.5.

By Proposition 7.1.2, if we let x be an indeterminate then we have a Lie algebra isomorphism

$$L(G_1) \cong \mathfrak{g}(\mathbb{F}) \otimes x\mathbb{F}[x]/x^k = \bigoplus_{i=1}^{k-1} x^i \mathfrak{g}(\mathbb{F}).$$

We now define certain Lie subalgebras of $L(G_1)$ arising from subgroups of $G(R)$. Let $H \leq G(R)$ and for $1 \leq i \leq k-1$, define

$$\mathfrak{h}_i = (H \cap G_i)G_{i+1}/G_{i+1} \cong (H \cap G_i)/(H \cap G_{i+1}),$$

which we view as an $\mathbb{F}_p$-subspace of $\mathfrak{g}(\mathbb{F})$. Then $[\mathfrak{h}_i, \mathfrak{h}_j]_{\mathbb{F}_p} \subseteq \mathfrak{h}_{i+j}$ if $i + j < k$, so

$$L(H) = \bigoplus_{i=1}^{k-1} (H \cap G_i)G_{i+1}/G_{i+1}$$

is a graded Lie subalgebra of $L(G_1)$. Observe that the group $H/(H \cap G_1) \cong G_1 H/G_1 \leq G(\mathbb{F})$ acts on each $\mathfrak{h}_i$ by conjugation.

Using the realization of $L(G_1)$ as $\mathfrak{g}(\mathbb{F}) \otimes x\mathbb{F}[x]/(x^k)$, we can write

$$L(H) = \bigoplus_{i=1}^{k-1} x^i \mathfrak{h}_i.$$

The dimension of $L(H)$ (as an $\mathbb{F}_p$-vector space) is naturally related to $|H \cap G_1|$, as seen in the following lemma.

Lemma 7.1.3. *Let $H \leq G(R)$ and consider the graded Lie algebras $L(G_1)$ and $L(H)$ defined above as vector spaces over $\mathbb{F}_p$. Then* $\dim(L(H)) = \log_p |H \cap G_1|$ *and* $\operatorname{codim}_{L(G_1)}(L(H)) = \log_p [G_1 : H \cap G_1]$.

Proof. The statement follows from a few computations. Note that $\dim(L(H)) = \sum_{i=1}^{k-1} \dim(\mathfrak{h}_i)$, so

$$|H \cap G_1| = \prod_{i=1}^{k-1} \frac{|H \cap G_i|}{|H \cap G_{i+1}|} = \prod_{i=1}^{k-1} p^{\dim(\mathfrak{h}_i)} = p^{\dim(L(H))},$$

proving the first claim.

Using $H = G_1$, it follows that $\dim(L(G_1)) = \log_p |G_1|$, so

$$\begin{aligned}
\log_p [G_1 : H \cap G_1] &= \log_p |G_1| - \log_p |H \cap G_1| \\
&= \dim(L(G_1)) - \dim(L(H)) \\
&= \operatorname{codim}_{L(G_1)} L(H).
\end{aligned}$$

□

7.2 Lower Bound Preliminaries

We now prove some lemmas that will enable us to use the graded Lie algebra constructed in the previous section to find lower bounds on the normal and non-normal residual finiteness growth of Chevalley groups.

We continue with the assumptions and notation of the previous section; in particular G is a simple simply connected Chevalley group with irreducible root system Φ. When bounding normal finiteness growth in the characteristic p case, we will want to guarantee $G_1 H \neq G(R)$ when H is a proper normal subgroup of $G(R)$. In almost all cases, $G(R)$ is perfect by Proposition 2.4.6, so the following lemma will apply.

Lemma 7.2.1. *Assume $G(R)$ is perfect and $H \trianglelefteq G(R)$. If $H \neq G(R)$, then $G_1 H \neq G(R)$.*

Proof. Recall that $R = \mathcal{O}/\mathfrak{m}^k$. For any $1 \leq i \leq k-1$, there is a natural surjective homomorphism

$$G_i/G_{i+1} \to G_i H/G_{i+1}H.$$

If $G_i H = G(R)$, then $G(R)/G_{i+1}H$ is the image of the abelian group G_i/G_{i+1}. Since $G(R)$ is perfect, $G(R)/G_{i+1}H$ must be trivial, so $G_i H = G_{i+1}H$.

In particular, if $G_1 H = G(R)$, then the above argument implies $G_k H = G(R)$. Since $G_k = 1$, we conclude that $H = G(R)$ if $G_1 H = G(R)$. □

We cannot apply the above lemma when $\mathcal{O} = \mathbb{F}_2[t]$ and Φ is type B_2 or G_2. We won't need this result in the G_2 case, so we now prove a similar result when Φ is type B_2. We write $G(R)'$ for the derived subgroup of $G(R)$.

Lemma 7.2.2. *Fix $k \geq 1$ and set $R = \mathbb{F}_2[t]/f(t)^k$ for some irreducible $f(t) \in \mathbb{F}_2[t]$. Let G be a simply connected Chevalley group of type B_2 and let H be a proper normal subgroup of $G(R)$. If $G(R)' \not\subseteq H$, then $G_1 H \neq G(R)$.*

Proof. Recall that the root system of type B_2 has roots $\{\pm\epsilon_1, \pm\epsilon_2, \pm(\epsilon_1 \pm \epsilon_2)\}$. We set up some notation to make the computations clearer. For $1 \leq j \leq k$, put $G(j) =$

$G(\mathbb{F}_2[t]/f(t)^j)$, and for $1 \leq i \leq j$, set $G(j)_i = \ker(G(j) \to G(i))$. We will continue writing G_i for $G(k)_i$. Note that $G(k) = G(R)$.

We first show $G_{k-1}H \neq G(k)$. If this is not the case, then

$$G(k)/H = G_{k-1}H/H \cong G_{k-1}/(H \cap G_{k-1})$$

is a nontrivial abelian quotient of $G(k)$, so $G(k)' \subseteq H$, a contradiction.

Recall that we can view $\mathfrak{h}_j = (H \cap G_j)G_{j+1}/G_{j+1}$ as a subspace of $\mathfrak{g}(\mathbb{F})$.

We now assume for the sake of contradiction that $G_1 H = G(k)$. Since $G_{k-1}H \neq G(k)$, there exists some $2 \leq j \leq k-1$ such that $G_{j-1}H = G(k)$ and $G_j H \neq G(k)$. Then

$$\mathfrak{g}(\mathbb{F})/\mathfrak{h}_{j-1} \cong G_{j-1}/(H \cap G_{j-1})G_j \cong G_{j-1}H/G_j H$$

is nontrivial, so $\mathfrak{h}_{j-1} \neq \mathfrak{g}(\mathbb{F})$.

Put $H(j) = G_j H/G_j$ and observe $H(j)$ is properly contained in $G(j)$. Since $G_{j-1}H = G(k)$, we have $G(j)_{j-1}H(j) = G(j)$, so $G(j)' \subseteq H(j)$. Hence

$$x_{\epsilon_1}(\pm f(t)^{j-1})x_{\epsilon_2+\epsilon_1}(\pm f(t)^{j-1}) = [x_{\epsilon_2}(1), x_{\epsilon_1-\epsilon_2}(f(t)^{j-1})] \in H(j) \cap G(j)_{j-1},$$

so $e_{\epsilon_1} + e_{\epsilon_2+\epsilon_1} \in \mathfrak{h}_{j-1}$.

The subspace $\mathfrak{h}_{j-1}$ is invariant under the action of $G_1H/G_1 = G(\mathbb{F})$ and is proper in $\mathfrak{f}(\mathbb{F})$, so $\mathbb{F}\mathfrak{h}_{j-1}$ is a proper ideal of $\mathfrak{g}(\mathbb{F})$ by Lemma 2.6.2. Then by Proposition 2.6.1, $\mathbb{F}\mathfrak{h}_{j-1}$ is the center of $\mathfrak{g}(\mathbb{F})$ or contains $\mathbb{F}e_\alpha$ for each short root α. Clearly $\mathbb{F}\mathfrak{h}_{j-1}$ is not the center, so it contains e_{ϵ_1} and thus also contains $e_{\epsilon_2+\epsilon_1}$. This then forces $\mathbb{F}\mathfrak{h}_{j-1}$ to be all of $\mathfrak{g}(\mathbb{F})$, a contradiction. □

If $g \in \mathcal{O}$, we write $G(\mathcal{O}, g) = \ker(G(\mathcal{O}) \to G(\mathcal{O}/g))$, and we denote the gcd of π and g as (π, g). We call $G(\mathcal{O}, g)$ a principal congruence subgroup; any subgroup of $G(\mathcal{O})$ containing a principal congruence subgroup is called a congruence subgroup.

If the rank of G is at least 2, then $G(\mathcal{O})$ has the congruence subgroup property: every finite index subgroup of $G(\mathcal{O})$ is a congruence subgroup (see Chapter 9 of [23] for details). We note that it is necessary that G be simply connected for this to be true.

Using the congruence subgroup property we will be able to reduce to the case of considering principal congruence subgroups. The next two statements will help us work with their images in $G(R)$.

Lemma 7.2.3. *Let $R = \mathcal{O}/\pi^k$ for some irreducible $\pi \in \mathcal{O}$ and set $\Delta = G(\mathcal{O}, g)$ for some $g \in \mathcal{O}$. Let $\overline{\Delta}$ be the image of Δ in $G(R)$.*

1. *If $(\pi, g) = 1$, then $\overline{\Delta} = G(R)$.*

2. *If $(\pi^k, g) = \pi^s$ with $s < k$, then $(\overline{\Delta} \cap G_i)G_{i+1}/G_{i+1} = \mathfrak{g}(\mathbb{F})$ for $s \leq i \leq k-1$.*

Proof. First assume $(\pi, g) = 1$. Then for any $f \in \mathcal{O}$, there exist $h_1, h_2 \in \mathcal{O}$ such that $h_1\pi^k + h_2 g = f$. Thus if $\alpha \in \Phi$,

$$x_\alpha(h_2 g) = x_\alpha(f)x_\alpha(-h_1\pi^k) \in \Delta,$$

so $x_\alpha(f \bmod \pi^k) \in \overline{\Delta}$. By Lemma 2.5.3, $G(R)$ is generated by $\{x_\alpha(f \bmod \pi^k) : f \in \mathcal{O}\}$, so $\overline{\Delta} = G(R)$.

Now assume $(\pi^k, g) = \pi^s$ with $s < k$. For notational convenience, write

$$\mathfrak{d}_i = (\overline{\Delta} \cap G_i)G_{i+1}/G_{i+1}.$$

Using similar reasoning as above, for any $\alpha \in \Phi$ and any $f \in \mathcal{O}$, $x_\alpha(\pi^s f \bmod \pi^k) \in \overline{\Delta}$. Hence for $s \leq i \leq k-1$,

$$\{x_\alpha(\pi^i f \bmod \pi^k) : \alpha \in \Phi, f \in \mathcal{O}\} \subseteq \overline{\Delta} \cap G_i,$$

so $\mathbb{F}e_\alpha \subseteq \mathfrak{d}_i$ for all $\alpha \in \Phi$.

Since the projection $G(\mathcal{O}) \to G(R)$ is surjective, $\overline{\Delta} \trianglelefteq G(R)$, so $G(R)/G_1 \cong G(\mathbb{F})$ acts on $\mathfrak{d}_i$. For $\alpha \in \Phi$, $t \in \mathbb{F}$,

$$x_\alpha(t)e_{-\alpha}x_\alpha(-t) = e_{-\alpha} + th_\alpha - t^2 e_\alpha.$$

Since $\mathbb{F}e_{-\alpha} \oplus \mathbb{F}e_\alpha \subseteq \mathfrak{d}_i$, we must have $th_\alpha \in \mathfrak{d}_i$. Since this is true for all $\alpha \in \Phi$ and $t \in \mathbb{F}$, in fact $\mathfrak{d}_i = \mathfrak{g}(\mathbb{F})$. □

Corollary 7.2.4. *With the same setup as in Lemma* 7.2.3*, let $H \leq \overline{\Delta}$ and fix $\alpha \in \Phi$, a short root if Φ is type C_l. Assume $(\pi^k, g) = \pi^s$ and $\mathbb{F}e_\alpha \not\subseteq \mathfrak{h}_j$ for some $1 \leq j \leq k-1$ such that $s < j/2$. If $s = 0$, then*

$$\operatorname{codim}_{L(\overline{\Delta})} L(H) \geq [\mathbb{F} : \mathbb{F}_p](j-1).$$

If $s \geq 1$, then

$$\operatorname{codim}_{L(\overline{\Delta})} L(H) \geq [\mathbb{F} : \mathbb{F}_p](j-2s+1).$$

Proof. Since $\mathbb{F}e_\alpha \not\subseteq \mathfrak{h}_j$ and $[\mathfrak{h}_i, \mathfrak{h}_{j-i}] \subseteq \mathfrak{h}_j$ for $1 \le i \le j-1$, we have $\mathbb{F}e_\alpha \not\subseteq [\mathfrak{h}_i, \mathfrak{h}_{j-i}]$. Put $\mathfrak{d}_i = (\overline{\Delta} \cap G_i)G_{i+1}/G_{i+1}$.

If $s = 0$, then $\overline{\Delta} = G(R)$ by Lemma 7.2.3, so $\mathfrak{d}_i = \mathfrak{g}(\mathbb{F})$ for $1 \le i \le k-1$. Then Corollary 4.1.4 implies

$$\operatorname{codim}_{\mathfrak{d}_i}(\mathfrak{h}_i) + \operatorname{codim}_{\mathfrak{d}_{j-i}}(\mathfrak{h}_{j-i}) \ge 2[\mathbb{F} : \mathbb{F}_p]$$

for $1 \le i \le j-1$. Hence

$$\operatorname{codim}_{L(\overline{\Delta})} L(H) \ge [\mathbb{F} : \mathbb{F}_p](j-1).$$

If $s \ge 1$, Lemma 7.2.3 gives that $\mathfrak{d}_i = \mathfrak{g}(\mathbb{F})$ for $s \le i \le k-1$. There are $j - 2s + 1$ integers in the interval $[s, j-s]$, so the previous reasoning yields the desired inequality. □

7.3 Lower Bounds In Characteristic 0

We continue with the notation of the previous section, with G remaining a simple simply connected Chevalley group with a fixed embedding into SL_d. Fix $\alpha \in \Phi$, a short root if G is type C_l. We first provide lower bounds for the normal and non-normal residual finiteness growth of $G(\mathbb{Z})$.

Lemma 7.3.1. *Let $R = \mathbb{Z}/p^k$ for a prime p, $k \ge 1$. Let $\Delta = G(\mathbb{Z}, N)$, let $\overline{\Delta}$ be the image of Δ in $G(R)$, and assume $(p^k, N) = p^s$. Let $r \ge N$ be sufficiently large and* *set*

$$L_r = (\mathrm{lcm}(1, 2, \cdots, r))^{3(\dim(G)+s)},$$

$$M_r = x_\alpha(L_r \bmod p^k).$$

If $H \leq \overline{\Delta}$ *and* $M_r \notin H$*, then* $[\overline{\Delta} : H] \geq \dfrac{1}{2} r^{a(G)}$*. If in addition* $H \trianglelefteq \overline{\Delta}$*, then* $[\overline{\Delta} : H] \geq \dfrac{1}{2d} r^{\dim(G)}$.

Proof. Let M_r, H be as in the statement and suppose $p^{m-1} || L_r$, by which we mean p^{m-1} divides L_r and p^m does not divide L_r. We have $m \leq k$ since $M_r \neq 1$, and $M_r \in \overline{\Delta}$ since $r \geq N$ implies $N | L_r$. The proof splits into three cases; we will consider H as an arbitrary subgroup and as a normal subgroup in each case.

Case 1: $k = 1$. Since $k = 1$, we have $R \cong \mathbb{F}_p$ and $p > r \geq N$, so $(p, N) = 1$. By Lemma 7.2.3, $\overline{\Delta} = G(\mathbb{F}_p)$, so H is a proper subgroup of $G(\mathbb{F}_p)$. Then by Lemma 2.5.5, $[G(\mathbb{F}_p) : H] \geq \frac{1}{2} p^{a(G)}$. Since M_r is nontrivial, p does not divide L_r, so by construction of L_r, $p > r$. Hence $[G(\mathbb{F}_p) : H] \geq \dfrac{1}{2} r^{a(G)}$, as desired.

If in addition H is normal, then $H \subseteq Z(G(\mathbb{F}_p))$ since $G(\mathbb{F}_p)/Z(G(\mathbb{F}_p))$ is simple. Thus by Lemma 2.5.5,

$$[G(\mathbb{F}_p) : H] \geq |G(\mathbb{F}_p)/Z(G(\mathbb{F}_p))| \geq \frac{1}{2d} p^{\dim(G)} > \frac{1}{2d} r^{\dim(G)}.$$

Case 2: $k \geq 2$, $m = 1$. Let G_1 be the kernel of the projection $G(R) \to G(\mathbb{F}_p)$, and recall the graded Lie algebras $L(G_1)$ and $L(H)$ defined in section 7.1. Since $m = 1$, p does not divide L_r, so again $p > r$ and $\overline{\Delta} = G(\mathbb{F}_p)$. We also have $M_r \notin G_1$. If in addition $G_1H \neq G(R)$, then the image of H in $G(R)/G_1 \cong G(\mathbb{F}_p)$ is proper, so

$$[G(R) : H] \geq \frac{1}{2}p^{a(G)} > \frac{1}{2}r^{a(G)}.$$

If H is normal, then by the same reasoning as before we see that $[G(R) : H] > \frac{1}{2d}r^{\dim(G)}$.

If $G_1H = G(R)$, then since $p > r$ is large, $G_1H/G_1 \cong G(\mathbb{F}_p)$ acts irreducibly on $\mathfrak{g}(\mathbb{F}_p)$ by Proposition 2.6.1, so for each j, $\mathfrak{h}_j$ is trivial or $\mathfrak{h}_j = \mathfrak{g}(\mathbb{F}_p)$. If all are $\mathfrak{g}(\mathbb{F}_p)$, this forces $H = G(R)$, contradicting $M_r \notin H$. Thus $\mathfrak{h}_j$ is trivial for some j and

$$\operatorname{codim}_{L(G_1)} L(H) \geq \operatorname{codim} \mathfrak{h}_j = \dim(G),$$

so $[G(R) : H] \geq p^{\dim(G)} > r^{\dim(G)}$.

Case 3: $k \geq 2$, $m \geq 2$. Since $M_r \notin H$ and $p^{m-1} || L_r$, we have $M_r \in G_{m-1} \setminus G_m$, so $\mathbb{F}_p e_\alpha \not\subseteq \mathfrak{h}_{m-1}$. If $p^l || \operatorname{lcm}(1, \cdots, r)$, then

$$m - 1 = 3(\dim(G) + s)l \text{ and } p^{(l+1)\dim(G)} > r^{\dim(G)}.$$

In particular, $s < j/2$, so by Corollary 7.2.4, if $s \geq 1$ then

$$\operatorname{codim}_{L(\overline{\Delta})}(L(H)) \geq m - 2s.$$

Since

$$m - 2s = 3(\dim(G) + s)l - 2s + 1 \geq \dim(G)(l + 1),$$

we conclude that

$$[\overline{\Delta} : H] \geq [\overline{\Delta} \cap G_1 : H \cap G_1] \geq p^{\dim(G)(l+1)} > r^{\dim(G)}.$$

A similar argument works when $s = 0$, using the corresponding inequality from Corollary 7.2.4. □

Theorem 7.3.2. *Let G be a simple Chevalley group of type Φ of rank at least 2, not necessarily simply connected, and let Δ be a finite index subgroup of $G(\mathbb{Z})$. Then $F^{\trianglelefteq}_{\Delta}(n) \succeq n^{\dim(G)}$ and $F^{\leq}_{\Delta}(n) \succeq n^{a(G)}$.*

Proof. Let G^{sc}_{Φ} be the simply connected Chevalley group of type Φ and let $\rho : G^{sc}_{\Phi}(\mathbb{Z}) \to G(\mathbb{Z})$ be the natural map; it is surjective and has finite kernel. Then $\rho^{-1}(\Delta)$ has finite index in $G^{sc}_{\Phi}(\mathbb{Z})$ and the map $\rho^{-1}(\Delta) \to \Delta$ is a surjection with finite kernel, so by Lemma 5.3.1, the residual finiteness growth of Δ is bounded below by that of $\rho^{-1}(\Delta)$. Thus we may assume from the start that G is simply connected.

Then $G(\mathbb{Z})$ satisfies the congruence subgroup property, so Δ contains some principal congruence subgroup. Since residual finiteness growth can only decrease by passing to a subgroup, we may assume $\Delta = G(\mathbb{Z}, N)$ for some $N \in \mathbb{Z}$. Let s be the largest power of a prime dividing N.

Fix $r \geq N$ sufficiently large and put $L_r = (\mathrm{lcm}(1, 2, \cdots, r))^{3(\dim(G)+s)}$. Fix some $\alpha \in \Phi$, a short root if G is type C_l. We show that $M_r = x_\alpha(L_r)$ is in every subgroup of $G(R)$ of sufficiently small index. First we need to determine the word length of M_r in Δ.

By Theorem A in [19], there exists a generating set X of $G(R)$ so that

$$||M_r||_X \leq C_1 \log |L_r|$$

for some $C_1 > 0$. By the prime number theorem, $\mathrm{lcm}(1, \cdots, r) \sim e^r$, so

$$\log |L_r| \leq C_2(\dim(G) + s)r$$

for some absolute constant C_2. Since Δ has finite index in $G(\mathbb{Z})$, we conclude that

$$||M_r||_Y \leq Cr$$

for some generating set Y of Δ and some constant C independent of r.

Now suppose $M_r \notin H \leq \Delta$. By the congruence subgroup property of $G(\mathbb{Z})$, $H \supseteq G(\mathbb{Z}, N')$ for some $N' \in \mathbb{Z}$. Set $R = \mathbb{Z}/N'$ and let $N' = \prod_{i=1}^{k} p_i^{k_i}$ be the prime factorization of N'. Write $G_{(i)} = G(\mathbb{Z}/p_i^{k_i})$ for each i. Then by the Chinese Remainder Theorem,

$$G(R) \cong \prod_{i=1}^{k} G_{(i)}.$$

Let $\pi_{N'}$ be the natural projection $G(\mathbb{Z}) \to G(R)$. Then $\pi_{N'}(M_r) \notin \pi_{N'}(H)$, so in some $G_{(i)}$, $\overline{M_r} = x_\alpha(L_R \bmod p_i^{k_i}) \notin \overline{H}$, where $\overline{M_r}$ and $\overline{H}$ are the images of M_r and H in $G_{(i)}$, respectively. So by Lemma 7.3.1,

$$[\Delta : H] \geq [\overline{\Delta} : \overline{H}] \geq \frac{1}{2} r^{a(G)},$$

and $[\Delta : H] \geq \dfrac{1}{2d} r^{\dim(G)}$ if $H \trianglelefteq \Delta$. Recalling that M_r has word length $n \leq Cr$ finishes the argument. □

7.4 Lower Bounds In Characteristic p

We continue using the same setup as in the previous section but now deal with the groups $G(\mathbb{F}_p[t]) \subseteq \mathrm{SL}_d(\mathbb{F}_p[t])$. Let $\alpha \in \Phi$ be a short root if G is of type C_l. Recall that the root system of type B_2 has roots $\{\pm\epsilon_1, \pm\epsilon_2, \pm(\epsilon_1 \pm \epsilon_2)\}$.

Lemma 7.4.1. *Let $R = \mathbb{F}_p[t]/f(t)^k$ for an irreducible polynomial $f(t)$, $k \geq 1$. Let $\Delta = \ker(G(\mathbb{F}_p[t]), g(t))$, let $\overline{\Delta}$ be the image of Δ in $G(R)$, and assume $(f(t)^k, g(t)) = f(t)^s$. Fix $r \geq \deg(g(t))$, and set*

$$L_r(t) = (\mathrm{lcm}\{h(t) \in \mathbb{F}_p[t] : \deg(h(t)) \leq r\})^{3(\dim(G)+s)}.$$

If $p = 2$ and G is of type B_2, let

$$M_r = x_{\epsilon_1}(L_r(t) \bmod f(t)^k) x_{\epsilon_1+\epsilon_2}(L_r(t) \bmod f(t)^k),$$

and otherwise set

$$M_r = x_\alpha(L_r(t) \bmod f(t)^k).$$

If $H \leq \overline{\Delta}$ and $M_r \notin H$, then $[\overline{\Delta} : H] \geq \frac{1}{2}p^{ra(G)}$. If in addition $H \trianglelefteq \overline{\Delta}$, then $[\overline{\Delta} : H] \geq \frac{1}{2d}p^{r\dim(G)}$.

Proof. Let M_r, H be as in the statement, put $q = p^{\deg(f(t))}$, and suppose $f(t)^{m-1}||L_r(t)$, where $m \leq k$ since $M_r \neq 1$. Observe that $M_r \in \overline{\Delta}$ since $g(t)|L_r(t)$. The argument splits into a few cases. As in the proof of Lemma 7.3.1, we will treat H as an arbitrary subgroup and then as a normal subgroup in each case. The arguments are similar to the characteristic 0 case, so details will sometimes be skipped.

Case 1: $k = 1$. Since $k = 1$, we have $R \cong \mathbb{F}_q$ and $\deg(f(t)) > r \geq \deg(g(t))$, so $f(t)$ and $g(t)$ are relatively prime. Then $\overline{\Delta} = G(\mathbb{F}_q)$ by Lemma 7.2.3, so H is a proper subgroup of $G(\mathbb{F}_q)$. By Lemma 2.5.5, $[G(\mathbb{F}_q) : H] \geq \frac{1}{2}q^{a(G)}$. Hence

$$[\overline{\Delta} : H] = [G(\mathbb{F}_q) : H] \geq \frac{1}{2}p^{ra(G)}.$$

If H is normal, then $H \subseteq Z(G(\mathbb{F}_q))$ since $G(\mathbb{F}_q)/Z(G(\mathbb{F}_q))$ is simple, so Lemma 2.5.5 gives

$$[\overline{\Delta} : H] \geq \frac{1}{2d}q^{\dim(G)} > \frac{1}{2d}p^{r\dim(G)}.$$

Case 2: $k \geq 2, m = 1$. Since $m = 1$, we again have $\deg(f(t)) > r$ and $\overline{\Delta} = G(R)$. Let G_1 be the kernel of the projection $G(R) \to G(\mathbb{F}_q)$, and define graded Lie algebras $L(G_1)$ and $L(H)$ as in section 7.1.

We first consider the case $H \trianglelefteq G(R)$. If $G(R)$ is perfect, then by Lemma 7.2.1, $G_1H \neq G(R)$. Hence the image of H in $G(\mathbb{F}_q)$ is proper and

$$[\overline{\Delta} : H] \geq \frac{1}{2d}p^{r\dim(G)}$$

as before. Otherwise, by Proposition 2.4.6 $p = 2$ and G is of type B_2 or G_2. In the former case,

$$M_r = [x_{\epsilon_1}(1), x_{\epsilon_1-\epsilon_2}(L_r(t) \bmod f(t)^k)] \in G(R)',$$

so $G(R)' \nsubseteq H$ and thus $G_1H \neq G(R)$ by Lemma 7.2.2, yielding the desired bound as shown above. If G is type G_2 and $G_1H = G(R)$, then $G_1H/G_1 \cong G(\mathbb{F}_q)$ acts irreducibly on each $\mathfrak{h}_i$ by Proposition 2.6.1 and Lemma 2.6.2. Hence some $\mathfrak{h}_i$ is trivial, so

$$\operatorname{codim}_{L(G_1)} L(H) \geq \dim(G) \deg(f(t))$$

and

$$[\overline{\Delta} : H] \geq p^{\dim(G) \deg(f(t))} \geq p^{r \dim(G)}.$$

If H is an arbitrary subgroup of $G(R)$, then the case $G_1 H \neq G(R)$ again reduces to a previous argument. So assume $G_1 H = G(R)$. Then $G_1 H / G_1 \cong G(\mathbb{F}_q)$ acts on each $\mathfrak{h}_i$, so for each i, either $\mathfrak{h}_i = \mathfrak{g}(\mathbb{F})$ or $\mathbb{F}\mathfrak{h}_i$ is a proper ideal, using Lemma 2.6.2. Since $\mathfrak{h}_i \subseteq \mathbb{F}\mathfrak{h}_i$, by examining Table 2.1 and Table 1.1 we see that each $\mathfrak{h}_\mathfrak{i}$ is all of $\mathfrak{g}(\mathbb{F}_q)$ or has codimension at least $a(G)\deg(f(t))$. Since H is proper, not all the $\mathfrak{h}_\mathfrak{i}$ can be $\mathfrak{g}(\mathbb{F}_q)$, so

$$\operatorname{codim}_{L(G_1)} L(H) \geq a(G) \deg(f(t)) > ra(G).$$

Thus $[G(R) : H] \geq p^{ra(G)}$.

Case 3: $k \geq 2, m \geq 2$. We handle H being normal and arbitrary simultaneously. Since $M_r \notin H$ and $f(t)^{m-1} || L_r(t)$, we have $M_r \in G_{m-1} \setminus G_m$, so $\mathbb{F}_q e_\alpha \not\subseteq \mathfrak{h}_j$ for some $m - 1 \leq j \leq k - 1$ ($\mathbb{F}_q(e_{\epsilon_1} + e_{\epsilon_1+\epsilon_2}) \not\subseteq \mathfrak{h}_j$ if G is type B_2, $p = 2$).

By the construction of $L_r(t)$, $f(t)^{m-1} || L_r(t)$ implies $m - 1 = 3(\dim(G) + s)l$ for some integer $l \geq 1$ satisfying $\deg(f(t))(l + 1) > r$. In particular, $s < j/2$, so by Corollary 7.2.4, if $s \geq 1$ then

$$\begin{aligned} \operatorname{codim}_{L(\overline{\Delta})}(L(H)) &\geq \deg(f(t))(j - 2s + 1) \\ &\geq \deg(f(t))(m - 2s). \end{aligned}$$

We have

$$m - 2s = 3(\dim(G) + s)l - 2s + 1 \geq \dim(G)(l + 1),$$

so

$$\deg(f(t))(m - 2s) \geq \dim(G)\deg(f(t))(l + 1) > r\dim(G),$$

and hence

$$[\overline{\Delta} : H] \geq [\overline{\Delta} \cap G_1 : H \cap G_1] \geq p^{r\dim(G)}.$$

A similar argument works when $s = 0$, using the corresponding inequality from Corollary 7.2.4.

We note that while Corollary 7.2.4 does not directly apply in the case G is of type B_2, $p = 2$, the same arguments in Corollary 4.1.4 work when using $e_{\epsilon_1} + e_{\epsilon_1+\epsilon_2}$ in place of e_α because

$$e_{\epsilon_1} + e_{\epsilon_1+\epsilon_2} = \pm[e_{\epsilon_1} + e_{\epsilon_2}, e_{\epsilon_1-\epsilon_2} + e_{\epsilon_2-\epsilon_1}]. \qquad \square$$

We can now prove the positive characteristic analogue of Theorem 7.3.2.

Theorem 7.4.2. *Let G be a simple Chevalley group, not necessarily simply connected, of rank at least 2, let p be a prime, and let Δ be a finite index subgroup of $G(\mathbb{F}_p[t])$. Then $F^{\trianglelefteq}_{\Delta}(n) \succeq n^{\dim(G)}$ and $F^{\leq}_{\Delta}(n) \succeq n^{a(G)}$.*

Proof. As in the proof of Theorem 7.3.2, we may assume G is simply connected and $\Delta = G(\mathbb{F}_p[t], g(t))$ for some $g(t) \in \mathbb{F}_p[t]$. Let s be the largest power of an irreducible polynomial dividing $g(t)$.

Fix $r \geq \deg(g(t))$ and set

$$L_r(t) = (\operatorname{lcm}\{h(t) : \deg(h(t)) \leq r\})^{3(\dim(G)+s)}.$$

Let Φ be the root system of G, and let $\alpha \in \Phi$, with the extra condition that α is a short root if Φ is of type $C_l, l \geq 2$. Set

$$M_r = \begin{cases} x_{\epsilon_1}(L_r(t))x_{\epsilon_1+\epsilon_2}(L_r(t)) & \text{if } \Phi = B_2, p = 2 \\ x_\alpha(L_r(t)) & \text{otherwise} \end{cases}.$$

By Theorem A in [19], there exists a generating set X of $G(\mathbb{F}_p[t])$ so that

$$||M_r||_X \leq C_1 \deg(L_r(t))$$

for some constant C_1. The degree of $\operatorname{lcm}\{h(t) \in \mathbb{F}_p[t] : \deg(h(t)) \leq r\}$ is at most $2p^r$, so $\deg(L_r(t)) \leq 6(\dim(G) + s)p^r$. Hence $||M_r||_X \leq C_2 p^r$ for some constant C_2. Since Δ has finite index in $G(\mathbb{F}_p[t])$, we conclude that M_r has word length $n \leq Cp^r$ for some constant C with respect to some generating set of Δ.

The remaining argument is the same as in the proof of Theorem 7.3.2. Substituting Lemma 7.4.1 for Lemma 7.3.1, one shows that if $M_r \notin H \leq \Delta$, then $[\Delta : H] \geq \dfrac{1}{2}p^{ra(G)}$, and if H is normal then $[\Delta : H] \geq \dfrac{1}{2d}p^{r\dim(G)}$. Then M_r having word length at most Cp^r implies $F^{\trianglelefteq}_{\Delta}(n) \succeq n^{\dim(G)}$ and $F^{\leq}_{\Delta}(n) \succeq n^{a(G)}$. □

Bibliography

[1] E. Abe and K. Suzuki, *On normal subgroups of Chevalley groups over commutative rings,* Tohoku Math. J. *28* (1976) no. 1, 185-198.

[2] Y. Barnea and R. Guralnick, *Subgroup growth in some pro-p groups*, Proceedings of the AMS. **130** (2001), 653-659.

[3] K. Bou-Rabee, *Quantifying residual finiteness*, J. of Algebra **323** (2010), 729-737.

[4] K. Bou-Rabee, *Approximating a group by its solvable quotients*, N.Y.J. of Math **17** (2011), 699-712.

[5] Bou-Rabee, Hagen, Patel, *Residual finiteness growths of virtually special groups*, Math. Z. **279** (2015) no. 1-2, 297-310.

[6] K. Bou-Rabee and T. Kaletha, *Quantifying residual finiteness of arithmetic groups*, Compos. Math. **148** (2012), 907-920.

[7] K. Bou-Rabee and D.B. McReynolds, *Asymptotic growth and least common multiples in groups*, Bull. Lond. Math. Soc. **43** (2011), 1059-1068.

[8] K. Bou-Rabee and D.B. McReynolds, *Extremal behavior of divisibility functions*, Geometriae Dedicata, **175** (2015), 407-415.

[9] N. Buskin, *Economical separability in free groups*, Sib. Math. J. **50** (2009) no. 4, 603-608.

[10] R. Carter, *Simple Groups of Lie Type*, Pure Appl. Math., vol 28, Wiley, London (1972).

[11] G. M. D. Hogeweij, *Almost classical Lie algebras: I, II,* Indag. Math. **44** (1982) no. 4, 441-460.

[12] M. Fried and M. Jarden, *Field Arithmetic*, Springer-Verlag, Berlin, (1986).

[13] R. Grigorchuk, *Degrees of growth of finitely generated groups and the theory of invariant means*, (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **48** (1984), 939-985.

[14] J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer, New York (1972).

[15] M. Kassabov and F. Matucci, *Bounding the residual finiteness of free groups,* Proc. Am. Math. Soc. **139** (2011), 2281-2286.

[16] P. Kleidman and M. Liebeck, *The subgroup structure of the finite classical groups*, Cambridge University Press, 1990.

[17] G. Kozma and A. Thom, *Divisibility and laws in finite simple groups*, Math. Ann. **364** (2016) no. 1-2, 79-95.

[18] S. Lang, *Algebraic Number Theory*, Springer-Verlag, 1970.

[19] Lubotzky, Mozes, Raghunathan, *The word and Riemannian metrics of semisimple groups*, Publ. Math. Inst. Hautes Etudes Sci. **91** (2000), 5-53.

[20] A. Lubotzky and D. Segal, *Subgroup Growth*, Progress in Mathematics, 212. Birkhauser Verlag, Basel, 2003.

[21] J. Milne, *Lie Algebras, Algebraic Groups, and Lie Groups*, 2013. (available at www.jmilne.org/math/CourseNotes/)

[22] V. Murty and J. Scherk, *Effective versions of the Chebotarev density theorem for function fields,* C. R. Acad. Sci. Paris Sér. I Math. **319** (1994), 523-528.

[23] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory.* Translated from the 1991 Russian original by Rachel Rowen. Pure and Applied Mathematics, 139. Academic Press, Inc., Boston, MA, 1994.

[24] S. Roman, *Field Theory*, Springer-Verlag, 1995.

[25] J.-P. Serre, *Lectures on $N_x(p)$*, Research Notes in Mathematics 11, CRC Press, 2012.

[26] R. Steinberg, *Lectures on Chevalley groups,* Yale University, 1968.

[27] A. Thom, *About the length of laws for finite groups*, http://arxiv.org/abs/1508.07730.

[28] A. V. Vasilyev, *Minimal permutation representations of finite exceptional groups of types G_2 and F_4*, Algebra and Logic. **35** (1996) no. 6, 371-383.

[29] A. V. Vasilyev, *Minimal permutation representations of finite exceptional groups of types E_6, E_7, and E_8*, Algebra and Logic. **36** (1997) no. 5, 302-310.