THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Invariant theory of maximal unipotent subgroups of reductive Groups

Stefano Giannini

Supervisor: Masoud Kamgarpour

Co-Supervisor: Sam Jeralds

Bachelor of Mathematics (Honours)

June 2022

The University of Queensland

School of Mathematics and Physics

Acknowledgements

I owe my gratitude to many people throughout my honours journey.
Firstly, I would like to thank my supervisor Masoud Kamgarpour for giving me the opportunity to study under him during my honours year. It has been an honour to work under him and I am truly grateful for his patience and guidance; his knowledge and passion for mathematics has inspired me to work hard.

I would also like to thank my co-supervisor Sam Jeralds who has taught me a great deal and for always making himself available to answer all my questions. I owe a special thanks to Matthew Spong who helped me tremendously throughout the beginning of my project.

I am grateful to my friends and colleagues at The University of Queensland who made this experience more enjoyable. I must express some gratitude to my friend Toan Pham for the many helpful conversations we shared and for teaching me to be a better student of mathematics.

Finally, I am profoundly grateful to my family for their constant patience, love and support.

Contents

1. Introduction

Background: In the early 20th century, Hilbert posed a series of 23 mathematical problems to the International Congress of Mathematicians. This series of open problems is generally regarded as the most influential set of mathematical problems stated by a mathematician. Hilbert's 14th problem concerns itself with finitely generated polynomial invariants. An important sub-problem is the description of the generators of the ring of invariants under the action of a linear algebraic group.

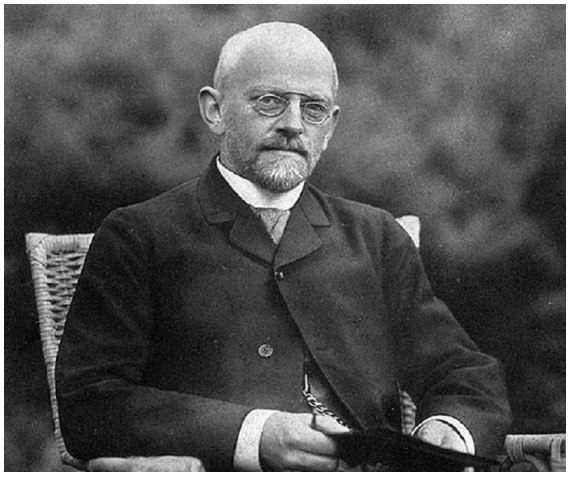


FIGURE 1. David Hilbert

For example, consider the action of the complex general linear group $\mathrm{GL}_n(\mathbb{C})$ on itself by conjugation. Then there is an induced action on the algebra of functions $\mathbb{C}[\mathrm{GL}_n]$. One observes that the trace and determinant polynomials are *invariant* under this action. That is, they remain unchanged under the conjugation action since $\mathrm{tr}(g^{-1}hg) = \mathrm{tr}(h)$ and $\det(g^{-1}hg) = \det(h)$ for all $h, g \in \mathrm{GL}_n(\mathbb{C})$. If $n = 2$, the polynomials tr and det generate the invariant ring while, as we shall see, if $n > 2$ more generators are needed.

In classical invariant theory, an important problem is to determine whether the subalgebra of invariant polynomials,

$$k[G]^G = \{f \in k[G] \,|\, g.f = f\},$$

is finitely generated over the field k. In other words, to determine whether every invariant polynomial can be described as a sum or product of elements from a finite set of invariant polynomials $(f_1, \ldots, f_r)$, called the *generators*.

In the case of Hilbert, he considered a linear algebraic group G acting on a finite dimensional vector space V over a field k [1]. Hilbert asked whether the algebra of invariant polynomials $k[V]^G$ is always finitely generated. Today, the problem has been generalised and reformulated as follows:

Hilbert's 14th Problem: *Let k be a field and K a subfield of the field of rational function $k(x_1, \ldots, x_n)$ in n variables. Is the algebra $K \cap k[x_1, \ldots, x_n]$ finitely generated over k?*

This more general statement does not necessarily deal with invariant polynomials directly. However, if one considers the case where K is the subfield of invariant functions in n variables (under some action of a group G) then $K \cap k[x_1, \ldots, x_n] = k[x_1, \ldots, x_n]^G$, and we retrieve the problem of classical invariant theory.

In general, the study of invariant objects appears in a wide range of areas of mathematics such as topology, geometry and algebra. It has found applications in taking quotients of varieties in algebraic geometry [2], in the description of Betti numbers (numbers which encode topological information about a space) for compact groups [3] and in representation

theory [4]. We will focus on the problem of understanding invariants of polynomials rings under the action of a group.

Classical invariant theorists studied this problem for certain actions of $\mathrm{SL}_n(\mathbb{C})$ on binary forms (i.e., homogeneous polynomials in two variables). For instance, Gordan, in 1868 solved this problem of finite generation of invariants for $\mathrm{SL}_2(\mathbb{C})$, then Hilbert solved this for arbitrary n [5],[1]. In fact, Hilbert's proof shows that $\mathbb{C}[G]^G$ is finitely generated for the broader class of *reductive groups* over $\mathbb{C}$; this result today is known as *Hilbert's finiteness theorem*. Further results by Hurwitz and Maurer led Hilbert to his 14th problem [6],[7].

As indicated above, when Hilbert's 14th problem was posed, some cases (e.g. $G = \mathrm{SL}_n(\mathbb{C})$) had already been solved. The situation was promising as there had been some progress in the "positive" direction. In 1962, Nagata was first to find a counterexample for this conjecture deeming it false in general [8]. Nagata constructed an action of k^3 on k^{32} whose algebra of invariant functions is not finitely generated (here k can have arbitrary characteristic). This led invariant theorists to the following natural question:

Suppose a linear algebraic group G acts on a variety X over a field k. What is the space $k[X]^G$ of invariant functions and when is it finitely generated?

Mumford conjectured that the family of reductive groups satisfy this finite generation property [9]. Nagata showed finite generation for geometrically reductive groups (i.e., groups whose rational representations are completely reducible) and Haboush showed that the notions of geometric reductivity and usual reductivity of algebraic groups coincide, thereby affirming Mumford's conjecture [10],[11]. Positive results for non-reductive groups, however, are more difficult to obtain. Grosshan showed that for a reductive group G, if U is a maximal unipotent subgroup, then $k[G]^U$ is finitely generated [12], [13]. In this paper we are concerned with the non-reductive case; in particular, the invariant ring of unipotent groups.

Our goal: Let U be a maximal unipotent subgroup of a reductive group G over an algebraically closed field k. For instance, if $G = \mathrm{GL}_n$ then U is the group

$$\mathrm{UL}_n = \left\{ \begin{pmatrix} 1 & * & * & \dots & * \\ & 1 & * & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & 1 & * \\ & & & & 1 \end{pmatrix} \right\}$$

consisting of upper triangular unipotent matrices. Let $k[U]$ denote the algebra of functions on U. For instance, for $G = \mathrm{GL}_n$, $k[U] = k[x_1, \dots, x_{n(n-1)/2}]$. We will study the action of U on itself by conjugation and study the corresponding ring of invariants $k[U]^U$. Our goal is to prove the following theorem.

Theorem 1. *Let k be an algebraically closed field of characteristic zero, then $k[U]^U$ is a free polynomial algebra in n variables, where n is the number of simple roots of G.*

This is stated in the literature [14], [15] but there is no transparent proof provided. Our goal is to provide a transparent proof in the spirit of Chevalley's restriction theorem.

Chevalley's restriction theorem states the following: Let G be a semisimple algebraic group with corresponding Lie algebra $\mathfrak{g}$ over an algebraically closed field k of characteristic zero. Further, fix a Cartan subalgebra $\mathfrak{h}$ and Weyl group W. Then the restriction map $k[\mathfrak{g}] \to k[\mathfrak{h}]$ induces the following isomorphism

$$k[\mathfrak{g}]^G \cong k[h]^W.$$

There are various ways of showing this. For example, in [16] techniques in geometric representation theory are used, whereas in [17], the proof utilises tools in analysis. The proof presented here (adopted from [18]) uses the theory of weights in Lie theory. More explicitly, since the set of weights Λ span $\mathfrak{h}^*$, the polynomials in $\lambda \in \Lambda$ span $k[h]$. Then if a function f is G-invariant, it is invariant under inner automorphisms. These automorphisms, when restricted to $\mathfrak{h}$, act by reflections which generate the Weyl group W. Hence, if we restrict f to the Cartan, we obtain $f|_{\mathfrak{h}} \in k[\mathfrak{h}]^W$. That is, we obtain a map $\theta : k[\mathfrak{g}]^G \to k[\mathfrak{h}]^W$. To show the restriction θ is surjective one shows that the elements $\operatorname{sym} \lambda^k := \sum_{\omega \in W} \omega.\lambda^k$ ($\lambda \in \Lambda^+$, $k \in \mathbb{Z}^+$) lie in the image of θ. To show this, use that for any $\lambda \in \Lambda^+$, and for all $\omega \in W$, $\omega.\lambda < \lambda$ and induct on the partial ordering on the weights. The idea of showing injectivity is to use that the semisimple regular elements of $\mathfrak{g}$ are a G-stable dense subset, where each element is conjugate to an element in $\mathfrak{h}$. Then one can show that $f \in \ker \theta$ must be zero on all of $\mathfrak{g}$.

Further, Chevalley in 1955 showed that $k[\mathfrak{h}]^W$ is a free polynomial algebra [19]. In particular, he showed the following: Let G be a finite reflection group in an n-dimensional vector space V over a field k of characteristic zero. Then $k[V]^G$ is a free polynomial algebra generated by n homogeneous elements. Thus, by Chevellay's restriction theorem, $k[\mathfrak{g}]^G$ is a finitely generated free polynomial algebra.

As alluded to before, our goal was to prove Theorem 1. We began by showing this for GL_n and then used similar techniques to show this for Sp_4. We then took these ideas and generalised them for general reductive groups using the language of roots in Lie theory. That is, our main result is the following.

Theorem 2. *Let U be a maximal unipotent subgroup of a reductive group G, over an algebraically closed field k of zero characteristic, and let $\mathfrak{n} = \mathrm{Lie}(U)$ be the Lie algebra of U. Furthermore, let Δ be the set of simple roots of G and define $\mathfrak{n}_\Delta := \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha$ to be the direct sum of the simple root spaces. Then the following isomorphism of k-algebras holds:*

$$k[\mathfrak{n}]^U \cong k[\mathfrak{n}_\Delta].$$

This result is presented in Section 10. Our approach relies on two key lemmas. The first lemma states that the projection onto the simple root spaces of $x \in \mathfrak{g}$ is invariant under the adjoint action of an element $A \in U$. It follows that the functions restricted to the simple root spaces are U-invariant. This allows us to conclude that the resriction of functions from $\mathfrak{n}$ to $\mathfrak{n}_\Delta$ yields a surjection $\psi : k[\mathfrak{n}]^U \to k[\mathfrak{n}_\Delta]$. To show injectivity we require the second key lemma which states: For a regular nilpotent element $x \in \mathfrak{n}$ there is an $A \in U$ such that the element upon taking the adjoint action of A on x resides in $\mathfrak{n}_\Delta$, i.e., $\mathrm{Ad}_A(x) \in \mathfrak{n}_\Delta$. Using this fact, if a function f whose restriction to $\mathfrak{n}_\Delta$ is zero, then it is zero on all the regular nilpotent elements. A similar argument to Chevellay's restriction theorem shows the kernel

is trivial. To show this holds at the group level, i.e., $k[U]^U \cong k[\mathfrak{n}_\Delta]$, we exploit the fact that the exponential map $\exp : \mathfrak{n} \to U$ is an isomorphism of varieties.

Outline of Thesis.

We begin in Section 2 by discussing classical algebraic geometry which will be the primary language used throughout this work. The main result is the contravariant equivalence of categories of affine varieties and rings of regular functions. In Section 2.1 we give a light treatment of the basics of the theory of Hopf algebras which "dualise" algebraic groups, which can be a helpful tool when studying these groups.

In Section 3 we study some of the basic properties of algebraic groups. In particular, we look at actions of algebraic groups, and state an important theorem (Theorem 62) which describes algebraic groups as subgroups of some GL_n .

In Section 4 we study Lie algebras of algebraic groups. We define tangent spaces to affine varieties as the set of derivations and relate this to the more familiar definition of differentials in calculus. We define the vector space of left-invariant derivations of G and show that this is isomorphic to the tangent space at the identity of G. We will use this to define the Lie algebra $\mathfrak{g}$ of G. We finish this section by providing examples of Lie algebras of some classical algebraic groups.

In Section 5 we explore some of the theory of semisimple and unipotent groups. We also study solvability of algebraic groups which leads into the important definition of a reductive group.

In Section 6 we study reductive groups, their characters and cocharacters. We compute the roots for $\mathrm{GL}_n(k)$ and $\mathrm{Sp}_4(k)$ which we use in the following sections to construct the root groups.

In Section 7 we will study some Lie theory and develop the necessary results to prove Chevalley's Restriction Theorem.

In Section 8 we prove that $k[U]^U$ is finitely generated and give an explicit description of the generators for the case where U is the maximal unipotent subgroup of the general linear group $\mathrm{GL}_n(k)$.

In Section 9 we look to the case where $G = \mathrm{Sp}_4(k)$, the symplectic group. We derive the (simple) root subspaces and use these to form a subspace H to prove the isomorphism $k[U]^U \cong k[H]$, where U is the maximal unipotent subgroup of G.

In Section 10 we present the main result of this paper (Theorem 142). We show the algebra of U-invariant functions on the Lie algebra, $\mathrm{Lie}(U)$, is isomorphic to the functions on the simple root spaces. We conclude this section by illustrating that this holds also at the group level by the exponential map.

2. Some Commutative Algebra and Algebraic Geometry

In this section we develop the basic language classical algebraic geometry, starting with Hilbert's Basis theorem which we state without proof but can be found in [20]. We then introduce some of the basic notions in algebraic geometry, that of affine varieties, defined by the zeros of a set of polynomials, from which we can form the Zariski topology. We will introduce the notion of morphisms between affine varieties, including some results which will be useful for following sections. We will see that for each open set in the Zariski topology there is an associated ring of functions attached. It turns out the study of affine varieties is equivalent to the study of these associated ring of functions.

Definition 3. A ring R is *Noetherian* if every ideal of R is finitely generated.

Theorem 4 (Hilbert's Basis Theorem)**.** *If a ring R is Noetherian and X is a finite set of indeterminates, then the polynomial ring $R[X]$ is Noetherian.*

Proof. See [20]. ■

Definition 5. Let I be an ideal of the commutative ring R, we define the *radical* of I as

$$\sqrt{I} = \{r \in R : r^n \in I \text{ for some } n \in \mathbb{Z}^+\}.$$

I is called a *radical ideal* if $\sqrt{I} = I$.

Definition 6. Let k be a fixed algebraically closed field. We define *affine n-space* over k, denoted $\mathbb{A}^n_k$, or simply $\mathbb{A}^n$, to be the set of all n-tuples of elements of k. For a *point* $P \in A^n$, if $P = (a_1, \ldots, a_n)$ with $a_i \in k$, then a_i is a *coordinate* of P.

Although some of the theory in the following sections hold for arbitrary fields, from here, unless stated otherwise, we will work with a field k which is algebraically closed field of characteristic zero.

Let $k[x_1, \ldots, x_n]$ be the polynomial ring in n variables over k. Elements of $k[x_1, \ldots, x_n]$ are functions from the affine n-space to k, by defining $f(P) = f(a_1, \ldots, a_n)$, where $f \in k[x_1, \ldots, x_n]$ and $P \in \mathbb{A}^n$. Consequently, we can study the *zeros* of f, namely $Z(f) = \{P \in \mathbb{A}^n : f(P) = 0\}$. More generally, if $T \subseteq k[x_1, \ldots, x_n]$, we define the *zero set* of T to be the common zeros of all elements of T, that is,

$$Z(T) = \{P \in \mathbb{A}^n : f(P) = 0 \text{ for all } f \in T\}.$$

If I is an ideal generated by T, then it is easy to see that $Z(I) = Z(T)$. Moreover, since $k[x_1, \ldots, x_n]$ is Noetherian, any ideal I is finitely generated, i.e., $I = (f_1, \ldots f_r)$. Therefore, the $Z(T)$ can be expressed as the common zeros of the finite set of generators $f_1, \ldots, f_r$.

Definition 7. A subset Y of $\mathbb{A}^n$ is an *algebraic set* if there exists a subset $T \subseteq k[x_1, \ldots, x_n]$ such that $Y = Z(T)$.

Proposition 8. *The union of two algebraic sets is an algebraic set. The intersection of any family of algebraic sets is an algebraic set. The empty set and the whole space are algebraic sets.*

Proof. If $Y_1 = Z(T_1)$ and $Y_2 = Z(T_2)$, then $Y_1 \cup Y_2 = Z(T_1T_2)$, where T_1T_2 is the product of all elements of T_1 with the elements of T_2. If $\{Y_\alpha = Z(T_\alpha)\}_{\alpha \in I}$ is any family of algebraic sets, then $\bigcap_{\alpha \in I} Y_\alpha = Z(\bigcup T_\alpha)$. The whole space $\mathbb{A}^n$ is the zero set of the constant function $f = 0$ and the empty set is the zero set of the constant function $f = 1$. ■

The preceeding proposition allows us to equip our affine n-space with a topology, known as the Zariski topology and thus motivates the following definition.

Definition 9. The topology whose open sets are the complements of algebraic sets is called the *Zariski topology.*

Definition 10. A nonempty open subset Y of a topological space X is *irreducible* if it cannot be expressed as the finite union of two proper subsets, each one of which is closed in Y. The empty set is not considered to be irreducible.

We observe that X is irreducible if and only if $X \neq C_1 \cup C_2$ for any two proper closed sets if and only if $\emptyset \neq X \backslash C_1 \cap X \backslash C_2$ if and only if any two nonempty open sets have a nonempty intersection then X is connected. Thus, irreducible implies connected. We will see for algebraic groups also the converse holds.

Example 11. The affine line $\mathbb{A}^1$ is irreducible. Its only proper closed subsets are finite sets, because the only polynomials which vanish on an infinite set is the zero polynomial. But $\mathbb{A}^1$ is infinite since k has zero characteristic and is algebraically closed so cannot be a finite union of finite sets.

Lemma 12. *Any nonempty open subset Y of an irreducible space X is irreducible and dense.*

Proof. Assume that Y is not dense. Then $X \neq \bar{Y}$ and $X = \bar{Y} \cup (X \backslash Y)$ contradicting irreducibility of X. So, Y is dense. Now suppose that $Y = Y_1 \cup Y_2$ where Y_1, Y_2 are closed in Y. Then $\bar{Y} = \bar{Y_1} \cup \bar{Y_2}$. But since Y is dense in X, this forces, $X = \bar{Y_1}$ (without loss of generality). However, the closure of Y_1 in Y means $Y_1 = Y \cap \bar{Y_1} = Y \cap X = Y$. Hence, Y is irreducible. ■

Definition 13. An *affine algebraic variety* (or simply, an *affine variety*) is an irreducible closed subset of $\mathbb{A}^n$ (with the induced topology).

Definition 14. For any subset $Y \subseteq \mathbb{A}^n$, let us define the *ideal* of Y in $k[x_1, \ldots, x_n]$ by $I(Y) = \{f \in k[x_1, \ldots, x_n] : f(P) = 0 \text{ for all } P \in Y\}$. We call $k[Y] = k[x_1, x_2, \ldots, x_n]/I(Y)$ the *coordinate ring* (or *affine k-algebra* or *affine algebra* when there is no confusion of fields) of Y

Example 15. Consider the subset of $X \subset \mathbb{A}^1$ given by $X = \mathbb{A}^1 \setminus \{0\}$. Then the associated affine algebra is given by $k[X] = k[x, y]/\langle xy - 1\rangle = k[x, x^{-1}]$. This illustrates that the affine coordinate ring is not necessarily a polynomial algebra.

We now have a function Z which maps subsets of $k[x_1, \ldots, x_n]$ to algebraic sets of $\mathbb{A}^n$, and a function I which maps algebraic sets to ideals of $k[x_1, \ldots, x_n]$. We want to study the extent of which these maps are inverses. This is studied in the following.

Proposition 16. *Let $Y_1 \subseteq Y_2 \subseteq \mathbb{A}^n$ and $T_1 \subseteq T_2 \subseteq k[x_1, \ldots, x_n]$. Then*

a) $Z(T_1) \supseteq Z(T_2)$.

b) $I(Y_1) \supseteq I(Y_2)$.

c) *For any two subsets X_1, X_2 of $\mathbb{A}^n$ we have $I(X_1 \cup X_2) = I(X_1) \cap I(X_2)$.*

d) *For any ideal $J \subseteq k[x_1, \ldots, x_n]$, $I(Z(J)) = \sqrt{J}$, the radical of J.*

e) *For any subset $Y \subseteq A^n$, $Z(I(Y)) = \bar{Y}$, the closure of Y.*

Proof. For (a) it is easy to see that if $P \in Z(T_2)$, then $P \in Z(T_1)$ since T_2 contains T_1. Similarly, for (b), if $f \in I(Y_2)$, then f vanishes on all of Y_2 which contains Y_1, so f vanishes on all of Y_1. For (c), if $f \in I(Y_1 \cup Y_2)$ then f vanishes on all of the union, so f vanishes on Y_1 and Y_2, i.e., $f \in I(Y_1) \cap I(Y_2)$. Conversely, if $f \in I(Y_1) \cap I(Y_2)$, then f vanishes on all of Y_1 and all of Y_2 so vanishes on their union. We illustrate (d) below (cf Corollary 18). For (e), observe that $Y \subset Z(I(Y))$, which is a closed set, hence $\bar{Y} \subseteq Z(I(Y))$. Conversely, if F is a closed set containing Y, then $F = Z(J)$ for some ideal $J \subseteq k[x_1, \ldots, x_n]$. Hence, $Z(J) \supseteq Y$, hence by (b), $I(Z(J)) \subseteq I(Y)$. But since $J \subseteq I(Z(J)) \subseteq I(Y)$, and by (a), $F = Z(J) \supseteq Z(I(Y))$. Thus, $Z(I(Y)) = \bar{Y}$. ■

Theorem 17 (Hilbert's Nullstellensatz)**.** *Let k be an algebraically closed field, let J be an ideal in $k[x_1, \ldots, x_n]$, and let $f \in k[x_1, \ldots, x_n]$ be a polynomial which vanishes at all points of $Z(J)$. Then $f^r \in J$ for some integer $r > 0$.*

Proof. See [21]. ■

Corollary 18. *There is a one-to-one inclusion-reversing correspondence between algebraic set in $\mathbb{A}^n$ and radical ideals in $k[x_1, \ldots, x_n]$, given by $Y \mapsto I(Y)$ and $J \mapsto Z(J)$. Furthermore, an algebraic set is irreducible if and only if its ideal is a prime ideal.*

Proof. If $f \in \sqrt{J}$, then $f^r \in J$ for some $r > 0$ and $J \subseteq I(Z(J))$, so one direction holds, i.e., $\sqrt{J} \subseteq I(Z(J))$. If $f \in I(Z(J))$ then f vanishes on all of $Z(J))$ which, by Hilbert's Nullstellensatz, implies $f^r \in J$, for some $r > 0$, i.e., $f \in \sqrt{J}$ and $I(Z(J)) = \sqrt{J}$
Now suppose that $Y \subset \mathbb{A}^n$ is irreducible and suppose $fg \in I(Y)$. Then $Y \subseteq Z(fg) = Z(f) \cup Z(g)$. Then we can write $Y = (Y \cap Z(f)) \cup (Y \cap Z(g))$. But both sets are closed in Y, so $Y = Z(f) \cap Y$ or $Y = Z(g) \cap Y$, so $f \in I(Y)$ or $g \in I(Y)$. Hence, $I(Y)$ is prime. Conversely, let J be a prime ideal of $k[x_1, \ldots, x_n]$ and suppose $Z(J) = Y_1 \cup Y_2$, with Y_1, Y_2 being closed. Then $I(Z(J)) = I(Y_1 \cup Y_2) = I(Y_1) \cap I(Y_2)$. This implies $\sqrt{J} = I(Y_1) \cap I(Y_2)$, but prime ideals are equal to their own radical. Hence, $J = I(Y_1) \cap I(Y_2)$. But this implies that $J = I(Y_1)$ or $J = I(Y_2)$ because $I(Y_1)I(Y_2) \subseteq I(Y_1) \cap I(Y_2) = J$ and so either $I(Y_1) \subseteq J$ or $I(Y_2) \subseteq J$. Thus, $Z(J) \subseteq Y_1$ or $Z(J) \subseteq Y_2$, so $Z(J)$ is irreducible.

■

Corollary 19. *Let $f \in k[x_1, \ldots, x_n]$ be an irreducible polynomial, then the associated variety, $V(f)$, is irreducible.*

Proof. Since $k[x_1, \ldots, x_n]$ is a unique factorisation domain, any irreducible polynomial is a prime element. Hence, a principal ideal generated by an irreducible polynomial, f, is a prime ideal, therefore, $V(f)$ is irreducible. ■

When $X \subset \mathbb{A}^n$ is irreducible $I(X)$ is a prime ideal so $k[X]$ becomes an integral domain (using the fact that for a commutative ring, R, an ideal, I, is prime iff R/I is an integral domain). Therefore, we can form the field of fraction for $k[X]$, being the *field of rational functions* on X, and we denote this by $k(X)$.

Corollary 20. *An algebraic set X is irreducible if an only if $k[X]$ is an integral domain.*

Example 21. By Corollary 20, we see that $\mathbb{A}^n$ is irreducible since its coordinate ring $k[x_1, \ldots, x_n]$ is an integral domain since k is.

We have seen that, in the Zariski topology, that the closed sets are the points vanishing on some set of polynomials. The closed sets that are defined by a single polynomial are the *principal open sets* and can be useful tools in the theory due to the following result.

Proposition 22. *Let $X \subset \mathbb{A}^n$ be an affine variety, then X has a basis (in the Zariski topology) consisting of* principal open subsets $X_f = \{x \in X \mid f(x) \neq 0\}$.

Proof. Let U be an Zariski open set, then there is a set of polynomials $\{f_1, \ldots, f_n\}$ such that $U = \{x \in X \mid f_i(x) \neq 0,\ \forall i \in \{1, \ldots, n\}\}$. If $x \in U$, then for some $j \in \{1, \ldots, n\}$, $f_j(x) \neq 0$, so we can form $X_{f_j} = \{x \in X \mid f_j(x) \neq 0\}$ so we can cover X with all these X_{f_j}'s. Moreover, if $x \in X_{f_1} \cap X_{f_2}$ for some $f_1, f_2 \in k[X]$, then let $g = f_1 f_2$ and we see that $x \in X_g \subseteq X_{f_1} \cap X_{f_2}$. ■

We want a description for the functions on our variety. This leads us to the definition of a *regular function*.

Definition 23. Let $U \subset X \subseteq \mathbb{A}^n$ be any open subset and $p \in U$ a point. We say that a function f on U is *regular at p* if in some neighbourhood V of p there exists polynomial functions $g, h \in k[x_1, \ldots, x_n]$ such that $f = g/h$ where $h(p) \neq 0$. We say that f is *regular on U* if it is regular at every point of U. We denote by $\mathcal{O}_{X,x}$ the ring of functions that are regular at $x \in X$, and call it the *local ring of X at x*, and $\mathcal{O}_X(U)$ the *ring of regular functions* on the open subset U.

The content of the following lemma describes the ring of regular functions $\mathcal{O}(x)$ at the point $x \in X$ as the localisation of the coordinate ring $k[X]_{M_x}$ corresponding to the maximal ideal containing x. This result validates the definition of regular functions and in a way confirms the way that the theory of algebraic geometry is developed is correct.

Lemma 24. *The following statements hold.*

(i) *Let X be an algebraic set and $x \in X$, then $\mathcal{O}_{X,x} \cong k[X]_{M_x}$, where $k[X]_{M_x}$ is the localisation of $k[X]$ at M_x, where M_x is the maximal ideal in the ring $k[X]$ corresponding to the point $x \in X$.*

(ii) *If X is an irreducible algebraic set, and $0 \neq f \in k[X]$, then $k[X]_f \cong \mathcal{O}_X(X_f)$, and in fact $\mathcal{O}_X(X) = k[X]$.*

Proof. See [22] ■

We now arrive to a key definition, that of a *morphism* between varieties.

Definition 25. If X and Y are two varieties, a *morphism* $\varphi : X \to Y$ is a continuous map such that for every open set $V \subseteq Y$, and for every regular function $f : V \to k$, the function $f \circ \varphi : \varphi^{-1}(V) \to k$ is regular.

The composition of morphisms is a morphism, so, in particular, we define an *isomorphism* of two varieties as a morphism $\varphi : X \to Y$ which has an inverse morphism $\varphi^{-1} : Y \to X$ such that composing these morphisms yields the identity morphism. One should view a morphism $\varphi : \mathbb{A}^m \to \mathbb{A}^n$ as an n-tuple of regular functions $(\varphi_1, \ldots, \varphi_n)$ in m variables sending polynomials to polynomials.

Example 26. Let $f(x, y) = y - x^2$, then $Z(f) = \{(t, t^2) \in \mathbb{A}^2 \mid t \in k\}$. Then the affine coordinate ring $k[x, y]\backslash\langle y - x^2\rangle$ is isomorphic to a polynomial ring of one variable. Since the coordinate ring is an integral domain, the algebraic set $Z(f)$ is irreducible, and therefore a variety. We consider the morphism $\varphi : Z(f) \to \mathbb{A}^1$ given by $(t^2, t) \mapsto t$. This morphism has inverse $\varphi^{-1} : \mathbb{A}^1 \to Z(f)$ given by $t \mapsto (t^2, t)$ which is a morphism. Hence, $Z(f) \cong \mathbb{A}^1$ as varieties.

Morphisms of varieties $\varphi : X \to Y$ correspond bijectively and contravariantly to ring homomorphisms of coordinate rings $\varphi^* : k[Y] \to k[X]$. We say that $\varphi^* : k[Y] \to k[X]$ is the *comorphism* (or the *pullback*) of the morphism φ, defined by $(\varphi^* f)(x) = f(\varphi(x))$, for all $x \in X$. Consequently, two varieties are isomorphic if and only if their affine k-algebras are isomomorphic. The correspondence comes from the equivalence of categories of algebraic sets and morphisms of algebraic set and the category of affine k-algebras and morphisms of k-algebras, see Theorem 28

Lemma 27. *If $\varphi : X \to Y$ is a morphism of varieties sets, then the pullback $\phi^* : k[Y] \to k[X]$ is an algebra homomorphism.*

Proof. The pullback φ^* is a ring homomorphism since

$$\varphi^*(f + g) = (f + g)(\varphi) = f \circ \varphi + g \circ \varphi = \varphi^* f + \varphi^* g,$$

and,

$$\varphi^*(fg) = (fg)(\varphi) = (f \circ \varphi)(g \circ \varphi) = (\varphi^* f)(\varphi^* g),$$

That ϕ^* maps the identity to the identity is immediate. Thus, φ^* is a ring isomorphism. ■

The fact that morphisms of varieties correspond to algebra homomorphisms is not a coincidence. Indeed, we have the following useful fact.

Theorem 28. *The contravariant functor which sends a variety $X \mapsto k[X]$ to its coordinate ring and for each morphism $(\varphi : X \to Y) \mapsto (\varphi^* : k[Y] \to k[X])$ is an isomorphism between the category of affine varieties and morphisms of affine varieties and the category of affine k-algebras and morphisms of affine k-algebras.*

Proof. See [22]. ■

This result allows one to study varieties via their associated k-algebra and vice versa, which is something we had already been doing. We now state some results which will be useful for some of the theory of algebraic groups.

Definition 29. A topological space X is called *Noetherian* if it satisfies the descending chain condition for closed subsets: That is, for any sequence $Y_1 \supseteq Y_2 \supseteq Y_3 \supseteq \dots$ of closed subsets, there is an integer n such that $Y_n = Y_{n+1} = \dots$.

Example 30. Affine n-space, $\mathbb{A}^n$, is Noetherian. This follows from Hilbert's Nullstellensatz (Theorem 18). Every closed subset $Y \subseteq \mathbb{A}^n$ corresponds to a radical ideal $I \subseteq k[x_1, \dots, x_n]$, so any chain of subsets in $\mathbb{A}^n$ corresponds to an ascending chain of ideals in $k[x_1, \dots, x_n]$ which stabilise because $k[x_1, \dots, x_n]$ is Noetherian (Theorem 4). Hence, any chain of subsets in $\mathbb{A}^n$ stabilises, therefore, $\mathbb{A}^n$ is Noetherian.

Proposition 31. *Let X be a Noetherian topological space. Then X has finitely many irreducible subsets, they are closed and cover X.*

Proof. See [23] ■

Definition 32. A subset of a topological space, X, is said to be *locally closed* if it is the intersection of an open set with a closed set. We call a subset *constructible* if it is a finite union of locally closed sets.

The next result will be useful when we consider the image of algebraic group homomorphisms. The proof is adopted from [24].

Proposition 33. *Let X be a topological space. If $Y \subset X$ is constructible then it contains a dense open subset of its closure $\overline{Y}$.*

Proof. Let $Y = \bigcup_{i=1}^n Y_i$, where each Y_i are locally closed and denote by $Z_i = \overline{Y_i} \backslash Y_i$ so that we can define $Z = \bigcup_i^n Z_i$. We claim that $W = \overline{Y} \backslash Z \subset Y$ and is dense in $\overline{Y}$. We observe that

$$Y \cup Z = \bigcup_{i=1}^{n} Y_i \cup \bigcup_{i}^{n} Z_i = \bigcup_{i}^{n} (Y_i \cup Z_i) = \bigcup_{i}^{n} \overline{Y_i} = \overline{\bigcup_{i}^{n} Y_i} = \overline{Y}.$$

But $W = \overline{Y} \backslash Z = (Y \cup Z) \backslash Z = Y \backslash Z \subset Y$, which gives the inclusion. To show that W is open in $\overline{Y}$, we notice that since Y_i is locally closed, Y_i is open in $\overline{Y_i}$ (this is an equivalent definition

for locally closed sets). Then Z_i, is closed in $\overline{Y_i}$ implies Z_i is closed in X. Since Z is closed, W is open in $\overline{Y}$. Now assume that W were not dense in $\overline{Y}$. By definition of density, there exists a nonempty open subset of Z which is open in $\overline{Y}$. Let n_0 be the smallest index such that $\bigcup_{i=1}^{n_0} Z_i$ contains a nonempty open subset, U, of $\overline{Y}$. $U \not\subset Z_i$, for any i, since if it were then U would be a nonempty open subset of $\overline{Y}_i$ with $U \cap Y_i = \emptyset$, which is absurd. Therefore, $n_0 > 1$ and $U \not\subset Z_{n_0}$. Then, we have $\emptyset \neq U \backslash Z_{n_0} \subset \bigcup_{i=1}^{n_0-1} Z_i$. But Z_{n_0} is closed, so $U \backslash Z_{n_0}$ is open in $\overline{Y}$, contradicting minimality of n_0. We conclude that W is dense in $\overline{Y}$. ■

Theorem 34 (Chevalley's Theorem)**.** *Let $\varphi : X \to Y$ be a morphism of varieties. Then φ preserves constructible sets, i.e., φ maps constructible sets to constructible sets.*

Proof. See [23]. ■

2.1. **Hopf Algebras.** We have seen that a variety is completely determined by its affine algebra. In the theory of affine algebraic groups, the underlying structure of the group is that of an affine variety. Therefore, it can be easier at times, to study (under the contravariant equivalence of categories) the corresponding affine algebra of the group as opposed to studying the group as a variety with morphisms. These affine algebras have the structure of *Hopf algebras* which will be studied in this section. We begin by stating the axioms of a unital associative algebra using commutative diagrams. We follow the approach of [23], [22] and [25].

Definition 35. An *algebra* with unit over a field k is a vector space A with two linear maps, multiplication $\mu : A \otimes A \to A$ and a unit map $\eta : k \to A$ such that the following diagrams commute:

$$\begin{array}{ccc} A \otimes A \otimes A & \xrightarrow{\mu \otimes id} & A \otimes A \\ \downarrow{\scriptstyle id \otimes \mu} & & \downarrow{\scriptstyle \mu} \\ A \otimes A & \xrightarrow{\mu} & A \end{array}$$

FIGURE 2. Associativity of multiplication in A.

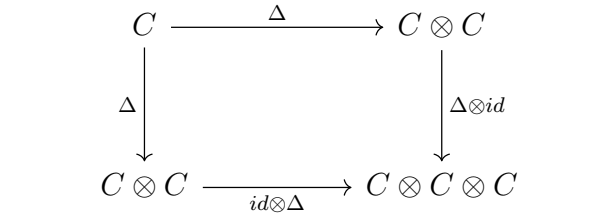
$$\begin{array}{ccccc} k \otimes A & \xrightarrow{\eta \otimes id} & A \otimes A & \xleftarrow{id \otimes \eta} & A \otimes k \\ & \searrow & \downarrow{\scriptstyle \mu} & \swarrow & \\ & & A & & \end{array}$$

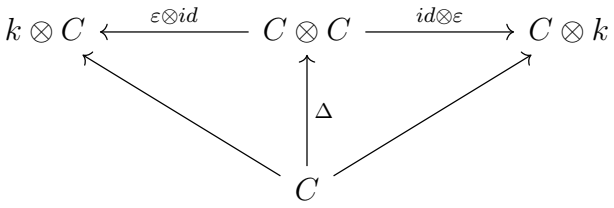
FIGURE 3. Unit map in A.

We may denote the algebra as a triple (A, μ, η), however, when there is no confusion we simply denote the algebra by A. Recall that if A and B are algebras, with multiplication maps μ_A, and μ_B, then $\phi : A \to B$ is an algebra homomorphism if $\phi \circ \mu_A = \mu_B(\phi \otimes \phi)$.

All we have done here is express the familiar axioms of algebras in terms of commutative diagrams. We will now dualise these definitions by reversing the arrows of the multiplication and unit maps to obtain an object we refer to as a *coalgebra*.

Definition 36. A *coalgebra* (C, Δ, ε) with counit is a vector space, C, together with two linear maps, comultiplication $\Delta : C \to C \otimes C$ and counit map $\varepsilon : C \to k$ such that the following diagrams commute:





If the diagrams commute then the comultiplication is said to be *coassociative* and the counit map is *counital*. We may denote the coalgebra as a triple (C, Δ, ε), however, when there is no confusion we simply denote the algebra by C. A *coalgebra homomorphism* $\psi : C \to D$ between two coalgebras C and D, with comultiplication Δ_C, Δ_D, and counit maps ε_C and ε_D, is a map which respects the structure of the coalgebra, i.e., it satisfies $(\psi \otimes \psi) \circ \Delta_C = \Delta_D(\psi \otimes \psi)$ and $\varepsilon_C = \varepsilon_D \circ \psi$.

Let us illustrate this with some examples.

Example 37. Consider the set of all $n \times n$ matrices over k, $\mathrm{Mat}_n(k)$. Then, take

$$k[\mathrm{Mat}_n] = k[x_{ij} : 1 \leq i, j \leq n].$$

The generators x_{ij} take a $M \in \mathrm{Mat}_n(k)$, with entries denoted by M_{ij}, and return the (i, j)-entry of M, that is, $x_{ij}(M) = M_{ij} \in k$. In other words x_{ij} are *coordinate functions*. It is enough to define the maps Δ and ε on the generators. For any $M, N \in \mathrm{Mat}_n(k)$,

$$x_{ij}(MN) = (MN)_{ij} = \sum_{k=1}^{n} m_{ik} n_{kj} = \sum_{k=1}^{n} x_{ik}(M) x_{kj}(N)$$

and this suggests that we should have $\Delta(x_{ij}) = \sum_{k=1}^{n} x_{ik} \otimes x_{kj}$. We also have that $x_{ij}(I_n) = \delta_{ij}$ which suggests $\varepsilon(x_{ij}) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. It is straigtforward check that the coassociativity and counitality axioms are indeed satisfied. That is, $(\mathrm{Mat}_n(k), \Delta, \varepsilon)$, with Δ and ε defined as such, form a coalgebra.

Example 38. Let kG denote the *group algebra* of G over a field k with basis $\{e_g\}_{g\in G}$. We can define comultiplication and counit maps as $\Delta(g) = g \otimes g$ and $\varepsilon(e_g) = 1$. These map are readily seen to satisfy the axioms.

In both the preceding examples, we observe that both vector spaces have the structure of an algebra and a coalgebra. We call such objects *bialgebras*.

Definition 39. Given a vector space B over a field k, we say that B is a *bialgebra* over k if (B, μ, η) is an algebra, (B, Δ, ε) is a coalgebra and the following conditions hold:

1. Δ and ε are morphisms of coalgebras;
2. μ and η are morphisms of algebras.

The bialgebra structure is usually denoted as a 5-tuple $(B, \mu, \eta, \Delta, \varepsilon)$.

Definition 40. Let A be an algebra and C a coalgebra and let $f, g \in \mathrm{Hom}(C, A)$. Then the *convolution* of f and g is given by

$$(f \star g)(c) = \mu \circ (f \otimes g) \circ \Delta(c) \qquad \forall c \in C.$$

Definition 41. Let $(H, \mu, \eta, \Delta, \varepsilon)$ be a bialgebra. An endomorphism, $S \in \mathrm{End}(H)$, of H is called an *antipode* for the bialgebra, H, if

$$S \star id_H = id_H \star S = \eta \circ \varepsilon.$$

A *Hopf algebra* is a bialgebra with an antipode.

Example 42. $k[\mathrm{Mat}_n(k)]$ in Example 37 is not an example a Hopf algebra because there is no antipode. However, if we take the general linear group, $\mathrm{GL}_n(k)$, then

$$k[\mathrm{GL}_n(k)] = k[x_{ij}, t : 1 \leq i, j \leq n]/\langle \det(x_{ij})t - 1\rangle,$$

where $t = \det^{-1}$ is indeed a Hopf algebra. The comultiplication and counit maps are the same from Example 37, however, we need to define $\Delta(\det)$. After a straigtforward but lengthy calculation using the comultiplication map, one obtains $\Delta(\det) = \det \otimes \det$. The antipode map follows from Cramer's rule $S(x_{ij}) = (-1)^{i+j} \det(M_{ij})t$.

Example 43. In Example 38 observed that the group algebra kG has a bialgebra structure. It is a Hopf algebra with antipode $S(e_g) = e_{g^{-1}}$. Indeed,

$$S \star id = \mu \circ (S \otimes id) \circ \Delta(e_g) = \mu \circ (S \otimes id)(e_g \otimes e_g) = \mu(e_{g^{-1}} \otimes e_g) = e_1$$

and

$$\eta \circ \varepsilon(e_g) = \eta(1) = e_1.$$

It is clear that $S \star id = id \star S$.

3. Algebraic Groups

Here, and for the next few sections, the objects of study are algebraic groups. We only present a light introduction of the theory, omitting some details. We refer the reader to [26] and [23] for a more thorough treatment of the theory.

Definition 44. Let G be a variety equipped with a group structure. If the group operation $m : G \times G \to G$, given by $m(g, h) = gh$, and $i : G \to G$, given by $i(g) = g^{-1}$ are morphisms of varieties, then G is an *algebraic group*.

We note here that the product $G \times G$ here is given Zariski topology and not the product topology. Unless otherwise specified, we will be working with *affine algebraic groups* where the underlying variety is an affine variety and the morphisms are morphisms of affine varieties.

We also have the notions of (normal) subgroups.

Definition 45. Let G be an algebraic group. An abstract subgroup $H \subset G$ (i.e., a subgroup in the group theory sense) that is closed (in the Zariski topology) is said to be an *algebraic subgroup* of G. An algebraic subgroup is *normal* if its normal in the abstract group sense.

Note that if $H \subset G$ is closed then $H \times H \subset G \times G$ is closed and the restriction of a morphism on $G \times G$ to $H \times H$ is a morphism. Therefore, H is an algebraic group in its own right.

Example 46. The *multiplicative group*, $\mathbb{G}_m$, over k is the affine closed subset $k^\times \subset \mathbb{A}^2$ cut off by the polynomial $f(x, y) = xy - 1$. The group operations are $m(x, y) = xy$ and $i(x) = x^{-1}$ which are clearly morphisms when we recall that the associated affine algebra is given by $k[\mathbb{G}_m] = k[x, x^{-1}]$, see Example 15. The comultiplication map $m^* : k[x, x^{-1}] \to k[x, x^{-1}] \otimes k[x, x^{-1}]$ is given by $m^*(x) = x \otimes x$, the coinverse map $i^* : k[x, x^{-1}] \to k[x, x^{-1}]$ is given by $i^*(x) = x^{-1}$ and the counit map $\eta^* : k[x, x^{-1}] \to k^*$ is given by $e^*(x) = 1$.

Example 47. We can identify $M_n(k)$, the set of $n \times n$ matrices, as the affine space, $\mathbb{A}^{n^2}$, in the natural way. The set of all $n \times n$ invertible matrices, $\mathrm{GL}_n(k)$, forms a group under matrix multiplication and is an open subset of $\mathbb{A}^{n^2}$ defined by nonvanishing points of the determinant map. As a closed subset of $\mathbb{A}^{n^2+1}$, we can describe $\mathrm{GL}_n(k)$ as the closed variety cut off by the polynomial $f(x) = \det(x)t - 1$, where $t = \det^{-1}$ and $x \in \mathrm{Mat}_n(k)$. That is, we have $\mathrm{GL}_n(k) \subset \mathbb{A}^{n^2+1}$ as $\mathrm{GL}_n(k) = \{((g)_{ij}, t) \in A^{n^2+1} | \det(g)t - 1 = 0\}$, where $t = \det(g)^{-1}$. The formulas for matrix multiplication are nothing but polynomials in the entries of the matrices. That the inverse formula is a morphism follows from Cramer's rule for the inverse of an $n \times n$ matrix $(g^{-1})_{ij} = (-1)^{i+j} \det(M_{j,i})/\det(g)$ where $M_{j,i}$ is the submatrix of g whose jth row and ith column have been deleted. We have also seen the comultiplication and counit maps for the associated Hopf algebra (see Example 37 and Example 42).

Any subgroup of the linear algebraic group GL_n which is closed in the Zariski topology is a linear algebraic group. For example:

(i) The *special linear group*, SL_n, is the group of $n \times n$ matrices whose determinants are 1. In other words, it is the kernel of the morphism

$$\det : \mathrm{GL}_n(k) \to k^\times.$$

The associated affine algebra is $k[\mathrm{SL}_n] = k[x_{ij} : 1 \leq i, j \leq n]/\langle \det(x_{ij}) - 1 \rangle$. The maps for the associated coalgebra are given by, passing to the quotient, of $k[\mathrm{GL}_n]$. This applies to all closed subgroups of GL_n.

(ii) The subgroup $B_n = \{f \in \mathrm{GL}_n \,|\, g_{ij} = 0 \text{ for } i < j\} \subset \mathrm{GL}_n(k)$ of invertible upper triangular matrices.

(iii) The *upper triangular unipotent group*

$$U_n = \{g \in B_n | \, g_{ii} = 1 \text{ for } 1 \leq i \leq n\}.$$

(iv) The (algebraic) *torus* $T_n = (\mathbb{G}_m)^n = k^* \times \cdots \times k^*$ (n copies). The map $\varphi : T_n \to D_n$ given by $(\lambda_1, \ldots, \lambda_n) \mapsto \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, where D_n is the subgroup of GL_n of diagonal matrices, is clearly an isomorphism of algebraic groups.

(v) Let $n = 2m$ be an even positive integer. The *symplectic group*, denoted $\mathrm{Sp}_n(k)$, is defined as the subgroup of $\mathrm{GL}_n(k)$ of matrices satisfying $g^t M g = M$ for a symplectic form $M = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}$. In other words it is the set

$$\mathrm{Sp}_n(k) = \{g \in \mathrm{GL}_n(k) : g^t M g = M\}.$$

where g^t denotes the transpose of g.

Definition 48. A *homomorphism* of algebraic groups $\varphi : G \to G'$ is a morphism of varieties that is also a homomorphism of groups. The notion of an *isomorphism* of algebraic groups is natural, i.e., an isomorphism of groups that is also an isomorphism of varieties.

Example 49. In this example we will classify the algebraic group homomorphisms $\varphi : \mathbb{G}_m \to \mathbb{G}_m$. We have seen that, under the contravariant equivalence of categories, a morphism $\varphi : \mathbb{G}_m \to \mathbb{G}_m$ corresponds to an algebra endomorphism $\varphi^* : k[t, t^{-1}] \to k[t, t^{-1}] \cong k[\mathbb{G}_m]$. If we restrict to the generators, $\varphi^*(t)$ and $\varphi^*(t^{-1})$ belong to $k[t, t^{-1}]$ and their product is 1. In preserving this property, it follows that $\varphi^*(t) = t^m$, where $m \in \mathbb{Z}$. Hence, $\mathrm{Hom}(\mathbb{G}_m, \mathbb{G}_m) = \{x \mapsto x^m : m \in \mathbb{Z}\} \cong \mathbb{Z}$.

Definition 50. For any topological space (X, τ), given an $x \in X$, the *irreducible component of x,* is the maximal irreducible subset of X that contains x. Here, the word 'maximal' refers to inclusion of sets.

For an algebraic group G, we denote by G^0 the *irreducible component of the identity* (or, the *identity component*) of G, that is, it is the maximal irreducible subset containing the identity. The identity component of a group turns out to be a useful tool in the study of algebraic groups; we state some useful results.

Proposition 51. *Let G be an algebraic group. Then*

(i) *G^0 the unique identity component*

(ii) *G^0 is a normal subgroup of finite index in G, whose cosets are the connected as well as irreducible components of G.*

(iii) *Each closed subgroup of finite index in G contains G^0.*

Proof.

(i) Suppose that X_1 and X_2 are two identity components. Then the image of the product morphism $m(X_1 \times X_2) = X_1X_2$ is irreducible since $X_1 \times X_2$ is. Since X_1 and X_2 both contain the identity $X_1 \supset X_1X_2$. On the other hand since X_2 contains the identity, $X_1X_2 \supset X_1$. Thus, $X_1 = X_1X_2$. The same argument shows $X_1X_2 = X_2$. Thus, $X_1 = X_2$.

(ii) Firstly, left translations $\lambda_x : G \to G$, $\lambda_x(y) = xy$ is an isomorphism of varieties with inverse morphism $\lambda_{x^{-1}}$, and similarly for right translation. For each $x \in G^0$, we have that $x^{-1}G^0$ contains the identity and is irreducible, hence $x^{-1}G^0 = G^0$, and in particular, $x^{-1} \in G^0$. Since $x \in G^0$ was arbitrary, we have that G^0 contains all its inverses. Moreover, $G^0G^0 = G^0$, therefore, G^0 is a closed subgroup of G. Now let $x \in G$ be arbitrary. Since xGx^{-1} is an irreducible component of G containing the identity, we must have that $xG^0x^{-1} = G^0$. Thus, G^0 is a normal subgroup. The cosets of G^0 are translations of G^0, i.e., homeomorphisms, and so are also irreducible components of G. There are a finite number of them since G is a Noetherian topological space, and since the G^0 cosets partition G they are disjoint, hence are the connected components of G.

(iii) Suppose H is a closed subgroup of finite index in G. We want to show $G^0 \subset H$. Each of the finitely many cosets are closed (using that translates are homeomorphisms) then the union of the cosets that are different to H are also closed, that is, $\bigcup_{x \notin H} xH$ is closed. Then the complement of this union, H, must be open. Then xH is also open and we have that G^0 is partitioned into a finite union of open sets. Since G^0 is connected and intersects H at the identity, we must have that $G^0 \subset H$.

■

Following from the preceeding proposition, we say an algebraic group G is *connected* if $G^0 = G$, and we see that the connected components are the irreducible components. The following results will be useful for later sections.

Example 52. Since the notions of irreducibility and connectedness of algebraic groups coincide, we immediately have that $\mathrm{GL}_n(k)$, is connected since it is an open subset (described by the principal open subset of non-vanishing determinant) of the irreducible space $\mathbb{A}^{n^2}$, hence irreducible by Lemma 12.

Lemma 53. *Let U, V be two dense open subsets of an algebraic group, G. Then $G = UV$.*

Proof. Since the inversion operation is a homeomorphism, V^{-1} is a dense open subset of G. Translations are also homeomorphisms, and, therefore, xV^{-1} is also a dense open subset of G for all $x \in G$. Therefore, by density of U and V, for every $x \in G$, $xV^{-1} \cap U \neq \emptyset$. Hence, U contains elements of the form xv^{-1}, where $v^{-1} \in V^{-1}$. Therefore, $x \in UV$. Thus, $G = UV$. ■

Proposition 54. *Let H be a subgroup of the algebraic group, G, and $\overline{H}$ its closure. Then*

(i) *$\overline{H}$ is a closed subgroup of G;*

(ii) *If H is constructible, then $H = \overline{H}$.*

Proof.

(i) Since inversion is a homeomorphism, we have that $\overline{H}^{-1} = \overline{H^{-1}} = \overline{H}$. Thus, $\overline{H}$ contains all of its inverses. Moreover, since translation is a homeomorphism, for $x \in H$, $x\overline{H} = \overline{xH} = \overline{H}$ for all $x \in H$. Now, for $x \in \overline{H}$, we have $Hx \subset H\overline{H}$, but now $\overline{H}x = \overline{Hx} \subset \overline{H\overline{H}} = \overline{H}$. $\overline{H}$ is closed under the group operation, hence $\overline{H}$ is a subgroup of G.

(ii) If H is constructible, by Proposition 33, H contains a dense open subset, U, of $\overline{H}$. Part (a) revealed that $\overline{H}$ is a group and, using Lemma 53, $\overline{H} = U \cdot U \subset H \cdot H = H$.

■

Corollary 55. *Let H, K be closed subgroups of an algebraic group G. If K normalizes H, then HK is a closed subgroup of G.*

Proof. Since K normalizes H, we have that $K \subset N_G(H)$ and it is an elementary exercise to check that HK is a subgroup. Since the H, K are closed subgroups of G, they are also open. Indeed, the complement of $G \backslash H = \bigcup_{x \notin H} xH$ is the union of closed sets, and similarly for H. Then $H = G \cap H$ is the intersection of an open and closed set, and similarly for K. Therefore, H and K are constructible. The image of $A \times B = AB$ under the product morphism (coming from the group operation) is constructible by Theorem 34. Using part (b) of Proposition 54. ■

Proposition 56. *Let $\varphi : G \to G'$ be a morphism of algebraic groups. Then*

(i) *$\ker \varphi$ is a closed subgroup of G*

(ii) *$\operatorname{im} \varphi$ is a closed subgroup of G'*

(iii) *$\varphi(G^0) = \varphi(G)^0$*

Proof.

(i) Since φ is continuous and $\ker \varphi$ is the inverse image of the closed set $\{1_G\}$, hence it is closed.

(ii) Since G is locally closed, it is constructible, therefore, by Chevalley's Theorem 34. Then $\varphi(G)$ is closed by Proposition 54.

(iii) Since $\varphi(G^0)$ is closed by part (b) and connected it is a subset of $\varphi(G)^0$. Conversely, by Proposition 51, $\varphi(G)^0 \subset \varphi(G^0)$, since $\varphi(G^0)$ has finite index in $\varphi(G)$, since a Noetherian topological space has finitely many maximal irreducible subspaces (see Proposition 31). The result follows immediately from Proposition 51 (iii).

■

3.1. **Actions of Algebraic Groups.** We can use algebraic groups to study the group of symmetries of geometric objects by studying their actions on varieties. This leads us to the next definition.

Definition 57. Let G be an algebraic group and X a variety. If the map $\varphi : G \times X \to X$ is a morphism of varieties such that the following properties are satisfied:

(i) $x_1 \cdot (x_2 \cdot y) = (x_1 \cdot x_2) \cdot y$ for all $x_i \in G$, $y \in X$;

(ii) $1_G \cdot y = y$ for all $y \in X$,

then G acts *morphically* on X, (or simply 'acts', where there is no confusion). In this situation, X is a *G-variety*, or a *G-space*.

The standard definitions of *stabilizers, orbits, fixed points, etc.* apply in the algebraic groups setting, see [27] for a review of these concepts.

Example 58. For any algebraic group G, we have G acting on itself by conjugation, i.e., $g.x = gxg^{-1}$ for all $g, x \in G$ is a morphism of varieties. The orbits are the conjugacy classes and the stabilizers of $g \in G$ are the centralisers of $g \in G$.

Definition 59. We say a group acts *linearly* on a vector space, V if the morphism $\varphi : G \times V \to V$ satisfies

(i) $g \cdot (v_1 + v_2) = g \cdot v_1 + g \cdot v_2$ for all $g \in G$ and $v_i \in V$;

(ii) $g \cdot (\lambda v) = \lambda(g \cdot v)$ for all $g \in G$ and $v \in V$.

The action of algebraic groups on an affine variety, X, induces a *linear* action on the affine algebra $k[X]$. If $x \in G$ and $f \in k[X]$, then $(xf)(y) = f(x^{-1}y)$ for all $x \in G$ and $y \in X$. We can express this in terms of the comorphism attached to the morphism $y \mapsto x^{-1} \cdot y$ for $x \in G$ and $y \in X$. For $f \in k[X]$, we have $(\tau_x f)(y) = f(x^{-1}y)$. Note that the map $\tau : G \to \mathrm{GL}(k[X])$ defined by $\tau(x) = \tau_x$ is a homomorphism. Indeed,

$$(\tau_{xy} f)(z) = f(y^{-1}x^{-1}z) = (\tau_y f)(x^{-1}z) = (\tau_x \tau_y f)(z).$$

Then τ_x is a k-algebra automorphism of $k[X]$. When G acts on itself, i.e., $X = G$ we will denote λ_x (resp. ρ_x) as the comorphisms of the left (resp. right) regular actions, $x \mapsto xy$ (resp. $x \mapsto yx^{-1}$). These comporhisms are the *left* (respectively *right*) *translation of functions* by x given by $(\lambda_x f)(y) = f(x^{-1}y)$ (resp. $(\rho_x f)(y) = f(yx)$).

Definition 60. Let V be a finite-dimensional vector space over a field k. A *rational representation* of G in V is a homomorphism of algebraic groups $\rho : G \to \mathrm{GL}(V)$. Equivalently,

V is said to be a *G-module*, and view V as an affine variety isomorphic to $\mathbb{A}^{\dim(V)}$, with the G-action given by $g.v = \rho(g)v$.

The following lemma will be useful when we study semisimple and unipotent parts of elements of a closed subgroup G of GL_n .

Lemma 61. *Let H be a closed subgroup of an algebraic group, G, and I the ideal of $k[G]$ vanishing on all of H. Then $H = \{x \in G | \rho_x(I) \subset I\}$*

Proof. If $x \in H$ and $f \in I$ then $(\rho_x f)(y) = f(yx) = 0$ for all $y \in H$. This shows the forward inclusion $H \subset \{x \in G | \rho_x(I) \subset I\}$. Conversely, let $x \in G$ have the property that $\rho_x(I) \subset I$. In other words, $0 = (\rho_x f)(y) = f(yx)$ for all $f \in I$ and $y \in H$. If we choose $y = 1_H$, then $f(x) = 0$, so $x \in H$. ■

We have seen in Example 47 that closed subgroups of GL_n are algebraic groups. It turns out that the converse holds, that is, any affine algberaic group is isomorphic to a closed subgroups of some $\mathrm{GL}_n(k)$.

Theorem 62. *Let G be an affine algebraic group. Then G is isomorphic to a closed subgroup of some* GL_n .

Proof. See [23]. ■

4. The Lie Algebra of an Algebraic Group.

4.1. **Tangent Spaces of Affine Varieties.** Here, we study/introduce some notions of tangent spaces of algebraic varieties to eventually describe the Lie algebra associated to an algebraic group. For this section we mainly follow the approach of [22] and [23].

Definition 63. Let A be a commutative ring and M a left A-module. A *derivation* $D : A \to M$ is an additive map such that for all $a, b \in A$ we have $D(ab) = a \cdot D(b) + b \cdot D(a)$. The set of all derivations from A to M, denoted $\mathcal{D}(A, M)$, is given a left A-module structure through the natural action.

If A is an algebra over k, then $\mathcal{D}_k(A, M) \subset \mathcal{D}(A, M)$ is the set of all linear derivations. For ease of notation, if $M = A$, $\mathcal{D}(A, A)$ is abbreviated to $\mathcal{D}(A)$.
If A is an associative algebra over a commutative ring R, an *augmentation*, ε, of A is a k-algebra homomorphism $\varepsilon : A \to k$. We say that an algebra together with an augmentation is an *augemented algebra,* denoted (A, ε). Then we can consider set of k-linear derivations on the augmented algebra (A, ε) denoted $\mathcal{D}_\varepsilon(A) \subset \mathcal{D}_k(A, k)$.

This set is defined as,

$$\mathcal{D}_\varepsilon(A) = \{\delta : A \to k : \ \delta(ab) = \varepsilon(a)\delta(b) + \varepsilon(b)\delta(a)\} \subset A^*.$$

The subset of these derivations are called the ε-derivations of A.

Example 64. Let $A = (k[t], \varepsilon_x)$, where ε_x is the evaluation at $t = x \in k$ is an augmentation algebra. The derivative map $\frac{d}{dt} : k[t] \to k$ is a derivation by the Leibniz rule $\frac{d}{dx}\big|_{t=x}(fg) = \varepsilon_x(g)\frac{d}{dt}\big|_{t=x} f + \big(\frac{d}{dt}\big|_{t=x} g\big)(\varepsilon_x(f)$.

Definition 65. Let A be a commutative algebra over a field k. The *module of differentials* Ω_A is defined as follows: let J be the kernel of the map $\mu : A \otimes_k A \to A$ defined by $a \otimes b \mapsto ab$. Then, $\Omega_A = J/J^2$.

We will show that the map $d : A \to \Omega_A$, defined as $d(a) = a \otimes 1 - 1 \otimes a + J^2$, is a derivation when we equip Ω_A with the structure of a left A-module by multiplication on the first tensorand. Indeed, let $a, b \in A$, then

$$a \cdot d(b) + b \cdot d(a) = (ab) \otimes 1 - a \otimes b + (ba) \otimes 1 - b \otimes a + J^2.$$

Notice that

$$J^2 \ni (a \otimes 1 - 1 \otimes a)(b \otimes 1 - 1 \otimes b) = (ab) \otimes 1 - a \otimes b - b \otimes a + 1 \otimes (ab)$$

Now subtracting this from the preceding equation, one obtains

$$a \cdot d(b) + b \cdot d(a) = ba \otimes 1 - 1 \otimes ab + J^2 = d(ab).$$

Definition 66. Let X be an algebraic variety. If $x \in X$, and $\varepsilon_x : \mathcal{O}_x \to k$ is the evaluation map (i.e., $\varepsilon_x(f) = f(x)$). Then the *tangent space* of X at x is defined as $T_x(X) = \mathcal{D}_{\varepsilon_x}(\mathcal{O}_x)$. That is, the tangent space at $x \in X$ is the space of point derivations of $\mathcal{O}_x$ at x. In particular,

$$T_x(X) = \{\delta : \mathcal{O}_x \to k \,|\, \delta(fg) = \varepsilon_x(f)\delta(g) + \delta(f)\varepsilon_x(g), \forall f, g \in \mathcal{O}_x\}$$

Definition 67. Let $f : X \to Y$ be a morphism of algebraic varieties and $f_x^* : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ the corresponding comorphism of algebras. We define the *differential* of f at the point x, denoted $d_x f$ as the map induced by f_x^*. More explicitly, $d_x f : T_x(X) \to T_{f(x)}(Y)$ is defined as $d_x f(\delta) = \delta \circ f_x^* \in T_{f(x)}(Y)$ for $T_x(X) \ni \delta : \mathcal{O}_{X,x} \to k$.

One has to verify that this definition is indeed correct. That is, if δ is a point derivation then so is $\delta \circ f_x^*$. Indeed, suppose $f : X \to Y$ is morphism with the induced comorphism $f_x^* : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$, then

$$\begin{aligned}\delta \circ f_x^*(gh) = \delta(gh \circ f_x) &= \delta(g(f_x)h(f_x)) = g(f_x)\delta(h(f_x)) + \delta(g(f_x))h(f_x) \\ &= (f_x^* g)\delta(f^* h) + (f_x^* h)\delta(f_x^* g) = (f_x^* g)(\delta \circ f_x^*)(h) + (f_x^* h)(\delta \circ f_x^*)(g).\end{aligned}$$

We see that the differential takes derivations on $x \in X$ to derivations on $f(x) \in Y$, i.e., "tangent vectors" on X to "tangent vectors" on Y, analogously to the differential in the context of Riemannian Geometry. Hence, we expect the differential map and the tangent space to satisfy properties similar to the properties in the calculus setting. This brings us to some important properties for later results when dealing with Lie algebras of algebraic groups.

Proposition 68. *The following properties hold:*

(i) *There is a chain rule for composition of morphisms, that is, let $\varphi : X \to Y$ and $\psi : Y \to Z$ be morphisms of algebraic varieties and d_x and $d_{\phi(x)}$ the corresponding differentials for the induced comorphisms ϕ^* and ψ^*. Then $d_x(\psi \circ \varphi) = d_{\varphi(x)}\psi \circ d_x\varphi$;*

(ii) *the differential of the identity morphism is the identity, that is, $d_x \operatorname{id}_X = \operatorname{id}_{T_x(X)}$.*

Proof. For (i) let $\delta \in \mathcal{D}_{\varepsilon_x}(\mathcal{O}_x)$, then

$$d_{\varphi(x)}\psi \circ d_x\varphi(\delta) = d_{\varphi(x)}\psi \circ (\delta \circ \varphi^*) = (\delta \circ \varphi^*\psi^*) = d_x(\psi \circ \varphi)(\delta).$$

For (ii) we have

$$d_x \operatorname{id}_x \circ(\delta)(f) = (\delta \circ id^*)(f) = \delta \circ f \circ \operatorname{id} = \delta(f) = (\operatorname{id}_{T_x(X)} \delta)(f).$$

■

We now prove a result which will allows us to describe the tangent space as the dual space of a quotient.

Lemma 69. *Let (A, ε) be a commutative augmented algebra and consider the linear map $\delta_A : A \to \ker(\varepsilon)/\ker(\varepsilon)^2$ defined as $\delta_A(a) = a - \varepsilon(a) + \ker(\varepsilon)^2$. Given any ε-derivation $\delta : A \to V$, where V is a vector space, there exists a unique linear map $L_\delta : \ker(\varepsilon)/\ker(\varepsilon)^2 \to V$ with the property that $L_\delta \circ \delta_A = \delta$.*

Proof. Given an ε-derivation δ, we define $L_\delta : \ker(\varepsilon)/\ker(\varepsilon)^2 \to V$ as $L_\delta(a + \ker(\varepsilon)^2) = \delta(a)$. This is well-defined since if $a + (\ker\varepsilon)^2 = b + (\ker\varepsilon)^2$, where $a, b \in A$, then $a - b \in \ker(\varepsilon)^2$, and in particular, $a - b = \alpha\beta$, where $\alpha, \beta \in \ker\varepsilon$. Then

$$L_\delta(a+(\ker\varepsilon)^2) - L_\delta(b+(\ker\varepsilon)^2) = \delta(a)-\delta(b) = \delta(a-b) = \delta(\alpha\beta) = \varepsilon(\alpha)\delta(\beta)+\varepsilon(\beta)\delta(\alpha) = 0.$$

Hence $L_\delta(a + \ker(\varepsilon)^2) = L_\delta(b + \ker(\varepsilon)^2)$. Linearity of L_δ follows from linearity of δ. Now, for all $a \in A$ we have

$$L_\delta \circ \delta_A(a) = L_\delta(a - \varepsilon(a) + (\ker\varepsilon)^2) = \delta(a - \varepsilon(a)) = \delta(a),$$

since $a \in \ker\varepsilon$. Lastly, suppose that L_δ and χ were linear maps with the property that $L_\delta \circ \delta_A = \delta$ and $\chi \circ \delta_A = \delta$. Then we have that $L_\delta \circ \delta_A(a) = \delta(a) = \chi\delta_A(a)$. Hence, $L_\delta = \chi$ because they agree on all of $\ker(\varepsilon)/\ker(\varepsilon)^2$ since δ_A is surjective, because for any $x \in \ker\varepsilon/(\ker\varepsilon)^2$, $x = a+(\ker\varepsilon)^2 = a-\varepsilon(a)1+(\ker\varepsilon)^2 = \delta_A(a)$. This proves the lemma. ■

Corollary 70. *Let X be an algebraic variety and $x \in X$. Then $T_x(X) \cong (\mathcal{M}_x \,/\, \mathcal{M}_x^2)^*$, with $\mathcal{M}_x = \ker(\varepsilon_x)$.*

Proof. Let X be an algebraic variety, then setting $A = \mathcal{O}_{X,x}$, $V = k$ and $\varepsilon = \varepsilon_x$ in Lemma 69 yields the desired isomorphism of vector spaces $\mathcal{D}_{\varepsilon_x}(\mathcal{O}_x) = T_x(X) \cong (\mathcal{M}_x \,/\, \mathcal{M}_x^2)^*$, where $\mathcal{M}_x = \ker(\varepsilon_x)$, because by Lemma 69, the maps $\delta \mapsto L_\delta$ and $L_\delta \mapsto \delta$ are inverses. Note that $\mathcal{M}_x$ is the maximal ideal of $\mathcal{O}_x$. ■

We have see the tangent space of an algebraic variety $X = \mathbb{A}^n$ at a point x described in terms of point derivations on the ring of regular functions $\mathcal{O}_{X,x}$. It turns out that this definition is equivalent to one which may seem more familiar in a calculus setting. In particular, that of taking the derivative of polynomials. We know that differential operator is a derivation, i.e., Leibniz rule is satisfied. Therefore, $\frac{\partial}{\partial x_i}|_{x=a} : \mathcal{O}_a \to k$ is a derivation. On the other hand, if we take a generator in $k[X]$, say x_j, then we know that $\frac{\partial}{\partial x_i}(x_j) = \delta_{ij}$, δ_{ij} being the kronecker delta. Hence, the $\{\frac{\partial}{\partial x_i} : 1 \leq i \leq n\}$ is a basis for $D_{\varepsilon_x}(\mathcal{O}_x)$.

We will now compute the differential of the determinant morphism which will be useful later.

Example 71. Let $\det : \mathrm{GL}_n(k) \to \mathbb{G}_m(k)$ be the determinant map, we will compute the differential at the identity. Note that we have already seen that this map is a morphism of varieties. We compute the differential as follows: let $A \in \mathrm{GL}_n(k)$ and I the identity matrix, then using the limit definition

$$\lim_{h\to 0} \frac{\det(hA+I) - \det(I)}{h} = \lim_{h\to 0} \frac{\sum_{\sigma\in S_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^n (hA+I)_{i,\sigma(i)}\right) - 1}{h}$$
$$= \lim_{h\to 0} \frac{\prod_{i=1}^n (hA+I)_{ii} - 1 + \sum_{\substack{\sigma\in S_n \\ \sigma\neq 1}} \prod_{i=1}^n (hA+I)_{i,\sigma(i)}}{h}.$$

We notice that all the terms in the summand in the last equality (when $\sigma \neq 1$) go to zero in the limit since $(hA+I)_{i,\sigma_i} = ha_{i,\sigma(i)}$, and the product of these vanish in the limit. In the first sum, all the order $o(h^2)$ terms in $\prod_{i=1}^n (hA+I)_{ii} - 1$ vanish in the limit and only the linear terms remain. Hence, the differential of the determiant is $d(\det)_A(I) = \sum_{i=1}^n a_{ii} = \operatorname{tr}(A)$.

4.2. **The space of left-invariant derivations.** In this section we introduce the notion of the associated Lie algebra to an algebraic group. We will define the Lie algebra of an algebraic group as the space of left invariant derivations on the associated affine algebra and show that this is equivalent to the tangent space at the identity of the algebraic group.

Definition 72. Let G be an algebraic group. We define the space of left invariant derivation on $k[G]$ as

$$\mathrm{Lie}(G) = \{\delta \in \mathcal{D}(k[G]) \mid \delta\lambda_x = \lambda_x\delta\}$$

and call it the *Lie algebra of* G, where λ_x is the left translation of functions.

Note that the space of derivations forms a Lie algebra since it is a vector space and the Lie bracket of two derivations is also a derivation. The Lie bracket here is $[\delta, \gamma] = \delta\gamma - \gamma\delta$. Also, the subspace of left-invariant derivations is indeed a subspace since if two derivations commute with λ_x then the Lie bracket of two derivations will also commute with λ_x. Moreover, one can check that the action on $\mathcal{D}(k[G])$ is a rational action [22].

Since an affine algebraic group has the structure of an affine variety, we can study the tangent space of an algebraic group, G, at a point $x \in G$. We will denote the tangent space at the identity of G as $\mathfrak{g} = T_e(G)$. Our goal is to show that the tangent space at the identity is isomorphic to the space of left invariant derivations. This leads us to the following theorem.

Theorem 73. *Let G be an algebraic group and* $\mathrm{Lie}(G)$ *the corresponding Lie algebra and define $\mathfrak{g} := T_e(G)$. Then $\theta : \mathrm{Lie}(G) \to \mathfrak{g}$, defined as $(\theta\delta)(f) = (\delta f)(e)$, is an isomorphism of vector spaces. If $\varphi : G \to G'$ is a morphism of algebraic groups, then $d\varphi_e : \mathfrak{g} \to \mathfrak{g}'$ is a homomorphism of Lie algebras, where $\mathfrak{g}' = T_{e'}(G')$.*

Proof. To prove the theorem we define an inverse map $\eta : \mathfrak{g} \to \mathrm{Lie}(G)$, and show this is the inverse of θ. The map η is defined as follows: Let $v \in \mathfrak{g}$ be a tangent vector, $x \in G$, $f \in k[G]$ and λ_x be the left translation by x. Then define the *right convolution by v* as

$$(f * v)(x) = v(\lambda_{x^{-1}} f).$$

We will show that $*v$ is a linear left invariant derivation. Linearity follows from the fact that v is linear and $\lambda_{x^{-1}}$ acts linearly on $k[G]$. Let $f, g \in k[G]$, then $*v$ is a derivation since

$$\begin{aligned}(fg * v)(x) = v(\lambda_{x^{-1}}(fg)) = v((\lambda_{x^{-1}} f)(\lambda_{x^{-1}} g)) = v(\lambda_{x^{-1}} f)g(x) + f(x)v(\lambda_{x^{-1}} g)\\ = (f * v)(x)g(x) + f(x)(g * v)(x) = ((f * v)g + f(g * v))(x).\end{aligned}$$

Furthermore, $*v$ is left invariant since

$$(\lambda_y(f * v))(x) = (f * v)(y^{-1}x) = v(\lambda_{x^{-1}y} f) = v(\lambda_{x^{-1}}(\lambda_y f)) = ((\lambda_y f) * v)(x).$$

Thus, $*v$ is a left invariant derivation. To show that η is the inverse of θ we will show that $\eta \circ \theta$ and $\theta \circ \eta$ are the identity maps. Firstly,

$$(f * \theta(\delta))(x) = \theta(\delta)(\lambda_{x^{-1}} f) = \delta(\lambda_{x^{-1}} f)(e) = \lambda_{x^{-1}}(\delta f)(e) = (\delta f)(x).$$

Hence, if $\delta \in \mathrm{Lie}(G)$, $\eta \circ \theta)(\delta) = \delta$. So, $\eta \circ \theta = \mathrm{id}_{\mathrm{Lie}(G)}$. Conversely,

$$\theta(*v)(f) = (f * v)(e) = v(\lambda_{e^{-1}} f) = v(f).$$

Hence, if $v \in \mathfrak{g}$, $(\theta \circ \eta)(v) = v$. Thus, η is the inverse of θ and we conclude that θ is a vector space isomorphism.

We will now show that the differential, $d\varphi : \mathfrak{g} \to \mathfrak{g}'$ of a morphism, $\varphi : G \to G'$, preserves the structure of the Lie bracket. For $u, v \in \mathfrak{g}$, let $u' = d\varphi(u)$, $v' = d\varphi(v)$ let $f' \in k[G']$, with $f = \varphi^* f'$ as the pullback along φ.

$$\begin{aligned}[d\varphi(u), d\varphi(v)] = [u', v']f' = u'v'f' - v'u'f'\\ = u'(f' * v') - v'(f' * u') = u(\varphi^*(f' * v')) - v(\varphi^*(f' * u')).\end{aligned}$$

Conversely,

$$d\varphi([u, v])f' = u(f * v) - v(f * u).$$

Hence, it is enough to show that $f * u = \varphi^*(f' * u')$, in other words, that $(\varphi^* f') * u = \varphi^*(f' * d\varphi(u))$. Since the terms on either side are functions on G, it suffices to show they agree on all of G. To this end, let $g \in G$, then

$$(\varphi^* f' * u)(g) = u(\lambda_{g^{-1}} \varphi^* f')$$

and on the other hand,

$$(f' * d\varphi(u))(\varphi(g)) = d\varphi(u)(\lambda_{\varphi(g)^{-1}} f') = u(\varphi^*(\lambda_{\varphi(g)^{-1}} f')).$$

All that is left to show is that $\lambda_{g^{-1}}\varphi^* f' = \varphi^*(\lambda_{\varphi(g)^{-1}} f')$, for which we will show they agree on arbitrary $h \in G$, that is,

$$(\lambda_{g^{-1}}\varphi^* f')(h) = \varphi^* f'(gh) = f'(\varphi(gh))$$

and conversely,

$$(\varphi^*(\lambda_{\varphi(g)^{-1}} f'))(h) = (\lambda_{\varphi(g)^{-1}} f')(\varphi(h)) = f'(\varphi(gh)),$$

using that φ is a morphism of algebraic groups. This proves the theorem. ■

An immediate corollary of this theorem is that the Lie algebra of an algebraic group, G, is finite-dimensional and in particular $\dim(\mathrm{Lie}(G)) = \dim(T_e(G))$.

We now consider the inner automorphism of G, denoted $\mathrm{Inn}\, x : G \to G$, given by $y \mapsto xyx^{-1}$, where $x, y \in G$. This inner automorphism is clearly a polynomial map and so a morphism of algebraic groups. We will note the differential of $\mathrm{Inn}\, x$ by $\mathrm{Ad}\, x$. By Theorem 73, $\mathrm{Ad}\, x$ is an automorphism for the Lie algebra $\mathfrak{g}$ corresponding to G. Indeed, $(\mathrm{Ad}\, x)(\mathrm{Ad}\, x^{-1}) = d(\mathrm{Inn}\, x) \circ d(\mathrm{Inn}\, x^{-1}) = d(\mathrm{Inn}\, x \circ \mathrm{Inn}\, x^{-1}) = d(\mathrm{Inn}\, e) = id_{\mathfrak{g}}$, using Proposition 68. Moreover, we have $(\mathrm{Ad}\, x)(\mathrm{Ad}\, y) = \mathrm{Ad}(xy)$, therefore, $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g}) \subset \mathrm{GL}(\mathfrak{g})$. We call Ad the *adjoint representation* of G. The following proposition describes the action of $\mathrm{Ad}\, x$ on $\mathrm{Lie}(G)$.

Proposition 74. *Let G be an algebraic group, $x \in G$, and $\mathrm{Ad}\, x$ as defined above. Then $\mathrm{Ad}\, x(\delta) = \rho_x \delta \rho_x^{-1}$, where ρ_x is right translation of functions.*

Proof. Firstly, for this to make sense $\rho_x \delta \rho_x^{-1}$ should lie in $\mathrm{Lie}(G)$, but this is true since left and right translations commute and δ is left invariant, hence $\rho_x \delta \rho_x^{-1}$ is left invariant. If we define $\mathrm{Ad}\, x(\delta) = \delta'$, and $\varphi := \mathrm{Inn}\, x$, then by definition of $\mathrm{Ad}\, x$ and the differential, $(\delta' f)(e) = \delta(\varphi^* f)(e)$ for every $f \in k[G]$. We can apply this to $\rho_x f$, so that $\varphi^*(\rho_x f)(y) = (\rho_x f)(xyx^{-1}) = f(xy) = (\lambda_{x^{-1}} f)(y)$. Since $f \in k[G]$, $y \in G$ were arbitrary, $\varphi^*(\rho_x f) = (\lambda_{x^{-1}} f)$. Now, using this fact, we can compute

$$\begin{aligned}(\delta' \rho_x)(f)(e) = (\mathrm{Ad}\, x(\delta)\rho_x)(f)(e) = (d_x\varphi)(\delta)\rho_x)f)(e) = \delta(\varphi^* \rho_x f)(e) = \delta(\lambda_{x^{-1}} f)(e) \\ = \lambda_{x^{-1}}(\delta f)(e) = (\delta f)(x) = \rho_x \delta(f)(e).\end{aligned}$$

We can conclude that $\delta' \rho_x = \rho_x \delta$, which implies $\delta' = \rho_x \delta \rho_x^{-1}$, as required.

■

We now compute some examples of induced Lie algebras by affine algebraic groups

Example 75. The additive group $\mathbb{G}_a$, over k, has $k[G_a] = k[t]$ as the associated affine coordinate ring. The Lie algebra is one-dimensional spanned by the left invariant derivation $\delta = \frac{\partial}{\partial t}$. It is enough to show this on the generator t. Let $x \in G$, then $\lambda_{x^{-1}}\delta(t) = \lambda_{x^{-1}} 1 = 1$. On the other hand, $\delta(\lambda_{x^{-1}} t) = \delta(t + x) = 1$.

Example 76. The multiplicative group G_m, over k, has $k[G_m] = k[t, t^{-1}]$ as its affine coordinate ring. The Lie algebra is one-dimensional spanned by the left invariant derivation $t\frac{\partial}{\partial t}$.

Example 77. To describe the Lie algebra for larger dimensional groups we look at introducing the Lie bracket on $T_e(G)$. Recall for an algebraic group G we have the multiplication map $\mu : G \times G \to G$ and we write $\mu^* f = \sum f_i \otimes g_i$ where $f, f_i, g_i \in k[G]$. Now suppose we have $u, v \in T_e(G)$, being derivations $k[G] \to k$. We can then construct the map $u \otimes v : k[G] \otimes k[G] \to k$ by $f \otimes g \mapsto (uf)(vg)$. We use this linear map to define the product $u \cdot v = (u \otimes v) \circ \mu^* : k[G] \to k$. To show that this satisfies $[u, v] = u \cdot v - v \cdot u$, and that it is associative, we will see that it identifies with $(*u)(*v)$ under the isomorphism $\theta : \mathrm{Lie}(G) \to \mathfrak{g}$ in Therorem 73. If we assume that $\mu^* f = \sum_i f_i \otimes g_i$, then for $x \in G$, we have following

$$(f * u)(x) = u(\lambda_{x^{-1}} f) = u\big(\sum_i f_i(x) g_i\big) = \sum_i f_i(x) u(g_i),$$

which implies $f * u = \sum f_i u(g_i)$. Using this fact we can compute,

$$((f * v) * u)(e) = \big(\sum_i f_i v(g_i)\big) * u)(e) = \big(\sum_i (f_i * u) v(g_i)\big)(e) = \sum_i u(f_i) v(g_i),$$

which is $u \cdot v$. Now, using this product of derivations we can compute the Lie algebra of $G = \mathrm{GL}_n(k)$. Since G is a principal open set defined by the polynomials in Example 47, it is an open subset of the affine space $\mathbb{A}^{n^2+1}$. We know that the tangent space is spanned by the mixed partial differential operators $\frac{\partial}{\partial x_{ij}}$ and evaluating at the identity. We use the tangent space in aid to compute the Lie algebra of G by fixing a tangent vector u, defined by $u(g_{ij}) = x_{ij}$. Here, g_{ij} are the indeterminates of the polynomial equations which define G, and x_{ij} are the n^2 numbers arranged in a $n \times n$ square matrix. Now, we invoke the product

$$(u \cdot v)(g_{ij}) = (u \otimes v)\mu^* g_{ij} = (u \otimes v)\big(\sum_h g_{ih} \otimes g_{hj}\big) = \sum_h x_{ih} y_{hj},$$

which is just matrix multiplication. It follows that the k-linear map $u \mapsto (x_{ij}) \in \mathfrak{g}$ is injective (since the kernel is trivial as the zero vector is the only vector which sends all g_{ij} to zero) and surjective since the dimension of the image is equal to the dimension of $\mathrm{Mat}_n(k)$. Thus, the Lie algebra of $\mathrm{GL}_n(k)$ is $\mathfrak{gl}_n(k)$.

We know that a linear algebraic group, G, is isomorphic to some closed subgroup of some GL_n. It is then convenient to describe the Lie algebra of an algebraic group by computing the Lie algebra as a subalgbera of $\mathfrak{gl}_n$.

Example 78. Let $U_n(k)$ be the subgroup of $\mathrm{GL}_n(k)$ of all upper triangular matrices with $g_{ii} = 1$ for all i and $g_{ij} = 0$ for $i < j$. As a variety, this is a closed subset of $\mathrm{Mat}_n(k)$ defined by the polynomials $g_{ii} - 1 = 0$, $g_{ij} = 0\,(i > j)$. Hence, its Lie algebra is translate of the tangent space at the identity to zero. In other words, the Lie algebra is the set of $n \times n$ matrices with zeros on and below the diagonal.

Example 79. Consider the special linear group, $\mathrm{SL}_n(k)$, defined in Example 47 as the zero set of $f(x_{ij}) = \det(x_{ij}) - 1$, where $1 \leq i, j \leq n$. We saw in Example 71 that the differential operator applied to the determinant map followed by evaluating at the identity is the trace. Therefore, the Lie algebra contains the set of all traceless matrices. On the other hand, a dimension count shows that the Lie algebra of $\mathrm{SL}_n(k)$, denoted $\mathfrak{sl}_n$, is the Lie algebra of all $n \times n$ traceless matrices over k.

5. Unipotent, Semisimple and Jordan Decomposition

We recall from Linear algebra that a linear operator, $x \in \mathrm{End}(V)$, on a finite dimensional vector space is *nilpotent* if there exists an $n \in \mathbb{Z}^+$ such that $x^n = 0$, *unipotent* if $x - 1$ is nilpotent, and *semisimple* if it is diagonalisable. Then the (additive) Jordan-Chevalley decomposition states that, for a linear operator x on a finite dimensional vector spaces, there is a basis for which x can be expressed uniquely as $x = x_s + x_n$, where x_s is semisimple and x_n is nilpotent, moreover, $x_s x_n = x_n x_s$. If x is invertible, then the eigenvalues of x are nonzero, and in particular, x_s is invertible (x_s is a diagonal matrix with the eigenvalues of x on the diagonal). Then we can write $x = x_s + x_n = x_s(1 + x_s^{-1}x_n)$, where $1 + x_s^{-1}x_n$ is unipotent. Thus, we have the (multiplicative) Jordan-Chevalley decomposition of a linear operator $x = x_s x_u$, where x_s is semisimple and x_u is unipotent since x_n is nilpotent and x_s^{-1} and x_n commute. Notice that the semisimple and unipotent parts commute.

For an arbitrary subgroup, G of GL_n it may certainly be the case where the semisimple and unipotent parts of the elements of G do not belong in G. However, if the subgroup G is Zariski closed, then $x = x_s x_u \in G$ also implies x_s and $x_x \in G$. This is due to the Lemma 61.

Definition 80. Let G be an affine algebraic group. An abstract subgroup $H \subset G$ is *unipotent* if $H \subset U_G = \{x \in G : x = x_u\}$.

In other words, a subgroup of an algebraic group is *unipotent* if all of its elements are unipotent.

We state some results about unipotent and semisimple elements without proof.

Proposition 81. *The following hold:*

(1) *Let $G \leq \mathrm{GL}_n$ be a unipotent group. Then there exists a $g \in \mathrm{GL}_n$ such that $gGg^{-1} \leq U_n$.*

(2) *If $\phi : G \to G'$ is a homomorphism of algebraic groups then $\phi(g_s) = \phi(g)_s$ and $\phi(g_u) = \phi(g)_u$ where $g = g_s g_u$.*

(3) *Let $g \in \mathrm{GL}_n(V)$, then there are unique elements $g_s, g_u \in G$ such that $g = g_s g_u$, where g_s and g_u are semisimple and unipotent respectively.*

Proof. See [26] and [28]. ■

6. Reductive Groups

Recall from group theory, the *commutator,* $[H, K]$, of two subgroups $H, K \leq G$ is the subgroup generated by the elements $hkh^{-1}k^{-1}$, $h \in H$, $k \in K$. This is also well-defined for algebraic groups.

Definition 82. Let G be an affine algebraic group and consider the sequence of closed normal subgroups

$$G^0 = G \supset G^1 \supset \cdots \supset G^n \supset \ldots,$$

defined as $G^1 = [G, G]$ and $G^{n+1} = [G^n, G^n]$ if $n > 1$. We say that G is *solvable* if there is an $n \geq 0$, such that $G^n = \{1\}$. Similarly, one can define a *lower central series* recurcively by $G^{(1)} = G$ and $G^{(i+1)} = [G^{(i)}, G]$ and we say that G is *nilpotent* if the lower central series terminates, i.e., $G^{(n+1)} = \{1\}$ for some n.

It is easy to see that a nilpotent group is solvable.

Example 83. The group of unipotent matrices $U_n \subset \mathrm{GL}_n$ is nilpotent, hence solvable. Therefore, any subgroup of U_n is solvable.

Lemma 84. *The following hold:*

(i) *Subgroups and homomorphic images of solvable groups are solvable.*

(ii) *If N is a normal solvable subgroup of G and the quotient G/N is also solvable. Then G is solvable.*

(iii) *If H and K are normal solvable subgroups of G then HK is solvable.*

Proof. These are standard results in elementary group theory. Part (iii) follows from the second isomorphism theorem of groups. ■

A priori an algebraic group need not have a *unique* largest normal solvable subgroup. However, using Corollary 55 and Lemma 84, we have uniqueness of a largest solvable normal subgroup of some algebraic group, G, since we can always take products of normal solvable subgroups. The Lemma guarantees that the product is indeed solvable and the Corollary guarantees closedness. Thus, G contains a unique largest normal solvable subgroup. The identity component of this largest normal solvable subgroup of G is called the *radical* of G, denoted $R(G)$. It is the largest connected normal solvable subgroup containing the identity. The subgroup of $R(G)$ containing all the unipotent elements is normal in G and we call it the *unipotent radical* of G, denoted $R_u(G)$. It is the largest connected normal unipotent subgroup of G. This motivates the following definition.

Definition 85. Let G be a nontrivial connected algebraic group. If $R(G) = \{1_G\}$ is trivial, we say that G is *semisimple*. If $R_u(G) = \{1_G\}$ is trivial, we say that G is *reductive*.

We also have a useful result, which can aid in computing the radical of a group, which we state without proof.

Theorem 86 (Lie-Kolchin)**.** *Let G be a connected solvable subgroup of* $\mathrm{GL}(V)$*, where $V \neq 0$ is finite-dimensional. Then G has a common eigenvector in V.*

Proof. See [23]. ■

For completeness, we also introduce the notion of a *Borel subgroup.*

Definition 87. A *Borel subgroup* of an algebraic group, G, is a maximal (properly contained in no other) closed connected solvable subgroup.

Theorem 88. *Let G be a connected algebraic group, then all the Borel subgroups of G are conjugate.*

Proof. See [23]. ■

Now we can reformulate the definition of the radical of a group (for connected algebraic groups) by providing an equivalent definition, using Borel subgroup, which leads to the following proposition.

Proposition 89. *Let G be a connected algebraic group. The radical (resp. unipotent radical) of G is the identity component of the intersection of all Borel subgroups (resp. of their unipotent parts).*

Proof. Since Borel subgroups are conjugate (Theorem 88), $\bigcap_{B\subset G} B$ is normal subgroup. By Proposition 51, $\left(\bigcap_{B\subset G} B\right)^0$ is also normal. Thus, $\left(\bigcap_{B\subset G} B\right)^0$ is a connected normal solvable subgroup. It follows by maximality, that $\left(\bigcap_{B\subset G} B\right)^0 \subseteq R(G)$. Conversely, $R(G)$ is a connected normal solvable subgroup, therefore, is contained in all the Borel subgroups. Therefore, $R(G) \subset \left(\bigcap_{B\subset G} B\right)$. Thus, $R(G) \subseteq \left(\bigcap_{B\subset G} B\right)^0$, proving equality. It follows that $R(G) = \left(\bigcap_{B\subset G} B_u\right)^0$. ■

Any unipotent group has a nontrivial unipotent radical since unipotent groups are their own unipotent radical. Examples of familiar reductive groups are the following.

Example 90 (GL_n and $\mathrm{SL}_n(k)$)**.** Let $G = \mathrm{GL}_n(k)$. By the Lie-Kolchin theorem, a Borel subgroup has the form of a triangular matrix. Let $B_1 \subset G$ be a Borel subgroup that is in upper-triangular form. Conjugating B_1 with a permutation matrix P of the form

$$P = \begin{pmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{pmatrix}$$ transposes B_1 into lower-triangular form. Since $R(G) = \left(\bigcap_{B\subset G} B\right)^0$, the radical is the intersection of a lower and upper-triangular matrix and so must be a diagonal matrix. But normality of $R(G)$ forces the diagonal entries to be identical. Thus, $R(G) = \{\lambda I_n : \lambda \in k^\times\}$, meaning that $\mathrm{GL}_n(k)$ is not semisimple. Taking the unipotent part, we obtain $R_u(G) = \{I_n\}$ which shows $\mathrm{GL}_n(k)$ is reductive.

A similar argument shows that for $G = \mathrm{SL}_n(k)$, $R(G) \subseteq T$, where T is a torus seen as the group of diagonal matrices. Again, normailty forces the diagonal entries to be equal, but since $\det(g) = 1$ for all $g \in G$, we have $R(G) = \{I_n\}$. Thus, $\mathrm{SL}_n(k)$ is semisimple and therefore reductive.

6.1. **Characters, Cocharacters and Roots.** In this section we explore some of the structure theory of reductive groups. In particular, we study characters, cocharacters and roots. We begin with perhaps a seemingly unrelated but important lemma.

Lemma 91 (Dedekind)**.** *Let G be an (abstract) group, $X(G)$ the set of all homomorphisms $G \to k^\times$. Then $X(G)$ is a linearly independent subset in $k[G]$.*

Proof. For a contradiction, asume the lemma is false. Choose n minimal so that for all $g \in G$, we have $\sum_{i=1}^{n-1} c_i\chi_i(g) + \chi_n(g) = 0$ with $c_i \in k$ not all zero and $\chi_i \in X(G)$ for all i. Since $\chi_1 \neq \chi_n$, there exists a $g \in G$ such that $\chi_1(g) \neq \chi_n(g)$. For $g, h \in G$, we obtain the following equations,

$$\sum_{i=1}^{n-1} c_i\chi_i(g)\chi_i(h) + \chi_n(g)\chi_n(h) = 0 = \sum_{i=1}^{n-1} c_i\chi_i(g)\chi_n(h) + \chi_n(g)\chi_n(h).$$

Subtracting these equations yields $\sum_{i=1}^{n-1} c_i(\chi_i(h) - \chi_n(h))\chi_i = 0$, contradiacting minimality of n. ∎

Recall a linear algebraic group G is a torus if it is isomorphic to a direct product $T = (\mathbb{G}_m)^n$, $n \in \mathbb{Z}^+$ of the multiplicative group $\mathbb{G}_m$. It follows that T connected and abelian and every element is semisimple. The converse also holds [29].

Definition 92. Let G be an algebraic group, an algebraic group homomorphism $\chi : G \to \mathbb{G}_m$ is called a *character* of G. The set of character of G, we denote by $X^*(G) = \mathrm{Hom}(G, \mathbb{G}_m)$.

For any $g = \mathrm{diag}(x_1, \ldots, x_n) \in T_n$, define the character $\chi_i(\mathrm{diag}(x_1, \ldots, x_n)) = x_i$, then this set of characters generate $X^*(T)$. Furthemore, $k[T_n] = k[\chi_i^{\pm} \mid i \in \{1, \ldots, n\}]$ and by Dedekind's Lemma [30] the character group is a linearly independent set in $k[G]$. Also, one can show that $\chi(T) \cong \mathbb{Z}^n$. Hence, $X^*(T_n)$ is a finitely generated abelian group, so it is sometimes customary to write the operations in $X^*(T)$ additively.

Example 93. Let $T_n \subset \mathrm{GL}_n(k)$ be the maximal torus in $\mathrm{GL}_n(k)$ over k with elements $t = \mathrm{diag}(t_1, \ldots, t_n)$. Then $\chi_i \in X^*(T)$, given by $\chi_i(t) = t_i$ is a character of T. These characters will be useful in the following sections.

For completeness, we mention that there is also a notion of a 'dual' of a character.

Definition 94. An algebraic group homomorphism $\lambda : \mathbb{G}_m \to G$ is called a *cocharacter* of G. The set of cocharacters of G is denoted $X_*(G)$.

Example 95. For the torus T_n, any algebraic group homomorphism $\lambda : \mathbb{G}_m \to T_n$ is given by $a \mapsto \mathrm{diag}(a^{r_1}, \ldots, a^{r_n})$. Hence, it follows that $X_*(T_n) \cong \mathbb{Z}^n$.

It turns out that if we compose a character χ and a cocharacter λ we obtain a homomorphism $\chi \circ \lambda : \mathbb{G}_m \to \mathbb{G}_m$. But we saw in Example 49 that any homomorphism $\mathbb{G}_m \to \mathbb{G}_m$ has the form $x \mapsto x^m$, where $m \in \mathbb{Z}$. This yields a pairing $\langle\ ,\ \rangle : X^* \times X_* \to \mathbb{Z}$ by $(\chi \circ \lambda)(x) = x^{\langle \chi, \lambda \rangle}$.

Theorem 96. *The following are equivalent statements for linear algebraic groups G.*

(i) *G is diagonalisable*

(ii) *$X^*(G)$ is an abelian group of finite type and forms a basis for $k[G]$*

(iii) *Any rational representation $\phi : G \to GL(V)$ of G decomposes into a direct sum of one-dimensional representations.*

Proof. (i)⇒ (ii). We have already seen that if G is diagonalisable, i.e., a subgroup of T_n, then $X^*(G)$ is abelian of finite type and is a linearly independent subset in $k[G]$. The character χ_i (defined above) are a spanning set for $k[G]$ because any polynomial functions is a combination of monomials χ_i and their inverses. (ii) ⇒ (iii). Let $\phi : G \to \mathrm{GL}(V)$ be a rational representation in a finite dimensional vector space V. If we fix a basis for ϕ, then we can write $\phi(g)$ as an $n \times n$ matrix $(\phi_{ij}(g))_{1\leq i,j\leq n}$ with $\phi_{ij} \in k[G]$. From (ii), we can re-write $\phi_{ij} = \sum_{\chi\in X^*(G)} a_{ij}\chi$. Thus, we can define n^2 linear maps $A_\chi : V \to V$ by $A_\chi = (a_{ij\chi})$ for this fixed basis of V. We then obtain the rational representation $\phi(g) = \sum_{\chi\in X^*(G)} A_\chi\chi(g)$. Then, we have

$$\phi(gh) = \sum_{\chi\in X^*(G)} A_\chi\chi(gh) = \phi(g)\phi(h) = \Big(\sum_{\chi\in X^*(G)} A_\chi\chi(g)\Big)\Big(\sum_{\chi\in X^*(G)} A_\chi\chi(h)\Big).$$

Now by Dedekinds lemma applied to $G \times G$, we have $A_\chi A_\alpha = \delta_{\chi,\alpha}A_\chi$ and $\sum_{\chi\in X^*(G)} A_\chi = \phi(1) = id_{\mathrm{GL}(V)}$. We define the subspace of $V = \operatorname{im} A_\chi$. It follows that V is a direct sum of the V_χ and $g \in G$ acts on $v \in V_\chi$ by

$$g.v = \phi(g)v = \sum_{\chi\in X^*(G)} A_\chi\chi(g)v = \chi(g)v.$$

(iii) ⇒ (i). This direction is immediate as soon as we view G as a closed subgroup of GL_n ■

This result motivates the following. If $\phi : T \to \mathrm{GL}(V)$ is a rational representation of a torus T. Then V is a direct sum of one-dimensional subspaces, V_χ, where T acts on each V_χ by multiplication of $\chi(t) \in X^*(T)$. The corresponding spaces

$$V_\chi = \{v \in V \mid \phi(t)v = \chi(t)v \,\forall t \in T\}$$

are called the *weight spaces* and the corresponding χ are the *weights*. When T is a torus contained in G acts on the Lie algebra $\mathfrak{g}$ of $G \supseteq T$ via the adjoint action, the nonzero weight spaces and weights are the *root spaces* and *roots* (relative to G) respectively. In fact, we have the following theorem which we state without proof.

Theorem 97. *Let G be a connected reductive group with T a maximal torus of G and $\mathfrak{g} = \mathrm{Lie}(G)$ with $\Phi = \Phi(G, T)$. Then*

(i) $\mathfrak{g} = \mathrm{Lie}(T) \oplus \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$, *where $\mathfrak{g}_\alpha$ denote the root space corresponding to the root α.*

(ii) $\dim(G) = \dim(\mathfrak{g}) = |\Phi| + \mathrm{rank}(G)$

(iii) *For each $\alpha \in \Phi$ there is a (unique up to scaling) homomorphism of algebraic group $u_\alpha : \mathbb{G}_a \to G$ which induces an isomorphism onto $u_\alpha(\mathbb{G}_a)$ such that $tu_\alpha(a)t^{-1} = u_\alpha(\alpha(t)a)$ for all $t \in T$ and $a \in k$. Furthermore,* $\mathrm{im}(du_\alpha) = \mathfrak{g}_\alpha$.

(iv) *U_α is the unique one-dimensional connected unipotent subgroup of G normalised by T with* $\mathrm{Lie}(U_\alpha) = \mathfrak{g}_\alpha$.

(v) $G = \langle T, U_\alpha \,|\, \alpha \in \Phi \rangle$.

Proof. See Theorem 8.17 in [28]. ■

Example 98 (Roots and root spaces of GL_n)**.** Let $G = \mathrm{GL}_n(k)$ and T the maximal torus in G, where $t \in T$ has the form $t = \mathrm{diag}(t_1, \ldots, t_n)$. For $1 \leq i \leq n$ let $\varepsilon_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{Z}^n$ be the vector with 1 in the i^{th} entry. Then for $m_i \in \mathbb{Z}$, define characters in $X^*(T)$ by $\chi_{\sum_{i=1}^n m_i \varepsilon_i}(t) = t_1^{m_1} \ldots t_n^{m_n}$. If we consider the adjoint action of T on $\mathfrak{g} = \mathrm{Mat}_n(k)$ on the basis elements, $tE_{ij}t^{-1} = t_i t_j^{-1} E_{ij}$, we see that $\chi_{\varepsilon_i - \varepsilon_j}$ is the root for the root space $\mathfrak{g}_\alpha$, $\alpha = \chi_{\varepsilon_i - \varepsilon_j}$, spanned by E_{ij}. Thus, the set of roots $\Phi(G,T) = \{\chi_{\varepsilon_i - \varepsilon_j} \in X^*(G) \,|\, 1 \leq i \neq j \leq n\}$. We can obtain the positive roots by choosing $\Phi^+ = \{\chi_{\varepsilon_i - \varepsilon_j} \,|\, 1 \leq i < j \leq n\}$. Then observe for any positive root

$$\chi_{\varepsilon_i - \varepsilon_j}(t) = t_i t_j^{-1} = t_i t_{i+1}^{-1} t_{i+1} t_{i+2}^{-1} t_{i+2} \ldots t_{j-1}^{-1} t_{j-1} t_j^{-1} = (\chi_{\varepsilon_i - \varepsilon_{i+1}} + \chi_{\varepsilon_{i+1} - \varepsilon_{i+2}} + \cdots + \chi_{\varepsilon_{j-1} - \varepsilon_j})(t).$$

It follows that the set of simple root $\Delta = \{\chi_{\varepsilon_i - \varepsilon_{i+1}} \,|\, 1 \leq i \leq n-1\}$. The corresponding simple root space, for the simple root $\alpha = \varepsilon_i - \varepsilon_{i+1} \in \Delta$, are $\mathfrak{g}_\alpha = \mathrm{span}\{E_{i,i+1}\}$, for all $1 \leq i \leq n-1$.

Example 99 (Roots and root spaces of Sp_4)**.** The symplectic group, $\mathrm{Sp}_m(k)$, is the subgroup defined as

$$\mathrm{Sp}_m(k) = \{g \in \mathrm{GL}_m(k) \,|\, g^t M g = M\},$$

where m is even, g^t denotes the transpose of g and M a non-degenerate bilinear form. Here, we consider the example when $m = 4$ with bilinear form

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

A diagonal matrix $t \in \mathrm{GL}_4(k)$ preserves the bilinear form if and only if $tMt = M$. Since

$$tMt = \begin{pmatrix} 0 & 0 & 0 & t_1 t_4 \\ 0 & 0 & -t_2 t_3 & 0 \\ 0 & t_3 t_2 & 0 & 0 \\ -t_4 t_1 & 0 & 0 & 0 \end{pmatrix}$$

we have that $t_4 = t_1^{-1}$ and $t_3 = t_2^{-1}$. We then have a maximal torus, T, of rank 2 given by

$$T := G \cap D_4 = \left\{ \begin{pmatrix} t_1 & 0 & 0 & 0 \\ 0 & t_2 & 0 & 0 \\ 0 & 0 & t_2^{-1} & 0 \\ 0 & 0 & 0 & t_1^{-1} \end{pmatrix} : t_1, t_2 \in k^\times \right\}.$$

We define characters similar to the preceding example. That is, define a character $\chi_{k_1\varepsilon_1 + k_2\varepsilon_2} : T \to \mathbb{G}_m$ by $\chi_{k_1\varepsilon_1 + k_2\varepsilon_2}(t) = t_1^{k_1} t_2^{k_2}$, where $t \in T$ is a diagonal matrix with entries t_i for $i \in \{1, 2\}$. The Lie algebra of Sp_4 is given by 4×4 matrices satisfying $g^t M + Mg = 0$. A general matrix in this Lie algebra takes the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & -a_{13} \\ a_{31} & a_{32} & -a_{22} & a_{12} \\ a_{41} & -a_{31} & a_{21} & -a_{11} \end{pmatrix}.$$

We will denote this Lie algebra of Sp_4 by $\mathfrak{g}$.

We denote the 4×4 matrix $t \in T$ by $t = \mathrm{diag}(t_1, t_2)$, where $t_1, t_2 \in k^\times$ are the $(1,1)$ and $(2,2)$ entries in the matrix respectively. We obtain the roots of $\mathfrak{g}$ via the adjoint action of T on $\mathfrak{g}$, in particular, on the span of a basis element as follows:

$$t(xE_{14})t^{-1} = t_1^2 xE_{14} = \chi_{2\varepsilon_1}(xE_{14}).$$

So, $2\chi_{\varepsilon_1}$ is a root. We also have

$$t(xE_{23})t^{-1} = t_1^2 xE_{23} = \chi_{2\varepsilon_2}(t)(xE_{23}),$$

$$t(xE_{13} - xE_{24})t^{-1} = t_1 t_2 (xE_{13} - xE_{24}) = \chi_{\varepsilon_1+\varepsilon_2}(xE_{13} - xE_{24}),$$

$$t(xE_{12} + xE_{34})t^{-1} = t_1 t_2 (xE_{12} + xE_{34}) = \chi_{\varepsilon_1-\varepsilon_2}(xE_{12} + xE_{34}).$$

For brevity we will omit $'\chi'$ and write $k_1\varepsilon_1 + k_2\varepsilon_2$ instead of $\chi_{k_1\varepsilon_1+k_2\varepsilon_2}$. Thus, we have obtained all roots $\Phi = \{\pm 2\varepsilon_1, \pm 2\varepsilon_2, \pm\varepsilon_1 + \varepsilon_2, \pm\varepsilon_1 - \varepsilon_2\}$. A choice of positive roots $\Phi^+ = \{\varepsilon_1 + \varepsilon_2, \varepsilon_1 - \varepsilon_2, 2_{\varepsilon_1}\}$ yields the simple roots $\Delta = \{\varepsilon_1 - \varepsilon_2, 2\varepsilon_2\}$. The corresponding simple root spaces are given by $\mathfrak{g}_{\varepsilon_1-\varepsilon_2} = \mathrm{span}\{E_{12} + E_{34}\}$ and $\mathfrak{g}_{2\varepsilon_2} = \mathrm{span}\{E_{23}\}$, where E_{ij} is the 4×4 matrix with 1 in the (i,j) entry and zeros elsewhere.

We will study roots and their properties more deeply in the next section.

7. Chevalley Restriction Theorem

In this section we discuss Chevalley's restriction theorem and give a proof. We follow Humphrey's approach [18] where the representation theory of weight modules of semisimple Lie algebras is utilised. We will attempt to provide the reader enough of the theory of the representation theory of Lie algebras, over algebraically closed field k of zero characteristic, to fully appreciate the main result of this section. However, for a review on the basic theory of Lie algebras and their representation theory we refer the reader to [18], [31] and [32] among the many references.

7.1. **Some Lie Theory.** Recall that a representation of a Lie algebra, $\mathfrak{g}$, is a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{gl}(V)$, where V is a finite dimensional vector space over an algebraically closed field k with characteristic zero. The notion of a representation of a Lie algebra $\mathfrak{g}$ and a $\mathfrak{g}$-module is interchangeable since we can retrieve one from the other. Hence, we may study the representation theory using the language of modules. The representation space V can be infinite-dimensional but for some results this will require more caution. We will consider the *adjoint representation* $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ which takes $x \in \mathfrak{g}$ to ad_x where ad_x acts on $\mathfrak{g}$ via $\mathrm{ad}_x(y) = [x, y]$, with $[\cdot, \cdot]$ being the *Lie bracket* on $\mathfrak{g}$. An *automorphism* of $\mathfrak{g}$ is an isomorphism of $\mathfrak{g}$ onto itself and we denote the group of all automorphisms by $\mathrm{Aut}(\mathfrak{g})$. For an element $x \in \mathfrak{g}$ such that ad_x is nilpotent, the exponential of this linear transformation terminates. More explicitly, we have $\exp(\mathrm{ad}_x) = 1 + \mathrm{ad}_x + \frac{1}{2}(\mathrm{ad}_x)^2 + \frac{1}{3!}(\mathrm{ad}_x)^2 + \cdots + \frac{1}{(k-1)!}(\mathrm{ad}_x)^{k-1}$, where $(\mathrm{ad}_x)^k = 0$. Then it is the case that $\exp(\mathrm{ad}_x) \in \mathrm{Aut}(\mathfrak{g})$. We call automorphisms of this form *inner automorphisms*. The subgroup $\mathrm{Inn}(\mathfrak{g}) \subset \mathrm{Aut}(\mathfrak{g})$ generated by these is a normal subgroup.

Example 100. If we consider $\mathfrak{g} = \mathfrak{sl}_2(k)$, with its standard basis $\{x, y, h\}$ given by

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Define $\sigma \in \mathrm{Inn}(\mathfrak{g})$ by $\sigma = \exp(\mathrm{ad}_x)\exp(\mathrm{ad}_{-y})\exp(\mathrm{ad}_x)$. One can compute the effect of σ on the basis by straight computations and we see that $\sigma(x) = -y$, $\sigma(y) = -x$, $\sigma(h) = -h$. Furthermore, σ has order 2.

This example will be important when we consider the action of the Weyl group later.

Let $\mathfrak{g}$ a semisimple Lie algebra and let $\mathfrak{h}$ be the *Cartan subalgebra* of $\mathfrak{g}$, i.e., a nilpotent subalgebra equal to its normaliser. It is a known fact that $\mathfrak{h}$ is abelian [18]. Then $\mathrm{ad}(\mathfrak{h})$ is a commuting family of semisimple endomorphisms of $\mathfrak{g}$ and hence $\mathrm{ad}_\mathfrak{h}$ is simultaneously diagonalisable. This gives rise to a decomposition of the Lie algebra $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus \mathfrak{g}_\alpha$, with $\mathfrak{g}_\alpha = \{x \in \mathfrak{g} \mid [h, x] = \alpha(h)x, \forall h \in \mathfrak{h}\}$ and $\alpha \in \mathfrak{h}^*$. The set of all $\alpha \in \mathfrak{h}^*$ such that $\mathfrak{g}_\alpha \neq 0$ is called the set of *roots* of $\mathfrak{g}$, whose collection as a set we denote by Φ, and the subspaces $\mathfrak{g}_\alpha$ are the *root spaces*. Then we have the well known root space decomposition $\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Phi} g_\alpha$.

More generally, let V be a $\mathfrak{g}$-module and $\lambda : \mathfrak{h} \to \mathbb{C}$. We denote by V_λ the set of eigenvectors with eigenvalue λ for the action of $\mathfrak{h}$ on V, induced by some representation. The eigenvalue λ is the *weight* and the eigenspace V_λ is the *weight space*.

Recall the *Killing form* on a semisimple Lie algebra $\mathfrak{g}$ is given by $\kappa(x, y) = \mathrm{tr}(\mathrm{ad}_x \mathrm{ad}_y)$. It is a symmetric, associative nondegenerate form on $\mathfrak{g}$. It is a standard result that the restriction of κ to $\mathfrak{h}$ is nondegenerate. This induces a vector space isomorphism $\mathfrak{h} \cong \mathfrak{h}^*$ which sends $\mathfrak{h}^* \ni \varphi \mapsto t_\varphi$ such that $\varphi(h) = \kappa(t_\phi, h)$ for all $h \in \mathfrak{h}$. Using the Killing form, one can show that embedded in $\mathfrak{g}$ are copies of $\mathfrak{sl}(2, k)$. This done by taking $x_\alpha \in \mathfrak{g}_\alpha$, $y_\alpha \in \mathfrak{g}_{-\alpha}$ and defining $h_\alpha = [x_\alpha, y_\alpha]$. There are some results that need to be shown, for example: If $\alpha \in \Phi$ then it is true that $-\alpha \in \Phi$ and $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = kt_\alpha$. It is then the case that x_α, y_α and h_α span a three-dimensional subalgebra isomorphic to $\mathfrak{sl}_2(k)$. Since h_α must be a multiple of t_α, we must have $h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$. It is also easy to see that $t_{-\alpha} = -t_\alpha$ and therefore, $-h_\alpha = h_{-\alpha}$. It is convenient to denote this subalgebra isomorphic to $\mathfrak{sl}_2(k)$ by S_α.

Definition 101. If $\alpha \in \Phi$ and $\beta \in \Phi \cup \{0\}$, the *$\alpha$-root string through β* is the S_α-module

$$M = \bigoplus_{c \in k} \mathfrak{g}_{\beta + c\alpha}\,.$$

It turns out that $c \in k$ can only take integer values. In fact, we have the following proposition.

Proposition 102. *If $\alpha \in \Phi$ then $\mathfrak{g}_\alpha$ is one-dimensional and the only multiples of α which are roots are $\pm\alpha$.*

Proof. We only provide a sketch of the proof. Consider the α-root string through 0

$$M = \mathfrak{h} \oplus \bigoplus_{c\alpha \in \Phi} \mathfrak{g}_{c\alpha}\,.$$

Note that we can decompose $\mathfrak{h} = \ker \alpha \oplus kh_\alpha$, where k is the field and use that $\ker \alpha$ is an S_α module. One can then take the quotient

$$M/\ker \alpha \cong W = kh_\alpha \oplus \bigoplus_{c\alpha} g_{c\alpha}$$

which is also an S_α module whose weight space with weight 0 is $W_0 = kh_\alpha$. This is one-dimensional so there can only be one irreducible V_{2n} ($n \in \mathbb{N}$) module, which is $W \supset S_\alpha \cong V_2$. This implies that $\mathfrak{g}_{2\alpha} = 0$ so twice a root is not a root, and so $\frac{1}{2}\alpha$ is not a root since α is a root. Thus, $W = V_2$ is irreducible, hence $\mathfrak{g}_\alpha$ is one-dimensional. ■

Corollary 103. *If $\alpha, \beta \in \Phi$ and $\beta \neq \pm\alpha$ then there is an α-root string through β which is irreducible and has the form*

$$V_m \cong \mathfrak{g}_{\beta+q\alpha} \oplus \mathfrak{g}_{\beta+(q-1)\alpha} \oplus \cdots \oplus \mathfrak{g}_{\beta-r\alpha}$$

and $m = q + r$, and $\beta(h_\alpha) = r - q \in \mathbb{Z}$.

Proof. The proof of this corollary follows from the preceeding proposition. If M is an α-root string through β then M is a direct sum of V_{2n+1} ($n \in \mathbb{N}$). Noting that a difference in h_α-weights corresponds to a difference of α in roots since $\alpha(h_\alpha) = 2$, then the root string only contains $\mathfrak{g}_{\beta+j\alpha}$ for integer j. One of these is M_1 which is one-dimensional, hence M irreducible. Then if $M \cong V_m$, then $V_m = \mathfrak{g}_{\beta+q\alpha}$ and $M_{-m} = \mathfrak{g}_{\beta-r\alpha}$ and the dimension of M is $1 + r + 1 = m + 1$. Then, we have

$$q + r = m = (\beta + q\alpha)(h_\alpha) = \beta(h_\alpha) + q(\alpha_h(h_\alpha)) = \beta(h_\alpha) + 2q.$$

It follows that $\beta(h_\alpha) = r - q \in \mathbb{Z}$. ■

We are now in a good position to put an inner product on $\mathfrak{h}^*$. Since $\beta(h_\alpha) \in \mathbb{Z}$, we can reformulate this using that $h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$, and by definition of t_β, we have

$$\beta(h_\alpha) = k(t_\beta, h_\alpha) = \frac{2\kappa(t_\beta, t_\alpha)}{\kappa(t_\alpha, t_\alpha)} = \frac{2(\beta, \alpha)}{(\alpha, \alpha)}.$$

Corollary 104. *If $\alpha, \beta \in \Phi$. Then*

$$\beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha \in \Phi$$

Proof. Since

$$\frac{2(\beta, \alpha)}{(\alpha, \alpha)} = \beta(h_\alpha) = q - r$$

this implies $\beta - (q - r)\alpha \in \Phi$ because $q \geq q - r \geq -r$, so $\beta - (q - r)\alpha$ occurs as a root in the α-root string through β. ■

One can describe a reflection about a hyperplane as a linear transformation which fixes pointwise vectors on a hyperplane and sends any vector orthogonal to the hyperplane to its negative. For a nonzero vector α the reflecting hyperplane is given by $P_\alpha = \{\beta \in V \mid (\alpha, \beta) = 0\}$, and the induced reflection about P_α can be written as $\sigma_\alpha(\beta) = \beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha$. It can be checked $\alpha \mapsto -\alpha$ and leaves P_α invariant. We will define $\langle \beta, \alpha \rangle := \frac{2(\alpha,\alpha)}{(\alpha,\alpha)}$. The subgroup of $\mathrm{GL}(V)$ generated by the reflections σ_α for $\alpha \in \Phi$ is called the *Weyl group* of Φ, we denote it by W.

We can summarise the above results as follows.

Theorem 105. *Let $\mathfrak{g}$, $\mathfrak{h}$ and Φ be as above and E a Euclidean vector space of dimension $|\Phi|$. Then*

(i) *If $\alpha \in \Phi$ then $-\alpha \in \Phi$ and no other multiple belongs to Φ.*

(ii) *If $\alpha, \beta \in \Phi$, then $\beta - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \Phi$.*

(iv) *Φ spans E and $0 \notin \Phi$.*

Proof. We have proved (i)-(iii) already. There is a discussion in [18] about (iv). ■

The above theorem motivates the following definition.

Definition 106. A subset Φ of the Euclidean space E is called a *root system* in E if the following hold:

(i) Φ spans E and $0 \notin \Phi$.

(ii) If $\alpha \in \Phi$ then $-\alpha \in \Phi$ and no other multiple belongs to Φ.

(iii) If $\alpha \in \Phi$, the reflection σ_α leaves Φ invariant.

(iv) If $\alpha, \beta \in \Phi$, then $\langle \beta, \alpha \rangle \in \mathbb{Z}$.

By (iii) of the above definition, the Weyl group W leaves Φ invariant. We also have the important definition.

Definition 107. A subset Δ of Φ is called a *base* if:

(i) Δ is a basis of E,

(ii) each root β can be written as $\beta = \sum c_\alpha \alpha$ for $\alpha \in \Delta$ and $c_\alpha \in \mathbb{Z}$ all integral coefficients, either all nonnegative or all nonpositive.

The roots in Δ are called *simple*. If all c_α are nonnegative, then β is a *positive root*, or if all the c_α are nonpositive then β is a *negative root*.

It is not obvious that every root system has a base. Nonetheless, this is true.

Theorem 108. *Φ has a base.*

Proof. See [18] Theorem 10.1. ■

Lemma 109. *Let α be a simple root, then σ_α permutes the positive roots except α.*

Proof. Let $\beta \in \Phi$ be a positive root. That is, $\beta = \sum_i c_i \alpha_i$, with $\alpha_i \in \Delta$, and at least one of the $c_i > 0$, say c_2. If we compute $\sigma_\alpha(\beta)$, where we label α as α_1, we obtain

$$\sigma_\alpha(\beta) = (c_1 - \langle \beta, \alpha \rangle)\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n.$$

Since $c_2 > 0$, this forces β to be positive. ■

For each $\alpha \in \Phi$, the hyperplanes P_α partition E into finitely many connected components $E \setminus \bigcup_\alpha P_\alpha$. These components are called the *Weyl chambers* of E. There is a nonempty set C containing all elements $x \in E$ such that $(x, \alpha) > 0$ for all positive roots $\beta \in \Phi$. The set C is a Weyl chamber called the *fundamental chamber*. It is easy to see that the Weyl group permutes the Weyl chambers. In fact, W acts transitively on the Weyl chambers [18]. Furthermore, it can be shown that each point in E is W-conjugate to a point in the closure of the fundamental Weyl chamber relative to a base Δ.

Root systems play a crucial role in the classification of semisimple Lie algebras. One can show, that for two semisimple Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ with isomorphic root systems Φ and Φ' then there Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ are isomorphic [18]. We can use this result to prove the existence of automorphisms of semisimple Lie algebras. For example, if we take the automorphism $\sigma \in \mathrm{Aut}(\Phi)$ which sends a root to its negative, then any automorphism of Φ can be extended to an automorphism of $\mathfrak{g}$. By construction, we have the following proposition.

Proposition 110. *Let $\mathfrak{g}$ be a semisimple Lie algebra over k, k algebraically closed with zero characteristic. Fix a nonzero $x_\alpha \in \mathfrak{g}_\alpha$ ($\alpha \in \Delta$) and let $y_\alpha \in \mathfrak{g}_{-\alpha}$ satisfy $[x_\alpha, y_\alpha] = h_\alpha$, i.e., $\{x_\alpha, y_\alpha, h_\alpha\}$ is a copy of $\mathfrak{sl}_2(k)$. Then there is an automorphism σ of $\mathfrak{g}$, of order 2, satisfying $\sigma(x_\alpha) = -y_\alpha$, $\sigma(y_\alpha) = -x_\alpha$, $\sigma(h) = -h$.*

Note that we discussed this in Example 100, with $\sigma = \exp(\mathrm{ad}_x)\exp(\mathrm{ad}_{-y})\exp(\mathrm{ad}_x)$ being the automorphism in the above theorem.

It turns out that Δ defines a partial ordering on E: we say that $\mu > \lambda$ if $\mu - \lambda$ is a sum of positive roots, or $\mu = \lambda$.

We now give a discussion on weights. Let Λ be the set of all $\lambda \in E$ for which $\langle \lambda, \alpha \rangle \in \mathbb{Z}$ ($\alpha \in \Phi$). An element $\lambda \in \Lambda$ is a *weight*. Fix a base $\Delta \subset \Phi$, then $\lambda \in \Lambda$ is *dominant* if all integers $\langle \lambda, \alpha \rangle$ ($\alpha \in \Delta$) are nonegative and *strongly dominant* if these integers are positive. The set of all dominant weights is denoted by Λ^+

Lemma 111. *Each weight is conjugate under W to one and only one dominant weight. If $\lambda \in \Lambda^+$, then $\sigma(\lambda) < \lambda$ for all $\sigma \in W$, and if λ is strongly dominant, then $\sigma\lambda = \lambda$ only when $\sigma = 1$.*

Proof. Firstly, we need to enlarge the partial ordering on the Euclidean space E by deeming that $\mu < \lambda$ if and only if $\lambda - \mu$ is a nonnegative $\mathbb{R}$-linear combination of simple roots. Now consider the fundamental dominant chamber $\mathcal{C}(\Delta) = \{\lambda \in \mathfrak{h}^* \mid (\lambda, \alpha_i) \geq 0,\ \forall i\}$. Take any weight μ and consider its W-orbit $\{w\mu\}_{w \in W}$. Choose the maximal $\lambda \in \{w\mu\}_{w \in W}$, we claim that $\lambda \in \mathcal{C}(\Delta)$. If not, then there is an i such that $\langle \lambda, \alpha_i \rangle < 0$. But

$$\sigma_{\alpha_i}(\lambda) = \lambda - \langle \lambda, \alpha_i \rangle \alpha_i = \lambda + c\alpha_i,$$

where $c_i \in \mathbb{R}_{>0}$, which belongs to $\{w\mu\}_{w \in W}$ and is larger in the ordering, contradicting maximality of λ. Since simple reflections generate W, this shows that any weight μ is W-conjugate to a dominant weight. Now suppose $\lambda = w\mu$ for some $w \in W$ and $\lambda, \mu \in \mathcal{C}(\Delta)$. We want to show that $\lambda = \mu$. For $\alpha_i \in \Delta$, consider

$$\sigma_{\alpha_i}(\lambda) = \lambda - \langle \lambda, \alpha_i \rangle \alpha_i$$

and

$$\langle \sigma_{\alpha_i}(\lambda), \alpha_i \rangle = \langle \lambda, \alpha_i \rangle - \langle \lambda, \alpha_i \rangle \langle \alpha_i, \alpha_i \rangle = -\langle \lambda, \alpha_i \rangle.$$

Hence $\sigma_{\alpha_i}(\lambda) \in \mathcal{C}(\Delta)$ if and only if $\langle \lambda, \alpha_i \rangle = 0$ if and only if $\sigma_{\alpha_i}(\lambda) = \lambda$. Since any $w \in W$ is a product of simple reflections this shows $w\lambda \in \mathcal{C}(\Delta)$ if and only if $w\lambda = \lambda$, i.e., each weight is conjugate to only one dominant weight. ■

Lemma 112. *Let $\lambda \in \Lambda^+$. Then the number of dominant weights $\mu < \lambda$ is finite.*

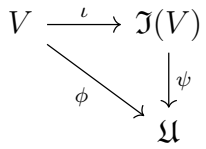
Proof. The dominant weight μ lies in the intersection of the compact set $\{x \in E \mid (x, x) \leq \lambda, \lambda)\}$ and Λ^+, the latter being a discrete set. Indeed, since $\lambda + \mu \in \Lambda^+$ and $\lambda - \mu$ is a sum of positive roots, we have $0 \leq (\lambda + \mu, \lambda - \mu) = (\lambda, \lambda) - (\mu, \mu)$. ■

Definition 113. We say that $\lambda \in \Lambda^+$ is minimal if for $\mu \in \Lambda^+$, with $\mu \leq \lambda$, we have that $\mu = \lambda$.

For finite dimensional $\mathfrak{g}$-modules, V, we know that $\mathfrak{h}$ acts diagonally on V and so we call the eigenspaces and eigenvectors *weight spaces* and *weights* respectively. If we allow V to be infinite-dimensional, then these notions are still well-defined. That is, whenever $V_\lambda \neq 0$, we call V_λ a *weight space* and λ its *weight* of $\mathfrak{h}$ on V. It is not necessarily true for V infinite-dimensional that V decomposes into a sum of its weight spaces. However, the sum of the weight spaces V_λ is always a direct sum. Furthermore, if V' is the direct sum of the weight spaces then V' is a $\mathfrak{g}$-submodule of V since $\mathfrak{g}_\alpha$ permutes the weight spaces. For example $h.x.v = x.h.v + [h, x].v = (\lambda(h) + \alpha(h))x.v$, meaning that V_λ is sent to $V_{\lambda+\alpha}$ under $\mathfrak{g}_\alpha$ for all $h \in \mathfrak{h}$, $x \in \mathfrak{g}_\alpha$ and $\lambda \in V_\lambda$.

We will call a nonzero vector $v \in V_\lambda$ *maximal* if for all $\mathfrak{g}_\alpha$, $\alpha \in \Delta$, $\mathfrak{g}_\alpha . v = 0$. It is not necessarily the case that maximal vectors exist when $\dim V = \infty$. To study finite-dimensional irreducible $\mathfrak{g}$-modules we look at g-modules generated by a maximal vector. Firstly, we recall some basic facts about universal enveloping algebras.

Let V be a fixed finite dimensional vector space over an arbitrary field k. Define the *tensor algebra*, $\mathfrak{I}(V)$, on V by $\mathfrak{I}(V) = \bigsqcup_i T^i V$, where $T^i V = V \otimes \cdots \otimes V$ (i-times) with $T^0 V = k$. The tensor algebra, $\mathfrak{I}(V)$, on V is the universal associative algebra satisfying the universal property: given any k-linear map $\phi : V \to \mathfrak{U}$, where $\mathfrak{U}$ is an associative unital algebra, there exists a unique k-algebra homomorphism $\psi : \mathfrak{I}(V) \to \mathfrak{U}$ such that the following diagram commutes:

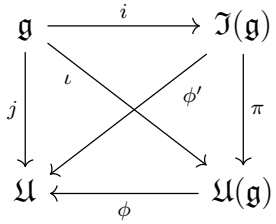


It is also interesting to take the ideal, I, generated by $I = \langle x \otimes y - y \otimes x \rangle$, $x, y \in V^*$, and to consider the quotient $\mathfrak{P} = \mathfrak{I}(V^*)/I$. This is nothing but a commutative associative algebra that is identifiable as a polynomial algebra in $n = \dim V$ variables.

The *universal enveloping algebra* of $\mathfrak{g}$ is a pair $(\mathfrak{U}, i)$, where $\mathfrak{U}$ is an associative algebra with 1 over k and $i : \mathfrak{g} \to \mathfrak{U}$ is a linear map satisfying

$$i([x, y]) = i(x)i(y) - i(y)i(x) \qquad (*)$$

for $x, y \in \mathfrak{g}$ and, moreover: for any associative, unital k-algebra $\mathcal{U}$ and any linear map $j : \mathfrak{g} \to \mathcal{U}$ satisfying the above $(*)$ equation, there exists a unique homomorphism of algebras $\phi : \mathfrak{U} \to \mathcal{U}$ such that $\phi \circ i = j$. The existence of such a pair can be seen as follows. Let $\mathfrak{I}(\mathfrak{g})$ be the tensor algebra on $\mathfrak{g}$, and let J be the two-sided ideal generated by all $J = \langle x \otimes y - y \otimes x - [x, y] \rangle$, where $x, y \in \mathfrak{g}$. Then define $\mathfrak{U} = \mathfrak{I}(\mathfrak{g})/J$ and $\pi : \mathfrak{I}(\mathfrak{g}) \to \mathfrak{U}(\mathfrak{g})$ to be the canonical homomorphism. The claim is that $(\mathfrak{U}, i)$ is a universal enveloping algebra where i is the restriction of π to $\mathfrak{g}$. Indeed, let $j : \mathfrak{g} \to \mathfrak{U}$ be as in the definition, i.e., satisfying $(*)$. The universal property of $\mathfrak{I}(\mathfrak{g})$ gives rise to an algebra homomorphism $\phi' : \mathfrak{I}(\mathfrak{g}) \to \mathfrak{U}$ such that $j = \phi' \circ i$. The property $(*)$ of j show that all elements $x \otimes y - y \otimes x - [x, y]$ lie in the kernel $\ker \phi'$, and so induces a homomorphism $\phi : \mathfrak{U}(\mathfrak{g}) \to \mathfrak{U}$ such that $\phi \circ i = j$. The uniqueness of ϕ follows from the fact that 1 and the image $\operatorname{im} i$ generate $\mathfrak{U}(\mathfrak{g})$. The proof of the claim can also be seen by the following diagram:



There is no significant amount of insight about the structure of a universal enveloping algebra for a finite-dimensional Lie algebra $\mathfrak{g}$ from these definitions. It turns out, however, that one can give an explicit definition for a basis of this enveloping algebra, this result is due to the Poincare-Birkov-Witt theorem (and its corollary) that a universal enveloping algebra on $\mathfrak{g}$ has *PBW-basis*:

Theorem 114. *Let $\{v_1, \ldots, v_n\}$ be a basis for the vector space $\mathfrak{g}$. Then the set*

$$\{v_1^{x_1} v_2^{x_2} \ldots x_n^{v_n} \mid x_i \in \mathbb{N}\},$$

forms a basis for $\mathfrak{U}(\mathfrak{g})$.

Proof. See Theorem 2.1.11 in [33]. ■

In particular, for a semisimple Lie algebra $\mathfrak{g}$, one obtains a useful characterisation of the universal enveloping algebra, $\mathfrak{U}(\mathfrak{g})$, on $\mathfrak{g}$ by reordering the PBW-basis in such a way that $\mathfrak{U}(\mathfrak{g}) = \mathfrak{U}(\mathfrak{n}^-) \otimes \mathfrak{U}(\mathfrak{h}) \otimes \mathfrak{U}(\mathfrak{n}^+)$, where $\mathfrak{n}^+$, $\mathfrak{n}^-$ are the span of the positive and negative root spaces and $\mathfrak{h}$ is the Cartan subalgebra [33].

We are now able to introduce the notion of a standard cyclic module for semisimple Lie algebras. That is, we say that, for a maximal weight vector, v^+, of weight λ, $V = \mathfrak{U}(\mathfrak{g})v^+$ is *standard cyclic* of weight λ and call λ the highest weight of V. Our goal is to describe such modules. Firstly, let $x_\alpha \in \mathfrak{g}_\alpha$ $\alpha \in \Phi^+$, and $y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $[x_\alpha, y_\alpha] = h_\alpha$. We have the following theorem from [18].

Theorem 115. *Let V be a standard cyclic module with maximal vector $v^+ \in V_\lambda$ and let $\Phi^+ = \{\beta_1, \ldots, \beta_m\}$. Then:*

(a) *V is spanned by the vectors $y_{\beta_1}^{i_1} \ldots y_{\beta_m}^{i_m}$, $i_k \in \mathbb{Z}^+$; in particular, V is the direct sum of its weight spaces.*

(b) *The weights of V are of the form $\mu = \lambda - \sum_{i=1}^{l} k_i \alpha_i$, $k_i \in \mathbb{Z}^+$, i.e., all weights satisfy $\mu < \lambda$.*

(c) *For each $\mu \in \mathfrak{h}^*$, $\dim V_\mu < \infty$ and $\dim V_\lambda = 1$.*

(d) *Every submodule of V is the direct sum of its weight spaces.*

(e) *V is an indecomposable $\mathfrak{g}$-module, with a unique maximal proper submodule and a corresponding unique irreducible quotient.*

Proof.

(a) Recall that we have the decomposition $\mathfrak{g} = \mathfrak{n}^- + \mathfrak{b}$, where $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+$ is the Borel subalgebra. Hence, from the PBW theorem $\mathfrak{U}(\mathfrak{g}).v^+ = \mathfrak{U}(\mathfrak{n}^-)\,\mathfrak{U}(\mathfrak{b}).v^+ = \mathfrak{U}(\mathfrak{n}^-)k.v^+$ since v^+ is a common eigenvector for $\mathfrak{b}$ and $b.v = (h+n).v^+ = h.v + n.v^+ = \lambda v^+$, where $h \in \mathfrak{h}$, $n \in \mathfrak{n}^+$. But $\mathfrak{U}(\mathfrak{n}^-)$ has a basis of vectors of the form $y_{\beta_1}^{i_1} \ldots y_{\beta_m}^{i_m}$, hence (a).

(b) The vector $y_{\beta_1}^{i_1} \dots y_{\beta_m}^{i_m}.v^+$ has weight $\lambda - \sum_{j=1}^m i_j\beta_j(h)$, which is a consequence of $h.y.v^+ = y.h.v^+ + [h,y].v^+ = (\lambda(h) - \alpha(h))y.v$ Hence, if we write β_j as nonnegative integer linear combinations of simple roots $\beta \in \Delta$, we obtain (b).

(c) There are only finite number of vectors of the form $y_{\beta_1}^{i_1} \dots y_{\beta_m}^{i_m}.v^+$ for which

$$\sum_{j=1}^{m} i_j\beta_j = \sum_{i=1}^{l} k_i\alpha_i,$$

$\alpha_i \in \Delta$, $k_i \in \mathbb{Z}^+$, and by (a) these span the weight space V_μ, if $\mu = \lambda - \sum k_i\alpha_i$. Furthermore, the only vector with weight λ is v^+, hence $\dim V_\lambda = 1$.

(d) Let W be a submodule of V. Write each $W \ni w = v_1 + \dots + v_n$ where $v_i \in V_{\mu_i}$ where μ_i's are distinct weights. We want to show that each v_i lie in W. Suppose not, then we can write $w = v_1 + \dots + v_n$ such that each $v_i \notin W$ and $n > 1$ is minimal. Then, since all weights are distinct, we can find a $h \in \mathfrak{h}$ so that $\mu_1(h) \neq \mu_2(h)$. Then $W \ni h.w = \sum_{i=1}^n \mu_i(h)v_i$ and so is

$$(h - \mu_1(h).1).w = \sum_{i=2}^{n} (\mu_i(h) - \mu_1(h))v_i.$$

But the choice of w forces $v_2 \in W$ which is a contradiction.

(e) From (c) and (d) we deduce that each proper submodule of V lies in the sum of weight spaces other than V_λ (since $\dim V_\lambda = 1$), so the sum of all such submodules is proper. Hence, V has a unique maximal submodule with a corresponding unique irreducible quotient. Recall that a submodule is maximal if and only if its quotient is irreducible. Hence, V cannot be the direct sum of two proper submodules since each of these is contained in W. Thus, V is indecomposable.

■

We do not yet know that standard cyclic modules of highest weight $\lambda \in \mathfrak{h}^*$ exist and are unique. Thus, we will show uniqueness and provide a method of construction for any $\lambda \in \mathfrak{h}^*$.

Theorem 116. *Let V, W be standard cyclic modules of highest weight λ. If V and W are irreducible then they are isomorphic.*

Proof. Consider the $\mathfrak{g}$-module defined as $X = V \oplus W$. If v^+, w^+ are the respective maximal vectors for V and W with weight λ, then it is easy to see that λ is the weight of the maximal vector $x^+ = v^+ + w^+$ in X. Now, define the submodule of X as the standard cyclic module Y generated by x^+. Let $p_V : Y \to V$ and $p_W : Y \to W$ be the projection maps which are also easily seen to be $\mathfrak{g}$-module homomorphisms. Furthermore, the images of the projection maps are $\operatorname{im} p_V = V$ and $\operatorname{im} p_W = W$. By the first isomorphism theorems $Y/W \cong V$ and $Y/V \cong W$, but since V and W are irreducible then so are Y/W and Y/V. But, by uniqueness of irreducible quotients of standard cyclic modules (Theorem 115 part(e)), we have that V and W are isomorphic. ■

Having shown uniqueness, we now turn to existence. We use a process of induction, similar to that of induced representation in the representation theory of finite groups. This is motivated by Theorem 115 (c) that a standard cyclic module, which we view as a $\mathfrak{b}$-module, contains a one-dimensional submodule spanned by the maximal vector. This leads us to the following: take a one-dimensional vector space D_λ spanned by a vector v^+, and define a $\mathfrak{b}$ action on D_λ by $\left(h + \sum_{\alpha\in\Phi^+} x_\alpha\right).v^+ = h.v^+ = \lambda(h)v^+$, for a fixed $\lambda \in \mathfrak{h}^*$. This yields a $\mathfrak{b}$-module structure on D_λ. This implies that we have a $\mathfrak{U}(\mathfrak{b})$-module structure on D_λ, hence we can form the tensor product $Z(\lambda) = \mathfrak{U}(\mathfrak{g}) \otimes_{\mathfrak{U}(\mathfrak{b})} D_\lambda$, which is now a $\mathfrak{U}(\mathfrak{g})$-module under the natural left action of $\mathfrak{U}(\mathfrak{g})$.

Theorem 117. *$Z(\lambda)$ is the standard cyclic module of weight λ.*

Proof. Clearly, $1 \otimes v^+$ generates $Z(\lambda)$. The vector $1 \otimes v^+$ is nonzero since $\mathfrak{U}(\mathfrak{g})$ is a free $\mathfrak{U}(\mathfrak{b})$-module with basis consisting of the monomials $y_{\beta_1}^{i_1} \dots y_{\beta_m}^{i_m}$, $\beta_j \in \Phi^+$ for all j and $i_j \in \mathbb{Z}^+$. Hence, $1 \otimes v^+$ is a maximal vector of weight λ. ■

Corollary 118. *Let $\lambda \in \mathfrak{h}^*$. Then there exists an irreducible standard cyclic module, $V(\lambda)$, of weight λ.*

Proof. The construction of $Z(\lambda)$ above yields a standard cyclic module. It has a unique maximal submodule, $Y(\lambda)$ by Theorem 115 (d), taking the quotient $V(\lambda) = Z(\lambda)/Y(\lambda)$ yields an irreducible standard cyclic module of weight λ by Theorem 115 (e). ■

Although it is not strictly relevant for the purposes of this section, in the study of these standard cyclic modules, there are two main questions one concernes themselves with: (1) To determine which of the $V(\lambda)$ are finite dimensional; and (2) To determine for a given $V(\lambda)$, the exact weights that occur and their multiplicities. We will state a sufficient result for this first question without proof. If the reader is interested and/or, is not familiar with the proof of these results, we refer them to Theorem 21.2 in [18].

If we assume that V is a finite dimensional irreducible $\mathfrak{g}$-module, then V has at least one maximal vector which generates all of V and hence is isomorphic to $V(\lambda)$. For each simple root $\alpha_i \in \Delta$, if we let $S_{\alpha_i} := S_i$ be a copy of $\mathfrak{sl}_2(k)$, then $V(\lambda)$ becomes a finite-dimensional module for S_i with a maximal vector which coincides with the maximal vector for $\mathfrak{g}$. Then, since there is a maximal vector of weight λ, the weight for the Cartan subalgebra $\mathfrak{h}_i \subset S_i$ is determined by value $\lambda(h_i)$, $h_i = h_{\alpha_i}$. But from the theory of highest weight representations for $\mathfrak{sl}_2$, we know that the highest weight is a nonnegative integer, hence $\lambda(h_i)$ is a nonnegative integer. In summary:

Theorem 119. *If V is a finite dimensional irreducible $\mathfrak{g}$-module of highest weight λ, then $\lambda(h_i)$ is a nonnegative integer for all $1 \le i \le |\Delta| = l$, where Δ is the set of simple roots.*

Moreover, it follows from the representation theory of $\mathfrak{sl}_2(k)$ that if μ is a weight of V, then $\mu(h_i) = \langle \mu, \alpha_i \rangle \in \mathbb{Z}$ for all $1 \le i \le l$. Thus, the weights occurring in a finite dimensional module are also abstract weights as defined above. This means that all the theory developed

for abstract weights is now at our disposal. We will call a linear functional $\lambda \in \mathfrak{h}^*$ *integral* whenever $\lambda(h_i)$ is integral. Then λ is *dominant integral* if $\lambda(h_i)$ is nonnegative. The set of dominant integral linear functionals we denote by Λ^+. We denote the set of all weights of V by $\Pi(V)$, if V is a $\mathfrak{g}$-module and by $\Pi(\lambda)$ if $V = V(\lambda)$. The next result is useful for Chevallay's restriction theorem which we prove in the next section.

Theorem 120. *If $\lambda \in \mathfrak{h}^*$ is dominant integral, then the irreducible $\mathfrak{g}$-module $V = V(\lambda)$ is finite-dimensional, and its set of weights $\Pi(\lambda)$ is permuted by W, the Weyl group, with $\dim V_\mu = \dim V_{\sigma\mu}$.*

Proof. See [18]. ■

We conclude this section with an interesting corollary.

Corollary 121. *The map $\lambda \mapsto V(\lambda)$ induces a one-to-one correspondence between Λ^+ and the isomorphism classes of finite dimensional irreducible $\mathfrak{g}$-modules.*

7.2. **Chevalley's Restriction Theorem.** Let V be a finite dimensional vector space and consider the symmetric algebra $\mathfrak{P}(V) := \mathfrak{I}(V^*)/I$ of polynomial functions in n variables, where $\mathfrak{I}(V^*)$ is the tensor algebra of V^* and $I = \langle x \otimes y - y \otimes x - [x, y] \rangle$. We will draw our attention to $\mathfrak{P}(\mathfrak{g})$ and $\mathfrak{P}(\mathfrak{h})$ in particular, which we will denote by $k[\mathfrak{g}]$ and $k[\mathfrak{h}]$ respectively.

Recall that the set of weights Λ span $\mathfrak{h}^*$ and hence polynomials in weights $\lambda \in \Lambda$ span $k(\mathfrak{h})$. Since the Weyl group acts on $\mathfrak{h}^*$ it acts on $k[\mathfrak{h}]$, and we denote by $k[h]^W$ the subalgebra of W-invariant functions on $\mathfrak{h}$. We also define $\operatorname{sym} f := \sum_{\sigma \in W} \sigma.f$ as the sum of all distinct W-conjugates, for $f \in k[h]$. The the set of all $\operatorname{sym} \lambda^n$, where $\lambda \in \Lambda^+$ and $n \in \mathbb{Z}^+$, span $k[\mathfrak{h}]^W$ since each $\lambda \in \Lambda$ is W-conjugate to a dominant linear function by Lemma 111. We know from elementary character theory that the set of characters of a representation spans the set of class functions. In particular, the set of all polynomial functions $x \mapsto \operatorname{tr}(\phi(x)^k)$, where $\phi : \mathfrak{g} \to \mathfrak{gl}(V)$ is any irreducible representation of $\mathfrak{g}$, is a spanning set for $k[\mathfrak{g}]^G$.

Now let $f \in k[\mathfrak{g}]$, if we restrict f to $\mathfrak{h}$, we obtain a polynomial function on $\mathfrak{h}$. If, moreover, f is G-invariant, then, in particular, f is fixed by all inner automorphisms τ_α, $\alpha \in \Phi$ (see Example 100). But $\tau_\alpha|_{\mathfrak{h}}$ is the reflection σ_α (see Proposition 110), and these reflections σ_α, for all α, generate the Weyl group W. Hence $f|_{\mathfrak{h}} \in k[h]^W$. Hence, the restriction yields an algebra homomorphism $\theta : k[\mathfrak{g}]^G \to k[\mathfrak{h}]^W$. The following theorem of Chevalley states that this map is surjective.

Theorem 122. *The map $\theta : k[\mathfrak{g}]^G \to k[\mathfrak{h}]^W$ is surjective.*

Proof. By the previous discussion, it will suffice to show that each $\operatorname{sym} \lambda^n$, where $\lambda \in \Lambda^+$ and $n \in \mathbb{Z}^+$ lies in the image of θ. To prove this we will use induction on the partial ordering on Λ^+. We know there is a finite number $\mu \in \Lambda^+$ such that $\mu < \lambda$, by Lemma 112. Proceeding by induction, let $\lambda \in \Lambda^+$ be minimal. Then there can be no $\mu \in \Lambda^+$ occurring as a weight of an irreducible representation ϕ whose highest weight is λ. By Theorem 115 and Theorem 120, the only weights of ϕ are the W-conjugates of λ each having multiplicity one. Now the map $x \mapsto \operatorname{tr}(\phi(x)^n)$ is a G-invariant polynomial function, f, whose restriction

to $\mathfrak{h}$ is $\operatorname{sym}\lambda^n$. Hence, $\operatorname{sym}\lambda^n = \theta(f)$. For the induction step, we fix a $\lambda \in \Lambda^+$ and $n \in \mathbb{Z}^+$. Let ϕ, again, be the irreducible representation whose highest weight is λ. Let f be the function which maps $x \mapsto \operatorname{tr}(\phi(x)^n)$. Then the restriction to $\mathfrak{h}$ yields the sum of the weights occurring in this representation, which thanks to Theorem 120, we can write as $f|_{\mathfrak{h}} = \operatorname{sym}\lambda^n + \sum_{\mu<\lambda} c_{\mu,n} \operatorname{sym}\mu^n$, where $c_{\mu,n}$ are constants depending on $\mu \in \Lambda^+$ and $n \in \mathbb{Z}^+$. Then the terms involving $\mu < \lambda$ all lie in the image by induction, i.e., for each $\mu < \lambda$ there is a $p_\mu|_{\mathfrak{h}} = \operatorname{sym}\mu^n$. Then

$$\operatorname{sym}\lambda^n = \theta(f) - \sum_{\mu<\lambda} c_{\mu,n}\theta(p_\mu) = \theta\bigg(f - \sum_{\mu<\lambda} c_{\mu,n} p_\mu\bigg).$$

Thus, $\operatorname{sym}\lambda^n$ lies in the image of θ, proving surjectivity. ■

For completeness we also provide the proof of injectivity of θ. To this end, we introduce some notions that will be useful.

Definition 123. An element $x \in \mathfrak{g}$ is called *regular* if $\dim Z_{\mathfrak{g}(x)} = \operatorname{rk}(\mathfrak{g})$.

Here, $Z_{\mathfrak{g}(x)}$ denotes the centraliser of $x \in \mathfrak{g}$ and $\operatorname{rk}(\mathfrak{g}) = \dim(\mathfrak{h})$ is the rank of $\mathfrak{g}$. We say that an element $x \in \mathfrak{g}$ is *semisimple* if the map $\operatorname{ad}_x : \mathfrak{g} \to \mathfrak{g}$ is diagonalisable. We denote by $\mathfrak{g}^{sr}$ the set of all regular semisimple elements of $\mathfrak{g}$. The following Lemma we take from [16] and [32].

Lemma 124. *Fix $\mathfrak{h} \subset \mathfrak{b}$ a Cartan and Borel subalgebra of $\mathfrak{g}$. Then*

(i) *Any element of $\mathfrak{h}$ is semisimple and any semisimple element of $\mathfrak{g}$ is G-conjugate to an element of $\mathfrak{h}$.*

(ii) *The set $\mathfrak{g}^{sr}$ is a G-stable dense subset of $\mathfrak{g}$.*

Theorem 125. *The map $\theta : k[\mathfrak{g}]^G \to k[\mathfrak{h}]^W$ is an isomorphism*

Proof. Let $f \in \ker\theta$, that is $f|_{\mathfrak{h}} = 0$. Since any element $x \in \mathfrak{g}^{sr}$ is G-conjugate to an element in $\mathfrak{h}$, we have that $f|_{\mathfrak{g}^{sr}} = 0$. But the set of semisimple regular elements $\mathfrak{g}^{sr} \subset \mathfrak{g}$ are dense in $\mathfrak{g}$, hence $f = 0$. Thus, θ is injective. ■

8. Invariant Polynomials on the Maximal Unipotent Group of GL_n

Let U denote the maximal unipotent subgroup of $G := \mathrm{GL}_n(k)$ of $n \times n$ upper triangular matrices with entries on the main diagonal as 1. Denote by $k[U]$ the algebra of polynomial functions on U and suppose G acts on G by conjugation. This induces an action on the algebra $k[G]$ by $g.f(x) = f(g^{-1}xg)$. Our goal is to describe the subalgebra, $k[U]^U$, of the invariant functions on $k[U]$. More specifically, we want to understand what are the generators of this subalgebra and whether or not it is finitely generated. We will show that this subalgebra of invariants is isomorphic to $k[H]$, where H is made up from the simple roots subgroup of G. The root subgroups we obtain by exponentiating the one-dimensional root subspaces, $\mathfrak{g}_\alpha$.

Lemma 126. *Let $A, B \in U$, then $(ABA^{-1})_{i,i+1} = B_{i,i+1}$ for $i \in \{1, 2, \ldots, n-1\}$.*

Proof. We compute $(AB)_{i,i+1} = (BA)_{i,i+1}$. The left hand side is given as

$$(AB)_{i,i+1} = \sum_{j=1}^{n} a_{ij} b_{j,i+1} = \sum_{j=i}^{i+1} a_{ij} b_{j,i+1}, = b_{i,i+1} + a_{i,i+1}.$$

On the other hand, the right hand side yields

$$(BA)_{i,i+1} = \sum_{j=1}^{n} b_{ij} a_{j,i+1} = \sum_{j=i}^{i+1} b_{ij} a_{j,i+1} = a_{i,i+1} + b_{i,i+1}.$$

Now for $B := CA^{-1}$, we have $(ACA^{-1})_{i,i+1} = (CA^{-1}A)_{i,i+1} = C_{i,i+1}$, proving the result. ■

Corollary 127. *The functions $x_{i,i+1} \in k[U]$ are U-invariant, where U acts by conjugation.*

Lemma 128. *Let $H \subset U$ be the subset of matrices of the form*

$$h = \begin{pmatrix} 1 & h_{12} & & & & \\ & 1 & h_{23} & & & \\ & & \ddots & \ddots & & \\ & & & 1 & h_{n-1,n} & \\ & & & & 1 & \end{pmatrix}.$$

Then the restriction $k[U] \to k[H]$ defines a surjection $\iota^* : k[U]^U \to k[H]$.

Proof. The generators of $k[H]$ are $\{x_{i,i+1}\}$ for $i \in \{1, \ldots, n-1\}$. But by Corollary 127, these generators belong to $k[U]^U$, thus restricting $k[U]$ to $k[H]$ induces a surjection $\iota^* : k[U]^U \to k[H]$. ■

Remark 129. *The subset $H \subset G$ in the lemma is formed by the simple root subgroups obtained by exponentiating the simple root spaces found in Example 98.*

Lemma 130. *Let $g \in U$ with $g_{i,i+1} \neq 0$ units for all $i \in \{1, 2 \ldots, n-1\}$. Then g is conjugate (in U) to*

$$g' = \begin{pmatrix} 1 & g_{12} & & & & \\ & 1 & g_{23} & & & \\ & & \ddots & \ddots & & \\ & & & 1 & g_{n-1,n} & \\ & & & & 1 & \end{pmatrix}$$

where empty entries are zero and the super-diagonal of g and g' coincide.

Idea of proof: Before going into the details of the proof, we explain the idea behind the proof. Since we are proving the existence of an $a \in U$, we let $g \in U$ be arbitrary and we solve entry-wise the equations induced by the matrix equation $ag = g'a$. More explicitly, after equating the matrix equation $ag = g'a$, we obtain system of linear equations in a_{ij} and we want to know whether for any $g \in U$, the system of linear equations is always, at least, a consistent, exactly determined set of equations. The main idea in showing this is to parameterise the first row of the matrix a, i.e., the $a_{1,j}$ terms, and to write each equation as a function of these $a_{1,j}$ terms. We will see that this is possible and that we obtain a system of consistent underdetermined equations, meaning that we can always find an $a \in A$ for each $g \in U$ (in fact, there are infinitely many such a). We show this algorithmically by computing and solving the linear equations for each diagonal starting from the super diagonal and working through each successive diagonal. It will be easy to see that adopting this method will yield a set of underdetermined and consistent linear equations.

Proof. Given a $g \in U$, we show that there exists an $a \in U$ such that $ag = g'a$. This matrix equation will induce a system of linear equations when we compare the (i,k) entries on both sides of the equals sign. To prove the lemma, it is enough to show that these linear equations can always be solved for the a_{ij} entries. To this end, we equate $(ag)_{ik} = (g'a)_{ik}$ for each entry (i,k), that is, we compute:

$$\begin{aligned} \sum_{j=1}^{n} g'_{ij}a_{jk} &= \sum_{j=1}^{n} a_{ij}g_{jk}, \\ g_{i,i+1}a_{i+1,k} &= g_{ik} + \sum_{j=i+1}^{k-1} a_{ij}g_{jk}, \end{aligned} \tag{1}$$

where the simplificaion to Equation(1) is due to the form of the matrices g and g'. Notice that we can always solve for $a_{i+1,k}$ since $g_{i,i+1}$ is invertible for all $i \in \{1,\dots,n-1\}$, by assumption. We can navigate through the different linear equations induced by $(ag)_{i,k} = (g'a)_{i,k}$ by varying through the (i,k) components.

Restricting to the entries on the super diagonal where $k \leq i+1$, for all $i \in \{1,\dots,n-1\}$, we see that these equations are trivially satisfied, that is, we get equations of the form $0 = 0$. This imposes no conditions on the entries of A. Restricting to the next diagonal, that is, setting $k = i+2$ in Equation (1), we have the following parameterisation on $a_{i+1,i+2}$ by $a_{i,i+1}$:

$$g_{i,i+1}a_{i+1,i+2} = g_{i,i+2} + a_{i,i+1}. \tag{2}$$

To reiterate, the term $a_{i+1,i+2}$ can be parameterised by the term $a_{i,i+1}$. But i runs through all values between $\{1,\dots,n-1\}$, meaning we can parameterise each entry on the diagonal $(i,i+2)$ by $a_{1,2}$. Indeed, we can write $a_{n-1,n}$ in terms of $a_{n-2,n-1}$ which can be written in terms of $a_{n-3,n-2}$ and so on until we write a_{23} in terms of our parameter a_{12}. In this way, we have a consistent, underdetermined system of $n-2$ equations and $n-1$ unknowns which we know we can solve for after a choice for our parameter a_{12}.
Next we restrict Equation (1) to the entry $(i,i+3)$, that is, the next diagonal up and we obtain

$$g_{i,i+1}a_{i+1,i+3} = g_{i,i+3} + a_{i,i+1}g_{i+1,i+3} + a_{i,i+2}g_{i+2,i+3} \tag{3}$$

If we fix the entry a_{13} as a parameter, and using the relations in Equation (2), we observe that we can parameterise $a_{i+1,i+3}$ in terms of a_{13} and a_{12} for every $i \in \{1, \dots, n-3\}$ because $a_{i,i+1}$ can be paramterised by a_{12} and by the same process of the previous diagonal, $a_{i,i+2}$ can be parameterised by $a_{1,3}$ for $i \in \{1, \dots, n-3\}$. Now, we have a consistent, underdetermined system of $n-4$ equations and $n-3$ unknowns, where all the $a_{i,i+2}$ are all parameterised by a_{12} and a_{13}. We can solve these set of equations since we have already specified a value for a_{12} and make a choice for a_{13}. More explicitly, making a choice for a_{13} allows us to solve for a_{24} which allows us to solve for a_{35} and so on. Essentially, we continue this process for each diagonal.

For the last diagonal where $k = n$, Equation (1) yields,

$$g_{i,i+1}a_{2,n} = g_{i,n} + \sum_{j=i+1}^{n-1} a_{ij}g_{j,n-1} = g_{i,n} + \sum_{j=i+1}^{n-2} a_{ij}g_{j,n-1} + a_{i,n-1}g_{n-1,n}. \tag{4}$$

We fix $a_{1,n-1}$ as a parameter. Each $a_{i,j}$, $j \in \{i+1, i+2, \dots, n-2\}$, can be parameterised by the entries $a_{1,j}$ for $j \in \{2, 3, \dots, n-2\}$ (by the iterative process) and so the $a_{2,n}$ entry can be paramaterised by the entries of the form a_{1j} for $j \in \{2, 3, \dots, n-1\}$. We then have a consistent underdetermined system of one equation with $n-1$ paramters. Overall, we have a consistent, underdetermined system of $\frac{(n-1)(n-2)}{2}$ equations with $n-1$ unknowns. Therefore, there always exists a solution (in fact, there are infinitely many) to the equations $(Ag)_{ij} = (g'A)_{ij}$ and we conclude that there always exists an $A \in G$ such that $AgA^{-1} = g'$. Thus, g is conjugate to g'.

■

Theorem 131. *The map $\iota^* : k[U]^U \to k[H]$ is an isomorphism.*

Proof. By Lemma 128, ι^* is a surjective homomorphism. We will show that $\ker \iota^* = \{0\}$. Suppose that $f \in \ker \iota^*$, then $f(g) = 0$ for all $g \in H$. Define the Zariski open dense subset of G by

$$X = \{g \in G \,|\, g_{i,i+1} \neq 0, \text{for all} \, 1 \leq i \leq n-1\}.$$

Now, let $x \in X$. By Lemma 130, there is an $A \in U$ such that $AxA^{-1} = x'$, where $x' \in H$. Then since f is U-invariant,

$$f(AxA^{-1}) = f(x') = 0,$$

by assumption. Since f is continuous and zero on a dense subset, $X \subset U$, we conclude that $f(U) = 0$, therefore, $f = 0$. Thus, ι^* is injective. ■

9. Polynomial Invariants of a Symplectic Group

In this section, we apply the results of the preceeding section to the symplectic group $\mathrm{Sp}_m(\mathbb{C})$, (even m), the group of invertible matrices preserving a non-degenerate bilinear form. That is,

$$\mathrm{Sp}_m(k) = \{g \in \mathrm{GL}_m(k) \,|\, g^t M g = M\}.$$

Here, we look at the case when $m = 4$ for the bilinear form

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

From here, in this section, we will write G for $\mathrm{Sp}_4(k)$. For a unipotent matrix $g \in \mathrm{GL}_4(k)$ that satisfies $g^t M g = M$, has the form

$$\begin{pmatrix} 1 & g_{12} & g_{12}g_{23} - g_{24} & g_{14} \\ 0 & 1 & g_{23} & g_{24} \\ 0 & 0 & 1 & g_{12} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Therefore, the maximal unipotent subgroup of G, consists of those matrices

$$U = G \cap U_4(k) := \left\{ \begin{pmatrix} 1 & g_{12} & g_{13} & g_{14} \\ 0 & 1 & g_{23} & g_{24} \\ 0 & 0 & 1 & g_{12} \\ 0 & 0 & 0 & 1 \end{pmatrix} : g_{ij} \in k,\, g_{13} = g_{12}g_{23} - g_{24} \right\}.$$

As in the GL_n case, G acts on itself by conjugation. Our goal in this section is to describe $k[U]^U$, and show that is isomorphic to $k[H]$ for a particular subset H, which we now construct.

The subset H is obtained by exponentiating the simple root spaces associated to $\mathfrak{g}$. For this we will compute these root groups.

Hence, the subvariety, H, will be obtained by exponentiating the root space $\mathfrak{g}_{\varepsilon_1 - \varepsilon_2}$ and $\mathfrak{g}_{2\varepsilon_2}$. The one-dimensional root spaces are the span of the eigenvectors under the conjugation action which we computed in Example 99. That is,

$$\mathfrak{g}_{\varepsilon_1 - \varepsilon_2} = \mathrm{span}_k \left\{ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\}$$

and similarly, for the other root spaces. By exponentiating these root spaces we obtain one-dimensional unipotent subgroups. Note that in general the exponential map is not polynomial, however, since the root spaces contain niltpotent matrices (the matrices in these

root spaces square to zero), exponentiation will map these spaces back into U. Computing the exponential for $\mathfrak{g}_{\varepsilon_1-\varepsilon_2}$, we have

$$U_{\varepsilon_1-\varepsilon_2} = \exp\begin{pmatrix} 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 \end{pmatrix} = I_4 + \begin{pmatrix} 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & x & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & x \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $x \in k$. Similarly, computing the exponential of $\mathfrak{g}_{2\varepsilon_2}$,

$$U_{2\varepsilon_2} = \exp\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We now take the subset $H \subset U$ which encodes the subgroups $U_{\varepsilon_1-\varepsilon_2}$ and $U_{2\varepsilon_2}$. That is,

$$H = \left\{ \begin{pmatrix} 1 & g_{12} & g_{13} & 0 \\ 0 & 1 & g_{23} & 0 \\ 0 & 0 & 1 & g_{12} \\ 0 & 0 & 0 & 1 \end{pmatrix} : g_{ij} \in k,\ g_{13} = g_{12}g_{23} \right\}.$$

Note that H itself is not a subgroup of U. To show that $k[U]^U \cong k[H]$, we use the same approach as we did for the case when $G = \mathrm{GL}_n(k)$.

Lemma 132. *The restriction map $k[U] \to k[H]$ defines a surjection $\Phi : k[U]^U \to k[H]$.*

Proof. We have seen in Corolllary 127 that the polynomials $x_{i,i+1}$ are U-invariant. Since the polynomial algebra $k[H] = k[x_{12}, x_{23}, x_{12}x_{23}]$ has generators x_{12}, x_{23}, the restriction to H defines a surjection from $k[U]^U \to k[H]$ by mapping invariant polynomials in $k[U]^U$ to the generators in $k[H]$. ■

Lemma 133. *Let $g \in U$ with entries $g_{ij} \in k$ where $g_{i,i+1} \neq 0$ for $i \in \{1, 2\}$. Then g is conjugate (in U) to*

$$g' = \begin{pmatrix} 1 & g'_{12} & g'_{13} & 0 \\ 0 & 1 & g'_{23} & 0 \\ 0 & 0 & 1 & g'_{12} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where the superdiagonals coincide, i.e., $g_{i,i+1} = g'_{i,i+1}$ for $i \in \{1, 2\}$ and $g'_{13} = g_{12}g_{23}$.

Proof. The proof uses the same technique as the proof in Lemma 130. It is enough to prove that there is a matrix $A \in U$ such that $Ag = g'A$. Let $A = (a_{ij}) \in U$, then $Ag = g'A$ if an

only if the a_{ij} entries satisfy the three equations:

$$a_{12}g_{23} + g_{13} = a_{23}g_{12} + g_{12}g_{23} \tag{5}$$

$$a_{23}g_{12} + g_{24} = a_{12}g_{23} \tag{6}$$

$$a_{12}a_{23}g_{12} + a_{12}g_{24} + g_{14} = a_{12}g_{12}g_{23} + 2a_{24}g_{12}. \tag{7}$$

Firstly, we notice that substituing Equation 5 into Equation 6 provides no further information. That is, the equation induced by this substitution is trivial. We keep the $a_{1,j}$ entries as parameters for $j \in \{2, 3, 4\}$. We then have the two conditions on a_{23} and a_{24} :

$$\begin{aligned} a_{23} &= g_{12}^{-1}(a_{12}g_{23} - g_{24}) \\ a_{24} &= \frac{1}{2}g_{12}^{-1}(a_{12}g_{24} + g_{14}) + \frac{1}{2}(a_{12}a_{23} - a_{12}g_{12}). \end{aligned}$$

Hence, the system of Equations 5-7 is solvable which implies that there is a matrix $A \in U$ with entries $a_{ij} \in k$ which satisfies $Ag = g'A$. Thus, g is conjugate to g'.

■

Theorem 134. *The map $\Phi : k[U]^U \to k[H]$ is an isomorphism of algebras.*

Proof. Let $f \in \ker \Phi$, then $f(g) = 0$ for all $g \in H$. Define the Zariski-open dense subset

$$X = \{g \in U : g_{i,i+1} \neq 0, \forall i \in \{1, 2\}\}.$$

By Lemma 133, For any $x \in X$, there is an $A \in U$ such that $AxA^{-1} = x'$, where $x' \in H$. By U-invariance of f,

$$f(x) = f(AxA^{-1}) = f(x') = 0.$$

Since f is continuous and zero on a dense subset, $X \subset U$, we conclude that $f = 0$. Thus, Φ is injective. ■

10. Polynomial Invariants on Maximal Unipotent Subgroups

In this section, we generalise the results in Section 8 and Section 9 to maximal unipotent subgroups of general reductive groups. First, we fix some notation.

Let G be a reductive group with corresponding Lie algebra $\mathfrak{g}$ and fix a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Let Φ denote the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h}$ and denote by Φ^+ and Δ a fixed set of positive and simple roots respectively. Further, fix a basis for the root spaces, that is, each root space $\mathfrak{g}_\alpha$ is spanned by the vector e_α for all $\alpha \in \Phi$. We present some results which are vital for our general theorem (cf: Theorem 142). The proof of the following lemma is adopted from [16].

Lemma 135. *Let G be an algebraic group with Lie algebra $\mathfrak{g}$. Let V be a finite-dimensional representation of G and $E \subset V$ a G-stable linear subspace. If G is connected, then for any $v \in V$, the following conditions are equivalent:*

(1) *The affine linear subspace $v + E \subset V$ is G-stable;*

(2) *We have $\mathfrak{g}.v \subset E$, that is the image of $\mathfrak{g}$ under the induced differential map $\mathfrak{g} \to V$, $x \mapsto x.v$ is contained in E.*

Proof. Suppose that (1) holds. Then we have $G \cdot v \subset v + E$. Taking the differential on both sides at the identity of G yields $\mathfrak{g}.v \subset E$, which is the second condition. Now suppose that condition (2) is true. We may identify the tangent space to $v + E$ at any point $u \in v + E$ as E. Let $x \in \mathfrak{g}$, and let γ_x be the vector field induced by the action of x on V. Then for any $u \in v + E$, we write

$$\gamma_x(u) = x \cdot u \in x(v + E) = x \cdot v + x \cdot E \subset \mathfrak{g} \cdot v + x \cdot E \subset E.$$

Note that $\mathfrak{g}.v \subset E$ by assumption and $\mathfrak{g}.E \subset E$ because E is G-stable. The above computation implies that the vector field γ_x is tangent to the subspace $v + E$ for any of its points (under our identification). Therefore, the subspace $v + E$ is stable under the action of a neighbourhood of the identity in G. But G is connected, therefore, $v + E$ is G-stable hence condition (1). ■

We also require the following result.

Theorem 136. *Let U be a unipotent algebraic group acting regularly on an irreducible variety V. Then all the U-orbits are closed*

Proof. See [34] for a poof of this result. ■

Now, define the *coordinate functions* $f_\alpha \in k[\mathfrak{n}]$ as follows: Let $x = \sum_{\alpha \in \Phi} a_\alpha e_\alpha$ where $a_\alpha \in k$ and e_α are the basis vectors of $\mathfrak{g}_\alpha$. Then $f_\alpha(x) = a_\alpha$. Define the set of *regular nilpotent elements* by

$$\mathfrak{n}^r := \{x \in \mathfrak{n} \mid f_\alpha(x) \neq 0, \forall \alpha \in \Delta\}.$$

The set $\mathfrak{n}^r$ is Zariski-open in $\mathfrak{n}$ being defined by nonvanishing polynomials on the root vectors e_α for all $\alpha \in \Delta$.

The proof of the following lemma has been extracted from Lemma 3.2.12 in [16] where the set of regular elements is a single B-orbit. Here, we only require that $x + [\mathfrak{n}, \mathfrak{n}]$ with $x \in \mathfrak{n}^r$ is a single U-orbit.

Lemma 137. *Let U be a unipotent algebraic group with Lie algebra $\mathfrak{n}$. The set $x + [\mathfrak{n}, \mathfrak{n}]$, where $x \in \mathfrak{n}^r$, is a single U-orbit.*

Proof. Since for any $x \in \mathfrak{n}$, we have that $[x, n] \subset [\mathfrak{n}, \mathfrak{n}]$, by Lemma 135, the set $x + [\mathfrak{n}, \mathfrak{n}]$ is stable under the adjoint U action. Then, since x is regular, $\dim(\mathrm{Ad}_U(x)) \geq \dim \mathfrak{n} - rk(\mathfrak{g}) = \dim[\mathfrak{n}, \mathfrak{n}]$. That is, the U-orbit of x is open in $x + [\mathfrak{n}, \mathfrak{n}]$. Moreover, we have, by Theorem 136, U-orbits are closed. Thus, $x + [\mathfrak{n}, \mathfrak{n}]$ is a single U-orbit, provided that x is regular. ■

Our next key lemma generalises Lemma 126 for general reductive groups.

Lemma 138. *Let $x \in \mathfrak{n} = \mathrm{Lie}(U)$, where U is a maximal unipotent subgroup of G. Then for all $A \in U$ and $\alpha \in \Delta$, the projection of x onto $\mathfrak{g}_\alpha$ is the same as the projection of $\mathrm{Ad}_A(x)$ onto $\mathfrak{g}_\alpha$.*

Proof. Since, by Theorem 97, U is generated by the root groups U_α, it is enough to show that $x \in \mathfrak{n}$ and $\mathrm{Ad}_{U_\alpha}(x)$ have the same projection onto the simple root spaces for some $\alpha \in \Phi^+$. But $U_\alpha \ni g_\alpha = \exp u_\alpha$, where $u_\alpha \in \mathrm{Lie}(U_\alpha)$, so we obtain

$$\mathrm{Ad}_{g_\alpha}(x) = \mathrm{Ad}_{\exp u_\alpha}(x) = \exp\left(ad(u_\alpha)(x)\right) = x + [u_\alpha, x] + \frac{1}{2}[u_\alpha, [u_\alpha, x]] + \cdots + \frac{1}{k!}ad(u_\alpha)^k(x).$$

Since $[\mathfrak{g}_\alpha, \mathfrak{g}_\gamma] \subset \mathfrak{g}_{\alpha+\gamma}$, and $\mathfrak{n}$ contains only the positive root spaces, the terms after x do not project onto the simple roots. ■

Corollary 139. *For all $\alpha \in \Delta$, the functions $f_\alpha \in k[\mathfrak{n}]$ are U-invariant.*

We can summarise Lemma 138 and Corollary 139 as follows.

Let

$$\mathfrak{n}_\Delta = \{x \in \mathfrak{n} \mid f_\alpha(x) = 0, \text{ for all } \alpha \notin \Delta\}$$

be the subset of all those elements in $\mathfrak{n}$ whose projection onto all non-simple root spaces is zero. That is, elements $x \in \mathfrak{n}_\Delta$ are spanned only by the simple root vectors. Then we have the following Proposition.

Proposition 140. *The restriction $k[\mathfrak{n}] \to k[\mathfrak{n}_\Delta]$ defines a surjection $\psi : k[\mathfrak{n}]^U \to k[\mathfrak{n}_\Delta]$.*

Proof. The generators of $k[\mathfrak{n}_\Delta]$ are those functions f_α, where $\alpha \in \Delta$, which are U-invariant. ■

Lemma 141. *Let $x \in \mathfrak{n}^r$ be a regular nilpotent element. Then there exists an $A \in U$ such that the projection of $\mathrm{Ad}_A(x)$ onto all non-simple root spaces is zero, i.e., $\mathrm{Ad}_A(x) \in \mathfrak{n}_\Delta$.*

Proof. Let $x = \sum_{\alpha\in\Delta} a_\alpha e_\alpha + \sum_{\alpha\notin\Delta} b_\alpha e_\alpha$ be a regular nilpotent element. Define $x' = \sum_{\alpha\in\Delta} a_\alpha e_\alpha$, then $x \in x' + [\mathfrak{n}, \mathfrak{n}]$ since $[\mathfrak{n}, \mathfrak{n}]$ contains all those elements whose projection onto the simple root spaces is zero. However, x' also lies in space $x' + [\mathfrak{n}, \mathfrak{n}]$. Then, since $x' + [\mathfrak{n}, \mathfrak{n}]$ is a single U-orbit (Lemma 137), whenever x' is regular, x and x' reside in the same orbit. This implies the existence of $A \in U$ such that $\mathrm{Ad}_A(x) = x'$. ■

Theorem 142. *The map $\psi : k[\mathfrak{n}]^U \to k[\mathfrak{n}_\Delta]$ is an isomorphism.*

Proof. We have already shown surjectivity. Suppose $f \in \ker\psi$. Then $f|_{\mathfrak{n}_\Delta} = 0$, i.e., $f(h) = 0$ for all $h \in \mathfrak{n}_\Delta$. For any $x \in \mathfrak{n}^r$, by Lemma 141, there exists a $g \in U$ such that $\mathrm{Ad}_g(x) = x'$ where $x' \in \mathfrak{n}^r \cap \mathfrak{n}_\Delta$. By assumption and since f is U-invariant, $f(x) = f(x') = 0$. Since $x \in \mathfrak{n}^r$ was arbitrary, we have that $f(\mathfrak{n}^r) = 0$. But, since f is continuous and zero on a dense subset of $\mathfrak{n}$, this implies $f = 0$ on $\mathfrak{n}$ and so is the constant zero function. Thus, $\ker\psi = \{0\}$, hence, ψ is injective. ■

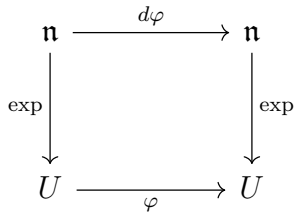
It turns out that Theorem 142 also holds at the group level.

Theorem 143. *The following isomorphism of k-algebras*

$$k[U]^U \cong k[\mathfrak{n}_\Delta]$$

holds.

Proof. We know that exponential map $\exp : \mathfrak{n} \to U$ is an isomorphism of affine varieties and we also have the following commutative diagram:



If we let φ be the conjugation action, then $d\varphi = \mathrm{Ad}$ is the adjoint action. Thus, if an element $x \in \mathfrak{n}$ is Ad_U-invariant, then, by chasing the commutative diagram, starting top left, we obtain two elements $\exp(x)$ and $g\exp(x)g^{-1}$ which sit underneath $x \in U$ (under the exponential map). In other words, $x \mapsto \exp(x)$ and $x \mapsto g\exp(x)g^{-1}$, hence by the commuting diagram, $\exp(x) = g\exp(x)g^{-1}$ which implies $\exp(x)$ is conjugation invariant. Similarly, if $\exp(x)$ is conjugation invariant, then we obtain $\exp(x) = \exp(\mathrm{Ad}_u(x))$ which, by injectivity of exp, implies $\mathrm{Ad}_u(x) = x$, i.e., x is Ad_U-invariant. Furthermore, since the exponential map is an isomorphism, the pullback along exp yields an isomorphism $k[\mathfrak{n}] \cong k[U]$ of the corresponding k-algebras. Thus, restricting to the elements that are Ad_U-invariant, and pulling back along the exponential map, exp, yields $k[\mathfrak{n}]^U \cong k[U]^U$. ■

References

[1] David Hilbert. Über die theorie der algebraischen formen. In *Algebra· Invariantentheorie· Geometrie*, pages 199–257. Springer, 1933.

[2] Igor Dolgachev. *Lectures on invariant theory*. Number 296. Cambridge University Press, 2003.

[3] Gurevich Grigorii Borisovich. *Foundations of the Theory of Algebraic Invarients*. P. Noordhoff, 1964.

[4] Peter J Olver. *Classical invariant theory*, volume 44. Cambridge University Press Cambridge, 1999.

[5] Paul Gordan. Beweis, dass jede covariante und invariante einer binären form eine ganze function mit numerischen coefficienten einer endlichen anzahl solcher formen ist. 1868.

[6] A. Hurwitz. über die erzeugung der invarianten durch integration. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1897:71–2, 1897.

[7] Ludwig Maurer. Über die endlichkeit der invariantensysteme. *Mathematische Annalen*, 57:265–313.

[8] Masayoshi Nagata, M Pavaman Murthy, et al. *Lectures on the fourteenth problem of Hilbert*, volume 31. Tata Institute of Fundamental Research Bombay, 1965.

[9] David Mumford, John Fogarty, and Frances Kirwan. *Geometric invariant theory*, volume 34. Springer Science & Business Media, 1994.

[10] Masayoshi Nagata. Invariants of group in an affine ring. *Journal of Mathematics of Kyoto University*, 3(3):369–378, 1963.

[11] William J Haboush. Reductive groups are geometrically reductive. *Annals of Mathematics*, 102(1):67–83, 1975.

[12] Frank D Grosshans. *Algebraic homogeneous spaces and invariant theory*. Springer, 2006.

[13] Frank Grosshans. Observable groups and hilbert's fourteenth problem. *American Journal of Mathematics*, 95(1):229–253, 1973.

[14] Yasmine Fittouhi and Anthony Joseph. Weierstrass sections for parabolic adjoint action in type *a*. *arXiv preprint arXiv:2001.00447*, 2020.

[15] Victoria Sevostyanova. The algebra of invariants for the adjoint action of the unitriangular group. *arXiv preprint arXiv:1605.00800*, 2016.

[16] Neil Chriss and Victor Ginzburg. *Representation theory and complex geometry*, volume 42. Springer, 1997.

[17] Nolan R Wallach. *Real reductive groups I*. Academic press, 1988.

[18] James E Humphreys. *Introduction to Lie algebras and representation theory*, volume 9. Springer Science & Business Media, 2012.

[19] Claude Chevalley. Invariants of finite groups generated by reflections. *American Journal of Mathematics*, 77(4):778–782, 1955.

[20] David Eisenbud. *Commutative algebra: with a view toward algebraic geometry*, volume 150. Springer Science & Business Media, 2013.

[21] Joe Harris. *Algebraic geometry: a first course*, volume 133. Springer Science & Business Media, 2013.

[22] Walter Ferrer Santos and Alvaro Rittatore. *Actions and invariants of algebraic groups*. CRC press, 2005.

[23] James E Humphreys. *Linear algebraic groups*, volume 21. Springer Science & Business Media, 2012.

[24] Jinpeng An. Rigid geometric structures, isometric actions, and algebraic quotients. *Geometriae Dedicata*, 157(1):153–185, 2012.

[25] Robert G Heyneman and Moss Eisenberg Sweedler. Affine hopf algebras, i. *Journal of Algebra*, 13(2):192–241, 1969.

[26] Tonny Albert Springer. Linear algebraic groups. In *Algebraic geometry IV*, pages 1–121. Springer, 1994.

[27] David S Dummit and Richard M Foote. *Abstract algebra*, volume 1999. Prentice Hall Englewood Cliffs, NJ, 1991.

[28] Gunter Malle and Donna Testerman. *Linear algebraic groups and finite groups of Lie type*, volume 133. Cambridge university press, 2011.

[29] Fiona Murnaghan. Linear algebraic groups. *Harmonic Analysis, the Trace Formula*, page 379, 2003.

[30] Keith Conrad. Linear independence of characters. *Online Notes*.

[31] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013.

[32] Jean-Pierre Serre. *Complex semisimple Lie algebras*. Springer Science & Business Media, 2000.

[33] Jacques Dixmier. *Enveloping algebras*. Number 11. American Mathematical Soc., 1996.
[34] Maxwell Rosenlicht. On quotient varieties and the affine embedding of certain homogeneous spaces. *Transactions of the American Mathematical Society*, 101(2):211–223, 1961.