A SHORT COURSE IN LIE THEORY

JAY TAYLOR

ABSTRACT. These notes are designed to accompany a short series of talks given at the University of Aberdeen in the second half session of the academic year 2009/2010. They will try to give a rough overview of material in the area of Lie theory. The main focus will be on structural properties of Lie theoretic objects such as abstract root systems, Weyl/Coxeter groups, Lie algebras, linear algebraic groups and finite groups of Lie type. Few proofs are given but references to proofs are given for all statements

Contents

1 Root Systems and Finite Reflection Groups 1

1.1 Motivation 1

1.2 Basic Definitions 2

1.3 A Presentation for W 4

1.4 Parabolic Subgroups and Chamber Systems 8

1.5 Coxeter Graphs, Dynkin Diagrams and Cartan Matrices 10

2 Lie Algebras 17

2.1 Basic Definitions and Introduction 17

2.2 Solvable and Nilpotent Lie Algebras 20

2.3 Testing for Solvability and Semisimplicity 22

2.4 The Lie Algebra $\mathfrak{sl}(n, \mathbb{C})$ 24

2.5 Semisimple Complex Lie Algebras 25

3 Linear Algebraic Groups 31

3.1 Some Definitions and The Zariski Topology 31

3.2 Regular Maps and Linear Algebraic Groups 34

3.3 The Lie Algebra of a Linear Algebraic Group 40

3.4 The Root Datum of a Linear Algebraic Group 48

3.5 The Classification of Simple Linear Algebraic Groups 53

4 Finite Groups of Lie Type 57

5 Bibliographic Remarks & Acknowledgments 61

References 62

1. Root Systems and Finite Reflection Groups

We will start these notes in a slightly unconventional place. We will start by introducing the notion of an abstract root system and show how this simple idea underpins a large area of pure mathematics. Standard references for the material in this section are [Bou02] and [Hum90], also a less verbose treatment of crystallographic root systems is given in [EW06, Chapter 11].

1.1. **Motivation.** We start with a little motivation for this slightly abstract concept. Most group theorists upon hearing the question "what is it that you actually do?" will reach for the little gem that is the dihedral group of order 8, which we shall refer to as D_8. We like this group simply because it is easy to draw a picture and explain the group actions via reflections and rotations. We often draw a picture as in figure 1.

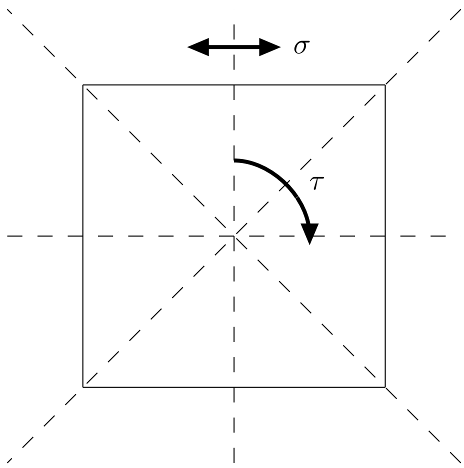


Figure 1. The Symmetry Group of the Square

We normally express D_8 as a group given by generators and relations. We pick a reflection and a rotation by $\pi/2$, as we have done in figure 1, then express D_8 as

$$\langle \sigma, \tau \mid \tau^4 = \sigma^2 = 1, \sigma\tau\sigma = \tau^{-1} \rangle.$$

However this group is generated simply by reflections. This is because the rotation τ can be achieved as a product of two reflections which are adjacent by an angle of $\pi/4$. It is an easy exercise to check that these two reflections will generate the whole of D_8. In fact in the dihedral group D_{2m} we can describe any rotation by $2\pi/m$ as the product of two reflections which are adjacent by an angle of π/m. Indeed it turns out that these two reflections will generate the whole of D_{2m}.

It is this property of D_8 that we are really interested in. We are interested in finite reflection groups or finite Coxeter groups.

Definition. We say a group G is a *Coxeter group* if it is isomorphic to a group given by a presentation of the form

$$\langle s_i \mid (s_i s_j)^{m_{ij}} = 1 \rangle,$$

where $m_{ii} = 1$ and $m_{ij} \geqslant 2$ for all $i \neq j$. If no relation occurs for a pair s_i, s_j then we define $m_{ij} = \infty$.

These groups, or more importantly a subclass of these groups known as *Weyl groups*, are a vital part of Lie theory. The class of Coxeter groups is quite large containing finite reflection groups, affine Weyl groups, Hyperbolic Coxeter groups and so on. However it turns out that every finite Coxeter group is a finite reflection group and indeed most of these are finite Weyl groups. It is the finite Coxeter groups which we shall be interested in.

The way in which we will classify the finite reflection groups is by what is known as a root system. We start by imagining the square in $\mathbb{R}^2$ such that it is centred at the origin. Now in our example of D_8 there are two reflections whose product give the rotation τ. In each of these reflecting lines lies a vector in $\mathbb{R}^2$, which will be our root. We describe this situation pictorially in figure 2.

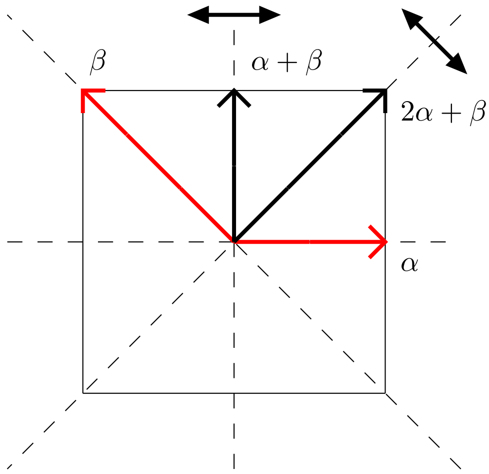


Figure 2. The Root System B_2.

1.2. **Basic Definitions.** Let V be a real Euclidean vector space endowed with Euclidean inner product $(\cdot,\cdot)$. Throughout this section W will refer to a finite reflection group acting on the vector space V. By this we mean W is a subgroup of $\mathrm{GL}(V)$ generated by reflections.

Definition. A *reflection* in V is a linear map $s_\alpha : V \to V$ which sends a non-zero vector α to its negative and fixes pointwise the hyperplane $H_\alpha = \{\lambda \in V \mid \lambda \perp \alpha\}$ orthogonal to α.

We have a simple formula that expresses the action of s_α on V. For all $\lambda \in V$ we have that

$$s_\alpha \lambda = \lambda - \langle \lambda, \alpha \rangle \alpha \qquad \text{where} \qquad \langle \lambda, \alpha \rangle = \frac{2(\lambda,\alpha)}{(\alpha,\alpha)}. \tag{1}$$

Indeed this makes sense as we have $s_\alpha \alpha = \alpha - 2\alpha = -\alpha$ and $s_\alpha \lambda = \lambda \Leftrightarrow (\lambda, \alpha) = 0 \Leftrightarrow \lambda \perp \alpha$. Note this now makes sense for all of V as $V = \mathbb{R}\alpha \oplus H_\alpha$. Let $\lambda, \mu \in V$ then we have

$$\begin{aligned}(s_\alpha\lambda, s_\alpha\mu) &= \left(\lambda - \frac{2(\lambda,\alpha)}{(\alpha,\alpha)}\alpha, \mu - \frac{2(\mu,\alpha)}{(\alpha,\alpha)}\alpha\right),\\ &= (\lambda,\mu) - \frac{2(\mu,\alpha)}{(\alpha,\alpha)}(\lambda,\alpha) - \frac{2(\lambda,\alpha)}{(\alpha,\alpha)}(\alpha,\mu) + \frac{4(\lambda,\alpha)(\mu,\alpha)}{(\alpha,\alpha)^2}(\alpha,\alpha),\\ &= (\lambda,\mu).\end{aligned}$$

Hence s_α is an orthogonal transformation and so an element of $\mathrm{O}(V) \leqslant \mathrm{GL}(V)$ the group of all orthogonal transformations of V. It is also clear to see that s_α is an element of order 2. Therefore any finite reflection group will be a subgroup of the orthogonal group of V.

Let α be a non-zero vector in V and s_α its associated reflection. Then s_α determines a reflection hyperplane H_α, as in the above definition of reflection, and a line $L_\alpha = \mathbb{R}\alpha$ orthogonal to this hyperplane.

Proposition 1.1. *If $t \in \mathrm{O}(V)$ and α is any non-zero vector in V then $ts_\alpha t^{-1} = s_{t\alpha}$. In particular if $w \in W$ then $s_{w\alpha}$ belongs to W whenever s_α does.*

Proof. We have that $(ts_\alpha t^{-1})(t\alpha) = ts_\alpha\alpha = t(-\alpha) = -t\alpha$ therefore certainly we have $ts_\alpha t^{-1}$ sends $t\alpha$ to its negative. Now $\lambda \in H_\alpha \Leftrightarrow t\lambda \in H_{t\alpha}$ and as $\lambda \in H_\alpha$ we have $0 = (\lambda,\alpha) = (t\lambda, t\alpha)$. Therefore we have $(ts_\alpha t^{-1})(t\lambda) = ts_\alpha\lambda = t\lambda$ whenever $\lambda \in H_\alpha$. □

This proposition tells us that W permutes the set of all lines L_α by $w(L_\alpha) = L_{w\alpha}$. For example in figure 2 we can see that the action of D_8 will permute the four reflection hyperplanes on the diagram. In turn the set of vectors $\{\pm\alpha, \pm\beta, \pm(\alpha+\beta), \pm(2\alpha+\beta)\}$ is stable under the action of D_8. It is this property of the vectors that we would like to encapsulate.

Definition. A *root system* is a finite set $\Phi \subset V$ of non-zero vectors such that:

(R1) Φ spans V,
(R2) if $\alpha \in \Phi$ then the only scalar multiples of α in Φ are $\pm\alpha$,
(R3) if $\alpha \in \Phi$ then $s_\alpha\Phi = \Phi$.

Remark. The definition of root system is indeed a very flexible thing. For the most general approach one may omit (R1) as in [Hum90, Section 1.2]. Often a fourth condition is added and this is what we will define in section 1.5 to be a crystallographic root system. Sometimes one may also remove (R2) from the definition of crystallographic root system, as in [Bou02]. However for now this definition is what is required.

Given a root system Φ of V we define the associated reflection group W to be the subgroup of $\mathrm{GL}(V)$ generated by all reflections s_α such that $\alpha \in \Phi$. Note that W is necessarily finite. This is because the only element to fix all elements of Φ is the identity. Therefore the natural homomorphism of W into the symmetric group on Φ has trivial kernel and so W must be finite.

So a root system is simply a collection of vectors in a Euclidean vector space. What would be useful is if we had some notion of a basis for Φ as the order of Φ could be very large even if the dimension of V is small. First we note that we can always endow a vector space with a total order. Let $\lambda_1, \ldots, \lambda_n$ be an ordered basis for V then we put a total order $<$ on V by the *lexicographical ordering*. This says that $\sum a_i\lambda_i < \sum b_i\lambda_i$ if and only if $a_k < b_k$ where k is the least index i for which $a_i \neq b_i$. We can now use this to define a notion of basis for Φ.

Definition. Let Φ be a root system for V then a subset $\Phi^+ \subset \Phi$ is called a *positive system* for Φ if it consists of all roots which are positive relative to some total ordering of V. Now roots come in pairs $\{\alpha, -\alpha\}$ therefore Φ is the disjoint union of Φ^+ and $\Phi^- = -\Phi$ and we call Φ^- a *negative system* for Φ. Call a subset $\Delta \subset \Phi$ a *simple system* for Φ if Δ is a vector space basis of V and every element of Φ can be expressed as a linear combination of Δ whose coefficients are either all positive or all negative. We call the roots belonging to Δ *simple roots* for Φ and the associated reflections *simple reflections*.

Example. We recall figure 2. A root system, positive system, negative system and simple system for $\mathbb{R}^2$ is given by

$$\begin{aligned}
\Phi &= \{\pm\alpha, \pm\beta, \pm(\alpha+\beta), \pm(2\alpha+\beta)\},\\
\Phi^+ &= \{\alpha, \beta, \alpha+\beta, 2\alpha+\beta\},\\
\Phi^- &= \{-\alpha, -\beta, -\alpha-\beta, -2\alpha-\beta\},\\
\Delta &= \{\alpha, \beta\}.
\end{aligned}$$

Note that this is not in any way the only root system for $\mathbb{R}^2$. In fact each of the dihedral groups D_{2m} comes from a different root system for $\mathbb{R}^2$.

Now it is clear that we can always form a positive and negative system for Φ. To do this we just fix a hyperplane of codimension 1 in V which contains no element of Φ and then label all roots on one side of the hyperplane positive and on the otherside negative. However apriori it is not clear that simple systems will exist but in fact they do.

Theorem 1.1. *Let Φ be a root system and Φ^+ be a positive system for Φ. Then Φ^+ contains a simple system for Φ and moreover this simple system is uniquely determined by Φ^+.*

Proof. See [Hum90, Theorem 1.3] □

Corollary 1.1. *Let Δ be a simple system for a root system Φ then $(\alpha, \beta) \leq 0$ for all $\alpha \neq \beta$ in Δ.*

From the proof of this theorem we obtain a very useful corollary, which is a constraint on the geometry of the roots in the simple system. It tells us that the angles between roots in a simple system are always obtuse. For example this tells us that in figure 2 we cannot have that $\{2\alpha+\beta, \alpha\}$ is a simple system for Φ. However $\{2\alpha+\beta, -\alpha\}$ is a simple system for Φ.

Now a simple system forms a basis for the vector space V and hence the cardinality of a simple system is an invariant of the root system.

Definition. Let Φ be a root system with simple system Δ. Then we say $|\Delta|$ is the *rank* of Φ.

For example in the case of D_8 we have the root system Φ which admits D_8 as a finite reflection group has rank 2. In fact all dihedral groups are finite reflection groups associated to root systems of rank 2. The converse statement is also true.

1.3. **A Presentation for W.** So far we have defined the finite reflection group W, associated to a root system Φ, to be the subgroup of $\mathrm{O}(V)$ generated by $\{s_\alpha\}_{\alpha\in\Phi}$. However this definition of W is not very practical as the cardinality of Φ can be very large with respect to the order of W. In fact in the case of the dihedral groups we have the cardinality of Φ is equal to the order of W. Therefore we would like to find a more efficient presentation of W. To do this we will first need a small lemma and theorem.

Lemma 1.1. *Let Φ be a root system with positive system Φ^+ and simple system Δ. Then for any $\alpha \in \Delta$ we have s_α permutes the set $\Phi^+ \setminus \{\alpha\}$.*

Proof. Let $\beta \in \Phi^+$ be such that $\beta \neq \alpha$. We know that $\beta = \sum_{\gamma\in\Delta} c_\gamma \gamma$ where $c_\gamma \geqslant 0$. Note that the only multiples of α in Φ are $\pm\alpha$ and so there is at least one $\gamma \in \Delta$ such that $c_\gamma > 0$. Now $s_\alpha\beta \in \Phi$ and by eq. (1) we have $\beta - \langle \alpha, \beta\rangle\alpha \in \Phi$. The coefficients c_γ for $\gamma \neq \alpha$ are unaffected and hence are still all positive. Therefore this forces $c_\alpha - \langle\alpha, \beta\rangle \geqslant 0$ and so $s_\alpha\beta \in \Phi^+$. Therefore $s_\alpha(\Phi^+ \setminus \{\alpha\}) = \Phi^+ \setminus \{\alpha\}$. □

Theorem 1.2. *Let Φ be a root system and let Φ_1^+, Φ_2^+ be two positive systems in Φ with associated simple systems Δ_1, Δ_2. Then there exists an element $w \in W$ such that $w(\Phi_1^+) = \Phi_2^+$ and hence $w(\Delta_1) = \Delta_2$.*

Proof. Let $x = |\Phi_1^+ \cap \Phi_2^-|$. We will prove the statement by induction on x. If $x = 0$ then we have $\Phi_1^+ = \Phi_2^+$ and so $w = 1$ is sufficient for the statement. Therefore we assume $x > 0$. Now we cannot have $\Delta_1 \subset \Phi_2^+$ as this would imply $\Phi_1^+ \subset \Phi_2^+$ but we assumed $x > 0$. Hence there exists a root $\alpha \in \Delta_1 \cap \Phi_2^-$.

Consider $s_\alpha(\Phi_1^+)$. By lemma 1.1 we have $s_\alpha(\Phi_1^+) = (\Phi_1^+ \setminus \{\alpha\}) \cup \{-\alpha\}$ and so we have $|s_\alpha(\Phi_1^+) \cap \Phi_2^-| = x - 1$. By induction there exists $w' \in W$ such that $w's_\alpha(\Phi_1^+) = \Phi_2^+$. Let $w = w's_\alpha$ then $w(\Phi_1^+) = \Phi_2^+$ as required. □

It is easy to verify the statement of lemma 1.1 in the case of D_8 using figure 2. For a given root system Φ with simple system Δ we let $W_0 = \langle s_\alpha \mid \alpha \in \Delta\rangle$. What we would like to show is that $W_0 = W$. For example, this would tell us that the dihedral groups are all generated by precisely two reflections. First however we require a definition.

Definition. Let Φ be a root system with simple system Δ. For any $\beta \in \Phi$ we have that $\beta = \sum_{\gamma\in\Delta} c_\gamma\gamma$ and we define the *height* of β to be $\mathrm{ht}(\beta) = \sum_{\gamma\in\Delta} c_\gamma$.

Theorem 1.3. *Let Φ be a root system with fixed simple system Δ. Then W is generated by simple reflections, in other words $W = W_0$.*

Proof. Let $\beta \in \Phi^+$ and consider the set $W_0\beta \cap \Phi^+$. This will be a subset of Φ^+ which is non-empty as it at least contains β. Choose an element $\gamma \in W_0\beta \cap \Phi^+$ such that $\mathrm{ht}(\gamma)$ is minimal amongst all such elements. We claim that $\gamma \in \Delta$ is a simple root. We have that $\gamma = \sum_{\alpha\in\Delta} c_\alpha\alpha$ such that $c_\alpha \geqslant 0$ and $0 < (\gamma, \gamma) = \sum_{\alpha\in\Delta} c_\alpha(\gamma, \alpha)$ and so $(\gamma, \alpha) > 0$ for at least one $\alpha \in \Delta$. If $\gamma = \alpha$ then we're done. If not then consider $s_\alpha\gamma$, which is positive by lemma 1.1. Recall that $(\gamma, \alpha) > 0$ and so by eq. (1) we obtain $s_\alpha\gamma$ from γ by subtracting a positive multiple of α. Therefore we have $\mathrm{ht}(s_\alpha\gamma) < \mathrm{ht}(\gamma)$. However, as $s_\alpha \in W_0$, we have $s_\alpha\gamma \in W_0\beta \cap \Phi^+$ and so $s_\alpha\gamma$ contradicts the minimality of γ.

We have just shown that for any positive root $\beta \in \Phi^+$ there exists a $w \in W_0$ and $\gamma \in \Delta$ such that $w\beta = \gamma \Rightarrow \beta = w^{-1}\gamma$. Therefore the set of positive roots is contained in the W_0 orbit of Δ, or in other words $\Phi^+ \subseteq W_0\Delta$. Assume $\beta \in \Phi^-$ then $-\beta \in \Phi^+$ and so there exists some $w' \in W_0$ and $\gamma' \in \Delta$ such that $-\beta = w'\gamma' \Rightarrow \beta = w'(-\gamma') = (w's_{\gamma'})\gamma'$. So we have $w's_{\gamma'} \in W_0$ and hence $\beta \in W_0\Delta \Rightarrow \Phi^- \subseteq W_0\Delta$. Therefore $\Phi \subseteq W_0\Delta$.

Let $\beta \in \Phi$ so $s_\beta \in W$. By the previous argument we have $\beta = w\alpha$ for some $w \in W_0$ and $\alpha \in \Delta$. By proposition 1.1 we have $ws_\alpha w^{-1} = s_{w\alpha} = s_\beta \in W_0$ and so $W = W_0$. □

Corollary 1.2. *Let Φ be a root system with fixed simple system Δ then given any $\beta \in \Phi$ there exists an element $w \in W$ such that $w\beta \in \Delta$.*

Remark. Let Φ be a root system with simple system $\Delta = \{\alpha_1, \dots, \alpha_n\}$. If we write $w = s_1 \cdots s_r$ for some $w \in W$ then we mean $s_i = s_{\alpha_i}$ for some $\alpha_i \in \Delta$. Note that we are allowing $s_i = s_j$ whenever $j \neq i+1$ or $i-1$.

Let Φ be a root system with simple system Δ of rank n and let $\{s_1, \dots, s_n\}$ be the set of all simple reflections. What we have just shown is that any element $w \in W$ can be expressed as a product of these simple reflections. This leads us to a quite natural function for W.

Definition. We define a function $\ell : W \to \mathbb{N}$ called the *length function* such that $\ell(1) = 0$ and for all $1 \neq w \in W$ we have $\ell(w)$ is the minimal number m such that w can be expressed as a product of m simple reflections. We call an expression for $w \in W$ *reduced* if it is a product of $\ell(w)$ simple reflections.

Remark. Just from this definition, and what we have already established, we can now list some basic properties of the length function.

- Clearly $\ell(w) = 1$ if and only if $w = s_i$ for some $1 \leqslant i \leqslant n$.
- We have $\ell(w) = \ell(w^{-1})$. This is because if $w = s_1 \cdots s_r$ then $w^{-1} = s_r \cdots s_1$ and vice versa.
- Recall that a reflection $s_\alpha : V \to V$ for some $\alpha \in \Delta$ is a linear operator on the vector space V with $\det(s_\alpha) = -1$. This means that for any $w \in W$ we have $\det(w) = (-1)^{\ell(w)}$.
- Indeed if $w \in W$ can be written as a product of r reflections then $\det(w) = (-1)^r$. This means that if $\ell(w)$ is even/odd then r must be even/odd. For $w, w' \in W$ we have

$$(-1)^{\ell(ww')} = \det(ww') = \det(w)\det(w') = (-1)^{\ell(w)}(-1)^{\ell(w')} = (-1)^{\ell(w)+\ell(w')}.$$

  Hence we have if $\ell(ww')$ is even/odd then $\ell(w) + \ell(w')$ is even/odd. In particular if $w \in W$ is such that $\ell(w) = r$ then for any $\alpha \in \Delta$ we have $\ell(s_\alpha w)$ is either $r+1$ or $r-1$.

Unfortunately as natural a concept as the length function is, it is not in general very practical for proofs. Instead for all $w \in W$ we define another integer $n(w)$ to be the number of positive roots $\alpha \in \Phi^+$ such that $w(\alpha) \in \Phi^-$. Recall that if $\alpha \in \Phi$ is a simple root then by lemma 1.1 we have $n(s_\alpha) = \ell(s_\alpha) = 1$, as s_α only sends α to its negative. In fact it is true in general that $n(w) = \ell(w)$ for all $w \in W$. We will not prove this but we mention the key result used in the proof of this statement known as the *deletion condition*.

Theorem 1.4 (Deletion Condition)**.** *Let $w = s_1 \cdots s_r$ be an expression of $w \in W$ as a product of simple reflections. Suppose $n(w) < r$. Then there exist integers $1 \leqslant j \leqslant k \leqslant r$ such that*

$$w = s_1 \cdots \hat{s_j} \cdots \hat{s_k} \cdots s_r,$$

where $\hat{}$ denotes an omitted simple reflection.

Proof. See [Car05, Theorem 5.15]. $\square$

Remark. The deletion condition tells us that we can obtain a reduced expression for $w \in W$ from any given expression simply by removing simple reflections. There is a more enlightening version of the deletion condition known as the *exchange condition*. This can be found in [Hum90, Section 1.7] and also a format of it is stated in [Gec03, Corollary 1.6.8].

The exchange condition says the following. Let $w \in W$ have a reduced expression $w = s_1 \cdots s_r$. Consider a simple reflection $s = s_\alpha$, for some $\alpha \in \Delta$, such that $\ell(sw) < \ell(w)$ then $sw = s_1 \cdots \hat{s_i} \cdots s_r$ for some $1 \leqslant i \leqslant r$ and so $w = ss_1 \cdots \hat{s_i} \cdots s_r$. This says that in the reduced expression for w a factor s is exchanged for a factor s_i.

Corollary 1.3. *We have $n(w) = \ell(w)$ for all $w \in W$.*

Proof. By definition we have $n(w) \leqslant \ell(w)$ because every time we apply the reflection in a reduced expression for w we either increase $n(w)$ by 1 or do nothing. Therefore assume for a contradiction that $n(w) < \ell(w)$, then by the deletion condition we could find an expression for w of length $\ell(w) - 2$. However $\ell(w)$ was defined to be the length of a reduced expression for w hence this is a contradiction. $\square$

Remark. As a consequence of this corollary we have that for any $w \in W$ we have

$$w\Delta = \Delta \Leftrightarrow w\Phi^+ = \Phi^+ \Leftrightarrow \ell(w) = n(w) = 0 \Leftrightarrow w = 1.$$

This means that the reflection group W acts *simply transitively* on the collection of simple systems for Φ. In other words the action is transitive and free.

So what we now know is that the length of any element in W is in fact equal to the number of positive roots made negative by w. Indeed it makes sense to now consider which elements are maximal with respect to this property. Therefore these are all elements $w \in W$ such that $w(\Phi^+) = \Phi^-$. In fact what can be shown is that given any root system with a fixed positive system there is a unique element with this property, which we call the *longest element* or *longest word*.

Proposition 1.2. *Let Φ be a root system with associated reflection group W and simple system Δ. The following are true.*

(a) *The maximal length of any element of W is $|\Phi^+|$.*
(b) *W has a unique element w_0 with $\ell(w_0) = |\Phi^+|$.*
(c) *$w_0(\Phi^+) = \Phi^-$.*
(d) *$w_0^2 = 1$.*

Proof. By corollary 1.3 we have $\ell(w) = n(w)$ and so $\ell(w) \leqslant |\Phi^+|$ by definition of $n(w)$. We have that Δ a simple system implies $-\Delta$ is a simple system coming from the positive system Φ^-. By theorem 1.2 we have there exists $w_0 \in W$ such that $w_0(\Phi^+) = \Phi^-$ and so $\ell(w) = n(w) = |\Phi^+|$. Hence w_0 is an element of maximal length.

Assume $w_0' \in W$ is such that $\ell(w_0') = n(w_0') = |\Phi^+|$. Then $w_0'(\Phi^+) = \Phi^-$. Let $w = (w_0')^{-1}w_0$ then $w(\Phi^+) = \Phi^+$, which means $\ell(w) = n(w) = 0$ and so $w = 1 \Rightarrow (w_0')^{-1}w_0 = 1 \Rightarrow w_0 = w_0'$. Thus w_0 is a unique element of maximal length. Finally we have $w_0^2(\Phi^+) = \Phi^+$, which means $\ell(w_0^2) = n(w_0^2) = 0$ and so $w_0^2 = 1$ as required. $\square$

Remark. A word of caution. Just because the longest element is an element of order 2 in W does not mean it is a reflection. This can even be seen not to be true in our example of D_8.

Finally with all these ingredients it is now possible to prove the following theorem that gives a presentation for the finite reflection group W. We will not prove this theorem here but instead reference locations to the standard proof which is attributed to Steinberg. The following theorem tells us that W is in fact a *finite Coxeter group*.

Theorem 1.5. *Let Φ be a root system with fixed simple system $\Delta = \{s_1, \ldots, s_n\}$ and associated reflection group W. Let m_{ij} denote the order of the element $s_i s_j$ for $1 \leqslant i, j \leqslant n$ and $i \neq j$. Then we have*

$$W \cong \langle s_i \mid s_i^2 = (s_i s_j)^{m_{ij}} = 1 \text{ for } i \neq j \rangle$$

Proof. See either [Hum90, Theorem 1.9] or [Car05, Theorem 5.18]. □

Example. We consider all the information of this section for the reflection group D_8. Let us fix the root system Φ and simple system $\Delta = \{\alpha, \beta\}$ as was specified in figure 2. Recall that the product $s_\beta s_\alpha$ was the same as the rotation τ from figure 1, hence this product has order 4. Thus theorem 1.5 tells us that

$$D_8 = \langle s_\alpha, s_\beta \mid (s_\alpha s_\beta)^4 = s_\alpha^2 = s_\beta^2 = 1 \rangle.$$

In table 1 we give the description of the elements of D_8 in terms of the reflection and rotation specified in figure 1 and also the presentation given above in terms of simple reflections from figure 2. We have also listed the order of the element and if the element was a reflection we have listed the unique positive root sent to its negative by that reflection.

Table 1. The Various Presentations of D_8.

o(w)	1	2	2	4	4	2	2	2
Φ	—	α	β	—	—	$2\alpha+\beta$	$\alpha+\beta$	—
$\langle s_\alpha, s_\beta\rangle$	1	s_α	s_β	$s_\alpha s_\beta$	$s_\beta s_\alpha$	$s_\alpha s_\beta s_\alpha$	$s_\beta s_\alpha s_\beta$	$s_\alpha s_\beta s_\alpha s_\beta$
$\langle \sigma, \tau\rangle$	1	σ	$\tau\sigma$	τ^3	τ	$\tau^3\sigma$	$\tau^2\sigma$	τ^2

From this table it is clear to see that the element $s_\alpha s_\beta s_\alpha s_\beta$ or τ^2 is the longest element. Indeed it is easy to check that this element sends every positive root to its negative. Note that we also have that $s_\beta s_\alpha s_\beta s_\alpha$ is a reduced expression for the longest element.

1.4. **Parabolic Subgroups and Chamber Systems.** So far we have defined a certain class of groups known as finite reflection groups. These are groups that act in a geometric way on a finite collection of roots in a vector space. In theorem 1.5 we have given a presentation for these groups, which shows that they are finite Coxeter groups. Using this presentation we would like to understand a little more about their subgroup structure and extract some information from the underlying geometry. We start by introducing a special class of subgroups.

Definition. Let Φ be a root system with simple system Δ and associated set of simple reflections S. For any subset $I \subseteq S$ we define $W_I \subseteq W$ to be the subgroup of W generated by all the simple reflections $s_\alpha \in I$ and let $\Delta_I := \{\alpha \in \Delta \mid s_\alpha \in I\}$. We call a subgroup of the form W_I a *parabolic subgroup* of W.

Remark. Note that if Δ is a simple system for Φ then by theorem 1.2 any other simple system is of the form $w\Delta$ for some $w \in W$. Hence if we have a parabolic subgroup W_I with respect to some simple system Δ then replacing Δ by $w\Delta$ will just replace W_I by wW_Iw^{-1}.

Example. Consider our running example of D_8. The root system of this group has a simple system given by $\Delta = \{\alpha, \beta\}$. Therefore the possible subsets I are given by $\varnothing$, $\{s_\alpha\}$, $\{s_\beta\}$, $\{s_\alpha, s_\beta\}$. So in turn the possible parabolic subgroups are $\{1\}$, $\langle s_\alpha \rangle$, $\langle s_\beta \rangle$ and D_8 itself.

Parabolic subgroups are important as they are themselves finite reflection groups. For example, it will be shown later that the symmetric group is a finite reflection group and every parabolic subgroup is a direct product of symmetric groups.

Lemma 1.2. *Let Φ be a root system with fixed simple system Δ and let S be the set of associated simple reflections. Take a subset $I \subset S$ and define $V_I := \operatorname{span}_{\mathbb{R}}(\Delta_I)$ and $\Phi_I = \Phi \cap V_I$. Then Φ_I is a root system in V_I with simple system Δ_I and with corresponding reflection group W_I.*

Proof. It's clear by definition that Φ_I spans V_I and hence (R1) is satisfied. Now as Φ_I is a subset of Φ then it's clear that (R2) is satisfied. Clearly we have W_I stabilises V_I and so we have that (R3) is satisfied. Hence Φ_I is a root system for V_I. Indeed as Δ is a vector space basis for V we have Δ_I will be a vector space basis for V_I and is clearly a simple system for Φ_I. Hence W_I will be the associated reflection group of Φ_I. □

Now having introduced these important subgroups of the reflection group we would like to consider more closely the description of W acting on V. Let Φ be a root system of V and fix a positive system Φ^+ with simple system Δ. Now we recall that each root $\alpha \in \Phi$ defines a corresponding hyperplane H_α. In general $V - \cup_{\alpha\in\Phi} H_\alpha$ is not a connected topological space and we call its connected components the *chambers* of W. Let C be a chamber then the *walls* of C are all the hyperplanes H_α such that $\overline{C} \cap H_\alpha$ has codimension 1 in V, where $\overline{C}$ denotes the topological closure of C.

To each hyperplane H_α we can associate two open half spaces A_α and A'_α where

$$A_\alpha = \{\lambda \in V \mid (\lambda, \alpha) > 0\}$$

and $A'_\alpha = -A_\alpha$. Given a fixed simple system we can define a canonical chamber $C_1 = \cap_{\alpha\in\Delta} A_\alpha$, which we refer to as the *fundamental chamber*. We have C_1 is open and convex as it is the intersection of open convex sets. Let $D_1 = \overline{C_1}$ be the topological closure of C_1, this is the intersection of closed half-spaces $H_\alpha \cup A_\alpha$. In other words

$$D_1 = \{\lambda \in V \mid (\lambda, \alpha) \geqslant 0 \text{ for all } \alpha \in \Delta\}.$$

Recall from the remark after corollary 1.3 that W acts simply transitive on the collection of simple systems for Φ. If we fix a simple system Δ for Φ then we have a corresponding chamber C_1. If we then replace Δ by $w\Delta$ for some $w \in W$ we replace C_1 by $C_w := w(C_1)$. This gives us a geometric interpretation of the simply transitive action on simple systems.

Theorem 1.6. *Let Φ be a root system with associated reflection group W. Then W acts simply transitively on the collection of chambers. Furthermore the topological closure of a chamber is a fundamental domain for the action of W on V.*

Proof. See [Hum90, Theorem 1.12]. □

By a *fundamental domain* we mean that each $\lambda \in V$ is conjugate under the action of W on V to one and only point in D_1. Given a fixed simple system Δ we call the collection of chambers a *chamber system* for Φ.

Example. We consider our running example of D_8. In figure 3 we indicate the chamber system and the closure of the canonical chamber D_1 with respect to our chosen simple system in figure 2.

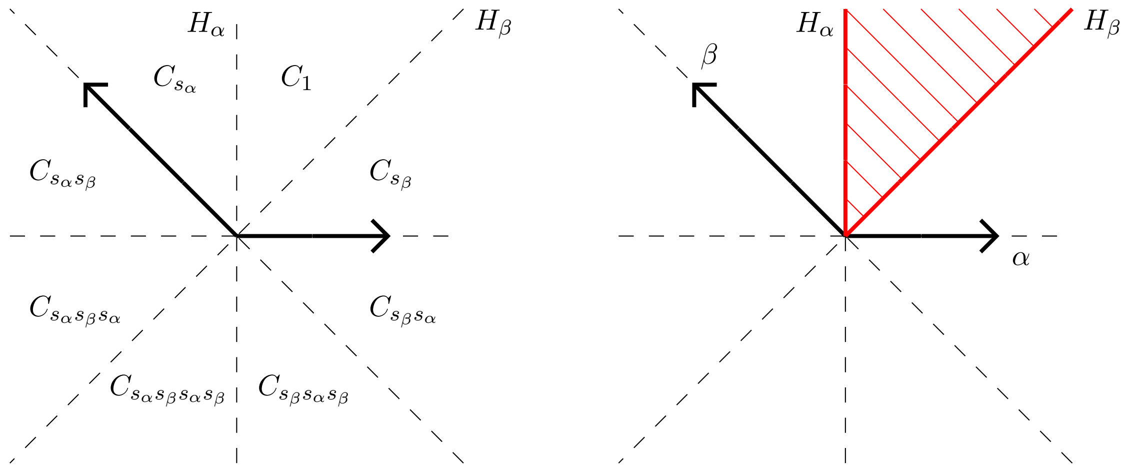
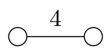


FIGURE 3. The Chamber System of B_2 and Fundamental Domain D_1.

1.5. **Coxeter Graphs, Dynkin Diagrams and Cartan Matrices.** Let Φ be a root system of rank n with simple system $\Delta = \{\alpha_1, \dots, \alpha_n\}$. Let W be the associated reflection group with set of generators $S = \{s_1, \dots, s_n\}$ corresponding to the simple roots and let m_{ij} be the order of the element $s_i s_j$ for $i \neq j$. Given the presentation of W given in theorem 1.5 we can see that W is determined up to isomorphism explicitly by the set S and the numbers m_{ij}.

We wish to encode this information visually using a graph. We let the vertices of our graph be in bijective correspondence with Δ and join a pair of vertices corresponding to distinct $\alpha_i, \alpha_j \in \Delta$ by an edge whenever $m_{ij} \geqslant 3$. We will label such an edge by m_{ij}. We call this labelled graph the *Coxeter graph* of W.

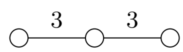
Example. Consider D_8. This group has a presentation given by $\langle s_1, s_2 \mid s_1^2 = s_2^2 = (s_1 s_2)^4 = 1\rangle$. Therefore the associated graph of D_8 is given by



Consider $\mathfrak{S}_4$, the symmetric group on 4 symbols. It can be shown that this symmetric group has the following Coxeter presentation

$$\mathfrak{S}_4 = \langle (12), (23), (34) \mid ((12)(34))^3 = ((23)(34))^3 = 1\rangle.$$

Therefore we have the following associated Coxeter graph of $\mathfrak{S}_4$.



The classification of the possible finite reflection groups relies heavily upon the associated Coxeter graphs. It can be shown that if two finite reflection groups have the same Coxeter graphs then there is an isometry of their underlying vector spaces which induces an isomorphism of the reflection groups. Hence the Coxeter graph of a finite reflection group determines it uniquely up to isomorphism, (see [Hum90, Proposition 2.1]).

Therefore we reduce the problem of classifying the finite reflection groups to one of classifying the associated Coxeter graphs. We note that we need a notion of irreducibility before we can carry out such a classification. In the spirit of using the Coxeter graph we make the following definition.

Definition. A finite reflection group W is called *irreducible* if its associated Coxeter graph is connected.

Theorem 1.7. *If W is an irreducible finite reflection group then its associated Coxeter graph is one from the list in figure 4. The label n on the first three types denotes the number of vertices on the graph. To remove redundancies we assume in the case of $\mathrm{I}_2(m)$ that $m \geqslant 5$. Also any unlabelled edge is assumed to have label 3.*

Proof. A nice succinct proof of this theorem can be found in [Hum90, Theorem 2.7]. Alternatively one can look at [Bou02, Chapter VI – Section 4 – Theorem 1] □

Indeed to each of these Coxeter graphs there does exist a root system and finite reflection group. The finite reflection groups of types $\mathrm{A}_n, \mathrm{B}_n, \mathrm{D}_n$ and $\mathrm{I}_2(m)$ have nice descriptions in terms of groups that we already know. We usually refer to these as finite reflection groups of *classical type*. However the groups $\mathrm{E}_6, \mathrm{E}_7, \mathrm{E}_8, \mathrm{F}_4, \mathrm{H}_3$ and H_4 have slightly more quirky descriptions and we refer to these as finite reflection groups of *exceptional type*. Lots of information on these groups can be found in [Hum90, Chapter 2] and the plates in the back of [Bou02].

In table 2 we give some basic information about the finite reflection groups of classical type. As an example we will also construct the root system in type A_n and show how the associated finite reflection group is given. Note that we denote the cyclic group of order 2 by $\mathbb{Z}_2$.

TABLE 2. Information for Finite Reflection Groups of Classical Type.

	A_n	B_n	D_n	$\mathrm{I}_2(m)$
W	$\mathfrak{S}_{n+1}$	$\mathfrak{S}_n \ltimes \mathbb{Z}_2^n$	$\mathfrak{S}_n \ltimes \mathbb{Z}_2^{n-1}$	D_{2m}
$\|W\|$	$n!$	$2^n n!$	$2^{n-1} n!$	$2m$
$\|\Phi\|$	$n(n+1)$	$2n^2$	$2n(n-1)$	$2m$

Example (Type A_n $(n \geqslant 1)$)**.** To gather information about this case we need to construct a root system of rank n. This will be a root system of the real Euclidean vector space $V \subset \mathbb{R}^{n+1}$ where

$$V = \left\{ (x_1, \dots, x_{n+1}) \;\middle|\; \sum_{i=1}^{n} x_i = 0 \right\}.$$

It is readily checked that this vector space is of dimension n. Let $\varepsilon_1, \dots, \varepsilon_{n+1}$ be the standard orthonormal basis of $\mathbb{R}^{n+1}$ then it is a quick check to verify that $\Delta = \{\varepsilon_1 - \varepsilon_2, \varepsilon_2 - \varepsilon_3, \dots, \varepsilon_n -$

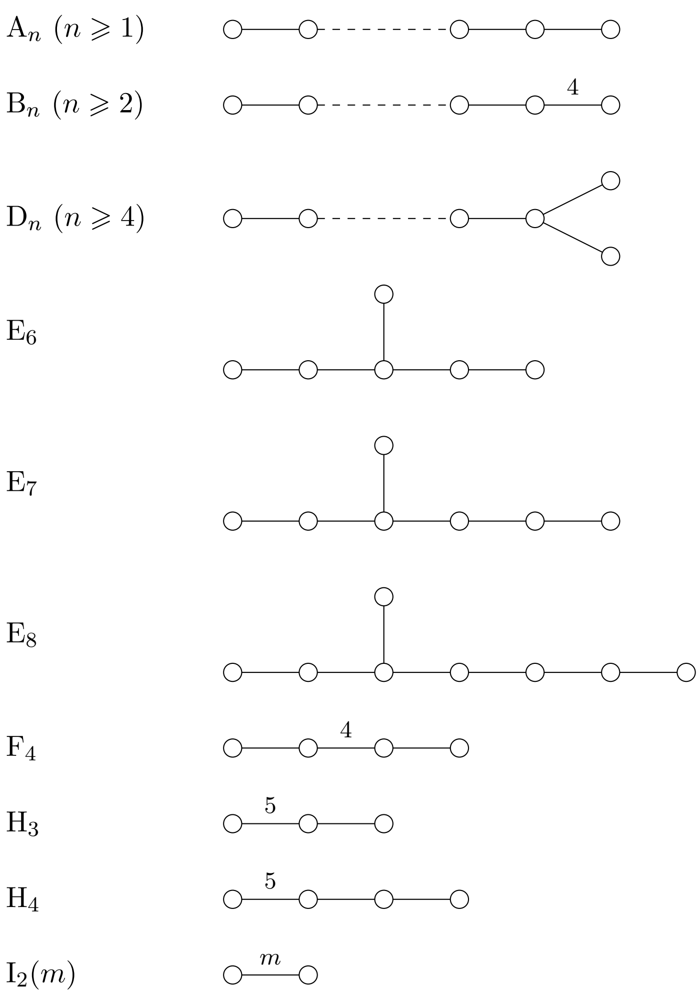


FIGURE 4. The Coxeter Graphs of Irreducible Finite Reflection Groups.

$\varepsilon_{n+1}\}$ is a basis for V. We define Φ to be the set of all $\lambda \in V \cap (\mathbb{Z}\varepsilon_1 \oplus \cdots \oplus \mathbb{Z}\varepsilon_{n+1})$ such that $||\lambda||^2 = 2$. In fact this implies that every element of Φ is of the form $\varepsilon_i - \varepsilon_j$ such that $1 \leqslant i \neq j \leqslant n+1$. Hence we have $|\Phi| = n(n+1)$ just as in table 2.

It's easy to see that Φ will be a root system for V. Indeed it is a root system of rank n with simple system given by Δ. In the case of A_2 we have the associated finite reflection group is $\mathfrak{S}_3 \cong D_6$. Note that the Coxeter graph of type A_2 is the same as the Coxeter graph of $I_2(3)$. Indeed this example tells us that we can realise the group D_6 either as the symmetry group of an equialateral triangle in $\mathbb{R}^2$ or as a symmetry group in $\mathbb{R}^3$.

It is now time to discuss our definition of root system introduced in section 1.2. So far we have not mentioned anything about the lengths of the vectors in our root system. For example in our case of D_8 it is clear from figure 2 that the roots in our simple system are of different lengths. To classify the finite reflection groups we associate to each Coxeter graph a symmetric $n \times n$ matrix. During this process we force the length of each root to be of unit length. Of course rescaling the vectors in a root system does not stop it being a root system and indeed does not change the associated finite reflection group. However we would like to make a distinction between root systems where roots have different lengths.

Let V be a real Euclidean vector space and consider a subgroup $G \leqslant \mathrm{GL}(V)$. We say G is *crystallographic* if it stabilises a lattice L in V. By a *lattice* L we mean the $\mathbb{Z}$ span of a basis of a real vector space. This leads us to the following definition.

Definition. Let V be a Euclidean vector space and Φ a root system for V. We say Φ is a *crystallographic root system* if Φ also satisfies:

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

Remark. By requiring that the values $\langle \beta, \alpha \rangle$ are integers we are ensuring, by eq. (1), that all roots in Φ are $\mathbb{Z}$ linear combinations of the simple system Δ. Therefore the $\mathbb{Z}$ span of Δ is a lattice which is stable under the action of the associated finite reflection group.

Definition. Let Φ be a crystallographic root system. Then we call the associated finite reflection group W the *Weyl group* of Φ.

Proposition 1.3. *Let Φ be a crystallographic root system and let $\Delta = \{\alpha_i\}_{1 \leqslant i \leqslant n}$ be a simple system for Φ. Also let $S = \{s_i\}_{1 \leqslant i \leqslant n}$ be the associated simple reflections. We have for any distinct $\alpha, \beta \in \Phi$ that the angle θ between these two roots is such that*

$$\theta \in \{\pi/2, \pi/3, 2\pi/3, \pi/4, 3\pi/4, \pi/6, 5\pi/6\}. \tag{2}$$

Consequently we have that the order of each element $s_i s_j$ is such that $m_{ij} \in \{2, 3, 4, 6\}$ for all $i \neq j$.

Proof. Let $\alpha, \beta \in \Phi$ be any two roots. Recall that as V is a Euclidean vector space we have $(\alpha, \beta) = ||\alpha|| \cdot ||\beta|| \cos\theta$, where θ is the angle between the two vectors. By eq. (1) we have that

$$\langle \beta, \alpha \rangle = 2\frac{(\beta, \alpha)}{(\alpha, \alpha)} = 2\frac{||\beta||}{||\alpha||}\cos\theta \Rightarrow \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos^2\theta.$$

By (R4) we have that $4\cos^2\theta$ is a non-negative integer. However $0 \leqslant \cos^2\theta \leqslant 1$ and so we must have $\cos^2\theta \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\} \Rightarrow \cos\theta \in \{0, \pm\frac{1}{2}, \pm\frac{1}{\sqrt{2}}, \pm\frac{\sqrt{3}}{2}, \pm 1\}$. It's easy to verify that the only possible values of $0 < \theta \leqslant \pi$ are the ones appearing in the list in eq. (2).

Consider $\alpha_i, \alpha_j \in \Delta$ with $i \neq j$. We know that the reflection $s_i s_j \neq 1$ acts on the plane spanning α_i and α_j as a rotation through the angle $\theta = 2\pi/m_{ij}$. Indeed the angle between these two vectors is given by π/m_{ij}. By the list given in eq. (2) we can see that m_{ij} must indeed be either $2, 3, 4$ or 6. □

Corollary 1.4. *The root systems of types $\mathrm{H}_3, \mathrm{H}_4$ and $\mathrm{I}_2(m)$ for $m = 5$ or $m \geqslant 7$ are not crystalographic.*

Proof. Clear from the list in figure 4 using proposition 1.3. □

This corollary tells us that the dihedral groups are not crystallographic except for D_2, D_4, D_6, D_8 and D_{12}. These Weyl groups correspond to root systems of types A_1, $\mathrm{A}_1 \times \mathrm{A}_1$, A_2, B_2 and $\mathrm{I}_2(6)$. In figure 5 we indicate a simple system for the root systems associated to the dihedral groups D_6, D_8 and D_{12}. Note that the entire root system can be obtained from these diagrams by letting W act via reflections.

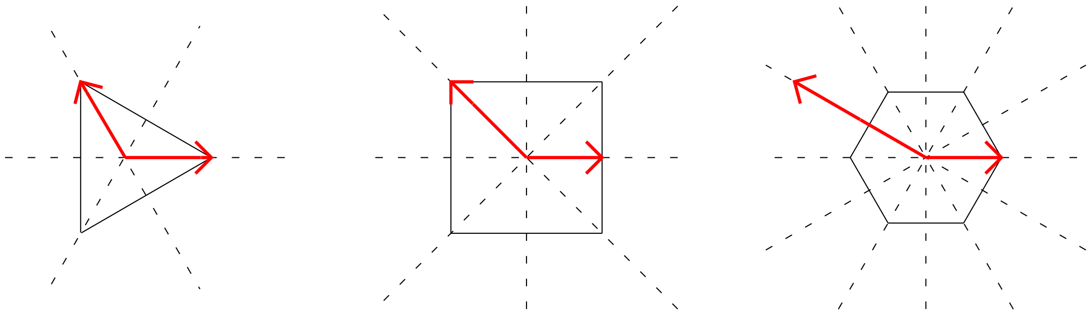
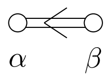


FIGURE 5. The Crystallographic Root Systems of D_6, D_8 and D_{12}.

Now we would like to classify the crystallographic root systems in a similar fashion to how we classified the finite reflection groups. Although similar, this classification will be slightly different. The Coxeter graphs will not be sufficient for this task as we would like to indicate when two roots are of different lengths. To encode this information we will need the *Dynkin diagram* of Φ. In any irreducible crystallographic root system it turns out that there are at most two different root lengths. Therefore when there are two roots of different lengths we refer to them as *long* and *short* roots.

Also we know by proposition 1.3 that the only possible labels on edges we require are $3, 4$ or 6. Therefore in the Dynkin diagram of Φ instead of using labels we indicate these numbers by joining two vertices with a single, double or triple bond. If two vertices joined by an edge represent simple roots of different lengths then we indicate the distinction by placing an arrow on the edge pointing to the short root.

Example. Consider our recurring example of D_8. Now the root system Φ which has D_8 as its associated Weyl group has a simple system $\Delta = \{\alpha, \beta\}$ as specified in figure 2. Now it's easy to see from the diagram that β is the long root and α is the short root. Therefore we have the associated Dynkin diagram of the root system is



Theorem 1.8. *Let Φ be an irreducible crystallogrphic root system then its associated Dynkin diagram is one from the list given in figure 6. The label n on the first four types denotes the number of vertices on the graph.*

Proof. See the proof in [Hum78, Section 11.4] or alternatively see the proof given in [EW06, Chapter 13]. □

We mentioned that to classify the finite reflection groups we associated to each Coxeter graph a symmetric $n \times n$ matrix. Indeed to each Dynkin diagram of a crystallographic root system we associate a uniquely specified matrix called the Cartan matrix. However we note that this matrix is *not* symmetric.

Definition. Let Φ be a crystallographic root system with simple system $\Delta = \{\alpha_1, \dots, \alpha_n\}$. Recall by (R4) that $\langle \alpha_i, \alpha_j \rangle$ is an integer for all i, j. We call the integers $\langle \alpha_i, \alpha_j \rangle$ *Cartan integers*. The *Cartan matrix* associated to Φ is the $n \times n$ matrix $(\langle \alpha_i, \alpha_j \rangle)_{1 \leqslant i,j \leqslant n}$.

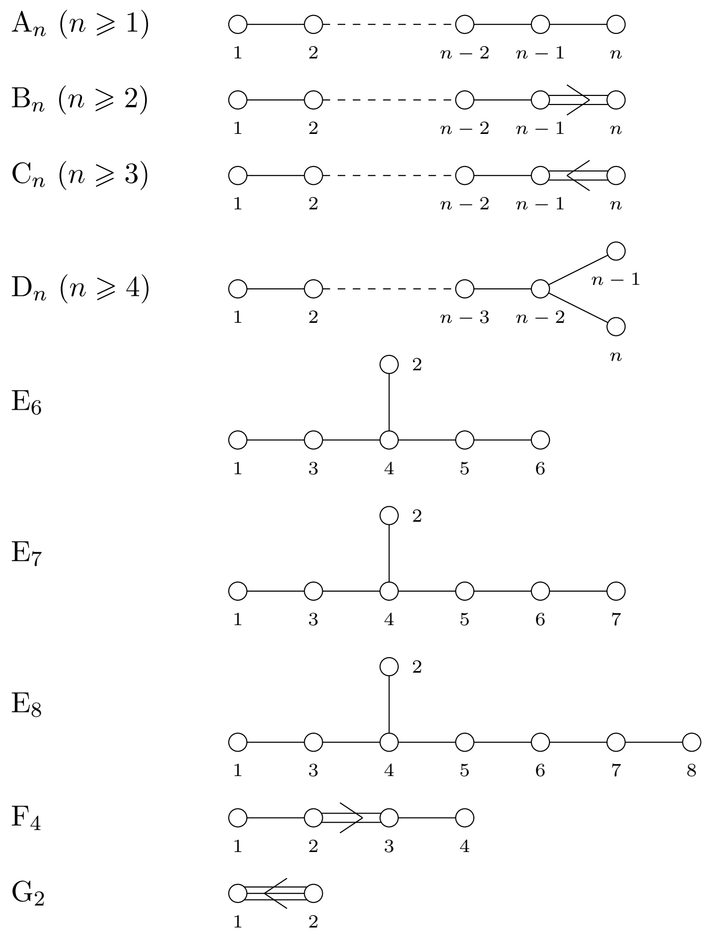


Figure 6. The Dynkin Diagrams of Irreducible Crystallographic Root Systems.

The Cartan matrix depends on the ordering of the simple roots but this is not such a hardship. In fact the important thing to realise is that the Weyl group acts simply transitively on the collection of simple systems for Φ and so the Cartan matrix is independent of the choice of Δ. In fact the Cartan matrix is a non-singular matrix which, like the Dynkin diagram, characterises Φ completely.

Example. We have the Cartan matrices of the Weyl groups D_6, D_8 and D_{12} are given by

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & -2 \\ -1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}.$$

It is the crystallographic root systems which underpin the structural information of Lie theory. It is such that these diagrams encapsulate the structural information of complex simple Lie algebras, simple Lie groups, simple linear algebraic groups and finite groups of Lie type. Again detailed information about the construction of each of these root systems can be found in [Hum90, Chapter 2] and the plates in the back of [Bou02].

Remark. A word of caution. Exactly how one draws and labels the Dynkin diagrams is very subjective. Here we have chosen to follow the description given in [Hum78, Theorem 11.4]. However this differs to the plates in [Bou02], which differs to the library in [GHL$^+$96] and so on.

2. Lie Algebras

2.1. **Basic Definitions and Introduction.** Throughout this section we will assume, so as not to over complicate things, that all vector spaces are finite dimensional.

Definition. A *Lie algebra* $\mathfrak{g}$ is a vector space over a field $\mathbb{K}$ with an associated bilinear map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, such that the following hold:

- $[x, x] = 0$ for all $x \in \mathfrak{g}$,
- $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in \mathfrak{g}$.

Remark. We call the latter axiom of the above definition the Jacobi Identity. The idea of this axiom is to be a replacement for associativity, as we do not have that a Lie algebra is an associative algebra. We refer to the bilinear map $[\cdot,\cdot]$ as the *Lie bracket* of $\mathfrak{g}$.

Example.

(a) Let $\mathfrak{g}$ be any vector space over any field $\mathbb{K}$. Then we can endow $\mathfrak{g}$ with the trivial bracket operation $[x, y] = 0$ for all $x, y \in \mathfrak{g}$. We refer to this as an *abelian Lie algebra.*

(b) Let $\mathbb{K} = \mathbb{R}$ and let $\mathfrak{g} = \mathbb{R}^3$. We define a product structure on $\mathfrak{g}$ using the standard vector product $x \wedge y$ for all $x, y \in \mathfrak{g}$. In other words if $x, y \in \mathfrak{g}$ such that $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ then

$$[x, y] = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

(c) Let V be any finite-dimensional vector space over a field $\mathbb{K}$. We define the *general linear Lie algebra* $\mathfrak{gl}(V)$ to be the vector space of all linear maps from V to V, endowed with the commutator bracket

$$[x, y] = x \circ y - y \circ x \qquad \text{for all } x, y \in \mathfrak{gl}(V).$$

(d) We now define a matrix analogue for the Lie algebra in example (c). Let $\mathbb{K}$ be any field and let $\mathfrak{gl}(n, \mathbb{K})$ be the vector space of all $n \times n$ matrices defined over $\mathbb{K}$. Then $\mathfrak{gl}(n, \mathbb{K})$ is a Lie algebra with Lie bracket given by

$$[x, y] = xy - yx \qquad \text{for all } x, y \in \mathfrak{gl}(n, k),$$

i.e. the commutator bracket. Note that a basis for $\mathfrak{gl}(n, \mathbb{K})$ as a vector space is given by the $n \times n$ unit matrices e_{ij} which have entry 1 in the ijth position and zeros elsewhere. We then see that the commutator bracket is given by

$$[e_{ij}, e_{k\ell}] = \delta_{jk}e_{i\ell} - \delta_{i\ell}e_{kj},$$

where δ_{ij} is the Kronecker delta.

(e) Let $\mathbb{K}$ be any field and $\mathfrak{sl}(2, \mathbb{K}) = \{x \in \mathfrak{gl}(2, \mathbb{K}) \mid \mathrm{tr}(x) = 0\} \subset \mathfrak{gl}(2, \mathbb{K})$ be the vector subspace of $\mathfrak{gl}(2, \mathbb{K})$ whose elements have trace 0. Now if $x, y \in \mathfrak{sl}(2, \mathbb{K})$ then we will have $[x, y] = xy - yx \in \mathfrak{sl}(2, \mathbb{K})$, hence the commutator bracket gives $\mathfrak{sl}(2, \mathbb{K})$ a Lie algebra structure. Assume $\mathbb{K} = \mathbb{C}$ then as a vector space it can be shown that $\mathfrak{sl}(2, \mathbb{K})$ has a basis given by

$$e = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad f = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

These elements have Lie bracket relations $[e, f] = h$, $[h, f] = -2f$, $[h, e] = 2e$.

(f) Let A be an associative algebra over a field $\mathbb{K}$. Clearly A is a vector space over $\mathbb{K}$ and we can give it the structure of a Lie algebra by endowing it with the commutator bracket $[x,y] = xy - yx$ for all $x, y \in A$.

Definition. Let $\mathfrak{g}$ be a Lie algebra, then we define $Z(\mathfrak{g}) := \{x \in \mathfrak{g} \mid [x,y] = 0 \text{ for all } y \in \mathfrak{g}\}$ to be the *centre* of the Lie algebra.

Definition. Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ a vector subspace of $\mathfrak{g}$. We can consider the Lie bracket of $\mathfrak{h}$ to be the restriction of the Lie bracket of $\mathfrak{g}$ to $\mathfrak{h}$.

- We say $\mathfrak{h}$ is a *subalgebra* of $\mathfrak{g}$ if $[\mathfrak{h},\mathfrak{h}] \subseteq \mathfrak{h}$, where $[\mathfrak{h},\mathfrak{h}] := \operatorname{span}\{[h_1,h_2] \mid h_1,h_2 \in \mathfrak{h}\}$.
- We say $\mathfrak{h}$ is an *ideal* of $\mathfrak{g}$ if $[\mathfrak{h},\mathfrak{g}] \subseteq \mathfrak{h}$, where $[\mathfrak{h},\mathfrak{g}] := \operatorname{span}\{[h,g] \mid h \in \mathfrak{h} \text{ and } g \in \mathfrak{g}\}$.

Example. For any field $\mathbb{K}$ we have that $\mathfrak{sl}(2,\mathbb{K}) \subset \mathfrak{gl}(2,\mathbb{K})$ is a Lie subalgebra and an ideal of $\mathfrak{gl}(2,\mathbb{K})$. This is because for all $x,y \in \mathfrak{gl}(2,\mathbb{K})$ we have $\operatorname{tr}([x,y]) = \operatorname{tr}(xy - yx) = \operatorname{tr}(xy) - \operatorname{tr}(yx) = 0$ and hence $[x,y] \in \mathfrak{sl}(2,k)$.

Proposition 2.1. *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ an ideal of $\mathfrak{g}$. Consider the vector space of cosets $\mathfrak{g}/\mathfrak{h} = \{x + \mathfrak{h} \mid x \in \mathfrak{g}\}$. We endow this vector space with the map $[\cdot,\cdot] : \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$ where*

$$[x + \mathfrak{h}, y + \mathfrak{h}] = [x,y] + \mathfrak{h}. \tag{3}$$

Then $\mathfrak{g}/\mathfrak{h}$ is again a Lie algebra called the quotient algebra *with respect to $\mathfrak{h}$.*

Proof. It is easy to check that the axioms of a Lie algebra hold for $\mathfrak{g}/\mathfrak{h}$. Therefore we only have to check that the Lie bracket given in eq. (3) is well defined. Assume $x + \mathfrak{h} = x' + \mathfrak{h}$ and $y + \mathfrak{h} = y' + \mathfrak{h}$ for some $x, x', y, y' \in \mathfrak{g}$. Now this implies that there exists $a, b \in \mathfrak{h}$ such that $x' = a + x$ and $y' = b + y$. Then

$$[x',y'] = [a + x, b + y] = [a,b] + [a,y] + [x,b] + [x,y] \in [x,y] + \mathfrak{h}$$

as $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ and hence $[a,b],[a,y],[x,b] \in \mathfrak{h}$. Therefore $[x,y] + \mathfrak{h} = [x',y'] + \mathfrak{h}$ and the Lie bracket is well defined. □

Definition. Let $\mathfrak{g}$ be a Lie algebra, then a *derivation* $\delta : \mathfrak{g} \to \mathfrak{g}$ is a linear map which satisfies the Leibniz rule

$$\delta([x,y]) = [\delta(x),y] + [x,\delta(y)] \qquad \text{for all } x,y \in \mathfrak{g}.$$

Now $\operatorname{Der}(\mathfrak{g})$, the vector space of all derivations of $\mathfrak{g}$, is a Lie algebra whose Lie bracket is given by the commutator bracket $[\delta_1,\delta_2] = \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1$ for all $\delta_1,\delta_2 \in \operatorname{Der}(\mathfrak{g})$. We define a very important derivation known as the *adjoint operator*. Let $x \in \mathfrak{g}$ then we define a map $\operatorname{ad}_x : \mathfrak{g} \to \mathfrak{g}$ by $\operatorname{ad}_x(y) = [x,y]$ for all $y \in \mathfrak{g}$.

Claim. For any Lie algebra $\mathfrak{g}$ we have $\operatorname{ad}_x \in \operatorname{Der}(\mathfrak{g})$ for all $x \in \mathfrak{g}$.

Proof. First of all we must show that ad_x is linear. For any $\alpha,\beta \in \mathbb{K}$ and $y,z \in \mathfrak{g}$ we have

$$\operatorname{ad}_x(\alpha y + \beta z) = [x, \alpha y + \beta z] = \alpha[x,y] + \beta[x,z] = \alpha \operatorname{ad}_x(y) + \beta \operatorname{ad}_x(z).$$

Hence the map is linear. We now show that this map satisfies the Liebniz rule. For all $y,z \in \mathfrak{g}$ we have

$$\begin{aligned}\mathrm{ad}_x([y,z]) = [x,[y,z]] &= -[y,[z,x]] - [z,[x,y]],\\ &= [y,[x,z]] + [[x,y],z],\\ &= [\mathrm{ad}_x(y),z] + [y,\mathrm{ad}_x(z)].\end{aligned}$$ □

Definition. For any Lie algebra $\mathfrak{g}$ we call a derivation $\delta \in \mathrm{Der}(\mathfrak{g})$ an *inner derivation* if there exists an element $x \in \mathfrak{g}$ such that $\delta = \mathrm{ad}_x$. Any derivation of $\mathfrak{g}$ which is not an inner derivation is called an *outer derivation.*

Note that the derivation ad_x is *not* to be confused with the *adjoint homomorphism.* We define the adjoint homomorphism to be the map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ given by $x \mapsto \mathrm{ad}_x$ for all $x \in \mathfrak{g}$. However, for this to make sense we must define what we mean by a Lie algebra homomorphism.

Definition. Let $\mathfrak{g}_1, \mathfrak{g}_2$ be Lie algebras defined over a common field $\mathbb{K}$. Then a homomorphism of Lie algebras $\varphi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a linear map of vector spaces such that $\varphi([x,y]) = [\varphi(x), \varphi(y)]$ for all $x, y \in \mathfrak{g}_1$. In other words φ preserves the Lie bracket.

Claim. The map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a homomorphism of Lie algebras.

Proof. Clearly this map is linear by the linearity properties of the Lie bracket. Hence to show this is a homomorphism we must show that $\mathrm{ad}_{[x,y]} = [\mathrm{ad}_x, \mathrm{ad}_y] = \mathrm{ad}_x \circ \mathrm{ad}_y - \mathrm{ad}_y \circ \mathrm{ad}_x$ for all $x, y \in \mathfrak{g}$. We do this by showing equivalence for all $z \in \mathfrak{g}$

$$\begin{aligned}\mathrm{ad}_{[x,y]}(z) = [[x,y],z] &= -[z,[x,y]],\\ &= [x,[y,z]] + [y,[z,x]],\\ &= \mathrm{ad}_x([y,z]) - \mathrm{ad}_y([x,z]),\\ &= (\mathrm{ad}_x \circ \mathrm{ad}_y - \mathrm{ad}_y \circ \mathrm{ad}_x)(z).\end{aligned}$$ □

Definition. A *representation* of a Lie algebra $\mathfrak{g}$ is a pair (V, ρ) where V is a vector space over $\mathbb{K}$ and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie algebra homomorphism.

Example.

(a) Take V to be any vector space over $\mathbb{K}$ and $\rho = 0$ to be the zero map. We call this the *trivial representation* of $\mathfrak{g}$.

(b) The adjoint homomorphism of $\mathfrak{g}$ is a representation of $\mathfrak{g}$ with $V = \mathfrak{g}$ and $\rho = \mathrm{ad}$. We call this the *adjoint representation* of $\mathfrak{g}$. Note that $x \in \ker \mathrm{ad} \Leftrightarrow \mathrm{ad}_x(y) = 0$ for all $y \in \mathfrak{g} \Leftrightarrow [x,y] = 0$ for all $y \in \mathfrak{g} \Leftrightarrow x \in Z(\mathfrak{g})$. Hence the adjoint representation is *faithful* if and only if the *centre* of $\mathfrak{g}$ is trivial.

Alternatively instead of thinking of representations we can also consider modules for a Lie algebra $\mathfrak{g}$.

Definition. Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{K}$. A $\mathfrak{g}$-*module* is a pair $(V, \cdot)$ where V is a vector space and $\cdot : \mathfrak{g} \times V \to V$ is a map satisfying the following conditions for all $x, y \in \mathfrak{g}$, $v, w \in V$ and $\lambda, \mu \in \mathbb{K}$.

- $(\lambda x + \mu y) \cdot v = \lambda(x \cdot v) + \mu(y \cdot v)$,
- $x \cdot (\lambda v + \mu w) = \lambda(x \cdot v) + \mu(x \cdot w)$,
- $[x,y] \cdot v = x \cdot (y \cdot v) - y \cdot (x \cdot v)$.

Example. Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{K}$. Then the pair $(\mathfrak{g}, [\cdot,\cdot])$ is itself a $\mathfrak{g}-module$. Also given any submodule $\mathfrak{h} \subseteq \mathfrak{g}$ we have $[\mathfrak{h}, \mathfrak{g}] \subseteq \mathfrak{g}$ and so the pair $(\mathfrak{g}, [\cdot,\cdot])$ is also a $\mathfrak{h}$-module.

Remark. As per usual with representation theory the language of modules and the language of representations are just two ways of talking about the same thing. Given a representation (V, φ) of a Lie algebra $\mathfrak{g}$ we have $(V, \cdot)$ is a $\mathfrak{g}$-module where we define $x \cdot v := \varphi(x)(v)$ for all $x \in \mathfrak{g}$ and $v \in V$. Likewise given any $\mathfrak{g}$-module $(V, \cdot)$ we have (V, ψ) is a representation where we define $\psi(x)$ to be the map $v \mapsto x \cdot v$ for all $x \in \mathfrak{g}$ and $v \in V$.

2.2. **Solvable and Nilpotent Lie Algebras.** When we are dealing with groups we introduce notions of simplicity, solvability, etc. to allow us to find some fundamental 'building blocks'. In the theory of Lie algebras we do a very similar thing and use the same terms as in group theory.

Definition. Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{K}$. We define a sequence of subspaces $\mathfrak{g}^{(m)} \subseteq \mathfrak{g}$ called the *derived series* inductively by

$$\mathfrak{g}^{(0)} = \mathfrak{g} \qquad \mathfrak{g}^{(m+1)} = [\mathfrak{g}^{(m)}, \mathfrak{g}^{(m)}] \text{ for } m \geqslant 0.$$

We then say that the Lie algebra $\mathfrak{g}$ is *solvable* if for some $m \geqslant 0$ we have $\mathfrak{g}^{(m)} = \{0\}$.

Definition. Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{K}$. We define a sequence of subspaces $\mathfrak{g}^{m} \subseteq \mathfrak{g}$ called the *lower central series* inductively by

$$\mathfrak{g}^{0} = \mathfrak{g} \qquad \mathfrak{g}^{m+1} = [\mathfrak{g}, \mathfrak{g}^{m}] \text{ for } m \geqslant 0.$$

We then say that the Lie algebra $\mathfrak{g}$ is *nilpotent* if for some $m \geqslant 0$ we have $\mathfrak{g}^{m} = \{0\}$.

Remark. Note the analogy with groups where we take the bracket to be the group commutator bracket, i.e. $[x, y] = x^{-1}y^{-1}xy$ for all $x, y \in G$ for some group G.

Let $\mathfrak{g}$ be a Lie algebra with ideals $\mathfrak{h}, \mathfrak{f} \subseteq \mathfrak{g}$. Then their product $[\mathfrak{h}, \mathfrak{f}] \subseteq \mathfrak{g}$ is also an ideal of $\mathfrak{g}$. This is because for all $[x, y] \in [\mathfrak{h}, \mathfrak{f}]$ and $z \in \mathfrak{g}$ we have

$$[[x, y], z] = [x, [y, z]] + [y, [z, x]] \in [\mathfrak{h}, \mathfrak{f}].$$

Hence every subspace $\mathfrak{g}^{(m)}$, $\mathfrak{g}^{\mathfrak{m}}$ with $m \geqslant 0$ in the derived series and lower central series are ideals. Therefore in both cases we have a descending series of ideals

$$\begin{aligned} \mathfrak{g} = \mathfrak{g}^{(0)} \supseteq \mathfrak{g}^{(1)} \supseteq \mathfrak{g}^{(2)} \supseteq \mathfrak{g}^{(3)} \supseteq \cdots, \\ \mathfrak{g} = \mathfrak{g}^{0} \supseteq \mathfrak{g}^{1} \supseteq \mathfrak{g}^{2} \supseteq \mathfrak{g}^{3} \supseteq \cdots. \end{aligned}$$

Example. Consider the subalgebra $\mathfrak{b}(n, \mathbb{K}) := \{(x_{ij}) \in \mathfrak{gl}(n, \mathbb{K}) \mid x_{ij} = 0 \text{ for } i > j\} \subseteq \mathfrak{gl}(n, \mathbb{K})$ of all upper triangular matrices. This Lie algebra is solvable but not nilpotent. Consider the subalgebra $\mathfrak{n}(n, \mathbb{K}) := \{(x_{ij}) \in \mathfrak{gl}(n, \mathbb{K}) \mid x_{ij} = 0 \text{ for } i \geqslant j\} \subseteq \mathfrak{gl}(n, \mathbb{K})$ of all strictly upper triangular matrices. This Lie algebra is nilpotent and solvable.

Example. Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{K}$. Consider the centre of the Lie algebra $Z(\mathfrak{g})$. It's clear to see that the centre of a Lie algebra is abelian and so we have $Z(\mathfrak{g})^{(1)} = Z(\mathfrak{g})^{1} = [Z(\mathfrak{g}), Z(\mathfrak{g})] = \{0\}$ and so the centre is both solvable and nilpotent.

Proposition 2.2. *Every nilpotent Lie algebra is solvable.*

Proof. Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{K}$. It's clear that $\mathfrak{g} = \mathfrak{g}^{(0)} \subseteq \mathfrak{g}^0 = \mathfrak{g}$. Assume for induction that $\mathfrak{g}^{(n)} \subseteq \mathfrak{g}^n$ for some $n \geqslant 0$, then it's clear that

$$\mathfrak{g}^{(n+1)} = [\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}] \subseteq [\mathfrak{g}, \mathfrak{g}^n] = \mathfrak{g}^{n+1}.$$

Therefore if $\mathfrak{g}^k = \{0\}$ for some $k \geqslant 0$ then we have $\mathfrak{g}^{(k)} = \{0\}$, which gives the desired result. □

Theorem 2.1. *Let $\mathfrak{g}$ be a Lie algebra.*

(a) *If $\mathfrak{g}$ is solvable and $\mathfrak{h} \subseteq \mathfrak{g}$ is a subalgebra of $\mathfrak{g}$, then $\mathfrak{h}$ is solvable.*
(b) *Let $\mathfrak{h}$ be a Lie algebra. If $\mathfrak{g}$ is solvable and $\varphi : \mathfrak{g} \to \mathfrak{h}$ is a surjective Lie algebra homomorphism, then $\mathfrak{h}$ is solvable.*
(c) *Let $\mathfrak{h} \subseteq \mathfrak{g}$ be an ideal. Suppose that $\mathfrak{g}/\mathfrak{h}$ and $\mathfrak{h}$ are solvable, then $\mathfrak{g}$ is solvable.*
(d) *Let $\mathfrak{h}, \mathfrak{f} \subseteq \mathfrak{g}$ be solvable ideals of $\mathfrak{g}$, then $\mathfrak{h} + \mathfrak{f} := \{h + f \mid h \in \mathfrak{h}, f \in \mathfrak{f}\}$ is a solvable ideal of $\mathfrak{g}$.*

Proof.

(a) As $\mathfrak{h} \subseteq \mathfrak{g}$ it's clear that $\mathfrak{h}^{(i)} \subseteq \mathfrak{g}^{(i)}$ for all $i \geqslant 0$. For some $m \geqslant 0$ we have $\mathfrak{g}^{(m)} = \{0\} \Rightarrow \mathfrak{h}^{(m)} = \{0\}$. Therefore $\mathfrak{h}$ is solvable.
(b) We want to prove by induction that $\varphi(\mathfrak{g}^{(i)}) = \mathfrak{h}^{(i)}$ for all $i \geqslant 0$. If $i = 0$ then by the surjectivity of φ we have $\varphi(\mathfrak{g}) = \varphi(\mathfrak{g}^{(0)}) = \mathfrak{h}^{(0)} = \mathfrak{h}$. Assume this holds true for some $i \geqslant 0$. Now for any $x \in \mathfrak{h}^{(i+1)}$ we have $x = [h_1, h_2]$ for some $h_1, h_2 \in \mathfrak{h}^{(i)}$. By assumption we have there are $g_1, g_2 \in \mathfrak{g}^{(i)}$ such that $\varphi(g_1) = h_1$ and $\varphi(g_2) = h_2$. Therefore $x = [h_1, h_2] = [\varphi(g_1), \varphi(g_2)] = \varphi([g_1, g_2]) \in \varphi(\mathfrak{g}^{(i+1)})$. Therefore we have $\varphi(\mathfrak{g}^{(i+1)}) = \mathfrak{h}^{(i+1)}$ and we're done.

    Now if $\mathfrak{g}$ is solvable then there exists an $m \geqslant 0$ such that $\mathfrak{g}^{(m)} = \{0\}$. Hence by above we have that $\varphi(\mathfrak{g}^{(m)}) = \mathfrak{h}^{(m)} = \{0\}$ and so $\mathfrak{h}$ is solvable.
(c) We have $\mathfrak{g}/\mathfrak{h}$ is solvable and so there exists an $m \geqslant 0$ such that $(\mathfrak{g}/\mathfrak{h})^{(m)} = \{0\}$. Consider the canonical homomorphism $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$ defined by $x \mapsto x + \mathfrak{h}$. Now π is clearly surjective and so by part (b) we have $\mathfrak{g}^{(i)} + \mathfrak{h} = (\mathfrak{g}/\mathfrak{h})^{(i)}$ for all $i \geqslant 0$.

    Now $\mathfrak{g}^{(m)} + \mathfrak{h} = \mathfrak{h} \Leftrightarrow \mathfrak{g}^{(m)} \subseteq \mathfrak{h}$ and $\mathfrak{h}$ is solvable so there exists $n \geqslant 0$ such that $\mathfrak{h}^{(n)} = \{0\}$. Therefore $\mathfrak{g}^{(m+n)} = (\mathfrak{g}^{(m)})^{(n)} \subseteq \mathfrak{h}^{(n)} = \{0\}$ and so $\mathfrak{g}$ is solvable.
(d) It is an easy check to verify that $\mathfrak{h} + \mathfrak{f}$ is an ideal. By the second isomorphism theorem, (see [EW06, Theorem 2.2]), we have $(\mathfrak{h} + \mathfrak{f})/\mathfrak{h} \cong \mathfrak{f}/\mathfrak{h} \cap \mathfrak{f}$. As $\mathfrak{f}$ is solvable we have the image $(\mathfrak{h} + \mathfrak{f})/\mathfrak{h}$ is solvable by part (b). As $\mathfrak{h}$ is also solvable we have $\mathfrak{h} + \mathfrak{f}$ is solvable by part (c). □

Remark. Note that parts (a) and (b) of theorem 2.1 hold for nipotent Lie algebras but parts (c) and (d) do not hold. The following lemma is the closest approximation we can get.

Lemma 2.1. *Let $\mathfrak{g}$ be a Lie algebra such that $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent, then $\mathfrak{g}$ is nilpotent.*

Proof. A similar inductive proof as in part (b) of theorem 2.1 shows that $(\mathfrak{g}/Z(\mathfrak{g}))^i = \mathfrak{g}^i + Z(\mathfrak{g})$ for each $i \geqslant 0$. Therefore $(\mathfrak{g}/Z(\mathfrak{g}))^m = \{0\}$ for some $m \geqslant 0$ means $\mathfrak{g}^m \subseteq Z(\mathfrak{g})$ and so $\mathfrak{g}^{m+1} = [\mathfrak{g}, \mathfrak{g}^m] \subseteq [\mathfrak{g}, Z(\mathfrak{g})] = \{0\}$. □

Definition. Let $\mathfrak{g}$ be a Lie algebra. We say $\mathfrak{g}$ is *simple* if $\mathfrak{g}$ has no ideals other than $\{0\}$ and $\mathfrak{g}$.

Remark. Note that if $\mathfrak{g}$ is an abelian Lie algebra such that $\{0\}$ and $\mathfrak{g}$ are the only ideals then we must have $\mathfrak{g}$ is 1-dimensional. We refer to this as the *trivial simple Lie algebra.*

Example. The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is simple. Assume $\mathfrak{h} \subseteq \mathfrak{sl}(2, \mathbb{C})$ is a non-zero ideal. Recall from example (e) that $\{e, f, h\}$ is a basis for $\mathfrak{sl}(2, \mathbb{C})$, hence any non-zero element $x \in \mathfrak{h}$ has an expression of the form $x = ae + bh + cf$ for some $a, b, c \in \mathbb{C}$ not all zero. Now we can see that we have

$$\begin{aligned}
[h, [e, [f, ae + bh + cf]]] &= [h, [e, -ah + 2bf]] = [h, 2ae + 2bh] = 4ae,\\
[h, [e, ae + bh + cf]] &= [h, -2be + ch] = -4be,\\
[e, [e, ae + bh + cf]] &= [e, -2be + ch] = -2ce,\\
[f, [f, ae + bh + cf]] &= [f, -ah + 2bf] = -2af,\\
[h, [f, ae + bh + cf]] &= [h, -ah + 2bf] = 4bf,\\
[h, [f, [e, ae + bh + cf]]] &= [h, [e, -2be + ch]] = [h, 2bh + 2cf] = -4cf.
\end{aligned}$$

So as at least one of a, b, c is non-zero and after scaling we have that $\{e, f\} \subseteq \mathfrak{h}$. However clearly this means we have $[e, f] = h \in \mathfrak{h}$ and so as $\mathfrak{h}$ contains a basis for $\mathfrak{sl}(2, \mathbb{C})$ we must have $\mathfrak{h} = \mathfrak{sl}(2, \mathbb{C})$.

Definition. Let $\mathfrak{g}$ be a Lie algebra. We say $\mathfrak{g}$ is *semisimple* if it has no non-zero solvable ideals. Equivalently we could say $\operatorname{rad} \mathfrak{g} = \{0\}$, where $\operatorname{rad} \mathfrak{g}$ is the maximal solvable ideal of $\mathfrak{g}$ called the *radical*.

Proposition 2.3. *Let $\mathfrak{g}$ be a Lie algebra, then $\mathfrak{g}/\operatorname{rad} \mathfrak{g}$ is semisimple.*

Proof. Let $\overline{\mathfrak{h}} \subseteq \mathfrak{g}/\operatorname{rad} \mathfrak{g}$ be a solvable ideal. By the correspondence of ideals, (see [EW06, Theorem 2.2]), there exists an ideal $\operatorname{rad} \mathfrak{g} \subseteq \mathfrak{h} \subseteq \mathfrak{g}$ such that $\overline{\mathfrak{h}} = \mathfrak{h}/\operatorname{rad} \mathfrak{g}$. Now the radical is solvable by definition and we assumed $\mathfrak{h}/\operatorname{rad} \mathfrak{g}$ solvable. Therefore by part (c) of theorem 2.1 we have $\mathfrak{h}$ is solvable, however the radical is maximal so we have $\overline{\mathfrak{h}} = \operatorname{rad} \mathfrak{g}$. □

This proposition gives us an idea of how to tackle the structure of Lie algebras. Given any Lie algebra $\mathfrak{g}$ we have $\operatorname{rad} \mathfrak{g}$ is solvable and $\mathfrak{g}/\operatorname{rad} \mathfrak{g}$ is semisimple. Hence we can reduce the problem down to studying an arbitrary solvable Lie algebra and an arbitrary semisimple Lie algebra. In fact when our ground field is $\mathbb{C}$ we have nice answers to both of these questions.

2.3. **Testing for Solvability and Semisimplicity.** From the previous section we know that we would like to know when $\mathfrak{g}$ is semisimple or solvable. Working just from the definitions this seems quite difficult but when our ground field is $\mathbb{C}$ we have effective solutions. From now on all our Lie algebras are complex, so they are vector spaces over $\mathbb{C}$. It turns out that given a complex Lie algebra $\mathfrak{g}$ we can detect solvability simply by looking at traces of derivations ad_x for $x \in \mathfrak{g}$.

Definition. Let $\mathfrak{g}$ be a complex Lie algbera. We define a symmetric bilinear form $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$

$$\kappa(x, y) := \operatorname{tr}(\operatorname{ad}_x \circ \operatorname{ad}_y),$$

for all $x, y \in \mathfrak{g}$, and call it the *Killing form* of $\mathfrak{g}$.

Remark. Recall that the trace satisfies the property $\operatorname{tr} y(xz) = \operatorname{tr}(xz)y$ for all linear transformations x, y, z of a vector space. Therefore we have $\operatorname{tr}([x, y]z) = \operatorname{tr}(x[y, z])$ and so $\kappa([x, y], z) = \kappa(x, [y, z])$ for all $x, y, z \in \mathfrak{g}$. We refer to this property of the Killing form as the *associativity* of the Killing form.

Now this definition might seem a little strange but it comes from the fact that the theorems for solvability are most easily proved when our Lie algebra is a subalgebra of $\mathfrak{gl}(V)$ for some complex vector space V. Therefore to get an abstract Lie algebra to be a subalgebra of $\mathfrak{gl}(V)$ we use the adjoint representation, which is faithful when the Lie algebra is semisimple. In fact we get that $\mathfrak{g}$ is solvable if and only if $\operatorname{ad}\mathfrak{g}$ is solvable. Before stating the criterion for solvability we state a useful fact about the Killing form.

Lemma 2.2. *Let $\mathfrak{g}$ be a complex Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ an ideal of $\mathfrak{g}$. Let $\kappa_{\mathfrak{h}} = \kappa|_{\mathfrak{h}\times\mathfrak{h}}$ denote the restriction of the killing form to $\mathfrak{h}$ then we have $\kappa_{\mathfrak{h}}(x,y) = \kappa(x,y)$ for all $x,y \in \mathfrak{h}$.*

Proof. Recall that if V is a finite dimensional vector space and $W \subseteq V$ is a vector subspace then for any map $\varphi : V \to V$ such that $\operatorname{im}(\varphi) \subseteq W$ we have $\operatorname{tr}(\varphi) = \operatorname{tr}(\varphi|_W)$. Now if $x,y \in \mathfrak{h}$ then $\operatorname{ad}_x \circ \operatorname{ad}_y : \mathfrak{g} \to \mathfrak{g}$ is such that $\operatorname{im}(\operatorname{ad}_x \circ \operatorname{ad}_y) \subseteq \mathfrak{h}$ and so by the remark we have $\kappa(x,y) = \kappa_{\mathfrak{h}}(x,y)$. □

Theorem 2.2 (Cartan's First Criterion)**.** *Let $\mathfrak{g}$ be a complex Lie algebra, then $\mathfrak{g}$ is solvable if and only if $\kappa(x,y) = 0$ for all $x \in \mathfrak{g}$ and $y \in \mathfrak{g}' := [\mathfrak{g},\mathfrak{g}]$.*

Proof. See [EW06, Theorem 9.6] or [Hum78, Theorem 5.1]. □

The Killing form can also be used to see when a complex Lie algebra is semisimple. In fact this is known as Cartan's second criterion.

Theorem 2.3 (Cartan's Second Criterion)**.** *Let $\mathfrak{g}$ be a complex Lie algebra, then $\mathfrak{g}$ is semisimple if and only if the Killing form κ is non-degenerate. In other words we have $\kappa(x,y) = 0$ for all $y \in \mathfrak{g}$ if and only if $x = 0$.*

Proof. See [EW06, Theorem 9.9] or [Hum78, Theorem 4.3]. □

From the initial statement it may not seem that Cartan's second criterion is particularly useful but in fact it gives a very useful characterisation of semisimplicity. For example, let $\mathfrak{g}$ be a complex Lie algebra and fix a basis $\{x_1,\dots,x_n\} \subseteq \mathfrak{g}$. The Killing form is a symmetric bilinear form and hence we can calculate its associated symmetric matrix $X = (\kappa(x_i,x_j))_{1\leqslant i,j\leqslant n}$ with respect to the chosen basis of $\mathfrak{g}$. It's clear that κ is non-degenerate if and only if $\det(X) \neq 0$. Therefore we can determine the semisimplicity of a complex Lie algebra $\mathfrak{g}$ by calculating the determinant of X.

Lemma 2.3. *Let $\mathfrak{g}$ be a semisimple complex Lie algebra. Then $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$ where $\mathfrak{g}_i$ for $1 \leqslant i \leqslant r$ are simple ideals of $\mathfrak{g}$. In fact every simple ideal of $\mathfrak{g}$ coincides with one of the $\mathfrak{g}_i$.*

Proof. Let $\mathfrak{h} \subseteq \mathfrak{g}$ be an ideal, then it is a quick check to show that $\mathfrak{h}^\perp := \{x \in \mathfrak{g} \mid \kappa(x,y) = 0 \text{ for all } y \in \mathfrak{g}\}$ is also an ideal. We have that the restriction of the Killing form $\kappa_{\mathfrak{h}\cap\mathfrak{h}^\perp}$ is identically zero and so by theorem 2.2 we have $\mathfrak{h}\cap\mathfrak{h}^\perp \subseteq \mathfrak{g}$ is solvable and hence $\mathfrak{h}\cap\mathfrak{h}^\perp = \{0\}$ as $\mathfrak{g}$ is semisimple. Therefore as $\dim\mathfrak{h} + \dim\mathfrak{h}^\perp = \dim\mathfrak{g}$ we must have $\mathfrak{g} = \mathfrak{h}\oplus\mathfrak{h}^\perp$.

We prove this statement by induction on $\dim\mathfrak{g}$. Assume $\mathfrak{g}$ has no nonzero proper ideal, then $\mathfrak{g}$ is simple and we're done. Otherwise let $\mathfrak{g}_1 \subset \mathfrak{g}$ be a minimal proper non-zero ideal of $\mathfrak{g}$. By the previous remark we have $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_1^\perp$. Now we have $\mathfrak{g}_1$ is semisimple as any solvable ideal of $\mathfrak{g}_1$ is a solvable ideal of $\mathfrak{g}$. Hence $\mathfrak{g}_1$ is simple as it was chosen to be minimal. Clearly $\dim\mathfrak{g}_1 \neq 0$ and so $\dim\mathfrak{g}_1^\perp < \dim\mathfrak{g}$. Therefore by induction it is a direct sum of simple ideals, say $\mathfrak{g}_1^\perp = \mathfrak{g}_2 \oplus\cdots\oplus\mathfrak{g}_r$. Hence we have $\mathfrak{g} = \mathfrak{g}_1\oplus\mathfrak{g}_2\oplus\cdots\oplus\mathfrak{g}_r$ as required.

Finally we show that the simple ideals are unique. Let $\mathfrak{h} \subset \mathfrak{g}$ be a simple ideal then $[\mathfrak{h},\mathfrak{g}] \subseteq \mathfrak{h}$ is also an ideal of $\mathfrak{h}$, which is non-zero because $Z(\mathfrak{g}) \subseteq \operatorname{rad}\mathfrak{g} = \{0\}$, and so $[\mathfrak{h},\mathfrak{g}] = \mathfrak{h}$.

However $[\mathfrak{h},\mathfrak{g}] = [\mathfrak{h},\mathfrak{g}_1] \oplus \cdots \oplus [\mathfrak{h},\mathfrak{g}_r]$ but $[\mathfrak{h},\mathfrak{g}]$ is simple so we must have just one summand is non-zero, say this is $[\mathfrak{h},\mathfrak{g}_i]$. Then $\mathfrak{h} \subseteq \mathfrak{g}_i$ and $\mathfrak{h} = \mathfrak{g}_i$ because $\mathfrak{g}_i$ is simple. □

Example. Consider the following complex Lie algebra and its ideal

$$\begin{aligned}
\mathfrak{g} &= \left\{ \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \,\middle|\, A, C \in \mathfrak{sl}(2,\mathbb{C}) \text{ and } B \in \mathfrak{gl}(2,\mathbb{C}) \right\} \subseteq \mathfrak{gl}(4,\mathbb{C}), \\
\mathfrak{h} &= \left\{ \begin{bmatrix} 0 & B \\ 0 & 0 \end{bmatrix} \,\middle|\, B \in \mathfrak{gl}(2,\mathbb{C}) \right\} \subseteq \mathfrak{g}.
\end{aligned}$$

We would like to consider the structure of this Lie algebra. We can see that $\mathfrak{h}$ is in fact a subalgebra of $\mathfrak{b}(4,\mathbb{C})$ and hence, by part (a) of theorem 2.1, we have $\mathfrak{h}$ is solvable. It is an exercise to show that in fact $\mathfrak{h} = \operatorname{rad}\mathfrak{g}$. Therefore by proposition 2.3 we have $\mathfrak{g}/\mathfrak{h}$ is semisimple. In fact it's easy to see that

$$\begin{aligned}
\mathfrak{g}/\mathfrak{h} &\cong \left\{ \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} \,\middle|\, A, C \in \mathfrak{sl}(2,\mathbb{C}) \right\}, \\
&= \left\{ \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \,\middle|\, A \in \mathfrak{sl}(2,\mathbb{C}) \right\} \oplus \left\{ \begin{bmatrix} 0 & 0 \\ 0 & C \end{bmatrix} \,\middle|\, C \in \mathfrak{sl}(2,\mathbb{C}) \right\}, \\
&\cong \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C}).
\end{aligned}$$

We have already seen that $\mathfrak{sl}(2,\mathbb{C})$ is simple and so $\mathfrak{g}/\mathfrak{h}$ decomposes as a direct sum of simple ideals.

2.4. **The Lie Algebra $\mathfrak{sl}(n,\mathbb{C})$.** We consider in detail the structure of the Lie algebra $\mathfrak{sl}(n,\mathbb{C})$ as it will be archetypical of the structure of all semisimple complex Lie algebras. We first start by noting that $\mathfrak{sl}(n,\mathbb{C})$ is a simple Lie algebra.

We have already seen the simplicity of $\mathfrak{sl}(2,\mathbb{C})$ and in fact a generalisation of the argument used there works for $\mathfrak{sl}(n,\mathbb{C})$. We have $\{e_{ij} \mid 1 \leqslant i \neq j \leqslant n\} \cup \{e_{ii} - e_{i+1,i+1} \mid 1 \leqslant i \leqslant n-1\}$ is a basis for the Lie algebra $\mathfrak{sl}(n,\mathbb{C})$. Given any non-zero element in a non-zero ideal of $\mathfrak{sl}(n,\mathbb{C})$ we can multiply on the left and right by suitable basis elements of $\mathfrak{sl}(n,\mathbb{C})$ to show that each e_{ij} is in the ideal for $i \neq j$. Then by the relations of the Lie algebra we have the full basis of $\mathfrak{sl}(n,\mathbb{C})$ is contained in the ideal and so we have the ideal is in fact $\mathfrak{sl}(n,\mathbb{C})$.

Now consider the subalgebra $\mathfrak{h} = \operatorname{span}\{e_{ii} - e_{i+1,i+1} \mid 1 \leqslant i \leqslant n-1\}$ of all diagonal matrices, it's clear that $\dim\mathfrak{h} = n-1$. Also it's easy to see that $[\mathfrak{h},\mathfrak{h}] = 0$ and so $\mathfrak{h}$ is an abelian Lie algebra. What we would like to do is decompose $\mathfrak{sl}(n,\mathbb{C})$ as a direct sum of $\mathfrak{h}$-modules. We recall that $(\mathfrak{sl}(n,\mathbb{C}), [\cdot,\cdot])$ can be viewed as a $\mathfrak{h}$-module. Now consider $h = \lambda_1 e_{11} + \lambda_2 e_{22} + \cdots + \lambda_n e_{nn} \in \mathfrak{h}$ then $\lambda_1 + \cdots + \lambda_n = 0$. For any e_{ij} with $i \neq j$ we have

$$\operatorname{ad}_h(e_{ij}) = [h, e_{ij}] = \lambda_1[e_{11}, e_{ij}] + \cdots + \lambda_n[e_{nn}, e_{ij}] = (\lambda_i - \lambda_j)e_{ij}. \tag{4}$$

Therefore for each $1 \leqslant i \neq j \leqslant n$ we have the 1-dimensional vector subspace $\mathbb{C}e_{ij} \subset \mathfrak{sl}(n,\mathbb{C})$ gives us a $\mathfrak{h}$-submodule of $\mathfrak{sl}(n,\mathbb{C})$. So we obtain the following decomposition

$$\mathfrak{sl}(n,\mathbb{C}) = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{C}e_{ij}.$$

Looking at eq. (4) we can see that $\{e_{ij} \mid i \neq j\}$ will be a collection of eigenvectors for the linear map $\operatorname{ad}_h : \mathfrak{sl}(n,\mathbb{C}) \to \mathfrak{sl}(n,\mathbb{C})$ regardless of the choice of $h \in \mathfrak{h}$. However the associated eigenvalues $\lambda_i - \lambda_j$ will depend on the choice of $h \in \mathfrak{h}$. What we would like is to be able to talk about the collection of maps ad_h, for all $h \in \mathfrak{h}$, and their common eigenvectors. To do this we introduce the following definition.

Definition. Let V be a vector space over a field $\mathbb{K}$ and consider a Lie subalgebra $\mathfrak{g} \subseteq \mathfrak{gl}(V)$. We say a linear map $\lambda : \mathfrak{g} \to \mathbb{K}$ is a *weight* if the space

$$V_\lambda := \{v \in V \mid x(v) = \lambda(x)v \text{ for all } x \in \mathfrak{g}\}$$

is non-zero. We call the space V_λ the associated *weight space* of λ.

Let $h \in \mathfrak{h}$ be as before then we define coordinate maps $\varepsilon_i : \mathfrak{h} \to \mathbb{C}$ by $\varepsilon_i(h) = \lambda_i$ for each $1 \leqslant i \leqslant n$. Then rephrasing eq. (4) we have $\mathrm{ad}_h(e_{ij}) = (\varepsilon_i - \varepsilon_j)(h)e_{ij}$. Therefore $\varepsilon_i - \varepsilon_j$ is a weight of $\mathfrak{sl}(n, \mathbb{C})$ with associated weight space

$$\mathfrak{sl}(n, \mathbb{C})_{\varepsilon_i - \varepsilon_j} = \{x \in \mathfrak{sl}(n, \mathbb{C}) \mid \mathrm{ad}_h(x) = (\varepsilon_i - \varepsilon_j)(h)x \text{ for all } h \in \mathfrak{h}\} = \mathbb{C}e_{ij}.$$

Hence the decomposition given in eq. (4) is a decomposition of $\mathfrak{sl}(n, \mathbb{C})$ into weight spaces with respect to the adjoint action of the abelian Lie subalgebra $\mathfrak{h}$. The weights $\varepsilon_i - \varepsilon_j$ are elements of the dual vector space $\mathfrak{h}^\star$.

Let $\Phi = \{\varepsilon_i - \varepsilon_j \mid 1 \leqslant i \neq j \leqslant n\} \subseteq \mathfrak{h}^\star$ be the collection of all weights for $\mathfrak{h}$. We define specific elements $\alpha_i = \varepsilon_i - \varepsilon_{i+1} \in \Phi$ for each $1 \leqslant i \leqslant n-1$ and let $\Delta = \{\alpha_i \mid 1 \leqslant i \leqslant n-1\}$. Now it's clear that for any weight $\varepsilon_i - \varepsilon_j \in \Phi$ we have

$$\varepsilon_i - \varepsilon_j = \pm(\alpha_i + \alpha_{i+1} + \cdots + \alpha_{j-1}),$$

where the sign depends on whether $i < j$ or $i > j$. Let $\mathfrak{h}^\star_{\mathbb{R}} = \mathrm{span}_{\mathbb{R}}(\Phi)$ denote the real span of Φ in $\mathfrak{h}^\star$. It is true that this is a real Euclidean vector space. We then have Φ is a root system for $\mathfrak{h}^\star_{\mathbb{R}}$ in the sense of section 1.2 with simple system Δ.

Remark. Note the similarities between the construction given here and the construction of the root system of type A_n given in section 1.5

2.5. **Semisimple Complex Lie Algebras.** We use the discussion in section 2.4 to now give the general structure theory of semisimple complex Lie algebras. Note that throughout this section we will give very few proofs for the results that we state but we do indicate locations for these proofs in [EW06] or [Hum78]. Now we would like to generalise the large abelian Lie subalgebra defined in the case of $\mathfrak{sl}(n, \mathbb{C})$. To do this we need a small discussion about the *Jordan decomposition*.

Theorem 2.4 (Jordan Decomposition)**.** *Let V be a complex vector space and consider $x \in \mathrm{End}\, V$. Then there exists complex linear maps $x_s, x_n \in \mathrm{End}\, V$ such that $x = x_s + x_n = x_n + x_s$ where x_s is semisimple, (i.e. diagonalisable) and x_n is nilpotent.*

Proof. See [BR02, Theorem 4.3]. □

A very powerful result in the theory of semisimple complex Lie algebras is known as the *abstract Jordan decomposition*. Given a semisimple complex Lie algebra $\mathfrak{g}$ the adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ provides to each element $x \in \mathfrak{g}$ an endomorphism $\mathrm{ad}_x \in \mathfrak{gl}(\mathfrak{g})$. By the Jordan decomposition we can decompose $\mathrm{ad}_x = (\mathrm{ad}_x)_s + (\mathrm{ad}_x)_n$ into a semisimple and nilpotent part. What we would like to show is that this gives us a decomposition of the element $x \in \mathfrak{g}$.

Theorem 2.5 (Abstract Jordan Decomposition)**.** *Let $\mathfrak{g}$ be a semisimple complex Lie algebra. Each $x \in \mathfrak{g}$ can be written uniquely as $x = s + n$ for some $s, n \in \mathfrak{g}$ where $\mathrm{ad}_s \in \mathfrak{gl}(\mathfrak{g})$ is semisimple, $\mathrm{ad}_n \in \mathfrak{gl}(\mathfrak{g})$ is nilpotent and $[s, n] = 0$. We call such a decomposition the* abstract Jordan decomposition *of x. Furthermore let (V, φ) be any representation of $\mathfrak{g}$. Suppose $x \in \mathfrak{g}$ has abstract Jordan decomposition $x = s + n$, then the Jordan decomposition of $\varphi(x)$ is $\varphi(x) = \varphi(s) + \varphi(n)$.*

Proof. See [EW06, Theorems 9.15 and 9.16] or [Hum78, Section 5.4 and Theorem 6.4] □

Definition. Let $\mathfrak{g}$ be a semisimple complex Lie algebra and let $x \in \mathfrak{g}$ have abstract Jordan decomposition $x = s + n$. We say $s \in \mathfrak{g}$ is the *semisimple part* of x and $n \in \mathfrak{g}$ is the *nilpotent part* of x. If $n = 0$ then we say x is *semisimple* and if $s = 0$ we say x is *nilpotent*.

We now continue with our discussion of the structure of semisimple complex Lie algebras. The way we started to understand $\mathfrak{sl}(n, \mathbb{C})$ in section 2.4 was to identify a large abelian Lie subalgebra with certain desirable properties. We now define such a Lie subalgebra for any semisimple complex Lie algebra.

Definition. Let $\mathfrak{g}$ be a semisimple complex Lie algebra. Then a Lie subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is said to be a *toral subalgebra* if $\mathfrak{h}$ consists entirely of semisimple elements. We say $\mathfrak{h} \subseteq \mathfrak{g}$ is a *maximal toral subalgebra* of $\mathfrak{g}$ if $\mathfrak{h}$ is a toral subalgebra and maximal amongst all such toral subalgebras with respect to inclusion.

Remark. Note that the classic approach to Lie algebras is to define a *Cartan subalgebra* of $\mathfrak{g}$. This is a nilpotent subalgebra which is self normalising, i.e. is equal to its own normaliser in $\mathfrak{g}$. In fact it is true that if $\mathfrak{g}$ is a semisimple Lie algebra defined over an algebraically closed field of characteristic zero then Cartan subalgebras and maximal toral subalgebras are the same thing, (see [Hum78, Corollary 15.3]). We choose to work with maximal toral subalgebras as it will fit in better with the theory later on.

Lemma 2.4. *Any maximal toral subalgebra of a semisimple complex Lie algebra is abelian.*

Proof. See [Hum78, Lemma 8.1]. □

We would like to follow the decomposition that we obtained for $\mathfrak{sl}(n, \mathbb{C})$. Now for a semisimple complex Lie algebra $\mathfrak{g}$ we have that a maximal toral subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is abelian. Therefore by definition $\{\mathrm{ad}_h \mid h \in \mathfrak{h}\}$ is a collection of commuting semisimple linear transformations of $\mathfrak{g}$. By standard results in linear algebra, (see [BR02, Theorem 3.4]), we have that we can simultaneously diagonalise this collection of linear transformations. This therefore allows us to decompose $\mathfrak{g}$ into a direct sum of weight spaces with respect to the adjoint action of $\mathfrak{h}$.

$$\mathfrak{g} = \bigoplus_{\alpha \in \mathfrak{h}^\star} \mathfrak{g}_\alpha \qquad \text{where } \mathfrak{g}_\alpha := \{x \in \mathfrak{g} \mid \mathrm{ad}_h(x) = \alpha(h)x \text{ for all } h \in \mathfrak{h}\}.$$

Now we have the zero map in $\mathfrak{h}^\star$ and so one of these weight spaces is the zero weight space $\mathfrak{g}_0 = \{x \in \mathfrak{g} \mid \mathrm{ad}_h(x) = [h, x] = 0 \text{ for all } h \in \mathfrak{h}\}$. However this is the same as the centraliser of $\mathfrak{h}$ in $\mathfrak{g}$ which we denote $C_{\mathfrak{g}}(\mathfrak{h})$. Note that $\mathfrak{h} \subseteq C_{\mathfrak{g}}(\mathfrak{h})$ and in fact we have $\mathfrak{h} = C_{\mathfrak{g}}(\mathfrak{h})$. We do not prove this fact here but instead refer to the proof in [Hum78, Proposition 8.2]. Now we let $\Phi \subseteq \mathfrak{h}^\star$ be the set of non-zero weights in $\mathfrak{h}^\star$. We can now express the decomposition as

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha. \tag{5}$$

Lemma 2.5. *Let $\mathfrak{g}$ be a semisimple complex Lie algebra with weight space decomposition as given in eq. (5). Then for $\alpha, \beta \in \Phi$ we have*

(a) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.
(b) $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ *if* $\alpha + \beta \neq 0$.
(c) *the restriction $\kappa_{\mathfrak{h}}$ is non-degenerate.*

Proof.

(a) Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$. From the Jacobi identity we have

$$\begin{aligned}\mathrm{ad}_h([x,y]) = [h,[x,y]] = [[h,x],y] + [x,[h,y]] &= [\alpha(h)x, y] + [x, \beta(h)y],\\ &= \alpha(h)[x,y] + \beta(h)[x,y],\\ &= (\alpha+\beta)(h)[x,y].\end{aligned}$$

Therefore $[x,y]$ is an eigenvector for ad_h with weight $\alpha+\beta$ and so $[x,y] \in \mathfrak{g}_{\alpha+\beta}$.

(b) We have $\alpha + \beta \neq 0$ and so there exists $h \in \mathfrak{h}$ with $(\alpha+\beta)(h) \neq 0$. Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$ then we have

$$\alpha(h)\kappa(x,y) = \kappa([h,x],y) = -\kappa([x,h],y) = -\kappa(x,[h,y]) = -\beta(h)\kappa(x,y).$$

by the associativity of the Killing form. Therefore we have $(\alpha+\beta)(h)\kappa(x,y) = 0$. We chose h such that $(\alpha+\beta)(h) \neq 0$ and so we must have $\kappa(x,y) = 0$.

(c) Recall from theorem 2.3 that κ is non-degenerate. Recall $\mathfrak{h} = \mathfrak{g}_0$ and so by part (b) we have $\kappa(\mathfrak{h}, \mathfrak{g}_\alpha) = 0$ for any $\alpha \in \Phi$. If $x \in \mathfrak{h}$ is such that $\kappa(x, \mathfrak{h}) = 0$ then by eq. (5) we have $\kappa(x, \mathfrak{g}) = 0$ but κ non-degenerate so $x = 0$ as required. □

This lemma provides us with the essential information we need regarding the weight space decomposition given in eq. (5). We focus on part (c) of lemma 2.5, which tells us that $\kappa_{\mathfrak{h}}$ is non-degenerate. This allows us to use the Killing form to define an isomorphism between $\mathfrak{h}$ and $\mathfrak{h}^\star$. To $\varphi \in \mathfrak{h}^\star$ we identify the unique element $t_\varphi \in \mathfrak{h}$ satisfying $\varphi(h) = \kappa(t_\varphi, h)$ for all $h \in \mathfrak{h}$. In particular this says we can associate the set $\Phi \subseteq \mathfrak{h}^\star$ with the subset $\{t_\alpha \mid \alpha \in \Phi\} \subseteq \mathfrak{h}$.

Lemma 2.6. *Let $\mathfrak{g}$ be a semisimple complex Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ a maximal toral subalgebra. Let $\Phi \subseteq \mathfrak{h}^\star$ be the set of non-zero weights with respect to $\mathfrak{h}$. For each $\alpha \in \Phi$ let $e_\alpha \in \mathfrak{g}_\alpha$ be a non-zero element of the weight space then there exists a non-zero element $f_\alpha \in \mathfrak{g}_{-\alpha}$ such that* $\mathrm{span}\{e_\alpha, f_\alpha, [e_\alpha, f_\alpha]\} \subseteq \mathfrak{g}$ *is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$. In fact $h_\alpha = [e_\alpha, f_\alpha] \in \mathrm{span}\{t_\alpha\} \subseteq \mathfrak{h}$.*

Proof. See [EW06, Lemma 10.5 and 10.6] or [Hum78, Proposition 8.3]. □

Example. Consider $\mathfrak{sl}(3,\mathbb{C})$. Now we specify the information given in section 2.4 to the case $n = 3$. The non-zero weights of the maximal toral subalgebra $\mathfrak{h}$ are given as

$$\alpha = \varepsilon_1 - \varepsilon_2 \qquad \beta = \varepsilon_2 - \varepsilon_3 \qquad \alpha + \beta = \varepsilon_1 - \varepsilon_3$$

and their negatives. Consider the weights $\pm(\alpha+\beta) \in \mathfrak{h}^\star$. We have the corresponding weight spaces are $\mathfrak{g}_{\alpha+\beta} = \mathbb{C}e_{13}$ and $\mathfrak{g}_{-(\alpha+\beta)} = \mathbb{C}e_{31}$. Therefore we can choose the elements $e_{\alpha+\beta} \in \mathfrak{g}_{\alpha+\beta}$ and $f_{\alpha+\beta} \in \mathfrak{g}_{-(\alpha+\beta)}$ to be e_{13} and e_{31} respectively. Therefore we have

$$e_{\alpha+\beta} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad f_{\alpha+\beta} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad h_{\alpha+\beta} = [e_{\alpha+\beta}, f_{\alpha+\beta}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Table 3. The Simple Complex Lie Algebras.

Type	A_n	B_n	C_n	D_n	E_6	E_7	E_8	F_4	G_2
Description	$\mathfrak{sl}(n+1,\mathbb{C})$	$\mathfrak{so}(2n+1,\mathbb{C})$	$\mathfrak{sp}(2n,\mathbb{C})$	$\mathfrak{so}(2n,\mathbb{C})$	-	-	-	-	-
$\lvert\Phi\rvert$	$n(n+1)$	$2n^2$	$2n^2$	$2n(n-1)$	72	126	240	48	12
$\dim\mathfrak{h}$	n	n	n	n	6	7	8	4	2
$\dim\mathfrak{g}$	$n(n+2)$	$n(2n+1)$	$n(2n+1)$	$n(2n-1)$	78	133	248	52	14

and it's easy to see that $\operatorname{span}\{e_{\alpha+\beta}, f_{\alpha+\beta}, h_{\alpha+\beta}\} \subseteq \mathfrak{sl}(3,\mathbb{C})$ is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.

So every semisimple complex Lie algebra is built up out of subalgebras that are isomorphic to $\mathfrak{sl}(2,\mathbb{C})$. This makes the study of $\mathfrak{sl}(2,\mathbb{C})$ incredibly important to the study of semisimple complex Lie algebras. Indeed we can use the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ to deduce lots of results about the structure of arbitrary semisimple complex Lie algebras. It turns out that this fact allows us to show that one of the properties of the weight spaces from section 2.4 holds in general.

Proposition 2.4. *Let $\mathfrak{g}$ be a semisimple complex Lie algebra and let $\mathfrak{h} \subseteq \mathfrak{g}$ be a maximal toral subalgebra. Then for each non-zero weight $\alpha \in \Phi \subseteq \mathfrak{h}^\star$ we have $\mathfrak{g}_\alpha$ is 1-dimensional.*

Proof. See [EW06, Proposition 10.9] or [Hum78, Proposition 8.4]. □

Corollary 2.1. *If $\mathfrak{g}$ is a semisimple complex Lie algebra with weight space decomposition as in eq. (5) then we have $\dim\mathfrak{g} = \dim\mathfrak{h} + |\Phi|$.*

Proof. Clear by eq. (5) and proposition 2.4. □

Finally what we would like to do now is endow the vector space $\mathfrak{h}^\star$ with a bilinear form. Recall that we identified $\mathfrak{h}$ with $\mathfrak{h}^\star$ using the Killing form. This allows us to define, without ambiguity, a map $(\cdot,\cdot) : \mathfrak{h}^\star \times \mathfrak{h}^\star \to \mathbb{C}$ by

$$(\alpha,\beta) = \kappa(t_\alpha, t_\beta)$$

for all $\alpha,\beta \in \mathfrak{h}^\star$. Now as κ is symmetric and non-degenerate on $\mathfrak{h}$ we have $(\cdot,\cdot)$ is symmetric and non-degenerate on $\mathfrak{h}^\star$. Consider the set of all non-zero weights $\Phi \subseteq \mathfrak{h}^\star$. We denote by $\mathfrak{h}^\star_\mathbb{R}$ the real span of Φ in $\mathfrak{h}^\star$.

Proposition 2.5. *The form $(\cdot,\cdot)$ is a real valued inner product on $\mathfrak{h}^\star_\mathbb{R}$. Therefore $\mathfrak{h}^\star_\mathbb{R}$ is a real Euclidean vector space.*

Proof. See [EW06, Proposition 10.15]. □

This is the final piece of information that we need to put the puzzle together. We can now state the theorem which we have been suggestively hinting towards through notation throughout this entire section.

Theorem 2.6. *Let $\mathfrak{g}$ be a semisimple complex Lie algebra with maximal toral subalgebra $\mathfrak{h}$. Let $\Phi \subseteq \mathfrak{h}^\star$ be the set of all non-zero weights in $\mathfrak{h}^\star$. Then Φ is a crystallographic root system for $\mathfrak{h}^\star_\mathbb{R}$ as in the sense of section 1.5. We refer to the elements of Φ as the roots of $\mathfrak{g}$ and the 1-dimensional subspaces $\mathfrak{g}_\alpha$ as the associated root spaces. We refer to the decomposition given in eq. (5) as the root decomposition of $\mathfrak{g}$.*

Proof. See [Hum78, Theorem 8.5]. □

What this theorem tells us is that every semisimple complex Lie algebra contains a crystallographic root system. Therefore we would hope that we could try and classify these Lie algebras using the Dynkin diagrams we described in section 1.5.

Theorem 2.7. *Let $\mathfrak{g}$ be a semisimple complex Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ a maximal toral subalgebra of $\mathfrak{g}$. The Dynkin diagram determined by the crystallographic root system $\Phi \subseteq \mathfrak{h}^\star$ is independent of the choice of $\mathfrak{h}$. Furthermore $\mathfrak{g}$ is simple if and only if the Dynkin diagram of Φ is connected. In fact to each Dynkin diagram in figure 6 there exists a unique simple Lie algebra, up to isomorphism, giving rise to it.*

Proof. See [EW06, Chapter 14]. □

In table 3 we give information regarding each of the simple complex Lie algebras. Note that in the table $\mathfrak{h}$ denotes a maximal toral subalgebra of the simple complex Lie algebra $\mathfrak{g}$. As in section 1.5 we refer to the Lie algebras of types $\mathrm{A}_n, \mathrm{B}_n, \mathrm{C}_n$ and D_n as being of classical type. The Lie algebras of type $\mathrm{E}_6, \mathrm{E}_7, \mathrm{E}_8, \mathrm{F}_4$ and G_2 are referred to as the Lie algebras of exceptional type as they do not have any kind of description in terms of classical Lie algebras. For information regarding the constructions of all these Lie algebras see [EW06, Chapters 12 and 13.2]

Remark. We make one closing remark for this section. Everything we have done during this section has depended upon a choice of maximal toral subalgebra. However it is true that given two maximal toral subalgebras $\mathfrak{h}_1, \mathfrak{h}_2 \subseteq \mathfrak{g}$ of a semisimple complex Lie algebra $\mathfrak{g}$ there exists an inner automorphism φ of $\mathfrak{g}$ such that $\varphi(\mathfrak{h}_1) = \mathfrak{h}_2$. Hence our choice of maximal toral subalgebra is unique up to an inner automorphism of $\mathfrak{g}$. This is in general quite difficult to prove but a proof is given in [Hum78, Corollary 16.4].

3. Linear Algebraic Groups

Throughout this section $\mathbb{K}$ will denote an algebraically closed field, primarily either $\mathbb{C}$ or $\overline{\mathbb{F}_p}$, where $p > 0$ is a prime. We also note that all algebras and modules will be commutative. During the initial few sections almost no proofs will be given to the statements as, unfortunately, the basics of the Zariski topology and algebraic geometry are not our main goal. Instead we leave references to proofs in [Gec03]. The interested reader may also want to examine [Hum75, Chapter I] for a more direct treatment of the algebraic geometry.

3.1. **Some Definitions and The Zariski Topology.** We start by recalling a fundamental theorem from commutative ring theory about polynomial rings defined over Noetherian rings. Recall that a ring is Noetherian if and only if every ideal of the ring is finitely generated.

Theorem 3.1 (Hilbert's Basis Theorem)**.** *Let R be a commutative ring with 1. If R is Noetherian and X is an indeterminate over R then the polynomial ring $R[X]$ is also Noetherian. In particular if $\mathbb{K}$ is a field then $\mathbb{K}[X_1, \dots, X_n]$ is Noetherian.*

Proof. See [Gec03, Theorem 1.1.1]. □

Definition. Let $S \subseteq \mathbb{K}[X_1, \dots, X_n]$ be any subset of the polynomial ring in n variables. Then the set

$$\mathbf{V}(S) = \{x \in \mathbb{K}^n \mid f(x) = 0 \text{ for all } f \in S\} \subseteq \mathbb{K}^n$$

is called the *algebraic set* defined by S. Conversely for any subset $V \subseteq \mathbb{K}^n$ we call the set

$$\mathbf{I}(V) = \{f \in \mathbb{K}[X_1, \dots, X_n] \mid f(x) = 0 \text{ for all } x \in V\}$$

the *vanishing ideal* of V. The quotient $\mathbb{K}[V] = \mathbb{K}[X_1, \dots, X_n]/\mathbf{I}(V)$ is called the *affine algebra* of V. From this point on we denote the image of $f \in \mathbb{K}[X_1, \dots, X_n]$ in $\mathbb{K}[V]$, under the canonical projection, by $\overline{f}$.

Example.

(a) Consider the constant polynomials $1, 0 \in \mathbb{K}[X_1, \dots, X_n]$, i.e. these are the maps defined by $1(x) = 1$ and $0(x) = 0$ for all $x \in \mathbb{K}^n$. Then we have

$$\begin{aligned}\mathbf{V}(\{1\}) &= \{x \in \mathbb{K}^n \mid 1(x) = 0\} = \varnothing,\\ \mathbf{V}(\{0\}) &= \{x \in \mathbb{K}^n \mid 0(x) = 0\} = \mathbb{K}^n.\end{aligned}$$

Therefore $\varnothing$ and $\mathbb{K}^n$ are always algebraic sets. Note the choice of the constant polynomial 1 is arbitrary. Any non-zero constant polynomial will give the empty set as an algebraic set.

(b) Let $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ be a single point. Considering the polynomials $X_i - x_i$ for each $1 \leqslant i \leqslant n$ we have

$$\mathbf{V}(\{X_1 - x_1, \dots, X_n - x_n\}) = \{x\}$$

and so every singleton set $\{x\}$ is an algebraic set.

Proposition 3.1. *Let $\{S_i\}_{i \in I}$, with $S_i \subseteq \mathbb{K}[X_1, \dots, X_n]$, be a possibly infinite family of subsets of the polynomial ring. Then we have*

$$\mathbf{V}\left(\bigcup_{i\in I} S_i\right) = \bigcap_{i\in I} \mathbf{V}(S_i).$$

Also for any $S, T \in \mathbb{K}[X_1, \ldots, X_n]$ we have $\mathbf{V}(ST) = \mathbf{V}(S) \cup \mathbf{V}(T)$.

Proof. For the first statement we have $x \in \mathbf{V}(\cup_{i\in I} S_i) \Leftrightarrow f(x) = 0$ for all $f \in \cup_{i\in I} S_i \Leftrightarrow x \in \mathbf{V}(S_i)$ for all $i \in I \Leftrightarrow x \in \cap_{i\in I} \mathbf{V}(S_i)$. For the second statement we recall that $ST = \{st \mid s \in S, t \in T\}$ then

$$\begin{aligned}
\mathbf{V}(ST) &= \{x \in \mathbb{K}^n \mid f(x) = 0 \text{ for all } f \in ST\}, \\
&= \{x \in \mathbb{K}^n \mid g(x)h(x) = 0 \text{ for all } g \in S, h \in T\}, \\
&= \{x \in \mathbb{K}^n \mid g(x) = 0 \text{ or } h(x) = 0 \text{ for all } g \in S, h \in T\}, \\
&= \mathbf{V}(S) \cup \mathbf{V}(T).
\end{aligned}$$

□

Corollary 3.1. *The collection of algebraic sets in $\mathbb{K}^n$ form the closed sets of a topology called the* Zariski topology. *A subset $X \subseteq \mathbb{K}^n$ is* open *if $\mathbb{K}^n \setminus X$ is closed.*

Remark. It is easy to convince oneself that the following properties of $\mathbf{I}$ and $\mathbf{V}$ hold.

(a) For any subsets $S_1 \subseteq S_2 \subseteq \mathbb{K}[X_1, \ldots, X_n]$ we have $\mathbf{V}(S_2) \subseteq \mathbf{V}(S_1)$.
(b) For any $S \subseteq \mathbb{K}[X_1, \ldots, X_n]$ we have $S \subseteq \mathbf{I}(\mathbf{V}(S))$.
(c) For any subsets $V_1 \subseteq V_2 \subseteq \mathbb{K}^n$ we have $\mathbf{I}(V_2) \subseteq \mathbf{I}(V_1)$.
(d) For any subset $V \subseteq \mathbb{K}^n$ we have $V \subseteq \mathbf{V}(\mathbf{I}(V))$.

Now $\mathbf{V}$ and $\mathbf{I}$ can be seen as operators between the subsets of $\mathbb{K}^n$ and subsets of the polynomial ring $\mathbb{K}[X_1, \ldots, X_n]$. It would be tempting to say that these operators form bijections between subsets of $\mathbb{K}^n$ and subsets of $\mathbb{K}[X_1, \ldots, X_n]$ but this is not the case. However if we restrict our view point to a smaller class of subsets then we will in fact get a bijection.

Lemma 3.1. *Let $V \subseteq \mathbb{K}^n$ be any subset then we have $\mathbf{V}(\mathbf{I}(V)) = \overline{V}$ where $\overline{V}$ is the topological closure of V in the Zariski topology, in other words the intersection of all closed sets that contain V.*

Proof. We have that $V \subseteq \mathbf{V}(\mathbf{I}(V))$ for any subset $V \subseteq \mathbb{K}^n$. Therefore as the topological closure is minimal we have $V \subseteq \overline{V} \subseteq \mathbf{V}(\mathbf{I}(V))$. Conversely suppose $W \subseteq \mathbb{K}^n$ is any closed set containing V, then we write $W = \mathbf{V}(S)$ for some subset $S \subseteq \mathbb{K}[X_1, \ldots, X_n]$. Recall that $V \subseteq W \Rightarrow \mathbf{I}(W) \subseteq \mathbf{I}(V)$ and also $S \subseteq \mathbf{I}(\mathbf{V}(S)) = \mathbf{I}(W) \subseteq \mathbf{I}(V)$. Therefore $\mathbf{V}(\mathbf{I}(V)) \subseteq \mathbf{V}(S) = W$ as required. □

So for any algebraic set V we have $\mathbf{V}(\mathbf{I}(V)) = V$ and so we can see that the operator $\mathbf{I}$ takes injectively algebraic sets of $\mathbb{K}^n$ to ideals of $\mathbb{K}[X_1, \ldots, X_n]$. It is natural to now ask the question, given an ideal $S \subseteq \mathbb{K}[X_1, \ldots, X_n]$ what is $\mathbf{I}(\mathbf{V}(S))$? Or more to the point, when is $\mathbf{I}$ surjective? The answer to this question is not intuitively obvious and is a result attributed to Hilbert, (see [Gec03, Theorem 2.1.9]). He showed that given any ideal $S \subseteq \mathbb{K}[X_1, \ldots, X_n]$ we have $\mathbf{I}(\mathbf{V}(S)) = \sqrt{S}$ where $\sqrt{S} = \{f \in \mathbb{K}[X_1, \ldots, X_n] \mid f^m \in S \text{ for some } m \geqslant 1\}$ is the *radical* of S. Therefore we have the operators $\mathbf{I}$ and $\mathbf{V}$ give inverse bijections between the following sets

$$\{V \subseteq \mathbb{K}^n \mid V \text{ is an algebraic set}\} \longleftrightarrow \{S \subseteq \mathbb{K}[X_1, \ldots, X_n] \mid S = \sqrt{S} \text{ is an ideal}\}.$$

Definition. Let Z be a non-empty topological space. We say Z is *reducible* if we can write $Z = Z_1 \cup Z_2$ where $Z_1, Z_2 \subset Z$ are proper non-empty closed subsets of Z. If this is not the case then we say Z is irreducible. We say a non-empty subset $Y \subseteq Z$ is *irreducible* if Y is irreducible with the induced topology.

A topological space Z is called *Noetherian* if every chain of closed sets $Z_1 \supseteq Z_2 \supseteq \cdots$ terminates.

Remark. Note that if $V \subseteq \mathbb{K}^n$ is an algebraic set then the closed sets of the induced topology on V are just the closed sets of $\mathbb{K}^n$, in the Zariski topology, that are contained in V.

Proposition 3.2. *Let $V \subseteq \mathbb{K}^n$ be a non-empty algebraic set.*

(a) *Then V is a Noetherian topological space with respect to the Zariski topology. In particular we have a decomposition $V = V_1 \cup \cdots \cup V_r$, where the V_i are the maximal closed irreducible subsets of V.*
(b) *We have V is irreducible in the Zariski topology if and only if $\mathbf{I}(V)$ is a prime ideal.*

Proof. See [Gec03, Proposition 1.1.11 and 1.1.12]. □

Example.

(a) Consider any singleton set $\{(x)\} \subseteq \mathbb{K}^n$, then we have $\mathbf{I}(\{x\}) = \{f \in \mathbb{K}[X_1, \dots, X_n] \mid f(x) = 0\}$. Clearly if $fg \in \mathbf{I}(\{x\})$ we have $fg(x) = f(x)g(x) = 0 \Rightarrow f(x) = 0$ or $g(x) = 0$ and so $\mathbf{I}(V)$ is a prime ideal. Therefore $\{x\}$ is irreducible.
(b) Consider the algebraic set $\mathbb{K}^n = \mathbf{V}(\{0\})$. Then we have $\mathbf{I}(\mathbb{K}^n) = \mathbf{I}(\mathbf{V}(\{0\})) = \sqrt{\{0\}}$ by Hillbert's Nullstellensatz. Now clearly if $f \in \mathbb{K}[X_1, \dots, X_n]$ is such that $f^m = 0$ for some $m \geqslant 1$ then we have $f(x)^m = 0$ for all $x \in \mathbb{K}^n$ but by induction we must have $f(x) = 0$ and hence $f = 0$. Therefore we have $\sqrt{\{0\}} = \{0\}$ and clearly $\{0\}$ is a prime ideal so $\mathbb{K}^n$ is irreducible.
(c) Consider the polynomial $p(X_1, X_2) = X_1X_2$ and the algebraic set $V = \{(x_1, x_2) \in \mathbb{K}^2 \mid p(x_1, x_2) = 0\} \subseteq \mathbb{K}^2$. Recall that as $\mathbb{K}$ is a field we have $x_1x_2 = 0 \Rightarrow x_1 = 0$ or $x_2 = 0$. Now let $f_1(X_1, X_2) = X_1$ and $f_2(X_1, X_2) = X_2$ then we have

$$V = \{(x_1, x_2) \in \mathbb{K}^2 \mid f_1(x_1, x_2) = 0\} \cup \{(x_1, x_2) \in \mathbb{K}^2 \mid f_2(x_1, x_2) = 0\}$$

and so V is reducible.

Remark. We comment that any irreducible topological space is also connected. However a connected topological space need not be irreducible, an example of which is provided by example (c) above.

We are now in the situation where we have a bijection between algebraic sets in $\mathbb{K}^n$ and radical ideals of the associated polynomial ring. This is the view point of algebraic geometry. On one side we have a very geometric setup and on the other side we have a purely algebraic setup. So the idea is to try and work with the ideals in $\mathbb{K}[X_1, \dots, X_n]$ to glean geometric information about the algebraic sets of $\mathbb{K}^n$.

However it turns out to be quite inconvenient to try and work with the ideals of polynomials in $\mathbb{K}[X_1, \dots, X_n]$ directly. Instead, given an algebraic set $V \subseteq \mathbb{K}^n$ we often choose to work with the affine algebra $\mathbb{K}[V]$. Before we can progress further with this though we need some more terminology and machinery.

3.2. **Regular Maps and Linear Algebraic Groups.**

Definition. Let $V \subseteq \mathbb{K}^n$ and $W \subseteq \mathbb{K}^m$ be two algebraic sets. We call $\varphi : V \to W$ a *regular map*, (or *morphism*), if there exists $f_1, \dots, f_m \in \mathbb{K}[X_1, \dots, X_n]$ such that

$$\varphi(x) = (f_1(x), \dots, f_m(x))$$

for all $x \in V$. We say φ is an *isomorphism* of algebraic sets if φ is a bijective regular map whose inverse φ^{-1} is also a regular map.

Remark. Let $\varphi : V \to W$ be a regular map and $\mathbf{V}(S) \subseteq W$ be a closed set for some $S \subseteq \mathbb{K}[Y_1, \dots, Y_m]$. Then we have $\varphi^{-1}(\mathbf{V}(S)) = \mathbf{V}(\{\psi \circ \varphi \mid \psi \in S\})$ and so the preimage of any closed set is closed, which means φ is continuous in the Zariski topology. Now let $\varphi : V \to W$ be a regular map between two algebraic sets $V \subseteq \mathbb{K}^n$ and $W \subseteq \mathbb{K}^m$. Then in fact we get an induced $\mathbb{K}$-algebra homomorphism of the associated affine algebras $\varphi^* : \mathbb{K}[W] \to \mathbb{K}[V]$, given by $\varphi^*(\bar{f}) = \bar{f} \circ \varphi$ for all $f \in \mathbb{K}[W]$.

Let us consider the special case of a regular map $\varphi : V \to \mathbb{K}$ for some algebraic set $V \subseteq \mathbb{K}^n$. Then for some $f \in \mathbb{K}[X_1, \dots, X_n]$ we have $\varphi(x) = f(x)$ for all $x \in V$ but how unique is this f? Assume there is $g \in \mathbb{K}[X_1, \dots, X_n]$ such that $\varphi(x) = g(x)$ for all $x \in V$, then we have $f(x) = g(x) \Rightarrow (f - g)(x) = 0$ for all $x \in V$. In other words $f - g \in \mathbf{I}(V)$ and so the image $\bar{f} \in \mathbb{K}[V]$ is uniquely determined by φ. On the other hand given any $f + \mathbf{I}(V) \in \mathbb{K}[V]$ we get a well defined regular map $\varphi : V \to \mathbb{K}$ by setting $\varphi(x) = f(x)$ for all $x \in V$. Due to this correspondence we will often refer to $\mathbb{K}[V]$ as the *ring of regular functions*.

Proposition 3.3. *Let $\varphi : V \to W$ be a regular map, then the assignment $\varphi \mapsto \varphi^*$ gives a bijection*

$$\{\textit{regular maps } \psi : V \to W\} \xleftrightarrow{1-1} \{\mathbb{K}\textit{-algebra homomomorphisms } \psi^* : \mathbb{K}[W] \to \mathbb{K}[V]\}.$$

Indeed we have φ is an isomorphism of algebraic sets if and only if φ^* is an isomorphism of $\mathbb{K}$-algebras.

Proof. See [Gec03, Proposition 1.3.4]. □

Remark. Note that the assignment $\varphi \mapsto \varphi^*$ is *contravariant functorial* in the following sense. Let $\varphi_1 : V \to W$ and $\varphi_2 : W \to X$ be regular maps then $(\varphi_1 \circ \varphi_2)^* = \varphi_2^* \circ \varphi_1^*$ and when $V = W$ we have $\mathrm{id}_V^* = \mathrm{id}_{\mathbb{K}[V]}$.

Example. Let $\mathbb{K} = \overline{\mathbb{F}_p}$ for some prime $p > 0$. Consider the regular map $\varphi : \mathbb{K} \to \mathbb{K}$ given by $\varphi(x) = x^p$. Now φ is clearly a regular map but the corresponding $\mathbb{K}$-algebra homomorphism $\varphi^* : \mathbb{K}[X] \to \mathbb{K}[X]$ given by $X \mapsto X^p$ is not surjective as $X^{1/p} \notin \mathbb{K}[X]$. Therefore we have φ is not an isomorphism of algebraic sets, even though it is a field isomorphism.

Let $V \subseteq \mathbb{K}^n$ be an algebraic set then it's not too difficult to show that the affine algebra $\mathbb{K}[V]$ is generated by the restrictions $X_i|_V$ of the coordinate maps to the algebraic set V. Now to every point $x \in V$ we can associate a $\mathbb{K}$-algebra homomorphism

$$\varepsilon_x : \mathbb{K}[V] \to \mathbb{K}$$

given by $\varepsilon_x(f) = f(x)$ for all $f \in \mathbb{K}[V]$. We call ε_x the evaluation homomorphism. Indeed the assignment $x \mapsto \varepsilon_x$ gives us a bijection $V \to \mathrm{Hom}_{\mathbb{K}-\mathrm{alg}}(\mathbb{K}[V], \mathbb{K})$. Therefore V can be reconstructed from its affine algebra and so in some sense these objects are linked, this leads us to the following definition.

Definition. Let $V \subseteq \mathbb{K}^n$ be an algebraic set and $\mathbb{K}[V]$ be the associated affine algebra. We refer to the pair $(V, \mathbb{K}[V])$ as an *affine algebraic variety.*

Remark. From now on we will usually refer to a set V as an affine variety with it assumed that V is an algebraic subset of some sufficiently large $\mathbb{K}^n$ and $\mathbb{K}[V]$ its associated affine algebra.

Note that this definition has a disadvantage in that our algebraic set is implicitly defined as being embedded in some affine space, however this affine space is quite arbitrary. In fact the arbitrariness of the affine space we embed into will prove useful later on. There does exist a notion of an abstract affine variety which does not assume an embedding into affine space but we shall not use it here. A definition, due to Steinberg, can be found in [Gec03, Definition 2.1.6].

Let V be an irreducible affine variety, then we would like to have a notion of size for V. Now the affine algebra $\mathbb{K}[V]$ is an integral domain and hence has an associated field of fractions, which we denote $\mathbb{K}(V)$. We can consider $\mathbb{K} \subseteq \mathbb{K}(V)$ as a field extension and consider its transcendence degree, (in other words the minimal number of algebraically independent elements of $\mathbb{K}(V)$ over $\mathbb{K}$). We then say that the *dimension* $\dim V$ of V is equal to the transcendence degree of $\mathbb{K} \subseteq \mathbb{K}(V)$.

Example. Consider $\mathbb{K}^n$ as an affine variety. Its associated affine algebra is $\mathbb{K}[X_1, \ldots, X_n]$ and we have the field of fractions to be $\mathbb{K}(X_1, \ldots, X_n)$. Therefore $\mathbb{K} \subseteq \mathbb{K}(X_1, \ldots, X_n)$ has transcendence degree n, which means $\mathbb{K}^n$ is an affine variety of dimension n.

We have now defined our desired notion of affine variety, introduced a way of relating two such objects and given a notion of size. We would finally like to know a construction for putting two such objects together.

Definition. Let $V = \mathbf{V}(S) \subseteq \mathbb{K}^n$ and $W = \mathbf{V}(T) \subseteq \mathbb{K}^m$ be affine varieties. We identify $\mathbb{K}^n \times \mathbb{K}^m$ with $\mathbb{K}^{n \times m}$ then we have $V \times W \subseteq \mathbb{K}^{n+m}$ is an algebraic set as

$$V \times W = \{(v, w) \in \mathbb{K}^{n+m} \mid f(v) = g(w) = 0 \text{ for all } f \in S, g \in T\} = \mathbf{V}(S \cup T) \subseteq \mathbb{K}^{n+m}.$$

Hence the *direct product* is an affine algebraic variety.

We define a regular map $f \cdot g : V \times W \to \mathbb{K}$ by $(f \cdot g)(v, w) = f(v)g(w)$ for all $f \in \mathbb{K}[V]$ and $g \in \mathbb{K}[W]$. Therefore the assignment $(f, g) \mapsto f \cdot g$ from $\mathbb{K}[V] \times \mathbb{K}[W] \to \mathbb{K}[V \times W]$ is $\mathbb{K}$-bilinear. By the universal property of the tensor product we have an induced mapping $\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[W] \to \mathbb{K}[V \times W]$.

Proposition 3.4. *Let $V \subseteq \mathbb{K}^n$ and $W \subseteq \mathbb{K}^m$ be non-empty affine algebraic varieties.*

(a) *If V and W are irreducible, then $V \times W$ is irreducible.*
(b) *We have $\mathbb{K}[V \times W] \cong \mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[W]$.*
(c) *We have $\dim(V \times W) = \dim V + \dim W$.*

Proof. See [Gec03, Proposition 1.3.8]. □

We finally have enough tools to introduce the heart of our desired topic. However we make a brief comment about the direct product $V \times W$ before giving this definition. Now $V \times W$ is also a topological space, which we endow again with the Zariski topology, not the product topology as you might imagine. In general there will be more closed sets in the Zariski topology on $V \times W$ than in the product topology.

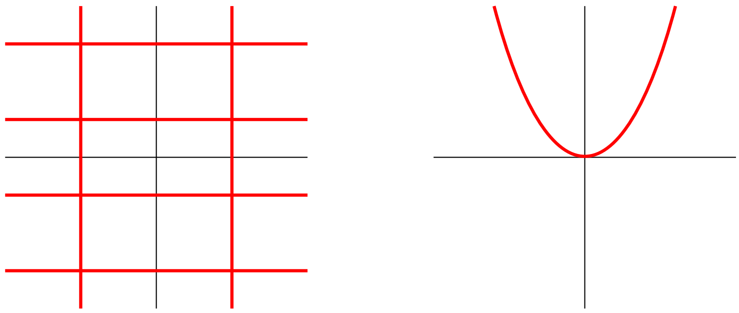


Figure 7. Product vs. Zariski Topology

Example. Consider affine 1-space, in this case $\mathbb{R}$. Now every closed set in the Zariski topology on $\mathbb{R}$ is a finite union of points. This is because every polynomial $f(X) \in \mathbb{R}[X]$ has a finite number of solutions and every closed set is a finite union of closed sets. Now consider the product topology on $\mathbb{R} \times \mathbb{R}$ which we identify with $\mathbb{R}^2$. Any curve, for example $g(X, Y) = Y - X^2 \in \mathbb{R}[X, Y]$, will define a closed set in the Zariski topology on $\mathbb{R}^2$.

In figure 7 we indicate a typical closed set in the product topology and a typical closed set in the Zariski topology on $\mathbb{R}^2$. It's easy to see that the curve cannot be constructed out of finite unions and arbitrary intersections of closed sets in the product topology and so there are more closed sets in the Zariski topology than the product topology.

Definition. A *linear algebraic group* G is an affine algebraic variety which also has the structure of a group, such that the multiplication $\mu : G \times G \to G$ and inversion $\iota : G \to G$ maps are regular.

A map $\varphi : G \to H$ is called a *homomorphism of algebraic groups* if φ is a regular map of affine varieties and a homomorphism of groups. Now φ is an *isomorphism of algebraic groups* if it is a bijective group homomorphism which is also an isomorphism of affine varieties.

Remark. It's very easy to see that any closed subgroup of a linear algebraic group will again be a linear algebraic group. This is because the restrictions of the regular maps μ, ι to any closed subset will again be regular.

Example.

(a) Consider the affine variety $\mathbb{K}$ as a group under addition, so the multiplication and inversion maps are given by $\mu(x, y) = x + y$ and $\iota(x) = -x$ for all $x, y \in \mathbb{K}$. Then clearly $\mathbb{K}$ is a linear algebraic group, which we call the *additive group* of $\mathbb{K}$ and denote $\mathbb{G}_a$. The associated affine algebra of $\mathbb{G}_a$ is the polynomial ring $\mathbb{K}[X]$ and the induced $\mathbb{K}$-algebra homomorphisms $\mu^* : \mathbb{K}[X] \to \mathbb{K}[X] \otimes \mathbb{K}[X] = \mathbb{K}[X_1, X_2]$ and $\iota^* : \mathbb{K}[X] \to \mathbb{K}[X]$ are given by

$$\mu^*(X) = X_1 + X_2 \qquad \text{and} \qquad \iota^*(X) = -X.$$

It's clear that $\dim \mathbb{G}_a = 1$.

(b) Consider the subset $\mathbb{K}^\times = \mathbb{K} \setminus \{0\} \subseteq \mathbb{K}$. Now at first this may not appear like a closed set but we can see this by utilising the arbitrariness of the affine space we embed our closed sets into. Identifying $\mathbb{K} \times \mathbb{K}$ with $\mathbb{K}^2$ we see that

$$\mathbb{K}^\times \cong \{(x, y) \in \mathbb{K} \times \mathbb{K} \mid xy = 1\} \subseteq \mathbb{K} \times \mathbb{K}$$

is a closed subvariety of $\mathbb{K}^2$. Now we can consider $\mathbb{K}^\times$ as a group under multiplication, so the multiplication and inversion maps are given by $\mu(x,y) = xy$ and $\iota(x) = x^{-1}$ for all $x, y \in \mathbb{K}^\times$. Then it's now clear that $\mathbb{K}^\times$ is a linear algebraic group, which we call the *multiplicative group* of $\mathbb{K}$ and denote $\mathbb{G}_m$. The associated affine algebra of $\mathbb{G}_m$ is $\mathbb{K}[X, X^{-1}]$ and the induced $\mathbb{K}$-algebra homomomorphisms $\mu^* : \mathbb{K}[X, X^{-1}] \to \mathbb{K}[X_1, X_2, X_1^{-1}, X_2^{-1}]$ and $\iota^* : \mathbb{K}[X, X^{-1}] \to \mathbb{K}[X, X^{-1}]$ are given by

$$\mu^*(X) = X_1X_2 \qquad \text{and} \qquad \iota^*(X) = X^{-1}.$$

(c) Consider the set of all $n \times n$ matrices over $\mathbb{K}$, which we denote $M_n(\mathbb{K}) = \mathbb{K}^n \times \mathbb{K}^n$. Identifying $\mathbb{K}^n \times \mathbb{K}^n$ with $\mathbb{K}^{n^2}$ we see this is an algebraic set with associated affine algebra $\mathbb{K}[M_n(\mathbb{K})] = \mathbb{K}[X_{ij} \mid 1 \leqslant i, j \leqslant n]$. Now let $(a_{pq}), (b_{rs}) \in M_n(\mathbb{K})$ be matrices then matrix multiplication is given by

$$\mu((a_{pq}), (b_{rs})) = \left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)$$

and so μ is clearly a regular map. The induced $\mathbb{K}$-algebra homomorphism $\mu^* : \mathbb{K}[M_n(\mathbb{K})] \to \mathbb{K}[X_{ij}] \otimes \mathbb{K}[X_{ij}]$ is given by

$$\mu^*(X_{ij}) = \sum_{\ell=1}^{n} X_{i\ell} \otimes X_{\ell j} \qquad \text{for all } 1 \leqslant i, j \leqslant n.$$

Now we cannot define an inversion map on $M_n(\mathbb{K})$ as not every matrix is invertible. Therefore $M_n(\mathbb{K})$ is *not* a linear algebraic group but instead we refer to it as a *linear algebraic monoid*.

(d) Consider $\mathrm{GL}_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid \det A \neq 0\} \subseteq \mathbb{K}^{n^2}$. Now again this does not look like a closed set but we can play the same game with $\mathrm{GL}_n(\mathbb{K})$ that we did with $\mathbb{K}^\times$. In other words we have

$$\mathrm{GL}_n(\mathbb{K}) \cong \{(A, y) \in M_n(\mathbb{K}) \times \mathbb{K} \mid \det(A)y - 1 = 0\} \subseteq \mathbb{K}^{n^2+1}.$$

Now this is a closed set as we recall the determinant of a matrix can be expressed as

$$\det = \sum_{\sigma \in \mathfrak{S}_n} (-1)^{\mathrm{sgn}(\sigma)} X_{1\sigma(1)} \cdots X_{n\sigma(n)} \in \mathbb{K}[X_{ij} \mid 1 \leqslant i, j \leqslant n],$$

where $\mathfrak{S}_n$ is the symmetric group on n letters. Hence det is a polynomial and so $\mathrm{GL}_n(\mathbb{K})$ is an algebraic set with associated affine algebra $\mathbb{K}[X_{ij}, Y]/\mathbf{I}(\mathrm{GL}_n(\mathbb{K})) \cong \mathbb{K}[\hat{X}_{ij}, 1/\det]$, where $\hat{X}_{ij}$ are the restrictions of the coordinate functions to $\mathrm{GL}_n(\mathbb{K})$.

Now we can define an inversion map $\iota : \mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$ by $\iota(A) = \det(A)^{-1}A'$ where A' is the matrix of cofactors. The matrix A' is defined entirely in terms of polynomials in the variables $\hat{X}_{ij}$ and hence ι is a regular map. Therefore $\mathrm{GL}_n(\mathbb{K})$ is a linear algebraic group, which is irreducible of dimension n^2.

(e) We consider the special linear group $\mathrm{SL}_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid \det A - 1 = 0\} = \mathbf{V}(\{\det -1\}) \subseteq M_n(\mathbb{K})$, which is clearly an affine variety. Therefore by our remark we have $\mathrm{SL}_n(\mathbb{K}) \leqslant \mathrm{GL}_n(\mathbb{K})$ is a linear algebraic group. It can be shown that when $\mathbb{K}$ is algebraically closed we have $\mathbf{I}(\mathrm{SL}_n(\mathbb{K})) = \langle \det -1 \rangle$ and $\det -1 \in \mathbb{K}[X_{ij}]$ is irreducible. Therefore $\mathrm{SL}_n(\mathbb{K})$ is irreducible and of dimension $n^2 - 1$.

(f) Let us assume for ease that char $\mathbb{K} \neq 2$, for details on why see [Gec03, Example 1.3.16]. We consider the following matrices of $M_n(\mathbb{K})$

$$J_n = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \qquad \Omega_{2n} = \begin{bmatrix} 0 & J_n \\ -J_n & 0 \end{bmatrix}.$$

Then we can define the following algebraic sets contained in $M_n(\mathbb{K})$ or $M_{2n}(\mathbb{K})$

$$\begin{aligned} \mathrm{O}_n(\mathbb{K}) &= \{A \in M_n(\mathbb{K}) \mid A^T J_n A = J_n\}, \\ \mathrm{SO}_n(\mathbb{K}) &= \mathrm{O}_n(\mathbb{K}) \cap \mathrm{SL}_n(\mathbb{K}), \\ \mathrm{Sp}_{2n}(\mathbb{K}) &= \{A \in M_{2n}(\mathbb{K}) \mid A^T \Omega_{2n} A = \Omega_{2n}\}. \end{aligned}$$

We have that these are called the *orthogonal*, *special orthogonal* and *symplectic* groups. It's easy to see that $\mathrm{O}_n(\mathbb{K})$ and $\mathrm{Sp}_{2n}(\mathbb{K})$ are algebraic sets as matrix multiplication is given by polynomials. Also $\mathrm{SO}_{2n}(\mathbb{K})$ is an algebraic set as the intersection of two algebraic sets is algebraic. Therefore as they are all closed subgroups of either $\mathrm{GL}_n(\mathbb{K})$ or $\mathrm{GL}_{2n}(\mathbb{K})$ we have that they are linear algebraic groups. It is also true that $\mathrm{SO}_n(\mathbb{K})$ and $\mathrm{Sp}_{2n}(\mathbb{K})$ are irreducible affine varieties.

The dimensions of these varieties are quite difficult to work out directly but it can be shown that $\dim \mathrm{SO}_{2n}(\mathbb{K}) = n(2n-1)$, $\dim \mathrm{SO}_{2n+1}(\mathbb{K}) = n(2n+1)$ and $\dim \mathrm{Sp}_{2n}(\mathbb{K}) = n(2n+1)$. One may notice some similarities between the dimensions of these linear algebraic groups and the dimensions given in table 3.

(g) Finally one can also be convinced that the following subgroups of $\mathrm{GL}_n(\mathbb{K})$ are also linear algebraic groups:
- $D_n \leqslant \mathrm{GL}_n(\mathbb{K})$ the subgroup of all diagonal matrices. This is isomorphic as an algebraic group to $\mathbb{G}_m \times \cdots \times \mathbb{G}_m$, ($n$ copies).
- $B_n \leqslant \mathrm{GL}_n(\mathbb{K})$ the subgroup of all upper triangular matrices. In other words $\{(a_{ij}) \in \mathrm{GL}_n(\mathbb{K}) \mid a_{ij} = 0 \text{ for } i > j\}$.
- $U_n \leqslant \mathrm{GL}_n(\mathbb{K})$ the subgroup of all unipotent matrices. In other words $\{(a_{ij}) \in \mathrm{GL}_n(\mathbb{K}) \mid a_{ij} = 0 \text{ if } i > j \text{ and } a_{ii} = 1 \text{ for } 1 \leqslant i \leqslant n\}$.

Having given a list of examples of linear algebraic groups we would like to now know a little bit more about their internal structure. In particular, if G is a disconnected linear algebraic group then what can we say about the connected component that contains the identity of the group? We will provide a nice answer to this question but first we need a small result about images of regular maps.

Lemma 3.2. *Let $V \subseteq \mathbb{K}^n$, $W \subseteq \mathbb{K}^m$ be two affine varieties and $\varphi : V \to W$ be a regular map. Then if V is irreducible we have $\overline{\varphi(V)} \subseteq W$ is also irreducible.*

Proof. To show $\overline{\varphi(V)}$ is irreducible is equivalent to showing $\mathbf{I}(\overline{\varphi(V)})$ is a prime ideal by part (b) of proposition 3.2. Consider $f, g \in \mathbb{K}[Y_1, \ldots, Y_m]$ such that $fg \in \mathbf{I}(\overline{\varphi(V)})$. Therefore for all $x \in V$ we have $f(\varphi(x))g(\varphi(x)) = 0$ and so $(f \circ \varphi)(g \circ \varphi) \in \mathbf{I}(V)$. However V is irreducible so $\mathbf{I}(V)$ is a prime ideal, which means either $f \circ \varphi \in \mathbf{I}(V)$ or $g \circ \varphi \in \mathbf{I}(V)$. Assume without loss of generality that $f \circ \varphi \in \mathbf{I}(V)$ then $\varphi(V) \subseteq \mathbf{V}(\{f\})$ but clearly, as $\mathbf{V}(\{f\})$ is closed, this means $\varphi(V) \subseteq \overline{\varphi(V)} \subseteq \mathbf{V}(\{f\}) \Rightarrow f \in \mathbf{I}(\mathbf{V}(\{f\})) \subseteq \mathbf{I}(\overline{\varphi(V)})$ and so we're done. □

Remark. Let G be a linear algebraic group, choose an element $x \in G$ and consider the maps $\lambda_x : G \to G$ and $\rho_x : G \to G$ given by $\lambda_x(g) = \mu(x,g)$ and $\rho_x(g) = \mu(g,x)$ for all $g \in G$. Clearly λ_x is a regular map and as ρ_x is a composition of the multiplication map with $\delta : G \times G \to G \times G$ given by $\delta(x,g) = (g,x)$ we have ρ_x is also a regular map.

In fact these are isomorphisms of the affine variety G with inverses $\lambda_{x^{-1}}$ and $\rho_{x^{-1}}$. As these maps are isomorphisms of the variety structure on G we can use these maps to transport geometric and topological information from a single point to anywhere else in G.

Theorem 3.2. *Let G be a linear algebraic group and let $G^\circ \subseteq G$ be an irreducible component containing the identity of G. Then the following hold.*

(a) *G° is a closed normal subgroup of G of finite index and the cosets of G° are precisely the irreducible components of G. In particular, the irreducible components of G are disjoint, and G° is uniquely determined.*
(b) *Every closed subgroup of G of finite index contains G°.*
(c) *G is irreducible if and only if G is connected in the Zariski topology.*

Proof.

(a) Let $X \subseteq G$ be an irreducible component of G which contains the identity. Then $\overline{XG^\circ} = \overline{\mu(X \times G^\circ)} \subseteq G$ is also a closed irreducible subset of G. Now $1 \in X \cap G^\circ$ means $X \subseteq \overline{XG^\circ}$ and $G^\circ \subseteq \overline{XG^\circ}$ but X, G° are maximal so we must have $X = G^\circ = \overline{XG^\circ}$. Hence we have G° is unique. By this argument we also have $G^\circ G^\circ \subseteq \overline{G^\circ G^\circ} = G^\circ$ and so G° is closed under multiplication. Finally recall that the inversion map is an isomorphism of affine varieties and so $\iota(G^\circ)$ is also irreducible and contains the identity, hence $\iota(G^\circ) = G^\circ$, so G° is a subgroup of G.

We need to show that G° is normal. Recall that for any $x \in G$ we have λ_x, ρ_x are isomorphisms of affine varieties. Hence $xG^\circ x^{-1} = (\lambda_x \circ \rho_{x^{-1}})(G^\circ)$ is an irreducible component of G containing the identity so must be G°. Finally let $X \subseteq G$ be an irreducible component and let $x \in G$ be such that $x^{-1} \in X$. Then we have $\rho_x(X) = Xx$ is an irreducible component containg the identity and so $X = G^\circ x^{-1}$.
(b) Let $H \leqslant G$ be a closed subgroup and let $g_1, \dots, g_r \in G$ be right coset representatives for G/H. As before $\rho_{g_i}(H) = Hg_i$ is a closed subset of G. Now $G^\circ = G^\circ \cap G = \cup_{i=1}^r (G^\circ \cap Hg_i)$ and G° irreducible means $G^\circ \cap Hg_i = G^\circ \Rightarrow G^\circ \subseteq Hg_i$ for some $1 \leqslant i \leqslant n$. However G° contains the identity so we must have $G^\circ \subseteq H$.
(c) Note it's clear that G irreducible implies G connected. Assume, if possible, that G is connected but not irreducible then we have G has a decomposition into the cosets of G°. However these cosets are disjoint and so this would imply G is not connected. □

If G is a linear algebraic group then we refer to the closed normal subgroup $G^\circ \leqslant G$ as *the connected component* of G. Note that there are many connected components of a linear algebraic group but only one contains the identity and hence is again a group. It is customary to refer to a linear algebraic group as connected rather than irreducible. For example, from our list above we have $\mathbb{G}_a$, $\mathbb{G}_m$, $\mathrm{GL}_n(\mathbb{K})$, $\mathrm{SL}_n(\mathbb{K})$, D_n, B_n and U_n are all connected. We have $\mathrm{O}_n(\mathbb{K})$ is connected if $\operatorname{char} \mathbb{K} \neq 2$. Also $\mathrm{SO}_n(\mathbb{K})$ and $\mathrm{Sp}_{2n}(\mathbb{K})$ are connected but this is quite hard to see directly. Indeed they can shown to be connected as a consequence of having a BN-pair, (see [Gec03, Section 1.7]).

We finally finish this section with a similar result which was obtained for Lie algebras. Recall from section 2.5 that any Lie algebra can be embedded into the general linear Lie algebra. In this way we defined the abstract Jordan decomposition for any Lie algebra. We now obtain a similar result for linear algebraic groups.

Theorem 3.3. *Let G be a linear algebraic group. Then there is an isomorphism of algebraic groups between G and some closed subgroup of $\mathrm{GL}_n(\mathbb{K})$, for some n.*

Proof. See [Gec03, Corollary 2.4.4]. □

Definition. Let G be a linear algebraic group and φ an isomorphism of G onto some closed subgroup of $\mathrm{GL}_n(\mathbb{K})$. Now for any element $x \in G$ we say

- x is *semisimple* if $\varphi(x)$ is diagonalisable.
- x is *unipotent* if all eigenvalues of $\varphi(x)$ are equal to 1.

Remark. The definitions given above are independent of the isomorphism φ.

Theorem 3.4 (Jordan Decomposition)**.** *Let G be a linear algebraic group. Then every element $x \in G$ has a unique decomposition $x = x_s x_u$ such that $x_s, x_u \in G$ are semisimple and unipotent respectively.*

Proof. See [Hum75, Theorem 15.3]. □

3.3. **The Lie Algebra of a Linear Algebraic Group.** Having introduced the notion of a linear algebraic group we would like to find a way to simplify the problems that arise in dealing with these objects. The way in which this simplification takes place is via linearisation. Our aim is to show that to each linear algebraic group G there is an associated Lie algebra which lives in the 'tangent space' of the identity. Indeed this Lie algebra is somehow a linear approximation of a linear algebraic group.

We try and motivate geometrically what we are about to do. Consider a polynomial $f(X_1, X_2) \in \mathbb{K}[X_1, X_2]$ which defines a curve V in $\mathbb{K}^2$. We would describe the tangent of this curve at a point $p = (p_1, p_2) \in \mathbb{K}^2$ as the set of solutions to the linear polynomial

$$\frac{\partial f}{\partial X_1}(p)(X_1 - p_1) + \frac{\partial f}{\partial X_2}(p)(X_2 - p_2).$$

Now this will be a straight line through p unless both partial derivatives vanish.

So in general let $V \subseteq \mathbb{K}^n$ be a non-empty affine variety with associated vanishing ideal $\mathbf{I}(V)$ and let us fix a point $p = (p_1, \dots, p_n) \in V$. Then for any $f \in \mathbb{K}[X_1, \dots, X_n]$ we define a linear polynomial

$$d_p(f) = \sum_{i=1}^{n} \frac{\partial f}{\partial X_i}(p) X_i \in \mathbb{K}[X_1, \dots, X_n].$$

For any point $v = (v_1, \dots, v_n) \in \mathbb{K}^n$ we consider the 'line through p in the direction of v' given by $L_v = \{p + tv \mid t \in \mathbb{K}\} \subseteq \mathbb{K}^n$. Then given our polynomial f we can consider the Taylor expansion of f, as a function of t, around a point $a \in \mathbb{K}$. In other words

$$\begin{aligned} f(p + tv) &= f(p + av) + f'(p + av)(p + tv - p - av) + O(t^2), \\ &= f(p + av) + f'(p + av)v(t - a) + O(t^2). \end{aligned}$$

Now considering this expansion around the point $a = 0$ and taking the derivative of f to be our associated linear polynomial $d_p(f)$ we get

$$f(p+tv) = f(p) + d_p(f)(v)t + O(t^2) \Rightarrow d_p(f)(v) \approx \frac{f(p+tv) - f(p)}{t}.$$

Now we say L_v is a *tangent line* at $p \in V$ if $d_p(f)(v) = 0$ for all $f \in \mathbf{I}(V)$.

Definition. Let $V \subseteq \mathbb{K}^n$ be a non-empty algebraic set. For a fixed point $p \in V$ we define

$$\begin{aligned} T_p(V) &= \{v \in \mathbb{K}^n \mid d_p(f)(v) = 0 \text{ for all } f \in \mathbf{I}(V)\}, \\ &= \{v \in \mathbb{K}^n \mid L_v \text{ is a tangent line at } p \in V\}, \end{aligned}$$

to be the *tangent space* at $p \in V$.

Example. Consider the circle $V \subseteq \mathbb{R}^2$ defined by the polynomial $f(X_1, X_2) = X_1^2 + X_2^2 - 1 \in \mathbb{R}[X_1, X_2]$. It can be shown that the vanishing ideal for this curve is just $\mathbf{I}(V) = \langle f \rangle \subseteq \mathbb{R}[X_1, X_2]$. Now consider the point $p = (0, 1) \in \mathbb{R}^2$ then we have

$$d_p(f) = \frac{\partial f}{\partial X_1}(0,1)X_1 + \frac{\partial f}{\partial X_2}(0,1)X_2 = 2X_2.$$

Now we have $d_p(f)(v) = 0$ if and only if $v = (v_1, 0) \in \mathbb{K}^2$. Hence $T_p(V)$ is the straight line defined by $X_2 = 0$, which is as we would expect.

Remark. Let $V \subseteq \mathbb{K}^n$ be a non-empty algebraic set and assume $\mathbf{I}(V)$ is generated by the polynomials $f_1, \dots, f_r \in \mathbb{K}[X_1, \dots, X_n]$. Then we claim for some point $p \in V$ that

$$T_p(V) = \{v \in \mathbb{K}^n \mid d_p(f_i)(v) = 0 \text{ for all } 1 \leqslant i \leqslant r\}.$$

Clearly we have $\subseteq$ holds. Now for the reverse inclusion assume $v \in \mathbb{K}^n$ is such that $d_p(f_i)(v) = 0$ for all $1 \leqslant i \leqslant r$. Now any $f \in \mathbf{I}(V)$ has an expression of the form $f = \sum_{j=1}^n h_j f_j$ for some $h_j \in \mathbb{K}[X_1, \dots, X_n]$. Therefore we have

$$d_p(f) = \sum_{j=1}^n d_p(h_j f_j) = \sum_{j=1}^n (f_j(p)d_p(h_j) + h_j(p)d_p(f_j)) = \sum_{j=1}^n h_j(p)d_p(f_j).$$

Hence we also have $d_p(f)(v) = 0$ for all $v \in V$.

Example.

(a) Let $V \subseteq \mathbb{K}^n$ be a linear subspace then V is defined by linear polynomials $f_1, \dots, f_m \in k[X_1, \dots, X_n]$ with constant term zero. Note we take constant term zero as we must have V contains $0 \in \mathbb{K}^n$. It can be shown that $\mathbf{I}(V) = \langle f_1, \dots, f_m \rangle$. The polynomials are linear and so $d_p(f) = f$ for all $f \in \mathbf{I}(V)$, which means $T_p(V) = V$ for all $p \in V$. In particular $\dim V = \dim T_p(V)$ for all $p \in V$.

In particular consider the case $V = \mathbb{K}^n$, then it can be shown that $\mathbf{I}(V) = \{0\}$ and so $T_p(\mathbb{K}^n) = \mathbb{K}^n$ for all $p \in \mathbb{K}^n$.

(b) Let $V = \{p\} \subseteq k^n$. Now it can be shown that $\mathbf{I}(V) = \langle X_1 - p_1, \dots, X_n - p_n \rangle$. In this case we only have to check the conditions for $T_p(V)$ on the generating polynomials. So for each $1 \leqslant i \leqslant n$ we have

$$d_p(X_i - p_i) = \frac{\partial}{\partial X_i}(X_i - p_i)(p)X_i = 1(p)X_i = X_i.$$

So if $d_p(X_i - p_i)(v) = 0$ for each $1 \leqslant i \leqslant n$ we must have $v = 0$. Therefore $T_p(\{p\}) = \{0\}$ for all singletons $\{p\} \subseteq k^n$.

Now we have introduced the tangent space we would like to indicate the algebraic structure that this space carries. We do this using the language of derivations.

Definition. Let A be a $\mathbb{K}$-algebra and M an A-module, let the action of A on M be denoted by $a \cdot m$ for all $a \in A$, $m \in M$. We say a $\mathbb{K}$-linear map $D : A \to M$ is a *derivation* if D satisfies the Liebniz rule

$$D(ab) = a \cdot D(b) + b \cdot D(a),$$

for all $a, b \in A$. We write $\mathrm{Der}_{\mathbb{K}}(A, M)$ for the space of all derivations from A to M.

Recall that we introduced the notion of a derivation for a Lie algebra in section 2.1. We comment that this definition encapsulates the definition given previously. This is because any Lie algebra over $\mathbb{K}$ can be turned into a $\mathbb{K}$-algebra and we can take our module M to be the Lie algebra itself.

Example. Consider the polynomial ring $R := \mathbb{K}[X_1, \dots, X_n]$ and an R-module M such that the action is given by $r \cdot m$ for all $r \in R$, $m \in M$. Now for any derivation $D \in \mathrm{Der}_{\mathbb{K}}(R, M)$ we have

$$D(f) = \sum_{i=1}^{n} \frac{\partial f}{\partial X_i} \cdot D(X_i).$$

To see this we note that, by the Liebniz rule, any derivation D is uniquely determined by its action on a set of generators for R. Checking on two arbitrary generators, for some $1 \leqslant i, j \leqslant n$ we have

$$\begin{aligned} D(X_k + X_\ell) &= \sum_{i=1}^{n} \frac{\partial}{\partial X_i}(X_k + X_\ell) \cdot D(X_i) = D(X_k) + D(X_\ell), \\ D(X_k X_\ell) &= \sum_{i=1}^{n} \frac{\partial}{\partial X_i}(X_k X_\ell) \cdot D(X_i) = X_k \cdot D(X_\ell) + X_\ell \cdot D(X_i). \end{aligned}$$

It is then an easy induction proof to show that this holds for any arbitrary collection of generators. Consider R as an R-module then we have $\{\partial/\partial X_i\}_{1 \leqslant i \leqslant n}$ is a linearly independent set of vectors in $\mathrm{Der}_{\mathbb{K}}(R, R)$. Hence $\mathrm{Der}_{\mathbb{K}}(R, R)$ is a free R-module of rank n with basis given by the partial derivatives.

Let $V \subseteq \mathbb{K}^n$ be an affine variety and pick a point $p \in V$. We turn the field $\mathbb{K}$ into a module for the affine algebra $\mathbb{K}[V]$ by defining the action to be $\cdot : (\overline{f}, \alpha) \mapsto f(p)\alpha$. We write $\mathbb{K}_p$ when we consider $\mathbb{K}$ as a $\mathbb{K}[V]$ module in this way.

Lemma 3.3. *Let $V \subseteq \mathbb{K}^n$ be a non-empty affine variety and $p \in V$ be any point. Then for any $v \in T_p(V)$ we have a well define derivation $D_v \in \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[V], \mathbb{K}_p)$ given by $D_v(\overline{f}) = d_p(f)(v)$ for any $f \in \mathbb{K}[X_1, \dots, X_n]$. Furthermore, the map $\Psi : T_p(V) \to \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[V], \mathbb{K}_p)$ given by $v \mapsto D_v$ is an isomorphism of $\mathbb{K}[V]$-modules.*

Proof. Let R denote the polynomial ring $\mathbb{K}[X_1, \dots, X_n]$ and let $\pi : R \to \mathbb{K}[V]$ be the canonical projection map. We start by confirming that the derivation D_v is well defined. Assume $f, g \in \mathbb{K}[V]$ such that $f - g \in \mathbf{I}(V)$ then for any $v \in T_p(V)$ we have

$$D_v(\overline{f}) - D_v(\overline{g}) = d_p(f)(v) - d_p(g)(v) = d_p(f - g)(v) = 0$$

by the definition of $T_p(V)$. Hence D_v is well defined.

Now we have $D_v(\overline{X_j}) = v_j$ for each $1 \leqslant j \leqslant n$ and so the map Ψ is injective. Let $D \in \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[V], \mathbb{K}_p)$ be a derivation then we have the composition $\tilde{D} = D \circ \pi \in \mathrm{Der}_{\mathbb{K}}(R, \mathbb{K}_p)$ is also a derivation. Let $v = (v_1, \dots, v_n) = (\tilde{D}(X_1), \dots, \tilde{D}(X_n)) \in \mathbb{K}^n$, then from the example above we see that for all $f \in R$ we have

$$D(\overline{f}) = \tilde{D}(f) = \sum_{i=1}^{n} \frac{\partial f}{\partial X_i} \cdot \tilde{D}(X_i) = d_p(f)(v).$$

For any $f \in \mathbf{I}(V)$ we have $\overline{f} = 0$ and so $D(\overline{f}) = d_p(f)(v) = 0$, which means $v \in T_p(V)$. Since $D_v(\overline{X_j}) = v_j = D(\overline{X_j})$ we have $D_v = D$. Therefore Ψ is surjective and we're done. $\square$

So we can identify the tangent space with the vector space of all derivations. This gives us a much more algebraic view point for the tangent space of a variety at a point. What we would now like to show is that $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[V], \mathbb{K}_1)$ has the extra structure of a Lie algebra. Then using the isomorphism just given we would like to pass this structure to the tangent space $T_1(V)$. From this point on we will always identify $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[V], \mathbb{K}_p)$ with the tangent space $T_p(V)$.

Definition. Let $V \subseteq \mathbb{K}^n$ and $W \subseteq \mathbb{K}^m$ be two non-empty affine varieties and let $\varphi : V \to W$ be a regular map. Consider a point $p \in V$ and then let $q := \varphi(p) \in W$. Then the map

$$d_p\varphi : \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[V], \mathbb{K}_p) \to \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[W], \mathbb{K}_q)$$

defined by $D \mapsto D \circ \varphi^*$ is called the *differential of* φ.

Remark. If $V = W$ and $\varphi = \mathrm{id}_V$ then $d_p(\mathrm{id}_V)$ is the identity map on $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[V], \mathbb{K}_p)$ for all $p \in V$. Also if we have another affine variety $X \subseteq \mathbb{K}^\ell$ and another regular map $\psi : W \to Z$ then we have $d_p(\psi \circ \varphi) = d_q\psi \circ d_p\varphi$. Therefore if φ is an isomorphism of affine varieties with inverse φ^{-1} we have $d_p\varphi$ is an isomorphism of vector spaces with inverse $d_q\varphi^{-1}$.

Example. Let G be a linear algebraic group. Recall that, for any $x \in G$, we have the regular map $\lambda_x : G \to G$ given by $g \mapsto xg$ for all $g \in G$ is an isomorphism of affine varieties. Therefore by the remark we have $d_g(\lambda_x) : T_g(G) \to T_{xg}(G)$ is an isomorphism of vector spaces. In particular taking $x = g^{-1}$ we have $T_g(G) \cong T_1(G)$ for all $g \in G$.

Theorem 3.5. *Let $G \subseteq M_n(\mathbb{K})$ be a linear algebraic group and $T_1(G) \subseteq M_n(\mathbb{K})$ be the corresponding tangent space at the identity.*

(a) *For $A, B \in T_1(G)$ we have $[A, B] = AB - BA \in T_1(G)$ and so $T_1(G)$ is a Lie subalgebra of $M_n(\mathbb{K}) = \mathfrak{gl}_n(\mathbb{K})$.*
(b) *If $H \subseteq M_n(\mathbb{K})$ is another linear algebraic group and $\varphi : G \to H$ is a homomorphism of algebraic groups, then $d_1\varphi : T_1(G) \to T_1(H)$ is a homomorphism of Lie algebras.*

Proof.

(a) We start by defining a product $\star : \mathbb{K}[G] \times \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[G], \mathbb{K}_1) \to \mathbb{K}[G]$. Consider $D \in \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[G], \mathbb{K}_1)$ and $f, g \in \mathbb{K}[G]$, then we define $f \star D \in \mathbb{K}[G]$ by

$$(f \star D)(x) = D(\lambda_x^*(f)) \qquad \text{for all } x \in G.$$

We claim that $\star D : \mathbb{K}[G] \to \mathbb{K}[G]$ is a derivation, in other words $fg \star D = f(g \star D) + g(f \star D)$. We show this by evaluating at $x \in G$

$$\begin{aligned}(fg \star D)(x) = D(\lambda_x^*(f)\lambda_x^*(g)) &= \lambda_x^*(f) \cdot D(\lambda_x^*(g)) + \lambda_x^*(g) \cdot D(\lambda_x^*(f)),\\ &= \lambda_x^*(f)(1)D(\lambda_x^*(g)) + \lambda_x^*(g)(1)D(\lambda_x^*(f)),\\ &= f(x)(g \star D)(x) + g(x)(f \star D)(x),\\ &= (f(g \star D) + g(f \star D))(x).\end{aligned}$$

Now for any $D, D' \in \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[G], \mathbb{K}_1)$ we define a map $[D, D'] : \mathbb{K}[G] \to \mathbb{K}$ by

$$[D, D'](f) = D(f \star D') - D'(f \star D) \qquad \text{for } f \in \mathbb{K}[G].$$

Using the fact that $\star D$ is a derivation it is easy to verify that $[D, D']$ is in fact an element of $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[G], \mathbb{K}_1)$. Therefore for all $f, g \in \mathbb{K}[G]$ we have

$$\begin{aligned}[D, D'](fg) &= D(fg \star D') - D'(fg \star D),\\ &= D(f(g \star D')) + D(g(f \star D')) - D'(f(g \star D)) - D'(g(f \star D)),\\ &= f(1)(D(g \star D')) - D'(g \star D)) + g(1)(D(f \star D') - D'(f \star D)),\\ &= f \cdot [D, D'](g) + g \cdot [D, D'](f).\end{aligned}$$

Now we wish to use the isomorphism Ψ from lemma 3.3, given by $P = (p_{ij}) \mapsto D_P$, to pass this work to $T_1(G)$. Now any derivation D_P is determined by $D_P(\overline{X_{ij}}) = p_{ij}$ for $1 \leqslant i, j \leqslant n$. We need to identify the matrix in $T_1(G)$ associated with $[D_P, D_{P'}]$ for $P, P' \in T_1(G)$. This will be achieved if we can show the following identity holds

$$[D, D'](\overline{X_{ij}}) = \sum_{i=1}^{n}(D(\overline{X_{il}})D'(\overline{X_{lj}}) - D'(\overline{X_{il}})D(\overline{X_{lj}})).$$

Now by the definition of $[D, D']$ we have that it's action on $\overline{X_{ij}}$ is determined entirely by $D(\overline{X_{ij}}), D'(\overline{X_{ij}}), \overline{X_{ij}} \star D, \overline{X_{ij}} \star D'$. Therefore to verify the identity it is enough to show that

$$\overline{X_{ij}} \star D = \sum_{l=1}^{n} \overline{X_{il}} D(\overline{X_{lj}}) \qquad \text{for all } D \in \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[G], \mathbb{K}_1).$$

Let $x \in G$ and recall that $(\overline{X_{ij}} \star D)(x) = D(\lambda_x^*(\overline{X_{ij}}))$. Evaluating the right hand side of the identity at x gives us $\sum_{l=1}^{n} \overline{X_{il}}(x)D(\overline{X_{lj}})$. Therefore the identity will hold if

$$\lambda_x^*(\overline{X_{ij}}) = \sum_{l=1}^{n} \overline{X_{il}}(x)\overline{X_{lj}} \qquad \text{for all } x \in G.$$

Evaluating the left hand side of this identity at $y \in G$ we see

$$\begin{aligned}\lambda_x^*(\overline{X_{ij}})(y) = \overline{X_{ij}}(xy) = \overline{X_{ij}}(\mu(x,y)) &= \mu^*(\overline{X_{ij}})(x,y),\\ &= \sum_{l=1}^{n}(\overline{X_{il}} \otimes \overline{X_{lj}})(x,y),\\ &= \sum_{l=1}^{n}\overline{X_{il}}(x)\overline{X_{lj}}(y).\end{aligned}$$

Therefore this identity holds and hence we're done. For the statements involving μ^* recall the description given in example (c) of section 3.2.

(b) We claim that for all $D \in \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[G], \mathbb{K}_1)$ and $h \in \mathbb{K}[H]$ we have

$$\varphi^*(h) \star D = \varphi^*(h \star d_1\varphi(D)).$$

To show that this identity holds we evaluate both sides at an element $x \in G$. On the left side we see $(\varphi^*(h) \star D)(x) = D(\lambda_x^*(\varphi^*(h)))$. On the right hand side we have

$$\varphi^*(h \star d_1\varphi(D))(x) = (h \star d_1\varphi(D))\varphi(x) = d_1\varphi(D)(\lambda_{\varphi(x)}^*(h)) = D(\varphi^*(\lambda_{\varphi(x)}^*(h))).$$

Hence it is enough to check that $\lambda_x^*(\varphi^*(h)) = \varphi^*(\lambda_{\varphi(x)}^*(h))$ holds for all $x \in G$. Evaluating this identity on both sides at some $y \in G$ we have

$$\begin{aligned}\lambda_x^*(\varphi^*(h))(y) &= (h \circ \varphi \circ \lambda_x)(y) = h(\varphi(xy)),\\ \varphi^*(\lambda_{\varphi(x)}^*(h))(y) &= (h \circ \lambda_{\varphi(x)} \circ \varphi)(y) = h(\varphi(x)\varphi(y)) = h(\varphi(xy)).\end{aligned}$$

Therefore the identity holds true. Finally we can now show that $d_1\varphi$ is indeed a homomorphism of Lie algebras. Let $D, D' \in \mathrm{Der}_{\mathbb{K}}(\mathbb{K}[G], \mathbb{K}_1)$ and $h \in \mathbb{K}[H]$ then we have

$$\begin{aligned}[d_1\varphi(D), d_1\varphi(D')](h) &= d_1\varphi(D)(h \star d_1\varphi(D')) - d_1\varphi(D')(h \star d_1\varphi(D)),\\ &= D(\varphi^*(h \star d_1\varphi(D'))) - D'(\varphi^*(h \star d_1\varphi(D))),\\ &= D(\varphi^*(h) \star D') - D'(\varphi^*(h) \star D),\\ &= [D, D'](\varphi^*(h)),\\ &= d_1\varphi([D, D'])(h).\end{aligned}$$

□

Remark. Throughout this section we have followed the format of [Gec03]. Although the author believes that this approach is the most concrete and easiest to understand it does not yield the nicest proofs. If the reader feels these proofs are too technical they might like to look at the proofs given in [Hum75, Section 9].

Example. Let $\mathbf{V}(S) \subseteq \mathbb{K}^n$ be a non-empty algebraic set such that $S \subseteq \mathbb{K}[X_1, \ldots, X_n]$. We always have that $S \subseteq \mathbf{I}(\mathbf{V}(S))$ and so it's clear to see that

$$T_1(\mathbf{V}(S)) \subseteq \{v \in \mathbb{K}^n \mid d_1(f)(v) = 0 \text{ for all } f \in S\}.$$

However to achieve the reverse inclusion usually involves some dimensional arguments and indeed is not always true. In the examples that follow the reverse inclusion is always true but we will not compute the dimensional arguments for simplicity.

(a) Consider the linear algebraic group $\mathbb{G}_m = \mathbb{K}^\times$. We recall that

$$\mathbb{K}^\times = \{(x, y) \in \mathbb{K} \times \mathbb{K} \mid xy - 1 = 0\} \subseteq \mathbb{K}^2.$$

In this case the tangent space will be the set of zeroes of the polynomial $d_{(1,1)}(XY-1)$. Using standard formulas for differentiation we see

$$d_{(1,1)}(XY - 1) = d_{(1,1)}(XY) = X(1, 1)d_{(1,1)}(Y) + Y(1, 1)d_{(1,1)}(X) = Y + X.$$

Therefore we have

$$T_1(\mathbb{K}^\times) = \{(x, y) \in \mathbb{K} \times \mathbb{K} \mid x + y = 0\} \cong \mathbb{K}.$$

Note that this isomorphism is an isomorphism of Lie algebras. Therefore we have $T_1(\mathbb{K}^\times)$ is the trivial 1-dimensional Lie algebra.

(b) Consider the linear algebraic group $\mathrm{SL}_n(\mathbb{K})$. We recall that

$$\mathrm{SL}_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid \det(A) - 1 = 0\} \subseteq \mathbb{K}^{n^2}.$$

Our aim is to calculate $d_1(\det -1)$ but it's easy to see that this will be equivalent to calculating $d_1(\det)$. First considering the partial derivatives of det we see

$$\begin{aligned}
\frac{\partial}{\partial X_{rs}}(\det) &= \sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) \frac{\partial}{\partial X_{rs}}(X_{1\pi(1)} \cdots X_{n\pi(n)}), \\
&= \sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) \left(\delta_{s\pi(r)} \prod_{\substack{i=1 \\ i \neq r}}^{n} X_{i\pi(i)} \right).
\end{aligned}$$

Therefore examining the derivative of det we see

$$\begin{aligned}
d_1(\det) &= \sum_{r,s=1}^{n} \frac{\partial}{\partial X_{rs}}(\det)(I_n) X_{rs}, \\
&= \sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) \sum_{r,s=1}^{n} \left(\delta_{s\pi(r)} \prod_{\substack{i=1 \\ i \neq r}}^{n} \delta_{i\pi(i)} \right) X_{rs}, \\
&= \sum_{r,s=1}^{n} \delta_{sr} X_{rs}, \\
&= \sum_{\ell=1}^{n} X_{\ell\ell}
\end{aligned}$$

We make a comment regarding the above calculation. The term $\prod_{i \neq r} \delta_{i\pi(r)}$ is non-zero if and only if π fixes $n-1$ points of $\{1, \ldots, n\}$. However if π fixes $n-1$ points then it must fix n points and so π must be the identity in $\mathfrak{S}_n$. Therefore $d_p(\det) = \text{trace}$ and so we have

$$T_1(\mathrm{SL}_n(\mathbb{K})) = \{A \in M_n(\mathbb{K}) \mid \text{trace}(A) = 0\} = \mathfrak{sl}(n, \mathbb{K})$$

(c) Consider the linear algebraic group $\mathrm{GL}_n(\mathbb{K})$. We recall that

$$\mathrm{GL}_n(\mathbb{K}) \cong \{(A, y) \in M_n(\mathbb{K}) \times \mathbb{K} \mid \det(A)y - 1 = 0\} \subseteq \mathbb{K}^{n^2+1}.$$

So we aim to calculate $d_{(I_n,1)}(\det \cdot Y - 1)$ but using formal formulas for differentiation it's easy to see that we have

$$\begin{aligned} d_{(I_n,1)}(\det \cdot Y - 1) &= d_{(I_n,1)}(\det \cdot Y), \\ &= \det(I_n, 1) d_{(I_n,1)}(Y) + Y(I_n, 1) d_{(I_n,1)}(\det), \\ &= Y + \text{trace}\,. \end{aligned}$$

Therefore we see that

$$T_1(\mathrm{GL}_n(\mathbb{K})) = \{(A, y) \in M_n(\mathbb{K}) \times \mathbb{K} \mid \text{trace}(A) + y = 0\} \cong M_n(\mathbb{K}) = \mathfrak{gl}(n, \mathbb{K}).$$

Note that the isomorphism between $T_1(\mathrm{GL}_n(\mathbb{K}))$ and $M_n(\mathbb{K})$ is an isomorphism of Lie algebras.

(d) Consider the linear algebraic group $U_n \leqslant \mathrm{GL}_n(\mathbb{K})$ of all unipotent matrices. We remember that this is the algebraic set defined by the polynomials

$$f_{ii}(X) = X_{ii} - 1 \qquad g_{ij}(X) = X_{ij} \qquad \text{for all } 1 \leqslant i < j \leqslant n.$$

We want to work out the tangent space $T_1(U_n)$ and so we calculate $d_1(f_{ii})$ and $d_1(g_{ij})$. Therefore

$$\begin{aligned} d_1(f_{ii}) &= \sum_{j=1}^{n} \frac{\partial}{\partial X_{jj}} (X_{ii} - 1)(1) X_{jj} = X_{ii} && 1 \leqslant i \leqslant n, \\ d_1(g_{ij}) &= \sum_{\ell,m=1}^{n} \frac{\partial}{\partial X_{\ell m}} (X_{ij})(1) X_{\ell m} = X_{ij} && 1 \leqslant i < j \leqslant n. \end{aligned}$$

Hence the tangent space $T_1(U_n)$ is the set of all matrices that vanish on these polynomials. This gives us

$$T_1(U_n) = \{A \in M_n(k) \mid a_{ij} = 0 \text{ for } 1 \leqslant j \leqslant i \leqslant n\} = \mathfrak{n}(n, \mathbb{K}).$$

(e) Consider the linear algebraic group $D_n \leqslant \mathrm{GL}_n(\mathbb{K})$ of all diagonal matrices. This is isomorphic as an algebraic group to $\mathbb{G}_m \times \cdots \times \mathbb{G}_m$, ($n$ copies). Therefore we will get $T_1(D_n) \cong T_1(\mathbb{G}_m) \oplus \cdots \oplus T_1(\mathbb{G}_m) = \mathbb{K} \oplus \cdots \oplus \mathbb{K}$. Therefore $T_1(D_n)$ is the abelian Lie algebra of dimension n.

3.4. **The Root Datum of a Linear Algebraic Group.** We wish study the whole collection of linear algebraic groups but this is too complicated, much like trying to study all finite groups is too complicated. Therefore we try and identify a class of linear algebraic groups that is not so small that our theorems are not powerful but is not so big that it's over complicated. One may think that just restricting to connected linear algebraic groups would be enough but this class is too big. Instead we restrict ourselves to what are known as the connected *reductive* linear algebraic groups.

We note that if G is a linear algebraic group then we say G is *solvable* if it is solvable as an abstract group. Likewise we say a closed subgroup $H \leqslant G$ is *normal* if it is normal as an abstract group.

Definition. Let G be a linear algebraic group. We define the *radical* of G to be the unique maximal closed connected solvable normal subgroup of G, which we denote $R(G)$. We define the *unipotent radical* of G to be the unique maximal closed connected normal subgroup of G all of whose elements are unipotent and we denote this $R_u(G)$.

Now it's clear that for any linear algebraic group G we have $R_u(G) \subseteq R(G)$, indeed $R_u(G)$ is just all unipotent elements in $R(G)$. We say G is *reductive* if $R_u(G) = \{1\}$ and *semisimple* if $R(G) = \{1\}$. Therefore any semisimple linear algebraic group is also reductive.

Example. We have $\mathrm{GL}_n(\mathbb{K})$ is a connected reductive linear algebraic, which is not semisimple. It is not so easy to see this with the information we have to hand. However it is known that any closed connected solvable subgroup of $\mathrm{GL}_n(\mathbb{K})$ is conjugate to a subgroup of $B_n \leqslant \mathrm{GL}_n(\mathbb{K})$, (this is known as the Lie-Kolchin theorem). Hence we have $R(\mathrm{GL}_n(\mathbb{K})) \leqslant B_n$. Now B_n is conjugate to $B_n^- \leqslant \mathrm{GL}_n(\mathbb{K})$ the subgroup of all lower triangular matrices. This means $R(\mathrm{GL}_n(\mathbb{K})) \leqslant B_n \cap B_n^- = D_n$. However the radical must be normal, which gives us that $R(\mathrm{GL}_n(\mathbb{K}))$ is the set of all scalar multiples of the identity. Therefore $R(\mathrm{GL}_n(\mathbb{K})) = Z(\mathrm{GL}_n(\mathbb{K})) \cong \mathbb{G}_m$. However the only unipotent element of $\mathbb{G}_m$ is the identity so we have $R_u(\mathrm{GL}_n(\mathbb{K})) = \{1\}$.

What we have avoided mentioning so far is the idea of a quotient group. Indeed to really understand the structure of linear algebraic groups we need to know that there is a sensible way to consider the quotient by a normal subgroup. We won't prove this here as it is not a trivial statement.

Lemma 3.4. *Let G be a linear algebraic group and $N \leqslant G$ a closed normal subgroup. Then the abstract group G/N has the structure of an affine variety and indeed is a linear algebraic group.*

Proof. See [Hum75, Theorem 11.5]. □

Definition. Let G be a linear algebraic group and $B \leqslant G$ a subgroup of G. We say B is a *Borel subgroup* if B is a closed connected solvable subgroup, which is maximal with respect to inclusion amongst all such subgroups of G.

Let T be a linear algebraic group, then we say T is a *torus* if it is isomorphic as an algebraic group to $\mathbb{G}_m \times \cdots \times \mathbb{G}_m$, for some finite number of copies. Let G be a linear algebraic group and $T \leqslant G$ a closed subgroup of G. Then we say T is a *maximal torus* of G if T is a torus and maximal with respect to inclusion amongst all such subgroups of G.

Remark. A maximal torus is abelian and any abelian group is solvable. Hence any maximal torus is a closed connected solvable subgroup and so must lie in some Borel subgroup.

Now we have these basic notions we can discuss what the internal structure of an arbitrary linear algebraic group is. We have the following chain of normal subgroups

$$G \xrightarrow{\text{(finite)}} G^\circ \xrightarrow{\text{(semisimple)}} R(G) \xrightarrow{\text{(torus)}} R_u(G) \xrightarrow{\text{(unipotent)}} \{1\}.$$

Now the labels on the lines describe the quotients. For example G/G° is a finite group, $G^\circ/R(G)$ is semisimple, etc. Note that $G^\circ/R_u(G)$ is a connected reductive group and so this picture suggests that an arbitrary connected reductive group is built out of a semisimple group and torus. Indeed this is the case as if G is a connected reductive group we have $G = G'Z(G)^\circ$ where G' is the derived subgroup of G and $Z(G)^\circ$ is the connected component of the centre. This picture suggests that we can break down the study of arbitrary linear algebraic groups to studies of these types of groups.

Example. Consider the linear algebraic group $\mathrm{GL}_n(\mathbb{K})$. We have already seen that this group is connected reductive. It is an exercise to show that $[\mathrm{GL}_n(\mathbb{K}), \mathrm{GL}_n(\mathbb{K})] = \mathrm{SL}_n(\mathbb{K})$ and that $Z(\mathrm{GL}_n(\mathbb{K}))^\circ = D_n$. Therefore we have $\mathrm{GL}_n(\mathbb{K}) = \mathrm{SL}_n(\mathbb{K})D_n$, where $\mathrm{SL}_n(\mathbb{K})$ is semisimple and clearly D_n is a torus.

Now much like a maximal toral subalgebra was important in the study of Lie algebras we will have the maximal torus defined here will be important to the study of linear algebraic groups. We give some very important facts about these subgroups, which we will not prove.

Theorem 3.6. *Let G be a connected linear algebraic group, $T \leqslant G$ a maximal torus of G and $B \leqslant G$ a Borel subgroup containing T.*

(a) *All Borel subgroups of G are conjugate.*
(b) *All maximal tori of G are conjugate.*
(c) $N_G(B) = B$.
(d) *We have $B = UT$, such that $U \cap T = \{1\}$, and $U = R_u(B)$. Hence B is a semidirect product of U and T.*
(e) *There exists a unique* opposite Borel subgroup *B^- such that $B \cap B^- = T$. We also have $B^- = U^-T$ where $U^- = R_u(B^-)$.*
(f) *We have $N_G(T)^\circ = C_G(T) \Rightarrow N_G(T)/C_G(T)$ is finite, (this is known as the* rigidity of tori*).*
(g) *We have $N_G(T)/C_G(T)$ is always a Weyl group.*
(h) *If G is reductive then $C_G(T) = T$.*

Proof. For part (a) see [Gec03, Theorem 3.4.3] or [Hum75, Theorem 21.3]. Note that this proof requires knowledge of projective varieties which we have not introduced here. Part (b) is a consequence of part (a) given in [Hum75, Corollary A - Section 21.3]. For part (c) see [Hum75, Theorem 23.1]. For part (d) see [Gec03, Theorem 3.5.6]. For parts (f)-(h) see [Hum75, Section 16]. □

Example. Consider the connected reductive group $G = \mathrm{GL}_n(\mathbb{K})$. Now a standard Borel subgroup of $\mathrm{GL}_n(\mathbb{K})$ is given by the subgroup B_n and we have $B_n = U_nD_n$ where

$$B_n = \left\{\begin{bmatrix} \star & \star & \star \\ & \star & \star \\ & & \star \end{bmatrix}\right\} \qquad U_n = \left\{\begin{bmatrix} 1 & \star & \star \\ & 1 & \star \\ & & 1 \end{bmatrix}\right\} \qquad D_n = \left\{\begin{bmatrix} \star & & \\ & \star & \\ & & \star \end{bmatrix}\right\}.$$

It's an exercise to show that $N_{\mathrm{GL}_n(\mathbb{K})}(B_n) = B_n$. Now the opposite Borel subgroup $B_n^- = U_n^- D_n$ is clearly going to be

$$B_n^- = \left\{ \begin{bmatrix} \star & & \\ \star & \star & \\ \star & \star & \star \end{bmatrix} \right\} \qquad U_n^- = \left\{ \begin{bmatrix} 1 & & \\ \star & 1 & \\ \star & \star & 1 \end{bmatrix} \right\}.$$

It's obvious that $B_n = B_n^- = D_n$.

We have already seen that $\mathrm{GL}_n(\mathbb{K})$ is a connected reductive group and $C_G(D_n) = D_n$ is a maximal torus of $\mathrm{GL}_n(\mathbb{K})$. We have $N_G(D_n)/D_n$ will be isomorphic to the group of all $n \times n$ permutation matrices, which in turn is naturally isomorphic to the symmetric group $\mathfrak{S}_n$ which we have already seen to be a Weyl group.

We recall that in the Lie algebra setting we fixed a maximal toral subalgebra and decomposed the Lie algebra with respect to the adjoint action. In this way we obtained a root system for the Lie algebra. Indeed this is what we plan to describe for a linear algebraic group but instead of a root system we will obtain a slightly more elaborate construction known as a root datum.

Let G be a connected linear algebraic group and $T \leqslant G$ a maximal torus. Then we consider $X(T) = \mathrm{Hom}_{\mathrm{alg}}(T, \mathbb{G}_m)$ to be the set of algebraic group homomorphisms from T to $\mathbb{G}_m$. Now we have $X(T)$ is naturally an additive group by letting

$$(\chi_1 + \chi_2)(t) = \chi_1(t)\chi_2(t) \qquad \text{for all } \chi_1, \chi_2 \in X(T), t \in T.$$

Now if $T \cong \mathbb{G}_m$, as algebraic groups, then we have $X(T) = \mathrm{Hom}_{\mathrm{alg}}(\mathbb{G}_m, \mathbb{G}_m) \cong \mathbb{Z}$, as abstract groups. This is because any algebraic group homomorphism from $\mathbb{G}_m \to \mathbb{G}_m$ must be of the form $\lambda \mapsto \lambda^m$ for some $m \in \mathbb{Z}$. Therefore, in general, as T is isomorphic to n copies of $\mathbb{G}_m$ we have

$$X(T) \cong \mathbb{Z} \oplus \cdots \oplus \mathbb{Z}.$$

In other words $X(T)$ is a free abelian group of rank n. We call $X(T)$ the *character group* of T.

Now we let $Y(T) = \mathrm{Hom}_{\mathrm{alg}}(\mathbb{G}_m, T)$. Then we also have $Y(T)$ is naturally an additive group by letting

$$(\gamma_1 + \gamma_2)(\lambda) = \gamma_1(\lambda)\gamma_2(\lambda) \qquad \text{for all } \gamma_1, \gamma_2 \in Y(T), \lambda \in \mathbb{G}_m.$$

By the same argument as before we see that in general $Y(T) \cong \mathbb{Z} \oplus \cdots \oplus \mathbb{Z}$ is a free abelian group of rank n. We call $Y(T)$ the *cocharacter group* of T. Now we can define a bilinear map

$$\langle -, - \rangle : X(T) \times Y(T) \to \mathbb{Z},$$

which puts $X(T)$ and $Y(T)$ into duality as abelian groups. We note that given $\chi \in X(T)$, $\gamma \in Y(T)$ we have $\chi \circ \gamma \in \mathrm{Hom}_{\mathrm{alg}}(\mathbb{G}_m, \mathbb{G}_m)$. Therefore for all $\lambda \in \mathbb{G}_m$ we have $(\chi \circ \gamma)(\lambda) = \lambda^m$ for some $m \in \mathbb{Z}$. We then define our map by $\langle \chi, \gamma \rangle = m$. This map induces the required isomorphisms of abelian groups

$$X(T) \cong \mathrm{Hom}(Y(T), \mathbb{Z}) \qquad \text{and} \qquad Y(T) \cong \mathrm{Hom}(X(T), \mathbb{Z}).$$

We now stipulate that G is a connected reductive linear algebraic group. We recall $W = N_G(T)/T$ is a Weyl group. Now W acts on T by conjugation and we set $t^{\dot{w}} = w^{-1}tw$ for all

$t \in T$, where $\dot{w}$ is a representative in W for $w \in N_G(T)$. Then this conjugation action induces an action of W on the character and cocharacter groups of T by

$$
\begin{aligned}
{}^{w}\chi(t) &= \chi(t^{w}) && \text{for all } \chi \in X(T), t \in T, w \in W,\\
\gamma^{w}(\lambda) &= \gamma(\lambda)^{w} && \text{for all } \gamma \in Y(T), \lambda \in \mathbb{K}^{\times}, w \in W.
\end{aligned}
$$

We have T is contained in a Borel subgroup B, which decomposes as $B = UT$. Also T acts on U by conjugation and we look for the minimal proper subgroups of U which are stable under this action. These will all be connected unipotent subgroups of dimension 1, which are isomorphic to the additive group $\mathbb{G}_a$. Now T acts on each of these by conjugation and hence we get a homomorphism $T \to \mathrm{Aut}_{\mathrm{alg}}(\mathbb{G}_a)$. Now the only algebraic group automorphisms of $\mathbb{G}_a$ are of the form $\lambda \mapsto \mu\lambda$ for $\mu \in \mathbb{K}^{\times}$, which means we have an isomorphism of abstract groups $\mathrm{Aut}_{\mathrm{alg}}(\mathbb{G}_a) \cong \mathbb{K}^{\times}$. Therefore the conjugation action of T gives us an element of $X(T)$ and we call these elements the *positive roots*.

Let $\Phi^+ \subseteq X(T)$ denote the set of all such elements. Then to each positive root $\alpha \in \Phi^+$ we have an associated 1–dim subgroup $U_\alpha \subseteq U$, which we call the *root subgroup* associated to α. We know that given any maximal torus T and Borel subgroup B containing T then there exists a unique opposite Borel subgroup B^- such that $B \cap B^- = T$. So what happens if we play the exact same game with B^-? We get a set of elements $\Phi^- \subseteq X(T)$ which we call the *negative roots*. It turns out that $\alpha \in \Phi^+$ if and only if $-\alpha \in \Phi^-$. We let $\Phi = \Phi^+ \cup \Phi^- \subseteq X(T)$ and call this the set of *roots*.

Example. We aim to illustrate these ideas with an example. Consider the linear algebraic group $\mathrm{GL}_3(\mathbb{K})$. We have a maximal torus $T \leqslant \mathrm{GL}_3(\mathbb{K})$ and Borel subgroup $B = UT \leqslant \mathrm{GL}_3(\mathbb{K})$ given by

$$
B = \left\{ \begin{bmatrix} \star & \star & \star \\ & \star & \star \\ & & \star \end{bmatrix} \right\} \qquad U = \left\{ \begin{bmatrix} 1 & \star & \star \\ & 1 & \star \\ & & 1 \end{bmatrix} \right\} \qquad T = \left\{ \begin{bmatrix} \star & & \\ & \star & \\ & & \star \end{bmatrix} \right\},
$$

where stars indicate appropriate elements of $\mathbb{K}$. It's easy to see that there are three 1–dimensional proper subgroups of U normalised by the action of T. These are

$$
U_\alpha = \left\{ \begin{bmatrix} 1 & \lambda & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \right\} \qquad U_\beta = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ & 1 & \lambda \\ & & 1 \end{bmatrix} \right\} \qquad U_{\alpha+\beta} = \left\{ \begin{bmatrix} 1 & 0 & \lambda \\ & 1 & 0 \\ & & 1 \end{bmatrix} \right\}
$$

where $\lambda \in \mathbb{K}$.

We indicate how the root $\alpha \in X(T)$ is obtained. Consider the conjugation action of T on U_α then we have

$$
\begin{bmatrix} \mu_1 & & \\ & \mu_2 & \\ & & \mu_3 \end{bmatrix}
\begin{bmatrix} 1 & \lambda & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix}
\begin{bmatrix} \mu_1^{-1} & & \\ & \mu_2^{-1} & \\ & & \mu_3^{-1} \end{bmatrix}
=
\begin{bmatrix} 1 & \mu_1\mu_2^{-1}\lambda & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix}.
$$

Therefore if $t = \mathrm{diag}(\mu_1, \mu_2, \mu_3) \in T$ then $\alpha \in X(T)$ is given by $\alpha(t) = \mu_1\mu_2^{-1}$. Indeed a similar calculation gives us $\beta(t) = \mu_2\mu_3^{-1}$ and hence $(\alpha + \beta)(t) = \alpha(t)\beta(t) = \mu_1\mu_3^{-1}$. If we define coordinate maps $\varepsilon_i : T \to \mathbb{G}_m$ by $\varepsilon_i(t) = \mu_i$ then we have the positive and negative roots are given by

$$\begin{aligned} \alpha &= \varepsilon_1 - \varepsilon_2 & -\alpha &= -\varepsilon_1 + \varepsilon_2, \\ \beta &= \varepsilon_2 - \varepsilon_3 & -\beta &= -\varepsilon_2 + \varepsilon_3, \\ \alpha + \beta &= \varepsilon_1 - \varepsilon_3 & -\alpha &= -\varepsilon_1 + \varepsilon_3. \end{aligned}$$

We then have that the corresponding negative root subgroups are

$$U_{-\alpha} = \left\{ \begin{bmatrix} 1 & & \\ \lambda & 1 & \\ 0 & 0 & 1 \end{bmatrix} \right\} \qquad U_{-\beta} = \left\{ \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & \lambda & 1 \end{bmatrix} \right\} \qquad U_{-\alpha-\beta} = \left\{ \begin{bmatrix} 1 & & \\ 0 & 1 & \\ \lambda & 0 & 1 \end{bmatrix} \right\}$$

Note the resemblance to the descriptions of root systems given in the examples of $\mathfrak{sl}(3, \mathbb{C})$ in section 2.5 and the general case of A_n in section 1.5.

We now return to the general situation of G a connected reductive linear algebraic group.

Proposition 3.5. *Given a pair of roots $\{\alpha, -\alpha\} \subseteq \Phi$ then the closed subgroup $\langle U_\alpha, U_{-\alpha} \rangle \leqslant G$ is isomorphic as an algebraic group to either $\mathrm{SL}_2(\mathbb{K})$ or $\mathrm{PGL}_2(\mathbb{K}) = \mathrm{GL}_2(\mathbb{K})/Z(\mathrm{GL}_2(\mathbb{K}))$.*

Proof. This is mostly contained in [Hum75, Corollary 32.3]. □

Therefore we can find a surjective homomorphism $\varphi : \mathrm{SL}_2(\mathbb{K}) \to \langle U_\alpha, U_{-\alpha} \rangle$ such that

$$\left\{ \varphi \begin{bmatrix} 1 & \star \\ 0 & 1 \end{bmatrix} \right\} = U_\alpha \qquad \left\{ \varphi \begin{bmatrix} 1 & 0 \\ \star & 1 \end{bmatrix} \right\} = U_{-\alpha} \qquad \left\{ \varphi \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix} \right\} = T.$$

Therefore there is a homomorphism $\alpha^\vee : \mathbb{G}_m \to T$ given by

$$\alpha^\vee(\lambda) = \varphi \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix},$$

for all $\lambda \in \mathbb{G}_m$. The element $\alpha^\vee \in Y(T)$ is called the *coroot* associated to α. Given a root α we have its coroot $\alpha^\vee$ is determined uniquely by the property $\langle \alpha, \alpha^\vee \rangle = 2$. We write $\Phi^\vee$ for the collection of all coroots, which is a finite subset of $Y(T)$.

Example. We continue with our example of $\mathrm{GL}_3(\mathbb{K})$. In this case it is easy to see the validity of proposition 3.5 just by inspection. For each root $\alpha \in \Phi$ we are looking for a homomorphism $\alpha^\vee \in Y(T)$ such that $\langle \alpha, \alpha^\vee \rangle = 2$. It is easy to verify this condition holds for the following elements of $Y(T)$ and hence these are the coroots of the roots given previously.

$$\alpha^\vee(\lambda) = \begin{bmatrix} \lambda & & \\ & \lambda^{-1} & \\ & & 1 \end{bmatrix} \qquad \beta^\vee(\lambda) = \begin{bmatrix} 1 & & \\ & \lambda & \\ & & \lambda^{-1} \end{bmatrix} \qquad (\alpha+\beta)^\vee(\lambda) = \begin{bmatrix} \lambda & & \\ & 1 & \\ & & \lambda^{-1} \end{bmatrix}$$

Remark. A word of caution. In the above example we have $(\alpha + \beta)^\vee = \alpha^\vee + \beta^\vee$ but this is *not true* in general. In fact in $\mathrm{Sp}_4(\mathbb{K})$ such a property does not hold.

Now to each root $\alpha \in \Phi$ we can associate an element of the Weyl group in the following way. Recall the map $\varphi : \mathrm{SL}_2(\mathbb{K}) \to \langle U_\alpha, U_{-\alpha} \rangle$ then we have the element

$$n_\alpha = \varphi \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \langle U_\alpha, U_{-\alpha} \rangle$$

is such that $n_\alpha \in N_G(T)$, which follows from the homomorphism property of φ. We obtain an element of W by considering the image of n_α under the canonical projection. Recall that W acts on the character group and cocharacter group of T and indeed the action of this element is given by

$$\begin{aligned} \dot{n}_\alpha(\chi) &= \chi - \langle \chi, \alpha^\vee \rangle \alpha && \text{for all } \chi \in X(T), \\ \dot{n}_\alpha(\gamma) &= \gamma - \langle \alpha, \gamma \rangle \alpha^\vee && \text{for all } \gamma \in Y(T). \end{aligned}$$

With all this construction in mind we can now give the definition of a root datum.

Definition. A quadruple $(X, \Phi, Y, \Phi^\vee)$ is called a *root datum* if the following conditions are satisfied.

(a) X and Y are free abelian groups of the same finite rank with a non-degenerate map $X \times Y \to \mathbb{Z}$ denoted by $(\chi, \gamma) \mapsto \langle \chi, \gamma \rangle$ which puts them into duality.
(b) Φ and $\Phi^\vee$ are finite subsets of X and Y respectively and there is a bijection $\alpha \mapsto \alpha^\vee$ from Φ to $\Phi^\vee$, such that $\langle \alpha, \alpha^\vee \rangle = 2$.
(c) For each $\alpha \in \Phi$ the maps $w_\alpha : X \to X$ and $w_{\alpha^\vee} : Y \to Y$ defined by

$$\begin{aligned} w_\alpha(\chi) &= \chi - \langle \chi, \alpha^\vee \rangle \alpha && \text{for all } \chi \in X(T), \\ w_\alpha(\gamma) &= \gamma - \langle \alpha, \gamma \rangle \alpha^\vee && \text{for all } \gamma \in Y(T). \end{aligned}$$

satisfy $w_\alpha(\Phi) = \Phi$ and $w_{\alpha^\vee}(\Phi^\vee) = \Phi^\vee$.

If G is a connected reductive linear algebraic group and $T \leqslant G$ is a maximal torus then the quadruple $(X(T), \Phi, Y(T), \Phi^\vee)$ as given above is a root datum. Note that everything we have done so far seems to depend upon a choice of maximal torus and Borel subgroup but theorem 3.6 assures us that this choice is made only up to conjugation. Therefore we will usually just write X and Y as the character and cocharacter groups unless we need to specifically refer to a maximal torus. We end this section by stating a theorem which shows why the root datum is of use.

Theorem 3.7. *Let G be a connected reductive linear algebraic group and $(X, \Phi, Y, \Phi^\vee)$ be a root datum for G. Then we have $G = \langle T, U_\alpha \mid \alpha \in \Phi \rangle$. If G is semisimple then $G = \langle U_\alpha \mid \alpha \in \Phi \rangle$.*

Proof. See [Hum75, Theorem 27.3 and 27.5]. □

3.5. The Classification of Simple Linear Algebraic Groups.

Definition. Let G be a linear algebraic group. We say G is *simple* if G contains no proper closed connected normal subgroups.

Remark. Note that this definition means that a simple linear algebraic group may not be simple as an abstract group. Now $\mathrm{SL}_n(\mathbb{K})$ is not always simple as an abstract group because it can have a non-trivial centre. However this centre is finite, and hence disconnected, and so $\mathrm{SL}_n(\mathbb{K})$ is always a simple linear algebraic group. We also comment that any simple algebraic group must be semisimple as the radical is a connected normal subgroup.

From what we have constructed in the previous section it may seem like there is a glimmer of a root system in our root datum but it is possibly not clear exactly why. Let G be a linear algebraic group with root datum $(X, \Phi, Y, \Phi^\vee)$, then we have X is a $\mathbb{Z}$-module. Indeed $\mathbb{R}$ is

Table 4. The Simple Linear Algebraic Groups.

Type	A_n	B_n	C_n	D_n	E_6	E_7	E_8	F_4	G_2
$\lvert\pi\rvert$	$n+1$	2	2	4	3	2	1	1	1
G_{sc}	$\mathrm{SL}_{n+1}(\mathbb{K})$	$\mathrm{Spin}_{2n+1}(\mathbb{K})$	$\mathrm{Sp}_{2n}(\mathbb{K})$	$\mathrm{Spin}_{2n}(\mathbb{K})$	–	–	–	–	–
G_{ad}	$\mathrm{PGL}_{n+1}(\mathbb{K})$	$\mathrm{SO}_{2n+1}(\mathbb{K})$	$\mathrm{PCSp}_{2n}(\mathbb{K})$	$\mathrm{SO}_{2n}(\mathbb{K})$	–	–	–	–	–

also a $\mathbb{Z}$-module and so we can consider the tensor product of these $\mathbb{Z}$-modules $X_{\mathbb{R}} := X \otimes_{\mathbb{Z}} \mathbb{R}$. We can then consider this as a real vector space, likewise we have $Y_{\mathbb{R}} := Y \otimes_{\mathbb{Z}} \mathbb{R}$ is a real vector space.

Above we described how the Weyl group W acts on the character group X and cocharacter group Y. This action can be extended to an action of W on the real vector spaces $X_{\mathbb{R}}$ and $Y_{\mathbb{R}}$ by setting $w(\chi \otimes \lambda) = {}^w\chi \otimes \lambda$ and $w(\gamma \otimes \lambda) = \gamma^w \otimes \lambda$, for all $\chi \in X$, $\gamma \in Y$, $\lambda \in \mathbb{R}$. The subset $\Phi \otimes \mathbb{Z}$, (resp. $\Phi^\vee \otimes \mathbb{Z}$), will form a W-stable lattice in $X_{\mathbb{R}}$, (resp. $Y_{\mathbb{R}}$). Indeed it turns out that $X_{\mathbb{R}}$, (resp. $Y_{\mathbb{R}}$), is a Euclidean vector space and $\Phi \otimes \mathbb{Z}$, (resp. $\Phi^\vee \otimes \mathbb{Z}$), is a crystallographic root system for $X_{\mathbb{R}}$, (resp. $Y_{\mathbb{R}}$).

As every linear algebraic group contains a crystallographic root system for a real vector space we would hope that we could classify the simple linear algebraic groups by the Dynkin diagrams given in figure 6. Indeed given any simple linear algebraic group G we have that there is a unique Dynkin diagram associated to G. However given a Dynkin diagram there can be more than one simple linear algebraic associated to it.

What we would like is a method, purely using the root system, to distinguish between the different simple linear algebraic groups that can arise. There is such a method and it is motivated from the theory of Lie groups.

Recall that to the root datum $(X, \Phi, Y, \Phi^\vee)$ we have an associated perfect pairing $\langle -, - \rangle : X \times Y \to \mathbb{Z}$. We can extend this perfect pairing to the vector spaces $X_{\mathbb{R}}$, $Y_{\mathbb{R}}$ by defining $\langle \chi \otimes \lambda, \gamma \otimes \mu \rangle = \lambda\mu\langle \chi, \gamma \rangle$ for all $\chi \in X$, $\gamma \in Y$ and $\lambda, \mu \in \mathbb{R}$. We then define

$$\Omega = \{\omega \in X_{\mathbb{R}} \mid \langle \omega, \alpha^\vee \rangle \in \mathbb{Z} \text{ for all } \alpha \in \Phi\}$$

to be the *weight lattice* of G, (where $\alpha^\vee$ is associated with its canonical image $\alpha^\vee \otimes 1$ in $Y_{\mathbb{R}}$). Let us denote the root lattice $\Phi \otimes \mathbb{Z}$ by $\mathbb{Z}\Phi$, then associating X with its canonical image in $X_{\mathbb{R}}$ we have a sequence of inclusions $\mathbb{Z}\Phi \subseteq X \subseteq \Omega$. It turns out that the position of X between the root lattice and the weight lattice uniquely determines a simple linear algebraic group.

Definition. Let $(X, \Phi, Y, \Phi^\vee)$ be a root datum for a semisimple algebraic group, then we define $\pi = \Omega/\mathbb{Z}\Phi$ to be the *fundamental group* of the root datum.

It can be shown that the order of the fundamental group is finite and so there are only finitely many simple linear algebraic groups associated to each Dynkin diagram. In each type we single out two specific cases. Let G be a simple linear algebraic group with root datum $(X, \Phi, Y, \Phi^\vee)$. If $X = \mathbb{Z}\Phi$ then we say G is of *adjoint* type and we denote this group by G_{ad}. If $X = \Omega$ then we say G is of *simply connected* type and we denote this group by G_{sc}. In table 4 we give the order of the fundamental group in each type and indicate the simply connected and adjoint groups in the classical cases.

Although we have classified the simple linear algebraic groups we have not shown that problems involving connected semisimple linear algebraic groups can be reduced to the case of simples.

Theorem 3.8. *Let G be a connected semisimple linear algebraic group then there exists finitely many closed normal subgroups, which we label $\{G_i\}_{1\leqslant i\leqslant k}$. Now*

(a) *Each G_i is simple.*
(b) *If $i \neq j$ then $[G_i, G_j] = \{1\}$.*
(c) $G = [G, G]$
(d) $G = G_1 G_2 \dots G_k$
(e) *$G_i \cap G_1 \dots G_{i-1} G_{i+1} \dots G_k$ is finite for each i.*

Proof. See [Hum75, Theorem 27.5]. □

Therefore every connected semisimple linear algebraic group is an almost direct product of simple linear algebraic groups. So we can largely reduce the problem of studying linear algebraic groups to simple linear algebraic groups.

4. Finite Groups of Lie Type

Throughout this section we have $\mathbb{K} = \overline{\mathbb{F}_p}$ for some prime $p > 0$.

In the previous section we have shown how we can obtain a large collection of infinite groups defined over algebraically closed fields of characteristic p. Now given the field $\mathbb{K}$ we have a field automorphism $\sigma_q : \mathbb{K} \to \mathbb{K}$ given by $\sigma(x) = x^q$, for some $q = p^a$. This automorphism is such that the fixed points $\mathbb{K}^{\sigma_q} = \{x \in \mathbb{K} \mid \sigma_q(x) = x\}$ is isomorphic, as a field, to the finite field $\mathbb{F}_q$. What we would like to do is use this field automorphism to produce finite groups from the linear algebraic groups we have constructed previously.

Definition. We define a regular map of the linear algebraic group $\mathrm{GL}_n(\mathbb{K})$ by

$$\begin{aligned} F_q : \mathrm{GL}_n(\mathbb{K}) &\to \mathrm{GL}_n(\mathbb{K}) \\ (a_{ij}) &\mapsto (a_{ij}^q), \end{aligned}$$

A *standard Frobenius* map of a linear algebraic group G is any map $F : G \to G$ such that there exists an embedding $\theta : G \to GL_n(\mathbb{K})$ which satisfies $\theta \circ F = F_q \circ \theta$, for some $q = p^a$. A morphism $F : G \to G$ is called a *Frobenius* map if for some $m \in \mathbb{N}$ we have F^m is a standard Frobenius map.

Remark. This definition of Frobenius map is attributed to Carter, (see [Car93]), and while practical gives quite a simplistic view of the Frobenius map. Indeed for a more sophisticated view point see [DM91, Chapter 3] or [Gec03, Section 4.1] .

Lemma 4.1. *Let G be a connected reductive linear algebraic group and $F : G \to G$ a Frobenius map. Then the fixed point group $G^F = \{g \in G \mid F(g) = g\}$ is finite.*

Proof. Consider the standard Frobenius map F_q of $\mathrm{GL}_n(\mathbb{K})$. We clearly have $\mathrm{GL}_n(\mathbb{K})^{F_q} = \mathrm{GL}_n(\mathbb{F}_q)$. In general given any closed subgroup $G \leqslant \mathrm{GL}_n(\mathbb{K})$ we will have $G^{F_q} \leqslant \mathrm{GL}_n(\mathbb{F}_q)$. Therefore if G is a linear algebraic group with standard Frobenius map $F : G \to G$ we have $G^{\theta \circ F} = G^{F_q \circ \theta} \leqslant \mathrm{GL}_n(\mathbb{F}_q)$. Therefore $G^{\theta \circ F}$ is finite, which gives us G^F is finite. The result is then clear. □

We now consider the simple linear algebraic groups that we constructed in the previous section over $\mathbb{K}$. The above suggests that by considering fixed points of Frobenius maps we will obtain finite groups. Indeed, if G is a connected reductive linear algebraic group and $F : G \to G$ a Frobenius map then we call the finite group G^F a *reductive group.*

However what we usually refer to are the *finite groups of Lie type.* These groups aren't in general fixed points of Frobenius maps but are obtained by taking a quotient of a reductive group by a subgroup of its centre. There are four classes of reductive groups, which are the *Chevalley groups*, *Steinberg groups*, *Ree groups* and *Suszuki group.* The Steinberg, Ree and Suzuki groups are usually refered to as the *twisted groups*, the reason for this will become clear shortly. By considering the fixed points of the simple linear algebraic groups in table 4, under the standard Frobenius maps, we obtain what are known as the Chevalley groups.

Example. Consider the finite groups $\mathrm{GL}_n(q)$, $\mathrm{SL}_n(q)$ and $\mathrm{PGL}_n(q)$ defined over the field $\mathbb{F}_q$. These are all reductive groups of type A_{n-1} as they are the fixed point groups of the Frobenius map F_q acting on $\mathrm{GL}_n(\mathbb{K})$, $\mathrm{SL}_n(\mathbb{K})$ and $\mathrm{PGL}_n(\mathbb{K})$. Now consider the group $\mathrm{PSL}_n(q)$, this is not a reductive group. This is because if you try and construct $\mathrm{PSL}_n(\mathbb{K})$ over an algebraically closed field, what you get is $\mathrm{PGL}_n(\mathbb{K})$. However we refer to $\mathrm{PSL}_n(q)$ as a finite group of Lie type.

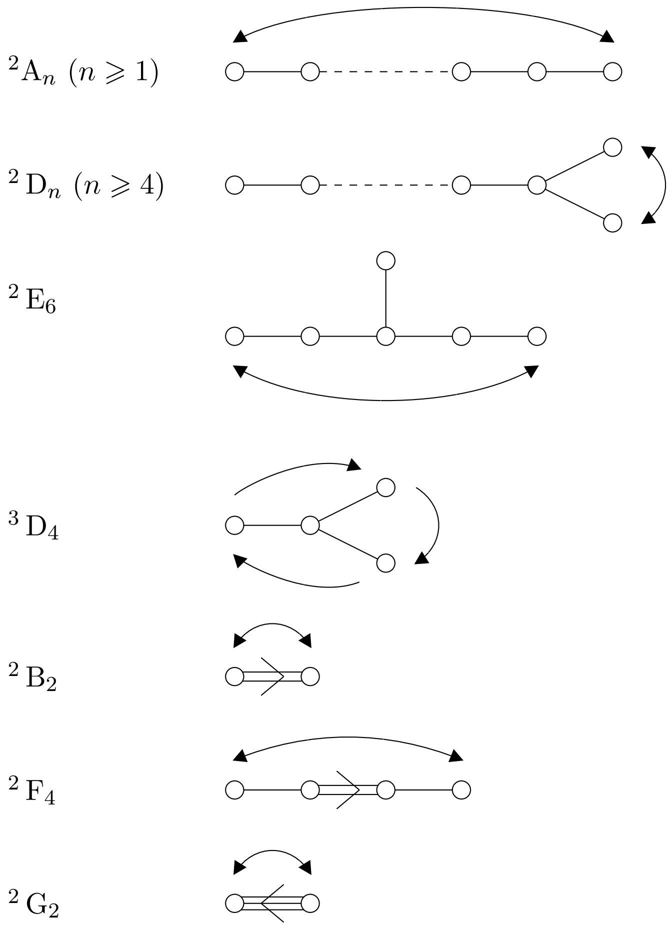


FIGURE 8. The Possible Graph Automorphisms of Dynkin Diagrams.

In fact Chevalley originally showed that to every complex simple Lie algebra there exists an infinite family of finite simple groups defined over arbitrary q, (modulo four exceptions when $q = 2$ or 3). He does this by using the root system of the Lie algebra to construct a special basis, known as a *Chevalley basis*, which he transports to the structure of a group using the exponential map. An example of such a finite simple group is $\mathrm{PSL}_n(q)$ and we consider these groups to be finite simple groups of Lie type. For more information on this see [Car72, Chapter 4].

In fact all the simple groups that Chevalley obtained can be obtained in the following way. Take a simple linear algebraic group of simply connected type and consider the finite fixed point group under a standard Frobenius map. This group in general will not be a finite simple group, (for example $\mathrm{SL}_n(q)$), but the quotient of this group by its centre is in general simple, (for example $\mathrm{PSL}_n(q)$).

At the moment we have only discussed the Chevalley groups but there are many more reductive groups. We recall that in the definition of Frobenius map we allowed for some power of the map to be a standard Frobenius map. This means, if G is a connected reductive linear algebraic group, we can take a standard Frobenius map and compose it with an automorphism

Table 5. The Twisted Reductive Groups.

Type	${}^2\mathrm{A}_n$	${}^2\mathrm{D}_n$	${}^2\mathrm{E}_6$	${}^3\mathrm{D}_4$	${}^2\mathrm{B}_2$	${}^2\mathrm{F}_4$	${}^2\mathrm{G}_2$
p	—	—	—	—	2	2	3
G_{sc}	$\mathrm{SU}_{n+1}(q^2)$	$\mathrm{Spin}_{2n}^-(q)$	—	—	—	—	—
G_{ad}	$\mathrm{PU}_{n+1}(q^2)$	$\mathrm{P}(\mathrm{CO}_{2n}^-(q)^0)$	—	—	—	—	—

of G, (which may only be an automorphism of the abstract group structure), to be obtain a new Frobenius map. We give now some different kinds of automorphisms of linear algebraic groups, and indicate when these are not automorphisms of varieties.

Inner: These are automorphisms given by conjugation, which are clearly regular maps as multiplication and inversion are regular maps.

Field: We have already discussed these previously. Let σ be an automorphism of the field $\mathbb{K}$ then we can extend this to an automorphism of the whole linear algebraic group. Note that these automorphisms are not automorphisms of varieties, as we have seen in the case of $F_q : \mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$.

Graph: This is an automorphism of the Dynkin diagram. We give a list of all possible graph automorphisms in figure 8. In the cases of A_n, D_n, E_6 and D_4 we have these graph automorphisms are automorphisms of varieties. However in the cases of B_2, F_4 and G_2 we have they are not. The distinction is that in the latter three cases we have the long and short roots are interchanged.

It can be shown that every Frobenius map is a composition of a standard Frobenius map and one of the graph automorphisms given in figure 8. The groups ${}^2\mathrm{A}_n$, ${}^2\mathrm{D}_n$, ${}^2\mathrm{E}_6$ and ${}^3\mathrm{D}_4$ are known as the Steinberg groups. The groups ${}^2\mathrm{F}_4$, ${}^2\mathrm{G}_2$ are known as the Ree groups and ${}^2\mathrm{B}_2$ is called the Suzuki group, (note that there are two sporadic simple groups referred to as Suzuki groups as well). We give partial information for these groups in table 5, namely associations with classical groups and restrictions on primes for which they are defined. For much more detailed information see [Car72] or [Car93, Section 1.19].

Example. We wish to indicate how the group $\mathrm{U}_3(q^2)$ is constructed from the group $\mathrm{GL}_3(\mathbb{K})$. We have already seen that $\mathrm{GL}_3(\mathbb{K})$ has a root system of type A_2, given relative to the maximal torus of diagonal matrices. Therefore we are looking to implement a graph automorphism τ such that $\tau(\alpha) = \beta$, where α and β are the two simple roots of $\mathrm{GL}_3(\mathbb{K})$.

We realise the graph automorphism τ in the following way. Recall the Weyl group $W \cong \mathfrak{S}_3$ of $\mathrm{GL}_3(\mathbb{K})$, then by proposition 1.2 there exists a unique element $w_0 \in W$ of longest length. This element has a representative in $N_{\mathrm{GL}_3(\mathbb{K})}(T)$ given by

$$\dot{w}_0 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Let $\tau : \mathrm{GL}_3(\mathbb{K}) \to \mathrm{GL}_3(\mathbb{K})$ be the map given by $\tau(x) = (x^{-t})^{\dot{w}_0}$, where x^{-t} denotes the inverse transpose of x and $x^{\dot{w}_0}$ denotes conjugation. If s_1, s_2 are abstract generators associated to the simple roots α, β of A_2 then we have $w_0 = s_1 s_2 s_1$. It's then easy to check using the diagram in corollary 1.4 that w_0 is such that ${}^{w_0}\alpha = -\beta$ and ${}^{w_0}\beta = -\alpha$.

Consider the following element of $\mathrm{GL}_3(\mathbb{K})$

$$t = \begin{bmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{bmatrix},$$

in the maximal torus $D_3 \leqslant \mathrm{GL}_3(\mathbb{K})$ for some $t_1, t_2, t_3 \in \mathbb{K}^\times$. Now what does it mean for t to be a fixed point of the Frobenius map $F = F_q \circ \tau$?

$$\begin{aligned} t = (F_q \circ \tau)(t) \Rightarrow \begin{bmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{bmatrix} &= F_q \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} t_1^{-1} & & \\ & t_2^{-1} & \\ & & t_3^{-1} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \\ \Rightarrow \begin{bmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{bmatrix} &= F_q \begin{bmatrix} t_3^{-1} & & \\ & t_2^{-1} & \\ & & t_1^{-1} \end{bmatrix}, \\ \Rightarrow \begin{bmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{bmatrix} &= \begin{bmatrix} t_3^{-q} & & \\ & t_2^{-q} & \\ & & t_1^{-q} \end{bmatrix}. \end{aligned}$$

So for t to be in the fixed point group of F we must have $t_1 = t_3^{-q}$, $t_2 = t_2^{-q}$ and $t_3 = t_1^{-q}$, which implies $t_1 = t_1^{q^2}$, $t_3 = t_3^{q^2}$ and $t_2^{q+1} = 1$. Therefore we have $t_1, t_3 \in \mathbb{F}_{q^2}^\times$ and $t_2 \in \mathbb{F}_{q^2}^\times \setminus \mathbb{F}_q^\times$, so the entries in these matrices will belong to the finite field $\mathbb{F}_{q^2}$.

In general the approach is to try and understand the reductive groups by obtaining information from the linear algebraic group through the Frobenius map. This allows us to obtain a lot of structural information from the linear algebraic group. For example, things like the Weyl group and root datum exist in the reductive groups. Also we often pass information regarding subgroups to the reductive groups. For example if G is a connected reductive linear algebraic group and T is a maximal torus of G we say the fixed point group T^F is a maximal torus of G^F. Now in general this will just be an abelian subgroup of G^F and in fact all such subgroups are not necessarily conjugate in G^F.

We can in fact study connected reductive linear algebraic groups and their associated finite groups by studying groups with a *BN-pair* or *Tits system*. It was Tits that introduced this notion. In a connected reductive linear algebraic group we have that B and N refer to a Borel subgroup and the normaliser of a maximal torus contained in B. For more information on BN-pairs see [Gec03, Section 1.6].

5. Bibliographic Remarks & Acknowledgments

The author makes the following bibliographic acknowledgments. Section 1 mostly follows the treatment given in [Hum90, Chapter 1 and 2] with some occasional proofs and odd things taken from [Car05]. Section 2 for the most part follows the treatment given by Humphreys in [Hum78] and Erdmann and Wildon given in [EW06]. Section 3.1 to 3.3 follows, almost to the letter, the format of [Gec03]. Section 3.4 and 3.5 follow the exposition given by Carter in [Car98] or [Car93]. Section 4 follows Carters summary of the material given in [Car93, Chapter 1].

I would like to thank Sejong Park for all his invaluable help in spotting errors in these notes.

References

[Bou02] Nicolas Bourbaki. *Lie Groups and Lie Algebras. Chapters 4–6.* Elements of Mathematics (Berlin). Springer-Verlag, 2002.

[BR02] Thomas S. Blyth and Edmund F. Robertson. *Further Linear Algebra.* Springer Undegraduate Mathematics Series. Springer-Verlag, 2002.

[Car72] Roger W. Carter. *Simple Groups of Lie Type*, volume 28 of *Pure and Applied Mathematics.* John Wiley & Sons, 1972.

[Car93] Roger W. Carter. *Finite Groups of Lie Type: Conjugacy Classes and Complex Characters.* John Wiley & Sons, Wiley Classics Library edition, 1993.

[Car98] Roger W. Carter. Introduction to Algebraic Groups and Lie Algebras. In Roger W. Carter and Meinolf Geck, editors, *Representations of Reductive Groups*, Publications of the Newton Institute, pages 1 – 20. Cambridge University Press, 1998.

[Car05] Roger W. Carter. *Lie Algebras of Finite and Affine Type.* Number 96 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2005.

[DM91] François Digne and Jean Michel. *Representations of Finite Groups of Lie Type.* Number 21 in London Mathematical Society Student Texts. Cambridge University Press, 1991.

[EW06] Karin Erdmann and Mark J. Wildon. *Introduction to Lie Algebras.* Springer Undegraduate Mathematics Series. Springer-Verlag, 2006.

[Gec03] Meinolf Geck. *An Introduction to Algebraic Geometry and Algebraic Groups.* Number 10 in Oxford Graduate Texts in Mathematics. Oxford University Press, first edition, 2003.

[GHL+96] M. Geck, G. Hiss, F. Lübeck, G. Malle, and G. Pfeiffer. CHEVIE – A system for computing and processing generic character tables for finite groups of Lie type, Weyl groups and Hecke algebras. *Appl. Algebra Engrg. Comm. Comput.*, 7:175–210, 1996.

[Hum75] James E. Humphreys. *Linear Algebraic Groups.* Number 21 in Graduate Texts in Mathematics. Springer-Verlag, 1975.

[Hum78] James E. Humphreys. *Introduction To Lie algebras and Representation Theory.* Number 9 in Graduate Texts in Mathematics. Springer-Verlag, 1978.

[Hum90] James E. Humphreys. *Reflection Groups and Coxeter Groups.* Number 29 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1990.

Institute of Mathematics, University of Aberdeen, King's College, Scotland, AB24 3UE
E-mail address: j.taylor@abdn.ac.uk
URL: http://www.abdn.ac.uk/~r01jmt8/