Undergraduate Thesis - Bachelor in Mathematics

An Introduction to BGG Category $\mathcal{O}$

David José de Melo Júnior

Universidade Federal do ABC

Title: An Introduction to BGG Category $\mathcal{O}$

Author: David José de Melo Júnior
Advisor: Dr. Luis Enrique Ramirez

Dissertation presented as a partial requisite for obtaining the title of Bachelor in Mathematics for the Federal University of ABC.

Examination Board:

Zhanna Gennadyevna Kuznetsova, PhD
Federal University of ABC

Gérman Alonso Benitez Monsalve, PhD
Federal University of Amazonas

Santo André, August 2021.

Contents

Acknowledgements

I would like to express my gratitude towards the Federal University of ABC (PROAP and Monitoring) and CNPq (PICME) for their funding who allowed me to finish my bachelors in the first place, without these funding projects I would not be able to stay in Santo André and finish my graduation.

I would also like to thank all of those who contributed to my academic development, especially my professors at the Federal University of ABC, with emphasis on Prof. Armando Caputi and Prof. Luis Enrique Ramirez who were my advisors on research projects that helped me immensely through my graduation, thank you for your patience, time, and the opportunities you provided me.

In addition, I would like to thank my family and friends who allowed me to finish my bachelors, helping me provide for myself during my stay in Santo André, tiping me into the workings of university life and for raising me into the person I am today.

Finally, I would like to thank my friends Nicholas St-Pierre and Christina Baker for their help with honing my English skills, they provided me with stimulating discussions as well as happy distractions to rest my mind on these pandemic times. Without them I would not be able to finish this work.

Abstract

This project aims to show initial results for representation theory on Lie algebras. In particular, we focus on tools necessary to introduce Category $\mathcal{O}$.

As a first step, we develop an introduction to the study of Lie algebras: theorems on solvable and nilpotent Lie algebras, the Cartan-Killing form, Cartan's Criteria and Weyl's theorem of complete reducibility.

After this introduction we start studying an analogue to Jordan decomposition on semi-simple Lie algebras in order to achieve an equivalent to Cartan subalgebras, the so called maximal toral subalgebras, with the aim of achieving root decomposition on semi-simple Lie algebras.

Some properties of root systems and the Weyl group are presented in an independent manner. The construction of a basis for root systems the action of the Weyl group on such bases as well as the classification of irreducible root systems.

We introduce the universal enveloping algebra as a tool to study representations on Lie algebras: The Poincaré-Birkhoff-Witt theorem and a first application on the study of highest weight modules.

Finally, we produce some initial results on category $\mathcal{O}$. Categorical properties of $\mathcal{O}$, some of its more important objects, and the classification of its irreducible objects as quotients of Verma modules.

Keywords: abstract-algebra, Lie-algebras, representation-theory

Introduction

Lie algebras were introduced to study the concept of infinitesimal transformation by Sophus Lie and Wilhelm Killing at the end of the 1800s, since then, the study of Lie algebras has developed into its own field, with results used throughout mathematics and physics, notably in quantum mechanics and particle physics.

This project serves as an introduction to all the main basic results present in Lie algebras with an emphasis on representation theory, serving as a gateway for further development on the field. More specifically, we want to produce results so as to reach a shallow understanding of BGG category $\mathcal{O}$ in the following way:

In the first chapter, we will present the construction of Lie algebras from an algebraic perspective, the structure of nilpotent, solvable and semi-simple Lie algebras as well as some properties on the representations of semi-simple Lie algebras, using [Hum72, chpt 1-6] as the main reference.

In the second chapter, we start studying the abstract Jordan decomposition (an analogue to the Jordan decomposition for operators in finite-dimensional vector spaces) as well as the study of maximal toral subalgebras (an equivalent structure to Cartan subalgebras) in a similar manner to [Hum72, chpt 5-8]. This study allows us to understand root decomposition, an essential part of the study of semi-simple Lie algebras.

In the third chapter, we will cover some important aspects of root systems and the Weyl group in an independent manner, similar to [Hum72, chpt 9-13]. These results are widely used in the classification of simple Lie algebras as well as in representation theory.

In the fourth chapter, we will cover the universal enveloping algebra, a fundamental tool for the study of representations. As a first application of them, we will present results for highest weight modules using as reference [Mar09, chpt 10].

In the fifth chapter, we shall introduce the Category $\mathcal{O}$ of representations, one of the most important categories of representations, using as main reference "Representations of Semisimple Lie algebras in BGG category $\mathcal{O}$" [Hum08]. We will present some of its categorical properties as well as a classification of its irreducible objects.

1 Lie Algebras: Basics

Lie theory is a huge field in modern mathematics with applications in quantum mechanics, particle physics and to the theory of differential equations, one can refer to [JGBS66] for some of these applications.

We are going to focus this dissertation on the theory of Lie algebras in an independent manner, one can refer to Appendix C for a simplified construction of Lie algebras from Lie groups. If the reader is unfamiliar with the concept of algebras or some of their basic properties, we refer them to Appendix A.

In this chapter, we are going to present the basic results on the study of Lie algebras, with a focus on criteria to discern when algebras are nilpotent, solvable or semi-simple as well as a big result on representation theory, *Weyl's theorem of complete reducibility*.

1.1. Lie Algebras: Axiomatic

This section is focused on defining the main language and tools we are going to use throughout the text, following standard modern notation seen in [Mar09] with proofs also based on [Hum72, chpt 1&2].

Fix an arbitrary field $\mathbb{F}$.

Definition 1.1.1 *A **Lie algebra** is an $\mathbb{F}$-algebra $\mathfrak{g}$ with a product (represented by $[\ ,\ ]$) satisfying, for all X, Y and $Z \in \mathfrak{g}$:*

(a) $[X, X] = 0$ *(Alternativity).*

(b) $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ *(Jacobi Identity).*

*A **subalgebra** is a subspace $\mathfrak{h}$ of $\mathfrak{g}$ such that for all $X, Y \in \mathfrak{h}$ we have $[X, Y] \in \mathfrak{h}$.*
*An **ideal** is a subspace $\mathfrak{h}$ of $\mathfrak{g}$ such that for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$ we have $[X, Y] \in \mathfrak{h}$.*

If $\mathbb{F}$ has characteristic different from 2 we can do better and require only anti-commutativity as an equivalent axiom to alternativity: (a') $[X, Y] = -[Y, X]$.

An important point to make is that Lie algebras are not generally associative, and that the only ones that are commutative (char $\mathbb{F} \neq 2$) are trivial in the sense that the bracket of any two elements is always 0, $[X,Y] = -[Y,X] = -[X,Y] \Rightarrow [X,Y] = 0$.

Let End(V) denote the space of linear transformations from V onto itself (also called **endomorphisms**), define $[X,Y] = X \cdot Y - Y \cdot X$ as the commutator, where $\cdot$ is the product given by composition of morphisms, this space with the defined bracket is a Lie algebra:

(a) $[X,X] = X^2 - X^2 = 0$.

(b) $[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]] = (XYZ-XZY-YZX+ZYX)+(YZX-YXZ-ZXY+XZY)+(ZXY-ZYX-XYZ+YXZ) = 0$.

The set of all $n \times n$ matrices over any field $\mathbb{F}$ is a Lie algebra with the commutator, denoted by $\mathfrak{gl}(n,\mathbb{F})$. With respect to the canonical basis $\{e_{ij} : 1 \leq i,j \leq n\}$ of $\mathfrak{gl}(n,\mathbb{F})$ the Lie bracket has an explicit form: $[e_{ij}, e_{kl}] = \delta_{jk} e_{il} - \delta_{il} e_{kj}$.

Note that there is no restriction on the dimension of V for End(V) to be a Lie algebra. We can extend this to V of arbitrary dimension, denoting the Lie algebra End(V) with product given by the commutator by $\mathfrak{gl}(\mathbf{V})$.

Example 1.1.2 *Matrix algebras:*

1. ***Special linear algebra:*** $\mathfrak{sl}(n,\mathbb{F}) = \{X \in \mathfrak{gl}(n,\mathbb{F}) \mid Tr(X) = 0\}$.
   $\mathfrak{sl}(n,\mathbb{F})$ *is a vector subspace of* $\mathfrak{gl}(n,\mathbb{F})$ *as the kernel of* $Tr : \mathfrak{gl}(n,\mathbb{F}) \to \mathbb{F}$ *with* $\dim \mathfrak{sl}(n,\mathbb{F}) = n^2 - 1$.
   It is also closed under the Lie bracket:

   $$Tr([X,Y]) = Tr(XY - YX) = Tr(XY) - Tr(YX) = 0,$$

   it follows that $\mathfrak{sl}(n,\mathbb{F})$ *is a Lie subalgebra of* $\mathfrak{gl}(n,\mathbb{F})$

2. *If* $B : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ *is a bilinear form, then the subspace:*

   $$\mathfrak{o}_B(n,\mathbb{F}) = \{X \in \mathfrak{gl}(n,\mathbb{F}) \,|B(Xv,w) + B(v,Xw) = 0 \textit{ for all } v,w \in \mathbb{F}^n\}$$

   is a Lie subalgebra of $\mathfrak{gl}(n)$.
   We shall check that it is closed under the Lie bracket (it is trivially a subspace):

   $$\begin{aligned} B((XY-YX)v,w) &= B(XYv,w) - B(YXv,w) = -B(Yv,Xw) + B(Xv,Yw) \\ &= B(v,YXw) - B(v,XYw) = -B(v,(XY-YX)w). \end{aligned}$$

We shall denote the matrix of this bilinear form with respect to the canonical basis with the same name as the bilinear form itself, that is: $B(v,w) = v^T Bw$.
Doing this allows us to simply write $\mathfrak{o}_B(n,\mathbb{F}) = \{X \in \mathfrak{gl}(n,\mathbb{F}) \mid X^T B + BX = 0\}$.

3. ***Orthogonal algebra**: A notable bilinear form is the inner product, whose matrix is the identity:*
$$\mathfrak{o}(n,\mathbb{F}) = \mathfrak{o}_I(n,\mathbb{F}) = \{X \in \mathfrak{gl}(n,\mathbb{F}) \mid X^T = -X\}$$
An easy calculation, see for instance [Hum72, p.3]*, shows that the dimension of this space depends on the parity of n:*
$$\dim \mathfrak{o}(2n,\mathbb{F}) = 2n^2 - n, \qquad \dim \mathfrak{o}(2n+1,\mathbb{F}) = 2n^2 + n.$$

4. ***Symplectic algebra**: For algebras with even dimension, another common bilinear form is the one defined by the matrix*
$$J = \begin{pmatrix} 0_{n\times n} & I_{n\times n} \\ -I_{n\times n} & 0_{n\times n} \end{pmatrix}.$$
The algebra induced by this bilinear form is denoted by $\mathfrak{sp}(2n,\mathbb{F})$ *and can be written explicitly as:*
$$\mathfrak{sp}(2n,\mathbb{F}) = \mathfrak{o}_J(2n,\mathbb{F}) = \{X \in \mathfrak{gl}(n,\mathbb{F}) \mid X^T J + JX = 0\}.$$
Another simple calculation on matrix blocks can be used to show that the dimension of $\mathfrak{sp}(2n,\mathbb{F})$ *is equal to* $2n^2 + n$*, see* [Hum72, p.3].

5. *The **upper triangular** and **strictly upper triangular** matrices are a Lie algebra. In fact, for the space spanned by* $\{e_{ij} \mid 1 \le i < j \le n\}$*:*
$$[e_{ij}, e_{kl}] = \delta_{jk} e_{il} - \delta_{il} e_{kj}.$$
$\delta_{jk} e_{il} \neq 0$ *only if* $j = k$*, but that implies* $i < j = k < l$*. Analogously* $\delta_{il} e_{kl} \neq 0$ *implies that* $k < j$*. The same is valid for the space spanned by* $\{e_{ij} \mid i \le j\}$.

If $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, then $\mathfrak{gl}(n,\mathbb{F})$, $\mathfrak{sl}(n,\mathbb{F})$, $\mathfrak{o}(n,\mathbb{F})$ and $\mathfrak{sp}(2n,\mathbb{F})$ are the Lie algebras of the classic Lie groups with the same name, these being, respectively, the general linear group, the special linear group, the orthogonal group, and the symplectic group. See, for instance [Hal04, p.39-41].

Example 1.1.3 *Trivial and low-dimensional algebras:*

1. *For any vector space V, the product $[X,Y]=0$ for any $X,Y\in V$ defines a Lie algebra structure on V, which is called the **trivial** or **abelian** Lie algebra on V.*

2. *If $\mathfrak{g}$ is a one-dimensional Lie algebra, through alternativity, it is abelian.*

3. *For non-trivial $\mathfrak{g}$ two-dimensional, there is a basis $\{X,Y\}$ in $\mathfrak{g}$ such that $[X,Y]=Y$. Let $\{X_0,Y_0\}$ be a basis of this two-dimensional Lie Algebra, and $[X_0,Y_0]=aX_0+bY_0$. Assuming $b\neq 0$ we find:*

$$\left[\frac{1}{b}X_0,\frac{a}{b}X_0+Y_0\right]=\frac{1}{b}[X_0,Y_0]=\frac{a}{b}X_0+Y_0,$$

then $X=\frac{a}{b}X_0$ and $Y=\frac{1}{b}X_0+Y_0$ satisfy the condition. If $b=0$, as the algebra is non-trivial, $a\neq 0$, and:

$$\left[-\frac{1}{a}Y_0,X_0\right]=X_0,$$

therefore, $X=-\frac{1}{a}Y_0$ and $Y=X_0$ satisfy the condition.

Definition 1.1.4 *Let $\mathfrak{g}$ be any Lie algebra:*

- *A **derivation** of $\mathfrak{g}$ is a linear transformation $D\in\mathfrak{gl}(\mathfrak{g})$ satisfying:*

$$D[X,Y]=[DX,Y]+[X,DY]\text{ for all } X,Y\in\mathfrak{g}.$$

- *The **adjoint** of an element $X\in\mathfrak{g}$ is defined as $ad(X)\in\mathfrak{gl}(\mathfrak{g})$ by:*

$$ad(X)Y=[X,Y],\ Y\in\mathfrak{g}.$$

- *A **representation** of $\mathfrak{g}$ in a vector space V is a Lie algebra morphism ρ from $\mathfrak{g}$ to $\mathfrak{gl}(V)$, meaning:*

$$\rho([X,Y])=\rho(X)\rho(Y)-\rho(Y)\rho(X).$$

Proposition 1.1.5 *The adjoint of an element is a derivation and the adjoint representation $ad:X\mapsto ad(X)$ is a representation from $\mathfrak{g}$ into itself.*

Proof: Using the Jacobi identity:

$$\begin{aligned}\mathrm{ad}(X)[Y,Z]&=[X,[Y,Z]]=-[Y,[Z,X]]-[Z,[X,Y]]\\&=[Y,[X,Z]]+[[X,Y],Z]=[Y,\mathrm{ad}(X)Z]+[\mathrm{ad}(X)Y,Z].\end{aligned}$$

$$\begin{aligned}\mathrm{ad}([X,Y])Z &= [[X,Y],Z] = [X,[Y,Z]] + [Y,[Z,X]] \\ &= [X,[Y,Z]] - [Y,[X,Z]] = \mathrm{ad}(X)\mathrm{ad}(Y)Z - \mathrm{ad}(Y)\mathrm{ad}(X)Z \quad \text{for all } Z \in \mathfrak{g}.\end{aligned}$$

□

With these tools defined, we can now start to deepen our study of the structure of Lie algebras. For this purpose, we will define two series of ideals that are indispensable to the definition of the classic types of Lie algebras. Before that, some clarification on notation:

Given subsets $\mathfrak{h}$ and $\mathfrak{k}$ of a Lie algebra $\mathfrak{g}$, let $[\mathfrak{h},\mathfrak{k}] := \{[H,K] \in \mathfrak{g} \mid H \in \mathfrak{h}, K \in \mathfrak{k}\}$. Note that if $\mathfrak{g}$ and $\mathfrak{k}$ are subspaces then so is $[\mathfrak{h},\mathfrak{k}]$, furthermore, $[\mathfrak{h},\mathfrak{k}] = [\mathfrak{k},\mathfrak{h}]$.

Definition 1.1.6 *The **lower central series** is defined recursively as:*

$$\begin{cases}\mathfrak{g}^1 = \mathfrak{g}, \\ \mathfrak{g}^n = [\mathfrak{g}, \mathfrak{g}^{n-1}].\end{cases}$$

*The **derived series** is also defined recursively:*

$$\begin{cases}\mathfrak{g}^{(1)} = \mathfrak{g}, \\ \mathfrak{g}^{(n)} = [\mathfrak{g}^{(n-1)}, \mathfrak{g}^{(n-1)}].\end{cases}$$

Proposition 1.1.7 *Let $\mathfrak{g}$ be a Lie algebra, then:*

(a) $\mathfrak{g} \supset \mathfrak{g}^2 \supset \cdots \supset \mathfrak{g}^n \supset \cdots$ *, and* $\mathfrak{g} \supset \mathfrak{g}^{(2)} \supset \cdots \supset \mathfrak{g}^{(n)} \supset \cdots$.

(b) All members of these series are ideals of $\mathfrak{g}$.

(c) For all n, $\mathfrak{g}^{(n)} \subset \mathfrak{g}^n$.

Proof:

(a) For the case of the lower central series, we shall consider $n > 2$ (the case when $n = 2$ follows directly from $\mathfrak{g}$ being a Lie algebra). Let $X \in \mathfrak{g}^{n+1}$ with $X = [Y,Z]$ for $Y \in \mathfrak{g}^n$ and $Z \in \mathfrak{g}$. By induction $Y \in \mathfrak{g}^{n-1}$, therefore $X \in [\mathfrak{g}^{n-1}, \mathfrak{g}] = \mathfrak{g}^n$.
For the case of the derived series, we also proceed by induction, the case $\mathfrak{g}^{(2)} = [\mathfrak{g},\mathfrak{g}] \subset \mathfrak{g}$ is direct, assuming $\mathfrak{g}^{(n)} \subset \mathfrak{g}^{(n+1)}$ for $X \in \mathfrak{g}^{(n+1)}$ with $X = [Y,Z]$ for $Y,Z \in \mathfrak{g}^{(n)}$, then since $Y,Z \in \mathfrak{g}^{(n-1)}$, the result follows.

(b) For the lower central series, the result is direct from the previous result, $[\mathfrak{g},\mathfrak{g}^n] = \mathfrak{g}^{n+1} \subset \mathfrak{g}^n$. For the derived series, we will use induction. The base case $\mathfrak{g}^{(1)} = \mathfrak{g}$ is trivially an ideal of $\mathfrak{g}$. Assume that $\mathfrak{g}^{(n)}$ is an ideal of $\mathfrak{g}$, let $X \in \mathfrak{g}$, and $Y \in \mathfrak{g}^{(n+1)}$. As Y can be written as $[Z,W]$ for some $Z,W \in \mathfrak{g}^{(n)}$, then:

$$[X,Y] = [X,[Z,W]] = -[Z,[W,X]] - [W,[X,Z]].$$

By induction $[W,X] = -[X,W] \in \mathfrak{g}^{(n)}$, therefore, $[X,Y] \in \mathfrak{g}^{(n+1)}$ as the sum of two elements in $\mathfrak{g}^{(n+1)}$, proving that it is an ideal.

(c) Since $\mathfrak{g}^{(n)} \subset \mathfrak{g}$, the results follows from induction on n. In fact, if $\mathfrak{g}^{(n)} \subset \mathfrak{g}^n$:

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

□

With these ideal series being defined, we can finally start to study the different types of Lie algebras which are naturally more interesting:

Definition 1.1.8 *(a) An algebra is called **nilpotent** if $\mathfrak{g}^n = 0$ for some n.*

*(b) An algebra is called **solvable** if $\mathfrak{g}^{(n)} = 0$ for some n.*

*(c) An algebra is **semi-simple** if it has no non-zero solvable ideals.*

*(d) An algebra is **simple** if it has no non-zero proper ideals.*

1.2. Nilpotent and Solvable Lie Algebras

This first section of results focuses on finite-dimensional nilpotent and solvable algebras with an emphasis on matrix algebras. The results here are self-contained and

play an important role in representation theory through a characterization of these first algebras. Proofs presented here are based on [Hum72, chpt 3&4].
In this section, we fix a finite-dimensional vector space V over an arbitrary field $\mathbb{F}$.

Definition 1.2.1 *An operator $X \in \mathfrak{gl}(V)$ is **nilpotent** if $X^n = 0$ for a positive integer n. An element $X \in \mathfrak{g}$, where $\mathfrak{g}$ is any Lie algebra is **ad-nilpotent** if $ad(X)^n = 0$ for some positive integer n.*

Proposition 1.2.2 *If $X \in \mathfrak{gl}(V)$ is nilpotent then X is ad-nilpotent.*

Proof: Let n be such that $X^n = 0$. Additionally, let $L_X, R_X \in \mathfrak{gl}(\mathfrak{gl}(V))$ be the left and right multiplication by X respectively, that is, $L_X(Y) = XY$ and $R_X(Y) = YX$, for $Y \in \mathfrak{gl}(V)$, note that:

$$L_X^n = R_X^n = 0, \quad \mathrm{ad}(X) = L_X - R_X, \text{ and } L_X R_X Y = XYX = R_X L_X Y \text{ for all } Y \in \mathfrak{gl}(V).$$

The result follows from the binomial theorem for commuting operators:

$$\mathrm{ad}(X)^{2n} = (L_X - R_X)^{2n} = \sum_{k=0}^{2n} (-1)^{2n-k} \binom{2n}{k} L_X^k R_X^{2n-k},$$

which is 0 term by term because $R_x^{2n-k} = 0$ for $k < n$, and $L_x^k = 0$ if $k \geq n$. □

Theorem 1.2.3 (Engel's Lemma) *Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$, if $\mathfrak{g}$ consists of nilpotent endomorphisms and $V \neq 0$, then there exists a non-zero vector $v \in V$ for which $Xv = 0$ for all $X \in \mathfrak{g}$.*

Proof: We will proceed by induction on the dimension of $\mathfrak{g}$, the case when $\dim \mathfrak{g} = 1$ with $\{X\}$ as a basis is satisfied by taking any vector $v \neq 0$ and considering the smallest value of n such that $X^n v = 0$, then $X(X^{n-1}v) = 0$ and $X^{n-1}v \neq 0$.
If $\mathfrak{h}$ is any proper and non-trivial subalgebra of $\mathfrak{g}$, since $\mathrm{ad}(H)$ is nilpotent for any $H \in \mathfrak{h}$ we find that it also acts nilpotently on the quotient space $\mathfrak{g}/\mathfrak{h}$ with well defined action. By the induction hypothesis there exists some $\overline{X_0} \in \mathfrak{g}/\mathfrak{h}$ such that $[\mathfrak{h}, X_0] \in \mathfrak{h}$, moreover, $[X_0, X_0] = 0$, from which $\mathfrak{h} \oplus \mathbb{F}X_0$ is a subalgebra of $\mathfrak{g}$.
Assuming $\dim \mathfrak{g} \geq 2$, we can repeat the previous argument to construct a subalgebra $\mathfrak{h}$ and $X_0 \in \mathfrak{g}$ in such a way that $\mathfrak{g} = \mathfrak{h} \oplus \mathbb{F}X_0$ and $[\{X_0\}, \mathfrak{h}] \subseteq \mathfrak{h}$. This $\mathfrak{h}$ is in fact an ideal of $\mathfrak{g}$: $[\mathfrak{g}, \mathfrak{h}] = [\mathfrak{h}, \mathfrak{h}] + [\{X_0\}, \mathfrak{h}] \subseteq \mathfrak{h}$.
Again by induction, $W = \{v \in V \mid Hv = 0 \text{ for all } H \in \mathfrak{h}\}$ is non-zero as $\dim \mathfrak{h} < \dim \mathfrak{g}$,

but since $\mathfrak{h}$ is an ideal, $\mathfrak{g}W \subseteq W$, in fact:

$$Y(Xw) = XYw - [X,Y]w = 0 \text{ for all } X \in \mathfrak{g},\ Y \in \mathfrak{h}, \text{ and } w \in W.$$

Finally, as X_0 acts on W and is nilpotent, there exists a non-zero vector $v \in W$ such that $X_0 v = 0$, therefore $Xv = 0$ for any $X \in \mathfrak{g}$. □

When studying the properties of nilpotent and solvable Lie algebras, the kernel of the adjoint representation plays an important role as a tool to move from matrix algebras to general Lie algebras.

Definition 1.2.4 *Let $\mathfrak{g}$ be a Lie algebra, we define the* ***center*** *of $\mathfrak{g}$ as the ideal given by:*

$$\mathfrak{z}(\mathfrak{g}) := \{Z \in \mathfrak{g} \mid [Z,X] = 0 \textit{ for all } X \in \mathfrak{g}\}.$$

Theorem 1.2.5 (Engel's Theorem) *If $\mathfrak{g}$ is a finite-dimensional Lie algebra and all elements of $\mathfrak{g}$ are ad-nilpotent, then $\mathfrak{g}$ is nilpotent.*

Proof: The image of the adjoint representation satisfies the conditions of Theorem 1.2.3 in $\mathfrak{gl}(\mathfrak{g})$, implying that there exists non-zero $X \in \mathfrak{g}$ such that $[X,\mathfrak{g}] = 0$, therefore $\mathfrak{z}(\mathfrak{g}) \neq 0$. Now $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ consists of ad-nilpotent elements and has smaller dimension than $\mathfrak{g}$, therefore it is a nilpotent algebra by induction on the dimension of $\mathfrak{g}$.
But if $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is nilpotent then $\mathfrak{g}$ is also nilpotent. In fact, if $\mathfrak{g}/(\mathfrak{z}(\mathfrak{g}))$ is nilpotent, then $\mathfrak{g}^n \subset \mathfrak{z}(\mathfrak{g})$ for some n, but that implies that $\mathfrak{g}^{n+1} \subset [\mathfrak{g}, \mathfrak{z}(\mathfrak{g})] = 0$. □

A similar result to Engel's Lemma (Theorem 1.2.3) is valid for solvable Lie algebras under more strict conditions, in fact:

Theorem 1.2.6 (Lie's Theorem) *Let V be a non-trivial finite-dimensional vector space over an algebraically closed field of characteristic 0, and $\mathfrak{g}$ a solvable Lie subalgebra of $\mathfrak{gl}(V)$. There exists a linear form $\lambda : \mathfrak{g} \to \mathbb{F}$ and a non-zero $v \in V$ such that $Xv = \lambda(X)v$ for all $X \in \mathfrak{g}$.*

Proof: The proof of this theorem follows the same steps as the proof of Engel's Lemma (Theorem 1.2.3).

1. Locate an ideal $\mathfrak{h}$ of $\mathfrak{g}$ with co-dimension one.

2. Common eigenvectors exist for $\mathfrak{h}$ by induction.

3. $\mathfrak{g}$ stabilizes the space consisting of these eigenvectors, if W is such space, this means that $Xw \in W$ for any $X \in \mathfrak{g}$ and $w \in W$.

4. Find an eigenvector in this space for a single $X_0 \in \mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{h} \oplus \mathbb{F}X_0$.

---

1. Since $\mathfrak{g}$ is solvable then $\mathfrak{g}^{(2)} = [\mathfrak{g}, \mathfrak{g}] \neq \mathfrak{g}$, therefore any co-dimension one vector space containing $\mathfrak{g}^{(2)}$ is an ideal, fix one of these as $\mathfrak{h}$.

2. Proceeding by induction on $\dim \mathfrak{g}$, if $\dim \mathfrak{g} = 0$, the result is trivial.
   Since $\dim \mathfrak{h} = \dim \mathfrak{g} - 1 < \dim \mathfrak{g}$, we assume the existence of a non-zero common eigenspace, that is, some $\lambda \in \mathfrak{h}^*$ such that

$$W := \{v \in V \mid Xv = \lambda(X)v \text{ for all } X \in \mathfrak{h}\} \neq 0.$$

3. To prove that $\mathfrak{g}$ stabilizes W, let $w \in W$, $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$ arbitrary. Then

$$YXw = XYw - [X,Y]w = \lambda(Y)Xw - \lambda([X,Y])w.$$

   Thus $\lambda([X,Y]) = 0$ implies our desired result.
   Let n be the largest integer such that $\{w, Xw, \cdots, X^n w\}$ is linearly independent, and V_i spanned by $\{X^j w \mid 0 \leq j \leq i\}$. $V_N = V_n$ for $N > n$ and $\dim V_n = n+1$.
   $\mathfrak{h}$ stabilizes each V_i. By induction, the base case being $Yw \in W$ for all $Y \in \mathfrak{h}$ is trivial.

$$YX^i w = YXX^{i-1}w = XYX^{i-1}w - [X,Y]X^{i-1}w = (XY - [X,Y])X^{i-1}w \in V_i.$$

   From this we can see that $\mathrm{Tr}_{V_n}(Y) = n\lambda(Y)$, but this is valid for the special element in $\mathfrak{h}$ of the form $[X,Y]$. Since $\mathrm{Tr}([X,Y]) = \mathrm{Tr}(XY) - \mathrm{Tr}(YX) = 0$, then $n\lambda([X,Y]) = 0$. As we assumed characteristic 0, this forces $\lambda[X,Y] = 0$.

4. Letting $X_0 \in \mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{h} \oplus \mathbb{F}X_0$, as $X_0 : W \to W$, there exists an eigenvector $v \in W$ for X_0. Extend $\lambda : \mathfrak{h} \to \mathbb{F}$ to include X_0 in its domain and define $\lambda(X_0)$ as the eigenvalue of X_0.

□

Example 1.2.7 *The condition that the field be of characteristic* 0 *is in fact necessary.*

Consider for instance $\mathbb{F}$ to be any field of characteristic p, and the $p \times p$ matrices:

$$X = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p-1 \end{pmatrix},$$

seeing as:

$$XY = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p-1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad YX = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p-2 \\ p-1 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

We can see that $[X, Y] = X$ and therefore $\{X, Y\}$ span a solvable Lie subalgebra of $\mathfrak{gl}(p, \mathbb{F})$. X and Y have no common eigenvectors as all eigenvectors of Y are multiples of canonical basis vectors which are not eigenvectors of X.

Remark 1.2.8 *Both previous results (Engel's Theorem and Lie's Theorem) can be used to classify all matrix nilpotent and solvable algebras. Although such classification is very interesting by its nature, it provides us no tools for our goal to study representations of semi-simple Lie algebras so it will not be covered in this thesis.*

1.3. Cartan's Criteria

In this section we will consider V to be an n-dimensional vector space over an algebraically closed field $\mathbb{F}$ of characteristic 0. The purpose of this section is to reach a criteria for an algebra to be semi-simple using a well known bilinear form, and to characterize semi-simple Lie algebras as a direct sum of simple algebras. The results presented here are based on [Hum72, sec 4.3-4.4].

These results make heavy use of Jordan Decomposition of operators on finite-dimensional vector spaces and the realization of its terms as polynomials applied to the operator, we refer readers to Appendix B or [Hum72, sec 4.2] if they are unfamiliar with such results.

Proposition 1.3.1 *If $X \in \mathfrak{gl}(V)$ is such that its Jordan decomposition (Proposition B.0.1)* *is $X = S + N$, then $ad(X) = ad(S) + ad(N)$ is the Jordan decomposition of adX in $\mathfrak{gl}(\mathfrak{gl}(V))$.*

Proof: Let $\{e_{ij}\}$ be a basis of $\mathfrak{gl}(V)$ so that $S = \mathrm{diag}\{\lambda_i\}$ in this basis, then $[S, e_{ij}] = (\lambda_i - \lambda_j)e_{ij}$ and $\mathrm{ad}(S)$ is **semi-simple** (diagonalizable). If N is nilpotent then $\mathrm{ad}(N)$ is nilpotent (Proposition 1.2.2). Since ad is a representation, then:

$$[\mathrm{ad}(S), \mathrm{ad}(N)]_{\mathfrak{gl}(\mathfrak{g})} = \mathrm{ad}([S,N]_{\mathfrak{g}}) = 0.$$

The result follows from uniqueness of the Jordan Decomposition in $\mathfrak{gl}(\mathfrak{gl}(V))$. □

Theorem 1.3.2 (Cartan's Lemma) *Let $A \subset B$ be two subspaces of $\mathfrak{gl}(V)$, and set $M = \{X \in \mathfrak{gl}(V) \mid [X,B] \subset A\}$. If $X \in M$ is such that $Tr(XY) = 0$ for all $Y \in M$, then X is nilpotent.*

Proof: This proof consists mostly of technicalities on proving that specific matrices are present in M.

Let $X = S + N$ be the Jordan decomposition of X, fix a basis $\{v_1, \cdots, v_n\}$ in which $S = \mathrm{diag}(a_1, \cdots, a_m)$. We want to prove that $a_1 = \cdots = a_m = 0$ since in that case $S = 0$ and X is nilpotent.

Let E be the subspace of $\mathbb{F}$ generated by $\{a_i \mid 1 \leq i \leq m\}$ over $\mathbb{Q}$. Since we are assuming char $\mathbb{F} = 0$, it is enough to show that $E = 0$, or equivalently $E^* = 0$.

Let $f \in E^*$ and $Y = \mathrm{diag}(f(a_1), \cdots, f(a_m))$. If $\{e_{ij}\}$ is the corresponding basis of $\mathfrak{gl}(V)$, then $\mathrm{ad}(S)e_{ij} = (a_i - a_j)e_{ij}$ and $\mathrm{ad}(Y)e_{ij} = (f(a_i) - f(a_j))e_{ij}$. Now let $r(x) \in \mathbb{F}[X]$ be the polynomial such that $r(a_i - a_j) = f(a_i) - f(a_j)$, the existence of which follows from the Lagrange polynomial construction and linearity, then $\mathrm{ad}(Y) = r(\mathrm{ad}(S))$. Now $\mathrm{ad}(S)$ is the semi-simple part of $\mathrm{ad}(X)$, so it can be written as a polynomial in $\mathrm{ad}(X)$ without constant term by (Proposition B.0.1(b)) and therefore $\mathrm{ad}(Y)$, as a polynomial in S, maps B to A, proving that Y is in M. Now

$$\mathrm{Tr}(XY) = \sum_{j=1}^{m} a_j f(a_j) = 0 \Rightarrow \sum_{j=1}^{m} f(a_j)^2 = 0.$$

Since we restricted ourselves to $\mathbb{Q}$, this implies that $f(a_j) = 0$ for all j and therefore $f = 0$. Since f is arbitrary then we find $E^* = 0$. □

This lemma, technical and apparently not very useful is essential to one of the most important facts about Lie algebras, as a tool on the characterization of semi-simple Lie algebras. In fact:

Theorem 1.3.3 (Cartan's Criterion) *Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$. If $Tr(XY) = 0$ for* *all $X \in [\mathfrak{g},\mathfrak{g}]$ and $Y \in \mathfrak{g}$, then $\mathfrak{g}$ is solvable.*

Proof: Let $A = [\mathfrak{g},\mathfrak{g}]$ and $B = \mathfrak{g}$. The hypothesis shows that $\mathrm{Tr}(XY) = 0$ for all $X \in A$, and $Y \in B$. We need a stronger statement to use Cartan's lemma (Theorem 1.3.2): For all $X \in A$ and $Y \in M$ it follows that $\mathrm{Tr}(XY) = 0$, where $M = \{X \in \mathfrak{gl}(V) | [X,B] \subset A\}$. Note that $\mathrm{Tr}([X,Y]Z) = \mathrm{Tr}(X[Y,Z])$ for all $X,Y,Z \in \mathfrak{gl}(V)$, since:

$$\mathrm{Tr}(XYZ - YXZ) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(YXZ) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(Y(XZ)) \tag{1.1}$$

$$= \mathrm{Tr}(XYZ - XZY) = \mathrm{Tr}(X[Y,Z]). \tag{1.2}$$

Let $[X,Y] \in A$ with $X,Y \in \mathfrak{g}$, and $Z \in M$, then: $\mathrm{Tr}([X,Y]Z) = \mathrm{Tr}(X[Y,Z]) = 0$ since $[Y,Z] \in A$ and $X \in B$. Therefore by Theorem 1.3.2 every element in $[\mathfrak{g},\mathfrak{g}]$ is nilpotent, which implies that $[\mathfrak{g},\mathfrak{g}]$ is a nilpotent algebra. Since $\mathfrak{g}^{(n)} \subset [\mathfrak{g},\mathfrak{g}]^{n-1} = 0$ for some n, we conclude that $\mathfrak{g}$ is solvable. □

This in turn allows us to classify algebras with respect to the trace form in its adjoint representation, in fact:

Corollary 1.3.4 *Let $\mathfrak{g}$ be any finite-dimensional Lie algebra over $\mathbb{F}$. If $Tr(ad(X)ad(Y)) = 0$ for all $X \in [\mathfrak{g},\mathfrak{g}]$ and $Y \in \mathfrak{g}$, then $\mathfrak{g}$ is solvable.*

Proof: Using the previous theorem, we know that $\mathrm{ad}\mathfrak{g} \simeq \mathfrak{g}/\ker \mathrm{ad}$ is solvable, since $\ker \mathrm{ad} = \mathfrak{z}(g)$ is solvable, then $\mathfrak{g}$ is solvable. □

With this corollary in mind, we are able to define a natural bilinear form on finite-dimensional Lie algebras, the **Cartan-Killing** form.

Definition 1.3.5 *Let $\mathfrak{g}$ be any finite-dimensional Lie algebra over $\mathbb{F}$. The bilinear form $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{F}$ defined as $\kappa(X,Y) = Tr(ad(X)ad(Y))$ is called the Cartan-Killing form of $\mathfrak{g}$.*

Lemma 1.3.6 *κ is symmetric, and associative, in the sense that: $\kappa(X,[Y,Z]) = \kappa([X,Y],Z)$. For any ideal $\mathfrak{h}$ of $\mathfrak{g}$, its restriction to $\mathfrak{h} \times \mathfrak{h}$ is equal to the Cartan-Killing form of $\mathfrak{h}$. Furthermore, its radical $Rad\ \kappa = \{X \in \mathfrak{g} \mid \kappa(X,Y) = 0$ for all $Y \in \mathfrak{g}\}$ is an ideal of $\mathfrak{g}$.*

Proof: Since $\mathrm{Tr}(\mathrm{ad}(X)\mathrm{Tr}(Y)) = \mathrm{Tr}(\mathrm{ad}(Y)\mathrm{ad}(X))$, the symmetry follows.
Associativity follows from (1.1) and (1.2) on the proof of Theorem 1.3.3.
If $\mathfrak{h}$ is an ideal, $\mathrm{ad}(X)\mathrm{ad}(Y)|_{\mathfrak{h}} = \mathrm{ad}(X)|_{\mathfrak{h}}\mathrm{ad}(Y)|_{\mathfrak{h}}$. Therefore, since the trace of an endomorphism is equal to the trace of this endomorphism restricted to its image, the result follows.
Finally, if $Y \in \mathfrak{g}$ and $X \in \mathrm{Rad}\ \kappa$ then, through associativity $\kappa([X,Y],Z) = \kappa(X,[Y,Z]) = 0$ for any $Z \in \mathfrak{g}$. This proves that $[X,Y] \in \mathrm{Rad}\ \kappa$ and that $\mathrm{Rad}\ \kappa$ is an ideal. □

With this in mind we can make the first step to classify semi-simple finite dimensional Lie algebras.

Theorem 1.3.7 *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{F}$. Then the following statements are equivalent:*

(a) $\mathfrak{g}$ *has no non-zero abelian ideals.*

(b) $\mathfrak{g}$ *is semi-simple (has no non-zero solvable ideals).*

(c) The Killing form $\kappa(X,Y) = Tr(ad(X)ad(Y))$ *is non-degenerate.*

(d) $\mathfrak{g}$ *decomposes uniquely as a direct sum of simple ideals.*

Proof:
$(a) \Rightarrow (b)$. Let $\mathfrak{h}$ be a solvable ideal of $\mathfrak{g}$ and n minimum such that $\mathfrak{h}^{(n)} = 0$, then $\mathfrak{h}^{(n-1)}$ is an abelian subalgebra of $\mathfrak{g}$, but it is also an ideal:

Lemma 1.3.8 *If* $\mathfrak{h}$ *is an ideal of* $\mathfrak{g}$, *then* $\mathfrak{h}^{(k)}$ *is an ideal of* $\mathfrak{g}$ *for all* k.

Proof: By induction, the basis case $k = 1$ being the hypothesis, let $[X,Y] \in \mathfrak{h}^{(n)}$ with $X,Y \in \mathfrak{h}^{(n-1)}$ and $Z \in \mathfrak{g}$, then $[Z,[X,Y]] = [[Y,Z],X] + [[Z,X],Y]$.
Using the induction hypothesis we know that $[Y,Z],[Z,X] \in \mathfrak{h}^{(n-1)}$, knowing that $X,Y \in \mathfrak{h}^{(n-1)}$ we find $[Z,[X,Y]] \in [\mathfrak{h}^{(n-1)},\mathfrak{h}^{(n-1)}] = \mathfrak{h}^{(n)}$. □
$(b) \Rightarrow (c)$. let $\mathfrak{h} = \text{Rad}\,\kappa$, since $\kappa(X,Y) = 0$ for all $X \in \mathfrak{h}$ and $Y \in \mathfrak{g}$ (in particular for $Y \in [\mathfrak{h},\mathfrak{h}]$), then $\mathfrak{h}$ is a solvable ideal of $\mathfrak{g}$ by Cartan's Criterion (Corollary 1.3.4). Therefore, if κ is degenerate, then $\mathfrak{g}$ has solvable ideals.
$(c) \Rightarrow (d)$. Given an ideal $\mathfrak{h}$, we define $\mathfrak{h}^\perp = \{X \in \mathfrak{g} \mid \kappa(X,Y) = 0 \text{ for all } Y \in \mathfrak{h}\}$, since κ is non-degenerate we have $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$. $\mathfrak{h}^\perp$ is in fact an ideal of $\mathfrak{g}$, given $X \in \mathfrak{h}^\perp$ and $Z \in \mathfrak{g}$ then for all $H \in \mathfrak{h}$ we get $\kappa([X,Z],H) = \kappa(X,[H,Y]) = 0$ since $\mathfrak{h}$ is an ideal.
Existence: If $\mathfrak{g}$ is simple, then $\mathfrak{g}$ is the only non-trivial ideal, and the result follows. Otherwise, $\mathfrak{g}$ has a proper ideal $\mathfrak{h}$, and therefore $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$. As $\mathfrak{h}^\perp$ is a non-zero ideal of $\mathfrak{g}$, the result follows by doing the same argument for $\mathfrak{h}$ and $\mathfrak{h}^\perp$ (see Lemma 1.3.6).
Uniqueness: Assume the existence of a decomposition $\mathfrak{g} = \bigoplus \mathfrak{g}_i$, and let $\mathfrak{h}$ be any simple ideal of $\mathfrak{g}$. In this case, $[\mathfrak{g},\mathfrak{h}] = \bigoplus [\mathfrak{g}_i,\mathfrak{h}]$ and $[\mathfrak{g},\mathfrak{h}] = \mathfrak{h}$ is simple and therefore all terms are 0 except one, proving that $\mathfrak{h} = \mathfrak{g}_j$ for some j.
$(d) \Rightarrow (a)$. Let $\mathfrak{g} = \bigoplus_{i \in I} \mathfrak{g}_i$ be the decomposition of $\mathfrak{g}$ in simple ideals, then let $\mathfrak{h}$ be any ideal of $\mathfrak{g}$, then $[\mathfrak{g},\mathfrak{h}] = \bigoplus_{i\in I}[\mathfrak{g}_i,\mathfrak{h}]$. For any specific i, $[\mathfrak{g}_i,\mathfrak{h}]$ is an ideal of $\mathfrak{g}_i$, therefore it should be 0 or $\mathfrak{g}_i$, implying that $\mathfrak{h}$ is a sum of $\mathfrak{g}_i$ where i is running on a subset of I, since all $\mathfrak{g}_i$ are simple, $\mathfrak{h}$ cannot be abelian. □

1.4. Weyl's Theorem

Weyl's Theorem of complete reducibility allows us to define a similar decomposition to the decomposition of semi-simple Lie algebras as simple ideals, not for the algebras themselves but for their representations. These results are based on [Hum72, chpt 6].
In this section $\mathbb{F}$ will be an algebraically closed field of characteristic 0 and $\mathfrak{g}$ will be an n-dimensional Lie algebra over $\mathbb{F}$.

$\mathfrak{g}$-Modules

In order to set an algebraic structure for representations, we think of them not as algebra morphisms but rather as vector spaces with an action of the represented algebra. This structure allows us to talk about natural compositions, reducibility, sub-structures and other useful algebraic properties.

Definition 1.4.1 *A $\mathfrak{g}$-**module** is a vector space V together with an operator $\cdot : \mathfrak{g} \times V \to V$ in such a way that, for all $X, Y \in \mathfrak{g}$ and $v, w \in V$:*

(a) $X \cdot (v + w) = X \cdot v + X \cdot w$.

(b) $(X + Y) \cdot v = X \cdot v + Y \cdot v$.

(c) $[X, Y] \cdot v = X \cdot (Y \cdot v) - Y \cdot (X \cdot v)$.

Proposition 1.4.2 *Every $\mathfrak{g}$-module $(V, \cdot)$ defines a representation and every representation ρ defines a $\mathfrak{g}$-module in such a way that $a \cdot v = \rho(a)v$ for every $a \in \mathfrak{g}$ and $v \in V$.*

Proof: Given V a $\mathfrak{g}$-module, for $a \in \mathfrak{g}$ define $\rho(a) \in \mathfrak{gl}(V)$ such that $\rho(a) : v \mapsto a \cdot v$, it is in $\mathfrak{gl}(V)$ if and only if the action satisfies (a), it is a linear morphism if and only if (b) is satisfied and it satisfies $\rho([X, Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X)$ if and only if (c). □

Remark 1.4.3 *The use of the word "module" here is not arbitrary, the universal enveloping algebra introduced in Chapter 4 allows us to speak of $\mathfrak{g}$-modules as modules in the more usual sense.*

Proposition 1.4.4 *Let V and W be $\mathfrak{g}$-modules, and $X \in \mathfrak{g}$ arbitrary, then the following are $\mathfrak{g}$-modules:*

(a) $V \oplus W$ *with the action* $X \cdot (v + w) = X \cdot v + X \cdot w$.

(b) V^* *with the action* $X \cdot f : v \mapsto -f(X \cdot v)$.

(c) $V \otimes W$ *with the action* $X \cdot (v \otimes w) = (X \cdot v) \otimes w + v \otimes (X \cdot w)$.

(d) *If* $W \subset V$, *then* V/W *with* $X \cdot \overline{v} = \overline{X \cdot v}$.

(e) $Hom_{\mathbb{F}}(V, W)$ *with the action* $X \cdot f : v \mapsto X \cdot f(v) - f(X \cdot v)$.

Remark 1.4.5 *All of these are compatible with each other, for example, if* $W = \mathbb{F}$ *on* (e) *we get the same result as* (b) *since* $X \cdot f(v) = 0$ *(actions on the field are trivial).*

Some basic and needed definitions on the algebraic structure of $\mathfrak{g}$-modules are given below, we shall omit the product to simplify notation:

Definition 1.4.6 *A sub-module of a module* V *is a subspace* W *of* V *such that* $\mathfrak{g}W \subset W$. *A linear transformation* T *between modules* V *and* W *is said to be a homomorphism of* $\mathfrak{g}$*-modules if* $T(Xv) = XT(v)$ *for any* $X \in \mathfrak{g}$.

Proposition 1.4.7 *Let* $T : V \to W$ *be a homomorphism of* $\mathfrak{g}$*-modules, then* $\ker T$ *is a sub* $\mathfrak{g}$*-module of* V.

Proof: $\ker T$ is a subspace of V, now for any $X \in \mathfrak{g}$ and $v \in \ker T$

$$T(Xv) = XT(v) = X0 = 0.$$

Therefore $Xv \in \ker T$ and the proof is done. □

Definition 1.4.8 *Regarding the sub-structure of a particular* $\mathfrak{g}$*-module* V*:*
V *is said to be* ***irreducible*** *if it has no proper and non-trivial sub-modules.*
V *is said to be* ***indecomposable*** *if it cannot be decomposed as a direct sum of two of its proper sub-modules, that is if* $V = M \oplus N$ *for* M *and* N *sub-modules then* $M = 0$ *or* $N = 0$.
V *is said to be* ***reducible*** *if it is not irreducible.*

Initial Results on Representations

Lemma 1.4.9 *Let* $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ *be a representation of a finite dimensional semi-simple Lie algebra* $\mathfrak{g}$, *then* $\rho(\mathfrak{g}) \subset \mathfrak{sl}(V)$.

Proof: As $\mathfrak{g}$ is semi-simple, $[\mathfrak{g}, \mathfrak{g}] = \bigoplus [\mathfrak{g}, \mathfrak{g}_i] = \mathfrak{g}$ by simplicity.
Now since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, we can represent an element of $\mathfrak{g}$ by $[X, Y]$ and therefore $\mathrm{Tr}(\rho[X, Y]) = \mathrm{Tr}(\rho(X)\rho(Y) - \rho(Y)\rho(X)) = 0$. In particular, if V is one-dimensional then $\rho(\mathfrak{g}) = 0$ □

Lemma 1.4.10 (Schur's Lemma) *Let V be an irreducible finite-dimensional $\mathfrak{g}$-module over an algebraically closed field $\mathbb{F}$, and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be the correspondent representation. The only endomorphism that commutes with $\rho(\mathfrak{g})$ are the scalars*

Proof: If T is a non-trivial endomorphism that commutes with $\rho(\mathfrak{g})$, then T is a homomorphism of modules, and therefore its kernel must be 0 since it is a submodule of V.
We also know that $T - \lambda I$ commutes with $\rho(\mathfrak{g})$ for all $\lambda \in \mathbb{F}$, therefore $T - \lambda I$ is either 0 or an isomorphism. Since the field is algebraically closed, T has an eigenvalue, and in that case the kernel cannot be 0, therefore $T - \lambda I = 0 \Rightarrow T = \lambda I$ for some λ. □

We will proceed to introduce the idea of a special element in $\mathfrak{gl}(V)$ with respect to a $\mathfrak{g}$-module V, classically called the Casimir element.

Let $\beta : \mathfrak{g} \times \mathfrak{g} \to \mathbb{F}$ be any non-degenerate form satisfying $\beta(X, [Y,Z]) = \beta([X,Y],Z)$ (associativity). By non-degeneracy, every basis $A = \{X_1, \cdots, X_n\}$ of $\mathfrak{g}$ has a dual basis $B = \{Y_1, \cdots, Y_n\}$ with respect to β, that is $\beta(X_i, Y_j) = \delta_{ij}$.

For any $X \in \mathfrak{g}$, setting $[X, X_i] = \sum_j a_{ij} X_j$ and $[X, Y_i] = \sum_j b_{ij} Y_j$, we find:

$$a_{ik} = \sum_{j=1}^{n} a_{ij}\delta_{jk} = \sum_{j=1}^{n} a_{ij}\beta(X_j, Y_k) = \beta([X, X_i], Y_k) = -\beta(X_i, [X, Y_k]) = -b_{ki}.$$

Now, if $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a one-to-one representation of a semi-simple $\mathfrak{g}$, then the trace-form $\beta(X, Y) = \mathrm{Tr}(\rho(X)\rho(Y))$ is non-degenerate since $\rho(\mathfrak{g})$ is isomorphic to $\mathfrak{g}$ and Rad β is an ideal of $\mathfrak{g}$.

Define $c_\rho \in \mathfrak{gl}(V)$ as $c_\rho = \sum_{i=1}^{n} \rho(X_i)\rho(Y_i)$, the **Casimir element** of ρ. Note that $\mathrm{Tr}(c_\rho) = \sum_{i=1}^{n} \mathrm{Tr}(\rho(X_i)\rho(Y_i)) = \sum_{i=1}^{n} \beta(X_i, Y_i) = n = \dim \mathfrak{g}$.

One of the important properties of c_ρ is that it commutes with $\rho(\mathfrak{g})$. To simplify the proof of this statement, let $X_i = \rho(X_i)$:

$$\begin{aligned} Xc_\rho - c_\rho X &= \sum_{i=1}^{n} XX_iY_i - X_iY_iX = \sum_{i=1}^{n} XX_iY_i - X_iXY_i + X_iXY_i - X_iY_iX \\ &= \sum_{i=1}^{n} [X, X_i]Y_i + X_i[X, Y_i] = \sum_{i,j=1}^{n} a_{ij}X_jY_i + \sum_{i,j=1}^{n} b_{ij}X_iY_j = 0. \end{aligned}$$

In case of a representation that is not one-to-one, we can consider another Lie algebra $\mathfrak{g}' = \mathfrak{g}/\ker\rho$ and a new representation $\rho' : \mathfrak{g}' \to \mathfrak{gl}(V)$ that satisfies the given conditions. $c_{\rho'}$ still commutes with $\rho(\mathfrak{g})$ because $\rho(\mathfrak{g}) = \rho'(\mathfrak{g}')$. With this discussion, we can summarize the properties of the Casimir element of a representation:

Proposition 1.4.11 (Casimir Element) *If $\mathfrak{g}$ is semi-simple and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation with V finite-dimensional, then there exists an element $c_\rho \in \mathfrak{gl}(V)$ that commutes with $\rho(\mathfrak{g})$ and has non-zero trace. Moreover, if ρ is irreducible, then c_ρ acts as a scalar (Schur's Lemma).*

Weyl's Theorem of Irreducibility

Theorem 1.4.12 *Let V be a finite-dimensional $\mathfrak{g}$-module, where $\mathfrak{g}$ is semi-simple. Every proper sub $\mathfrak{g}$-module $W \subset V$ admits a sub $\mathfrak{g}$-module $W' \subset V$ such that $V = W \oplus W'$.*

Proof: Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be the induced representation on V and c_ρ be one of its Casimir elements. This proof proceeds in several cases and the use of induction.

If W is an irreducible sub $\mathfrak{g}$-module of V with co-dimension one, then V/W is a $\mathfrak{g}$-module of dimension one. $\mathfrak{g}$ being semi-simple, $\mathfrak{g}V \subset W$ and c_ρ is a $\mathfrak{g}$-module homomorphism. $\ker c_\rho$ is a module with $\ker c_\rho \neq V$, since $\mathrm{Tr}(c_\rho) \neq 0$. Finally $c_\rho : V \to W$ (c_ρ being in the span of $\rho(\mathfrak{g})^2$) and acts as a scalar in W by Schur's Lemma, therefore, $\ker c_\rho$ is the desired complement to W.

If W is a reducible sub-module of V with co-dimension one, let $W' \subset W$ be a sub-module. V/W' is a sub $\mathfrak{g}$-module of W/W' with co-dimension one, by induction, it has a one-dimensional sub-module $\overline{W}/W'$ such that $V/W' = W/W' \oplus \overline{W}/W'$.
Now $W' \subseteq \overline{W}$ with co-dimension one, so induction provides a one-dimensional sub-module M of $\overline{W}$ such that $\overline{W} = W' \oplus M$, it follows that $V = W \oplus M$.

Now let W be any sub-module, consider $H = \mathrm{Hom}(V, W)$ viewed as a $\mathfrak{g}$-module, and let $\mathcal{V}$ be the subspace of H consisting of maps whose restriction to W is a scalar, letting $\mathcal{V}_a = \{f \in H \mid f(w) = a \cdot w \text{ for all } w \in W\}$ then $\mathcal{V} = \bigcup_{a \in \mathbb{F}} \mathcal{V}_a$. Let $\mathcal{W}$ be the ones whose restriction to W are 0, that is $\mathcal{W} = \mathcal{V}_0$. These are both sub-modules of H and clearly $\mathcal{W} \subset \mathcal{V}$ is a subspace of co-dimension one (since its complement is determined by a scalar). To prove the fact that they are sub-modules, let $f \in \mathcal{V}$, $w \in W$ and $X \in \mathfrak{g}$, then:

$$(Xf)(w) = X(f(w)) - f(Xw) = X(a \cdot w) - a \cdot (Xv) = 0.$$

Therefore $Xf \in \mathcal{W}$, and $\mathcal{W}$ has a complement in $\mathcal{V}$, let this complement be spanned by $h \in \mathcal{V}$. By Lemma 1.4.10, $\mathfrak{g}$ acts on h trivially and therefore $(Xh)(v) = 0 \Rightarrow X(h(v)) - h(Xv) = 0$, meaning that h is a module homomorphism. Finally, as h sends V into W and acts as a scalar in W, then $V = W \oplus \ker h$. □

2 Root Decomposition

This chapter presents the main results on the study of semi-simple Lie algebras from an algebraic perspective, providing results for the root decomposition of Lie Algebras and reaching the structure that leads to their classification in Chapter 3, the same approach can be found in [Hum72, sec 5.4 & chpt 8].

In this chapter, we fix an algebraic closed field $\mathbb{F}$ of characteristic 0 and let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{F}$.

2.1. Abstract Jordan Decomposition

There is a natural way to extend the Jordan decomposition of finite-dimensional operators to a given semi-simple Lie algebra by the adjoint representation, the main idea is proving the existence of elements in a Lie algebra such that their adjoint satisfies the conditions of the decomposition.

Lemma 2.1.1 *Every derivation in a semi-simple Lie algebra is inner, meaning that if D is a derivation, there exists $Y \in \mathfrak{g}$ such that $D = ad(Y)$.*

Proof: Given D a derivation, the linear form in $\mathfrak{g}^*$ given by $f(X) = \mathrm{Tr}(D\,\mathrm{ad}(X))$ has a dual element $Y \in \mathfrak{g}$ in such a way that $\kappa(X,Y) = f(X)$. We shall prove that $D = \mathrm{ad}(Y)$. Given the element $E = D - \mathrm{ad}(Y)$, we have $\mathrm{Tr}(E\mathrm{ad}(X)) = 0$ for any $X \in \mathfrak{g}$. Now taking $X, Z \in \mathfrak{g}$ arbitrary

$$[E, \mathrm{ad}(X)]Z = E(\mathrm{ad}(X)Z\ ) - \mathrm{ad}(X)EZ = E[X,Z] - [X, EZ] = [EX, Z] = \mathrm{ad}(EX)Z,$$

implying that:

$$\begin{aligned}\kappa(EX,Z) &= \mathrm{Tr}(\mathrm{ad}(EX)\mathrm{ad}(Z)) = \mathrm{Tr}([E,\mathrm{ad}(X)]\mathrm{ad}(Z)) \\ &= \mathrm{Tr}(E\mathrm{ad}(X)\mathrm{ad}(Z) - \mathrm{ad}(X)E\mathrm{ad}(Z)) = \mathrm{Tr}(E\mathrm{ad}[X,Z]) \\ &= 0.\end{aligned}$$

As Z is arbitrary, $EX = 0$ for any X, so $E = 0$. Finally $D = \mathrm{ad}(Y)$. □

Lemma 2.1.2 *[Mar09, p.77] If D is a derivation and $\lambda, \mu \in \mathbb{F}$ then for every $X, Y \in \mathfrak{g}$:*

$$(D-(\lambda+\mu)I)^n[X,Y] = \sum_{i=0}^{n} \binom{n}{i} [(D-\lambda I)^{n-i}X, (D-\mu I)^i Y]. \tag{2.1}$$

Corollary 2.1.3 *On an algebraically closed field, if $D = S+N$ is the Jordan decomposition of a derivation, then S and N are derivations.*

Proof: Let $\mathfrak{g} = \bigoplus_{\alpha\in\mathbb{F}} \mathfrak{g}_\alpha$ be the generalized eigenspace decomposition of $\mathfrak{g}$ with respect to D, (2.1) shows that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$, and therefore if $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$ then:

$$S[X,Y] = (\alpha+\beta)[X,Y] = \alpha[X,Y] + \beta[X,Y] = [SX,Y] + [X,SY].$$

Therefore S is a derivation since $\mathfrak{g}$ is the sum of eigenspaces. Finally $N = D - S$ is a derivation. □

Proposition 2.1.4 *For every $X \in \mathfrak{g}$ there exists unique $S, N \in \mathfrak{g}$ satisfying the following conditions:*

(a) $X = S + N$.

(b) $ad(S)$ *is diagonizable and* $ad(N)$ *is nilpotent.*

(c) $[S,N] = 0$.

Such a decomposition will be called ***abstract Jordan decomposition****, S will be called the* ***semi-simple*** *part of X and N will be called the* ***nilpotent*** *part.*

Proof: Since $\mathrm{ad}(X)$ is a derivation, its semi-simple part and nilpotent part are derivations and therefore are adjoints of elements in $\mathfrak{g}$, let those elements be S and N respectively, $\mathrm{ad}(X) = \mathrm{ad}(S) + \mathrm{ad}(N)$.

Since $\mathfrak{g}$ is semi-simple, the adjoint representation is one-to-one (its kernel is $\mathfrak{z}(\mathfrak{g}) = 0$) and therefore $\mathrm{ad}(X) = \mathrm{ad}(S+N) \Rightarrow X = S + N$.

Finally, as $[\mathrm{ad}(S), \mathrm{ad}(N)] = 0$, $\mathrm{ad}([S,N]) = 0$ and therefore $[S,N] = 0$. □

Proposition 2.1.5 *Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\varphi : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation. If $X \in \mathfrak{g}$ has Abstract Jordan Decomposition $X = S + N$, then the Jordan Decomposition of $\varphi(X)$ is $\varphi(X) = \varphi(S) + \varphi(N)$.*
In particular the representation of any semi-simple element is semi-simple and of every nilpotent element is nilpotent. See [Hum72, p.29-30].

2.2. Toral Subalgebras

The purpose of this section is to use the results from the previous section to construct a subalgebra of semi-simple elements. When considering the joint eigenspace decomposition with respect to the adjoint of these elements, one obtains exactly the structure of root systems that allows us to classify semi-simple Lie algebras. This approach is based on [Hum72, sec 8.1].

Definition 2.2.1 (Toral Subalgebras) *A subalgebra of a semi-simple Lie algebra that consists only of semi-simple elements (with respect to the abstract Jordan Decomposition) is called a* ***toral subalgebra****.*

Remark 2.2.2 *We must note the deviation from standard literature, normally, root decomposition is done through the introduction of* ***Cartan subalgebras****.*
Humphreys proved the equivalence of maximal toral subalgebras and Cartan subalgebras over fields of characteristic 0 *on [Hum72, p.80], we chose to use toral subalgebras instead as it allows us to mantain our algebraic focus.*

The existence of toral subalgebras on our settings follows directly from Engel's Theorem (Theorem 1.2.5) and the abstract Jordan decomposition (Proposition 2.1.4). If there were no semi-simple elements on $\mathfrak{g}$, then all elements in $\mathfrak{g}$ would be ad-nilpotent, and therefore $\mathfrak{g}$ would be a nilpotent algebra, which cannot be semi-simple.

Proposition 2.2.3 *If* $\mathfrak{h} \subset \mathfrak{g}$ *is a toral subalgebra, then* $\mathfrak{h}$ *is abelian.*

Proof: We shall prove that $\mathrm{ad}(X)|_{\mathfrak{h}} = 0$ for any $X \in \mathfrak{h}$. If this is not the case, since X is diagonalizable, there exists an eigenvector H of $\mathrm{ad}(X)|_{\mathfrak{h}}$ with non-zero eigenvalue α. But in that case $\mathrm{ad}(X)H = \alpha H \Rightarrow \mathrm{ad}(H)X = -\alpha H \Rightarrow \mathrm{ad}(H)^2X = 0$. Meaning that $\mathrm{ad}(H)X$ is an eigenvector of $\mathrm{ad}(H)$ with eigenvalue 0.
On the other hand, $\mathrm{ad}(H)$ is diagonalizable, therefore there is a basis $\{Y_i\} \subset \mathfrak{h}$ of eigenvectors. $X = \sum \beta_i Y_i$ in this basis, applying $\mathrm{ad}(H)$ to this relation we see that $\mathrm{ad}(H)X$ is a sum of non-zero eigenvectors or 0, contradicting the fact that $\mathrm{ad}(H)X$ is an eigenvector with eigenvalue 0 or the fact that $H \neq 0$. □

Proposition 2.2.4 *If* $\mathfrak{h}$ *is a maximal toral subalgebra (with respect to inclusion), then it* *is possible to decompose* $\mathfrak{g}$ *with respect to* $\mathfrak{h}^*$, $\mathfrak{g} = \bigoplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}_\alpha$ *where*

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} | [H, X] = \alpha(H)X \text{ for all } H \in \mathfrak{h}\}.$$

Then these spaces satisfy $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$ *for any* $\alpha, \beta \in \mathfrak{h}^*$ *and moreover, if* $\alpha + \beta \neq 0$ *then* $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$.

Proof: Since all elements of $\mathrm{ad}(\mathfrak{h})$ are commuting semi-simple endomorphisms, then they are simultaneously diagonalizable with respect to a basis of $\mathfrak{g}$, in this case we can do the decomposition:

$$\mathfrak{g} = \bigoplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}_\alpha.$$

Now fix $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$, and $H \in \mathfrak{h}$:

$$[H, [X, Y]] = [[H, X], Y] + [X, [H, Y]] = \alpha(H)[X, Y] + \beta(H)[X, Y] = (\alpha + \beta)(H)[X, Y],$$

which implies that $[X, Y] \in \mathfrak{g}_{\alpha+\beta}$. For the remaining assertion:

$$\kappa([H, X], Y) = \alpha(H)\kappa(X, Y)$$
$$\kappa(X, [H, Y]) = \beta(H)\kappa(X, Y)$$
$$\kappa([H, X], Y) = -\kappa(X, [H, Y]) \Rightarrow (\alpha + \beta)(H)\kappa(X, Y) = 0.$$

□

Corollary 2.2.5 *The restriction of* κ *to* $\mathfrak{g}_0$ *is non-degenerate.*

Proof: If it is degenerate, let $X \in \mathfrak{g}_0$ be such that $\kappa(X, \mathfrak{g}_0) = 0$. In that case, by the previous relation $\kappa(X, \mathfrak{g}_\alpha) = 0$ for any $\alpha \neq 0$. Since $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$, then $\kappa(X, \mathfrak{g}) = 0$. But the Killing form of $\mathfrak{g}$ is non-degenerate (Theorem 1.3.7), a contradiction. □

Theorem 2.2.6 *Let* $\mathfrak{h}$ *be a maximal toral subalgebra, then* $\mathfrak{h} = \mathfrak{g}_0$.

Proof: We will proceed in steps:

(1) $\mathfrak{g}_0$ contains the nilpotent and semi-simple parts of its elements.
$X \in \mathfrak{g}_0$ implies that $\mathrm{ad}(X)H = 0$ for all $H \in \mathfrak{h}$, by Jordan decomposition properties, $\mathrm{ad}(S)$ and $\mathrm{ad}(N)$ must map $\mathfrak{h}$ to 0 (they are polynomials in $\mathrm{ad}(X)$).

(2) All semi-simple elements of $\mathfrak{g}_0$ lie in $\mathfrak{h}$.
If $S \in \mathfrak{g}_0$ is semi-simple and $[\mathfrak{h}, S] = 0$, then $\mathfrak{h} + \mathbb{F}S$ is a toral subalgebra, therefore $S \in \mathfrak{h}$ by maximality of $\mathfrak{h}$.

(3) The restriction of κ to $\mathfrak{h}$ is non-degenerate.
Let $H \in \mathfrak{h}$ be such that $\kappa(H,\mathfrak{h}) = 0$, if $N \in \mathfrak{g}_0$ is nilpotent, then $[N,H] = 0$ and $\mathrm{ad}(N)$ is nilpotent. $\mathrm{ad}(N)\mathrm{ad}(H)$ is therefore nilpotent and $\kappa(N,H) = \mathrm{Tr}(\mathrm{ad}(N)\mathrm{ad}(H)) = 0$. But then, for all $X = S + N \in \mathfrak{g}_0$, we have $\kappa(H,X) = 0$ since $S \in \mathfrak{h}$ by (2), contradicting Corollary 2.2.5 as $\mathfrak{h} \subseteq \mathfrak{g}_0$.

(4) $\mathfrak{g}_0$ is a nilpotent algebra.
If $S \in \mathfrak{g}_0$ is semi-simple, then $S \in \mathfrak{h}$ and therefore $[S,\mathfrak{g}_0] = 0$, implying that $\mathrm{ad}(S)$ is nilpotent in $\mathfrak{g}_0$. Now if $X = S+N$ is any element of $\mathfrak{g}_0$, then $\mathrm{ad}(X)$ is the sum of commuting nilpotent endomorphisms. Therefore $\mathrm{ad}(X)$ is nilpotent and by Engel's Theorem (Theorem 1.2.5) $\mathfrak{g}_0$ is nilpotent.

(5) $\mathfrak{h} \cap [\mathfrak{g}_0,\mathfrak{g}_0] = 0$.
Since $[\mathfrak{h},\mathfrak{g}_0] = 0$, then $\kappa(\mathfrak{h},[\mathfrak{g}_0,\mathfrak{g}_0]) = 0$ by associativity. Therefore, if $H \in [\mathfrak{g}_0,\mathfrak{g}_0]$ and $H \in \mathfrak{h}$, then $H = 0$ by (3).

(6) $\mathfrak{g}_0$ is abelian.
Otherwise $[\mathfrak{g}_0,\mathfrak{g}_0] \neq 0$, since $\mathfrak{g}_0$ is nilpotent, then $[\mathfrak{g}_0,\mathfrak{g}_0] \cap \mathfrak{z}(\mathfrak{g}_0) \neq 0$. Let X be an element in this intersection, its nilpotent part N is non-zero and also lies in $\mathfrak{z}(\mathfrak{g}_0)$ ($\mathrm{ad}(N)$ is a polynomial in $\mathrm{ad}(X)$), since $\mathrm{ad}(N)$ is nilpotent and commutes with $\mathfrak{g}_0$, then $\kappa(N,\mathfrak{g}_0) = 0$, contradicting Corollary 2.2.5.

(7) $\mathfrak{h} = \mathfrak{g}_0$
Otherwise, there exists a non-zero nilpotent element $N \in \mathfrak{g}_0$, but in that case since $\mathfrak{g}_0$ is abelian then for every $X \in \mathfrak{g}_0$, $\mathrm{ad}(N)\mathrm{ad}(X) = \mathrm{ad}(X)\mathrm{ad}(N)$ is a nilpotent endomorphism and therefore has trace 0, implying that $\kappa(N,\mathfrak{g}_0) = 0$ contradicting Corollary 2.2.5.

□

2.3. Finite-Dimensional Representations of $\mathfrak{sl}(2)$

The purpose of this section is to both exemplify the decomposition defined in the previous section, as well as construct tools to expand on root decomposition on the next section. The results from this section follow [Hum72, chpt 7].
Let $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{F})$, a traditional basis of this algebra is the following:

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The commutators in this basis satisfy the following relations:

$$[X,Y]=H,\ \ [H,X]=2X,\ \ [H,Y]=-2Y.$$

H is diagonal and acts diagonally in any $\mathfrak{g}$-module V (Proposition 2.1.5).
Because of this and the fact that $\mathbb{F}$ is algebraically closed, we have well defined eigenspaces with respect to H, denoted as $V_\lambda = \{v \in V \mid Hv = \lambda v\}$. If V is finite-dimensional, then $V = \bigoplus_{\lambda\in\mathbb{F}} V_\lambda$. Whenever $V_\lambda \neq 0$, we call λ a **weight** and V_λ a **weight space**.

Lemma 2.3.1 *If $v \in V_\lambda$, then $Xv \in V_{\lambda+2}$ and $Yv \in V_{\lambda-2}$.*

Proof: Since V is a $\mathfrak{g}$-module then $[A,B]v = A(Bv) - B(Av)$ for any $v \in V$ and $A,B \in \mathfrak{g}$, therefore:

$$H(Xv) = [H,X]v + X(Hv) = 2Xv + X(\lambda v) = (\lambda+2)Xv,$$

$$H(Yv) = [H,Y]v + Y(Hv) = -2Yv + Y(\lambda v) = (\lambda-2)Xv.$$

□

If we let V be a non-trivial finite-dimensional $\mathfrak{sl}(2,\mathbb{F})$-module, we know that there exists λ such that $V_\lambda \neq 0$ and $V_{\lambda+2} = 0$, we will call one of these weights maximal and any vector in V_λ a maximal vector. A direct consequence of this definition is that if v is a maximal vector, then $Xv = 0$.
We can determine the action of $\mathfrak{g}$ in a special subset of vectors, inspired by the idea of determining the action of X by the action of Y:

Lemma 2.3.2 *Let v_0 be a maximal vector of weight λ. Set $v_{-1} = 0$ and $v_i = \frac{1}{i!}Y^i v_0$ for $i \geq 0$, then:*

(a) $Hv_i = (\lambda - 2i)v_i$,

(b) $Yv_i = (i+1)v_{i+1}$,

(c) $Xv_i = (\lambda - i + 1)v_{i-1}$ *for* $i \geq 0$.

Proof:

(a) Follows directly from Lemma 2.3.1.

(b) Is a consequence of the definition, in fact:

$$v_{i+1} = \frac{1}{(i+1)!}Y^{i+1}v_0 = \frac{1}{i+1}Y\left(\frac{1}{i!}Y^i v_0\right) = \frac{1}{i+1}Yv_i.$$

(c) Since v_0 is maximal then $Xv_0 = 0$ and the result is valid, proceeding by induction:

$$\begin{aligned}
Xv_i &= \frac{1}{i}X(Yv_{i-1})\\
iXv_i &= [X,Y]v_{i-1} + Y(Xv_{i-1})\\
&= Hv_{i-1} + Y(Xv_{i-1})\\
&= (\lambda - 2(i-1))v_{i-1} + Y(\lambda - i + 2)v_{i-2}\\
&= (\lambda - 2(i-1))v_{i-1} + (\lambda - i + 2)(i-1)v_{i-1}\\
&= i(\lambda - i + 1)v_{i-1}.
\end{aligned}$$

□

Now if we consider V to be irreducible, we can explicitly classify V with respect to the set of v_i which in turn are completely determined by λ.

Theorem 2.3.3 *If V is irreducible, then the set of $\{v_i\}$ form a basis of V, λ is a positive integer and the number of vectors $\{v_i\}$ is precisely $\lambda + 1$.*

Proof: Since each v_i is an eigenvector of H with different eigenvalues (Lemma 2.3.1(a)), the set $\{v_i\}$ is linearly independent.
The span of the set $\{v_i\}$ is closed under the action of $\mathfrak{g}$. As this span is non-zero then it must be the whole V since it is an irreducible module.
Let m be the largest value such that $v_m \neq 0$ but $v_{m+1} = 0$ (V is non-trivial and finite-dimensional). Letting $i = m+1$ in Lemma 2.3.2(c) we find:

$$Xv_{m+1} = (\lambda - (m+1) + 1)v_m \iff 0 = (\lambda - m)v_m,$$

therefore $\lambda = m$, which is a positive integer, furthermore, $\{v_0, v_1, \cdots, v_m\}$ is a basis of V, and $\dim V = \lambda + 1$. □
We can summarize a classification of finite-dimensional irreducible $\mathfrak{sl}(2)$-modules using the previous theorem:

Theorem 2.3.4 (Classification of irreducible $\mathfrak{sl}(2)$-modules) *If V is an irreducible $\mathfrak{sl}(2)$-module of dimension $m+1$, then:*

(a) $V = V_{-m} \oplus V_{-m+2} \oplus \cdots \oplus V_{m-2} \oplus V_m$, each weight space with dimension one.

(b) V has a unique maximal weight and a unique maximal vector up to scalar multiples.

(c) The $\mathfrak{sl}(2)$-action is defined by the formulas in Lemma 2.3.2. In particular, for each natural number n, there exists a unique (up to isomorphism) n-dimensional irreducible $\mathfrak{sl}(2)$-module.

Since every $\mathfrak{sl}(2)$-module is the sum of irreducible modules (Weyl's Theorem 1.4.12), then:

Corollary 2.3.5 *Let V be any finite-dimensional $\mathfrak{sl}(2,\mathbb{F})$ module, then all the eigenvalues of H on V are integers and each occurs along with its negative an equal number of times. Moreover, in any decomposition of V as a direct sum of irreducible sub-modules, the number of modules in this decomposition is* $\dim(V_0)+\dim(V_1)$.

Proof: The first result is direct. Since every irreducible module is a sum of weight spaces with distance 2 from each other, then each irreducible sub-module must have weight 0 or 1, but not both. □

2.4. Root Decomposition

This section is based on [Hum72][chpt 8]. Let $\mathfrak{g}$ be a finite-dimensional semi-simple Lie algebra over an algebraically closed field $\mathbb{F}$ of characteristic 0.
Fix $\mathfrak{h}$ to be maximal toral subalgebra of $\mathfrak{g}$, and for $\alpha \in \mathfrak{h}^*$, define

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \mid [H,X] = \alpha(H)X \text{ for all } H \in \mathfrak{h}\}.$$

We will call a root, any $\alpha \in \mathfrak{h}^*\backslash\{0\}$ such that $\mathfrak{g}_\alpha \neq 0$, and denote the set of all roots by Φ.
Since the Killing form restricted to $\mathfrak{h}$ is non-degenerate, then for every $\alpha \in \mathfrak{h}^*$ we can find some $T_\alpha \in \mathfrak{h}$ that represents it, meaning $\alpha(H) = \kappa(T_\alpha, H)$.

Orthogonality

Reminding ourselves of the property of orthogonality between roots, meaning that if $\alpha+\beta \neq 0$ for elements in $\mathfrak{h}^*$ then $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ (Proposition 2.2.4).
With this we can summarize the first properties of root decomposition:

Proposition 2.4.1 *(a) Φ spans $\mathfrak{h}^*$.*

(b) If $\alpha \in \Phi$, then $-\alpha \in \Phi$.

(c) Let $\alpha \in \Phi$, $X \in \mathfrak{g}_\alpha$, and $Y \in \mathfrak{g}_{-\alpha}$, then: $[X,Y] = \kappa(X,Y)T_\alpha$.

(d) If $\alpha \in \Phi$, then $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is one dimensional.

(e) If $\alpha \in \Phi$, then $\alpha(T_\alpha) = \kappa(T_\alpha, T_\alpha) \neq 0$.

(f) If $\alpha \in \Phi$ and $X_\alpha \in \mathfrak{g}_\alpha$ is any non-zero element, then there exists $Y_\alpha \in \mathfrak{g}_\alpha$ such that $\{X_\alpha, [X_\alpha, Y_\alpha], Y_\alpha\}$ spans a three dimensional subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}(2)$.

(g) Let $\alpha \in \Phi$ and X_α, Y_α as above. If $H_\alpha = [X_\alpha, Y_\alpha]$ we have $H_\alpha = -H_{-\alpha} = \frac{2T_\alpha}{\kappa(T_\alpha, T_\alpha)}$.

Proof:

(a) If Φ fails to span $\mathfrak{h}^*$, then by duality we get a non-zero element $H \in \mathfrak{h}$ such that $\alpha(H) = 0$ for all $\alpha \in \Phi$. But in that case for any $X_\alpha \in \mathfrak{g}_\alpha$ this means $[H, X_\alpha] = \alpha(H)X_\alpha = 0$. This implies that $H \in \mathfrak{z}(\mathfrak{g})$, a contradiction.

(b) Suppose $\alpha \in \Phi$ and $-\alpha \notin \Phi$, then there is no element $\beta \in \Phi$ such that $\alpha + \beta = 0$, meaning that $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}) = 0$, contradicting the non-degeneracy of κ.

(c) Let $H \in \mathfrak{h}$ be arbitrary, then:

$$\kappa(H,[X,Y]) = \kappa([H,X],Y) = \alpha(H)\kappa(X,Y) = \kappa(H,T_\alpha)\kappa(X,Y) = \kappa(H,\kappa(X,Y)T_\alpha)$$
$$\kappa(H,[X,Y] - \kappa(X,Y)T_\alpha) = 0.$$

This in turn implies that $\mathfrak{h}$ is orthogonal to $[X,Y] - \kappa(X,Y)T_\alpha$, forcing $[X,Y] = \kappa(X,Y)T_\alpha$ since $[X,Y] \in \mathfrak{g}_0 = \mathfrak{h}$ and κ is non-degenerate in $\mathfrak{h}$.

(d) (*c*) shows that T_α spans $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$, provided it is not 0, and this space cannot be 0 because otherwise $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}) = 0$ and therefore $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}) = 0$.

(e) If $\alpha(T_\alpha) = 0$ choose $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$ such that $\kappa(X,Y) = 1$ (due to (*d*)). The subspace S spanned by $\{X, T_\alpha, Y\}$ satisfies $[X,Y] = T_\alpha$, $[X, T_\alpha] = 0$, and $[Y, T_\alpha] = 0$. Therefore S is a three-dimensional solvable Lie algebra isomorphic to $\mathrm{ad}(S) \subset \mathfrak{gl}(\mathfrak{g})$. $\mathrm{ad}([S,S])$ is nilpotent and $\mathrm{ad}(T_\alpha)$ is both semi-simple and nilpotent, implying that $\mathrm{ad}(T_\alpha) = 0 \Rightarrow T_\alpha = 0$, absurd.

(f) Let $H_\alpha = \frac{2T_\alpha}{\kappa(T_\alpha, T_\alpha)}$, then find $Y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $\kappa(X,Y) = \frac{2}{\kappa(T_\alpha, T_\alpha)}$. Therefore $[X_\alpha, Y_\alpha] = \kappa(X,Y)T_\alpha = H_\alpha$. Now $[H_\alpha, X_\alpha] = \alpha(H_\alpha)X_\alpha = \frac{2\alpha(T_\alpha)}{\kappa(T_\alpha, T_\alpha)}X_\alpha = 2X_\alpha$, and similarly for $[H, Y_\alpha] = -2Y_\alpha$.

(g) Remember that $T_{-\alpha}$ is defined as an element that satisfies $\kappa(T_{-\alpha}, H) = -\alpha(H)$ for any $H \in \mathfrak{h}$, but $\kappa(-T_\alpha, H) = -\kappa(T_\alpha, H) = -\alpha(H)$. Therefore $T_\alpha = -T_{-\alpha}$ and $H_\alpha = -H_{-\alpha}$.

□

Integrality

For each pair of roots $\alpha, -\alpha$ let S_α be a subalgebra isomorphic to $\mathfrak{sl}(2)$ constructed in Theorem 2.4.1. Several properties of its representations have been established, using those properties we can analyze some modules contained in $\mathfrak{g}$, given rise to the following:

Proposition 2.4.2 *Let $\alpha \in \Phi$ and $M_\alpha = \bigoplus_{c \in \mathbb{F}} \mathfrak{g}_{c\alpha}$, then M_α is a S_α-module and:*

(a) $\dim \mathfrak{g}_\alpha = 1$.

(b) If $c\alpha$ is a root, then $c = \pm 1$.

Proof: Since $[\mathfrak{g}_{c\alpha}, \mathfrak{g}_{c'\alpha}] \subset \mathfrak{g}_{(c+c')\alpha}$ and $S_\alpha \subset M_\alpha$, then M_α is a S_α-module via the adjoint representation.
Now the weights of H_α in M are $2c\alpha$ since $\alpha(H_\alpha) = 2$ and $[H_\alpha, X] = c\alpha(H_\alpha)X$ for any $X \in \mathfrak{g}_{c\alpha}$.
Note that $\ker \alpha$ is a sub-module of M with codimension 1 in $\mathfrak{h}$ complementary to H_α, therefore the weight 0 only occurs in $\ker \alpha$ and S_α, but S_α is irreducible and therefore the only even weights in M are $0, \pm 2$. This proves that twice a root can never be a root, but then half a root cannot be a root either, therefore there is no weight 1 in M. This in turn implies that $M = \ker \alpha \oplus S_\alpha$, in particular $\dim \mathfrak{g}_\alpha = 1$ and the only multiples of α which are roots are $\pm\alpha$. □

Proposition 2.4.3 *Let $\alpha, \beta \in \Phi$ such that $\beta \neq \pm\alpha$ and let:*

$$K_\alpha(\beta) = \mathfrak{g}_{\beta - r\alpha} \oplus \mathfrak{g}_{\beta-(r-1)\alpha} \oplus \cdots \oplus \mathfrak{g}_\beta \oplus \mathfrak{g}_{\beta+\alpha} \oplus \cdots \mathfrak{g}_{\beta+q\alpha},$$

where r is the largest integer such that $\beta - r\alpha$ is a root and q is the largest integer such that $\beta + q\alpha$ is a root. Then K_α is an S_α-module and :

(a) $\beta(H_\alpha) \in \mathbb{Z}$ and $\beta - \beta(H_\alpha)\alpha \in \Phi$.

(b) If i is such that $-r \leq i \leq q$, then $\beta + i\alpha \in \Phi$ and $\beta(H_\alpha) = r - q$.

Proof: Each root space is one dimensional, and none of the $\beta + i\alpha$ can be equal to 0 since $\beta \neq \pm\alpha$ and no other multiple of α is a root.

Now if $X \in \mathfrak{g}_{\beta+i\alpha}$ then $[H_\alpha, X] = (\beta + i\alpha)(H_\alpha)X = (\beta(H_\alpha) + 2i)X$, and therefore the only distinct weights of H_α are $\beta(H_\alpha) + 2i$. Since all of those weights must be integers (Proposition 2.3.5), then $\beta(H_\alpha) \in \mathbb{Z}$. Obviously not both 0 or 1 can occurs as weights, and because every root space is one dimensional, then K_α is irreducible $(\dim(K_{\alpha_0}) + \dim(K_{\alpha_1}) = 1)$.
Since every weight occurs along with its negative, then:

$$\beta(H_\alpha) - 2r = -(\beta(H_\alpha) + 2q) \Rightarrow \beta(H_\alpha) = r - q.$$

□

Rationality

The Killing form is non-degenerate (Theorem 1.3.7) and symmetric (Lemma 1.3.6), it is natural to think if there is a way to construct an inner product out of it. There is in fact a natural way to construct such inner product on roots.
Let $(\gamma, \delta) = \kappa(T_\gamma, T_\delta)$ for all $\gamma, \delta \in \mathfrak{h}^*$, $\{\alpha_1, \cdots, \alpha_\ell\}$ a basis consisting of roots and $\beta = \sum_{i=1}^{\ell} c_i \alpha_i \in \mathfrak{h}^*$.

Proposition 2.4.4 *If $\beta \in \Phi$, then $c_i \in \mathbb{Q}$.*

Proof: We know that

$$\beta(H_\alpha) = \kappa(T_\beta, H_\alpha) = \frac{2\kappa(T_\beta, T_\alpha)}{\kappa(T_\alpha, T_\alpha)} = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z}.$$

Using the definition of β with this fact in mind we get:

$$\frac{2(\beta, \alpha_j)}{(\alpha_j, \alpha_j)} = \sum_{i=1}^{\ell} c_i \frac{2(\alpha_j, \alpha_i)}{(\alpha_j, \alpha_j)}$$
$$\beta(H_{\alpha_j}) = \sum_{i=1}^{\ell} c_i \alpha_i(H_{\alpha_j}).$$

Therefore since this is a linear equation in c_i with integer coefficients solvable in $\mathbb{F}$, then it is solvable in $\mathbb{Q}$, implying that $c_i \in \mathbb{Q}$. □

Let $E_\mathbb{Q}$ be the space in $\mathbb{Q}$ spanned by the roots, we just proved that $\dim_\mathbb{Q}(E_\mathbb{Q}) = \ell$. Moreover:

Proposition 2.4.5 *The form $(\cdot, \cdot)$ is naturally extended to $E_\mathbb{Q}$ and is positive definite.*

Proof: Notice that:

$$(\lambda,\mu)=\kappa(T_\lambda,T_\mu)=\mathrm{Tr}(\mathrm{ad}(T_\lambda)\mathrm{ad}(T_\mu))=\sum_{\alpha\in\Phi}\alpha(T_\lambda)\alpha(T_\mu)=\sum_{\alpha\in\Phi}(\alpha,\lambda)(\alpha,\mu).$$

In particular $(\beta,\beta)=\sum_{\alpha\in\Phi}(\alpha,\beta)^2$, multiplying this relation by $\frac{4}{(\beta,\beta)^2}$ we get

$$\frac{4}{(\beta,\beta)}=\sum_{\alpha\in\Phi}\left(\frac{2(\alpha,\beta)}{(\beta,\beta)}\right)^2\in\mathbb{Z}.$$

Therefore $(\beta,\beta)\in\mathbb{Q}$ and in turn, since $\frac{2(\alpha,\beta)}{(\beta,\beta)}\in\mathbb{Z}$ then $(\alpha,\beta)\in\mathbb{Q}$, proving that the form is well defined in $E_{\mathbb{Q}}$ now since $(\beta,\beta)=\sum_{\alpha\in\Phi}(\alpha,\beta)^2$ it is positive definite as the sum of squares of rational numbers. □

Summary

Let E be the real vector space extending the base field of $E_{\mathbb{Q}}$ from $\mathbb{Q}$ to $\mathbb{R}$, then the following properties are satisfied:

(a) Φ spans E, and $0\notin\Phi$.

(b) If $\alpha\in\Phi$, the only other multiple of α in Φ is $-\alpha$.

(c) If $\alpha,\beta\in\Phi$, then $\beta-\dfrac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha\in\Phi$.

(d) If $\alpha,\beta\in\Phi$, then $\frac{2(\beta,\alpha)}{(\alpha,\alpha)}\in\mathbb{Z}$.

The pair (E,Φ) is called a **root system**.

We want to study root systems independently of Lie algebras, using the summary of the previous chapter as our definition. The classification of root systems is what allowed us to classify all finite-dimensional simple Lie algebras over $\mathbb{C}$.
This independent study allows us to better understand its underlying properties, the geometrical implications of Rationality (Theorem 2.4.5). It also serves as an introduction to the abstract theory of weights used extensively in representatoin theory.

3.1. Axiomatic

This section focuses on providing some initial study to the structure of root systems, including a non-unique partition of the root system, as well as natural bases to work with them. The results presented here are based on [Hum72][chpt 9 & 10].

Definition 3.1.1 (Root System) *Given an euclidean space E with inner product denoted by $\langle\ ,\ \rangle$ and a finite subset $\Phi \subset E$, then the pair (E, Φ) is said to be a* ***root system*** *if:*

(a) Φ *spans* E*, and* $0 \notin \Phi$.

(b) If $\alpha \in \Phi$ *then the only other multiple of* α *in* Φ *is* $-\alpha$.

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

(d) If $\alpha, \beta \in \Phi$*, then* $\frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}$.

To reduce notation, we will define $\frac{2\alpha}{\langle \alpha, \alpha \rangle}$ *as* $\alpha^\vee$ *and furthermore we will define as* s_α*, the linear transformation*

$$s_\alpha(v) = v - \langle v, \alpha^\vee \rangle \alpha,$$

which is the reflection with respect to the hyperplane defined by α*. With this in mind we can reduce (c) and (d) to:*

(c'). If $\alpha, \beta \in \Phi$, then $s_\alpha(\beta) \in \Phi$.

(d'). If $\alpha, \beta \in \Phi$, then $\langle \beta, \alpha^\vee \rangle \in \mathbb{Z}$.

We call the operation $(\alpha, \beta) \mapsto \langle \beta, \alpha^\vee \rangle$ the **Cartan product**. It is important to note that axiom (c') has a nice geometric interpretation: Φ is closed under reflections with respect to hyperplanes defined by roots.

Axiom (d) allows us to deduce some simple integer restrictions to a root system:

Proposition 3.1.2 *Let $\alpha, \beta \in \Phi$ with $\|\beta\| \geq \|\alpha\|$ (inner product norm). If θ is the angle between them, then the Cartan product is restricted to those in the following table:*

Table 3.1.: Possible values of the Cartan product

$\langle \alpha, \beta^\vee \rangle$	$\langle \beta, \alpha^\vee \rangle$	0
0	0	90°
±1	±1	60° or 120°
±1	±2	45° or 135°
±1	±3	30° or 150°
±2	±2	0° or 180°

Proof: Consider that $\langle \alpha, \beta \rangle = \|\alpha\| \|\beta\| \cos\theta$ where θ is the angle between the roots, then since $\|\alpha^\vee\| = \frac{2}{\|\alpha\|}$ then:

$$\langle \alpha, \beta^\vee \rangle \langle \beta, \alpha^\vee \rangle = \frac{4\|\alpha\|\|\beta\|}{\|\beta\|\|\alpha\|} \cos^2\theta = 4\cos^2\theta$$

As $0 \leq \cos^2\theta \leq 1$, the only possibilities are those for which $\langle \alpha, \beta^\vee \rangle \langle \beta, \alpha^\vee \rangle \leq 4$.
Which are all the cases present in Table 3.1, excluding the cases $(0, n)$ with $n \neq 0$ and $(1, 4)$.
The case $(0, n)$, $n \neq 0$ can be excluded. If $\langle \alpha, \beta^\vee \rangle = 0$, then $\langle \alpha, \beta \rangle = 0$ and therefore $\langle \beta, \alpha^\vee \rangle = 0$.
The case $(1, 4)$ can also be excluded. In fact, $4\cos^2\theta = 4$ implies $\theta = 0$ or $\theta = 180°$. Therefore β is a multiple of α, meaning that $\beta = \pm\alpha$ and therefore $\langle \alpha, \beta^\vee \rangle = \pm 2$.
Notice that the last case $(\pm 2, \pm 2)$ only occurs on proportional roots. □
An important corollary of this restriction is as follows:

Corollary 3.1.3 *Let α, β be non-proportional roots. If $\langle \alpha, \beta \rangle > 0$, then $\alpha - \beta$ is a root and if $\langle \alpha, \beta \rangle < 0$, then $\alpha + \beta$ is a root.*

Proof: Since $\langle \alpha, \beta \rangle$ is positive, Table 3.1 shows that $\langle \alpha, \beta^\vee \rangle = 1$ or $\langle \beta, \alpha^\vee \rangle = 1$.
If $\langle \alpha, \beta^\vee \rangle = 1$, then $s_\beta(\alpha) = \alpha - \beta$ and $\alpha - \beta$ is a root (Axiom (c)).
If $\langle \beta, \alpha^\vee \rangle = 1$, then $\beta - \alpha$ is a root and therefore $-(\beta - \alpha) = \alpha - \beta$ is a root (Axiom (a)).
If $\langle \alpha, \beta \rangle$ is negative, then $\langle \alpha, -\beta \rangle$ is positive and therefore $\alpha - (-\beta) = \alpha + \beta$ is a root. □

Following this, we try to partition the space E based on Axiom (b).

Definition 3.1.4 *Elements of $R := \{v \in E | \langle \alpha, v \rangle \neq 0$ for all $\alpha \in \Phi\}$ are called **regular**.*
*We call elements in $\Phi^+(\gamma) := \{\alpha \in \Phi | \langle \alpha, \gamma \rangle > 0\}$ **positive roots**.*
*Similarly, elements in $\Phi^-(\gamma) := \{\alpha \in \Phi | \langle \alpha, \gamma \rangle < 0\}$ are **negative roots**.*
*Furthermore, with respect to this partition, we call a positive root **decomposable** if $\alpha = \beta_1 + \beta_2$ for $\beta_1, \beta_2 \in \Phi^+(\gamma)$, and **indecomposable** if it is not decomposable.*

Theorem 3.1.5 (Root System Basis) *The set $\Delta(\gamma)$ of indecomposable elements is a basis of E and every element in $\Phi^+(\gamma)$ is in the $\mathbb{Z}_+$-span of $\Delta(\gamma)$.*

Proof: We proceed by steps:

(1) Every element in $\Phi^+(\gamma)$ is in the $\mathbb{Z}_+$-span of $\Delta(\gamma)$.
Otherwise, let β be an element that cannot be written in this way with $\langle \gamma, \beta \rangle$ as small as possible. Obviously $\beta \notin \Delta(\gamma)$, in that case $\beta = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in \Phi^+(\gamma)$. Since $\langle \gamma, \beta \rangle = \langle \gamma, \beta_1 \rangle + \langle \gamma, \beta_2 \rangle$ and each of $\langle \gamma, \beta_1 \rangle$ and $\langle \gamma, \beta_2 \rangle$ are positive, then β_1, β_2 must be in the $\mathbb{Z}_+$-span of $\Delta(\gamma)$ to avoid contradicting the minimality of β. Then $\beta = \beta_1 + \beta_2$ is in the $\mathbb{Z}_+$-span of $\Delta(\gamma)$.

(2) $\Delta(\gamma)$ spans E.
The previous item shows that $\Delta(\gamma)$ spans $\Phi^+(\gamma)$, through axiom (b) also spans Φ, finally, through axiom (a) spans E.

(3) If α, β are distinct elements in $\Delta(\gamma)$, then $\langle \alpha, \beta \rangle \leq 0$.
Otherwise, since β clearly cannot be $-\alpha$, then $\alpha - \beta$ is in Φ (Corollary 3.1.3). Therefore either $\alpha - \beta$ or $\beta - \alpha$ are in Φ^+. If $\alpha - \beta \in \Phi^+$ then $\alpha = \beta + (\alpha - \beta)$ and if $\beta - \alpha \in \Phi^+$ then $\beta = \alpha + (\beta - \alpha)$, contradicting the fact that α and β are indecomposable.

(4) $\Delta(\gamma)$ is a linearly independent set.
Suppose this is not the case and let $\{r_\alpha\}$ be such that $\sum_{\alpha \in \Delta(\gamma)} r_\alpha \alpha = 0$, divide this sum for the cases in which $r_\alpha > 0$ (call it Δ_1) and $r_\alpha < 0$ (call it Δ_2), then:

$$\sum_{\alpha \in \Delta_1} r_\alpha \alpha = \sum_{\beta \in \Delta_2} -r_\beta \beta = \epsilon \neq 0,$$

but in this case $\langle \epsilon, \epsilon \rangle = \sum_{\alpha,\beta} -r_\alpha r_\beta \langle \alpha, \beta \rangle > 0$, $r_\alpha > 0$, $r_\beta < 0$, and $\langle \alpha, \beta \rangle \leq 0$, a contradiction.

□

Remark 3.1.6 *Part 4. of the proof of Theorem 3.1.5 implies that any set in an euclidean space that satisfy pairwise $\langle \alpha, \beta \rangle < 0$ is linearly independent.*

From this point forward we fix a basis $\Delta = \Delta(\gamma)$ and call its elements **simple**.

Lemma 3.1.7 *If α is a positive root but not simple, then $\alpha - \beta$ is a positive root for some $\beta \in \Delta$. In particular, every positive root can be written as $\alpha_1 + \cdots + \alpha_i$ with each $\alpha_k \in \Delta$ not necessarily distinct, in such a way that each partial sum is a root.*

Proof: If $\langle \alpha, \beta \rangle \leq 0$ for every $\beta \in \Delta$, Remark 3.1.6 applies and $\Delta \cup \{\alpha\}$ is linearly independent, which is absurd. Therefore, there exists a $\beta \in \Delta$ such that $\langle \alpha, \beta \rangle > 0$, and $\alpha - \beta \in \Phi$ (Corollary 3.1.3). $\alpha - \beta$ is a positive root, since at least one coordinate of α with respect to Δ remains positive after the subtraction of β ($\alpha \neq \beta$). □

Lemma 3.1.8 *Let α be a simple root, then s_α permutes all positive roots except α.*

Proof: Let $\beta \in \Phi^+ - \{\alpha\}$, then we can write

$$\beta = \sum_{\delta \in \Delta} k_\delta \delta, \quad k_\delta \in \mathbb{Z}_+, \ \ k_r \neq 0 \text{ for some } r \neq \alpha.$$

As $\alpha \in \Delta$, the coefficient k_r is unchanged in $s_\alpha(\beta) = \beta - \langle \beta, \alpha^\vee \rangle \alpha$. Therefore $s_\alpha(\beta)$ has at least one positive coefficient and is a positive root. Moreover since s_α is bijective and $s_\alpha(-\alpha) = \alpha$ then $s_\alpha(\beta) \neq \alpha$. □

Corollary 3.1.9 *Set $\rho := \frac{1}{2} \sum_{\beta \in \Phi^+} \beta$. Then $s_\alpha(\rho) = \rho - \alpha$ for all $\alpha \in \Delta$.*

Proof: Let $\Phi_\alpha = \Phi^+ - \{\alpha\}$ and remember that $s_\alpha(\alpha) = -\alpha$

$$\begin{aligned} s_\alpha(\rho) &= \frac{1}{2} \sum_{\beta \in \Phi^+} s_\alpha(\beta) \\ &= \frac{1}{2} \sum_{\beta \in \Phi_\alpha} s_\alpha(\beta) + \frac{1}{2} s_\alpha(\alpha) \\ &= \frac{1}{2} \sum_{\beta \in \Phi_\alpha} \beta - \frac{1}{2}\alpha \\ &= \rho - \frac{1}{2}\alpha - \frac{1}{2}\alpha = \rho - \alpha. \end{aligned}$$

□

3.2. Weyl Group

The structure of root systems is deeply connected to the structure of reflection groups, as within any root system there is a clear reflection group: the group generated by all root reflections.

Definition 3.2.1 *The group generated by all of the root reflections is called the* ***Weyl group****:* $\mathcal{W} = \langle s_\alpha | \alpha \in \Phi \rangle$.

It is clear that this group is finite as a subgroup of root permutations. Since reflections preserve the inner product, so does $\mathcal{W}$. Therefore $\mathcal{W}$ can be seen as a finite subgroup of orthogonal transformations of E.

The purpose of this section is to analyze the general structure of this group, its action on the root system, and to which degree this group defines a root system. This section follows [Hum72][sec 10.3].

Another important point to consider is that $\mathcal{W}$ is a normal subgroup of Aut $\Phi = \{T \in \mathfrak{gl}(V) | T(\Phi) = \Phi \text{ and } \langle \alpha, \beta^\vee \rangle = \langle T(\alpha), T(\beta)^\vee \rangle\}$. In fact:

Proposition 3.2.2 *If* $T \in Aut\ \Phi$*, then* $Ts_\alpha T^{-1} = s_{T(\alpha)}$.

Proof: Let $\beta \in \Phi$ be any root, then:

$$(Ts_\alpha T^{-1})(\beta) = Ts_\alpha(T^{-1}\beta) = T(T^{-1}\beta - \langle T^{-1}\beta, \alpha, \alpha^\vee \rangle) = \beta - \langle \beta, T(\alpha)^\vee \rangle T(\alpha) = s_{T(\alpha)}(\beta).$$

Since Φ spans E, the result follows. □

Lemma 3.2.3 *Let* $\alpha_1, \cdots, \alpha_n \in \Delta$ *(not necessarily distinct). Write* $s_{\alpha_i} = s_i$*, if* $s_1 \cdots s_{n-1}(s_n)$ *is a negative root, then there exists some* $1 \leq t < n$ *such that* $s_1 \cdots s_n = s_1 \cdots s_{t-1}s_{t+1} \cdots s_{n-1}$.

Proof: Write $\beta_i = s_{i+1} \cdots s_{n-1}(\alpha_n)$ for $1 \leq i \leq n-2$ and $\beta_{n-1} = \alpha_n$. Since β_0 is negative by hypothesis and β_{n-1} is positive, we can find the smallest index t such that β_t is positive.

But then $\sigma_t \beta_t = \beta_{t-1}$ which is negative, this forces $\beta_t = \alpha_t$ by Lemma 3.1.8, considering Proposition 3.2.2 for $T \in \mathcal{W}$ we have that:

$$s_{T(\alpha)} = Ts_\alpha T^{-1},$$

in particular, since $\alpha_t = s_{t+1} \cdots s_{n-1}(\alpha_n)$:

$$s_t = (s_{t+1} \cdots s_{n-1}) s_n (s_{n-1} \cdots s_{t+1}).$$

□

Corollary 3.2.4 *If $s \in \mathcal{W}$ is written as a product of simple roots $s = s_1 \cdots s_t$, with t as small as possible, then $s(\alpha_t)$ is negative.*

Recall that basis of Φ are determined by regular elements. That is, any basis of Φ is of the form $\Delta(\gamma)$ for some regular element γ.

The next section of results focuses on proving that the Weyl group acts transitively on bases of Φ. Moreover, we can restrict the set of generators of $\mathcal{W}$ to a basis $\Delta(\gamma)$ of Φ.

- For every two bases $\Delta(\gamma)$ and $\Delta(\gamma')$ of Φ there exists one, and only one, $s \in \mathcal{W}$ such that $s(\Delta(\gamma)) = \Delta(\gamma')$.
- If $\gamma \in R$, then $\mathcal{W} = \langle s_\alpha : \alpha \in \Delta(\gamma) \rangle$.

For this purpose, fix $\gamma \in R$ and let $\mathcal{W}'(\gamma) = \langle s_\alpha : \alpha \in \Delta(\gamma) \rangle$.

Theorem 3.2.5 *If $\gamma' \in R$ is another regular element, then there exists $s \in \mathcal{W}'(\gamma)$ such that $\langle s(\gamma'), \alpha \rangle > 0$ for all $\alpha \in \Delta(\gamma)$. In particular there exists $s' \in \mathcal{W}'(\gamma)$ such that $s'(\Delta(\gamma)) = \Delta(\gamma')$.*

Proof: Consider ρ as defined in Corollary 3.1.9 and choose $s \in \mathcal{W}'$ such that $\langle s(\gamma'), \rho \rangle$ is as big as possible. For any $\alpha \in \Delta(\gamma)$, we have $s_\alpha s \in \mathcal{W}'$, and the choice of s implies:

$$\begin{aligned} \langle s(\gamma'), \rho \rangle \geq \langle s_\alpha s(\gamma'), \rho \rangle &= \langle s(\gamma'), s_\alpha(\rho) \rangle \\ &= \langle s(\gamma'), \rho - \alpha \rangle \\ &= \langle s(\gamma'), \rho \rangle - \langle s(\gamma'), \alpha \rangle. \end{aligned}$$

$\langle s(\gamma'), \alpha \rangle \geq 0$, but it cannot be equal to 0: $\langle s(\gamma'), \alpha \rangle = \langle \gamma', s^{-1}(\alpha) \rangle$ and γ' is regular.
Let $s' = s^{-1}$, we want to prove that $s'(\Delta(\gamma))$ is irreducible with respect to $\Phi^+(\gamma')$.
Since $\langle \gamma', s'(\alpha) \rangle > 0$ for any $\alpha \in \Delta(\gamma)$, the same is true for $\Phi^+(\gamma)$, proving that $\Phi^+(\gamma') = s'(\Phi^+(\gamma))$ (they have the same cardinality).
Now suppose that for some $\alpha \in \Delta(\gamma)$, $s'(\alpha)$ is not irreducible with respect to γ'. Then there exists $\beta_1, \beta_2 \in \Phi_+(\gamma)$ such that $s'(\alpha) = s'(\beta_1) + s'(\beta_2)$.
This implies that $\alpha = \beta_1 + \beta_2$, contradicting the fact that α is an irreducible root. □

Lemma 3.2.6 *Given any root α, there exists a regular element γ' such that $\alpha \in \Delta(\gamma')$.*

Proof: For a root $\beta \in \Phi$, denote by $\beta^\perp = \{v \in E : \langle \beta, v \rangle = 0\}$, we want to find an element on the hyperplane $\alpha^\perp$ that is not on any other hyperplane $\beta^\perp$, $\beta \neq \pm\alpha$. If this element does not exist, we have:

$$\alpha^\perp = \alpha^\perp \cap \bigcup_{\beta \in \Phi \setminus \{\pm\alpha\}} \beta^\perp = \bigcup_{\beta \in \Phi \setminus \{\pm\alpha\}} (\alpha^\perp \cap \beta^\perp),$$

but this in turn implies that the finite union of proper subspaces $\alpha^\perp \cap \beta^\perp$ is the space $\alpha^\perp$ which is an absurd as $\mathbb{R}$ is infinite (see B.0.2). We can find γ' close enough to γ in such a way that $\langle \alpha, \gamma' \rangle > |\langle \beta, \gamma' \rangle| > 0$ for every root $\beta \neq \pm\alpha$, which in turn implies that $\alpha \in \Delta(\gamma')$. □

Corollary 3.2.7 *Given any root α, there exists $s \in \mathcal{W}'$ such that $s(\alpha) \in \Delta(\gamma)$.*

Proof: Let $\gamma' \in R$ be such that $\alpha \in \Delta(\gamma')$ (Lemma 3.2.6) and let $s \in \mathcal{W}'$ be such that $s(\Delta(\gamma')) = \Delta(\gamma)$ (Theorem 3.2.5), then $s(\alpha) \in \Delta(\gamma)$. □

Proposition 3.2.8 $\mathcal{W} = \mathcal{W}'(\gamma)$ *for any* $\gamma \in R$.

Proof: We will prove that s_α for $\alpha \in \Phi$ is an element of $\mathcal{W}'(\gamma)$, using Corollary 3.2.7 then there exists $s \in \mathcal{W}'$ such that $s(\alpha) \in \Delta(\gamma)$, call $s(\alpha) = \beta$, then $s^{-1}(\beta) = \alpha$, but on the other hand (Proposition 3.2.2) $s_\alpha = s_{s^{-1}\beta} = s^{-1} s_\beta s \in \mathcal{W}'$. □

Proposition 3.2.9 *The only $s \in \mathcal{W}$ such that $s(\Delta) = \Delta$ in $\mathcal{W}$ is* 1.

Proof: Suppose $s(\Delta) = \Delta$ but $s \neq 1$. By Corollary 3.2.4, $s(\alpha_k)$ is negative for some k, a contradiction. □

3.3. Classification of Root Systems

The properties of the Weyl group proved in the previous section allows us to talk about an independence of bases, and reach a classification of root systems. This section follows [Hum72, chpt 11]

Definition 3.3.1 (Irreducible Root Systems) *A root system Φ is called reducible if there exists two proper subsets Φ_1 and Φ_2 of Φ with $\langle \Phi_1, \Phi_2 \rangle = 0$ and $\Phi_1 \cup \Phi_2 = \Phi$.*
A root system will be called irreducible if it is not reducible.

One important property of the irreducibility on root systems is that this property can be determined by any basis:

Proposition 3.3.2 *Let Φ be a root system with basis Δ: If Δ cannot be partitioned into two proper orthogonal subsets, then Φ is irreducible.*

Proof: Let $\Phi = \Phi_1 \cup \Phi_2$, then $\Delta = (\Delta \cap \Phi_1) \cup (\Delta \cap \Phi_2)$ which is a partition by orthogonal subsets, proper unless Δ is contained in one of them, let this be Φ_1 without loss of generality, but that implies:

$$\langle \Delta, \Phi_2 \rangle = 0 \Rightarrow \langle E, \Phi_2 \rangle = 0 \Rightarrow \Phi_2 = 0,$$

a contradiction, showing that if the basis is irreducible, so is the root system.
The converse is true, although the proof is not entirely trivial and shall not be covered for the purposes of this section. □
We already saw that every semi-simple Lie algebra has an associated root system, less clear though, is the fact that every root system defines a Lie algebra, and that the reducibility of this root system is directly related to the simplicity of this algebra, we shall summarize these results of equivalence by the following:

Definition 3.3.3 *Two root systems (Φ, V) and (Ψ, W) are said isomorphic if there is a linear isomorphism $T : V \to W$ satisfying $T(\Phi) = \Psi$ and $\langle T(\alpha), T(\beta)^\vee \rangle = \langle \alpha, \beta^\vee \rangle$, for all $\alpha, \beta \in \Phi$.*

Theorem 3.3.4 ([Hum72] p.96-101) *Every root system Φ with basis $\Delta = \{\alpha_1, \cdots, \alpha_\ell\}$ defines a Lie algebra $\mathfrak{g}$ generated by $\{X_i, H_i, Y_i, \alpha_i \in \Delta\}$ satisfying the relations*

(a) $[H_i, H_j] = 0$;

(b) $[X_i, Y_i] = H_i, [X_i, Y_j] = 0$ *if* $i \neq j$;

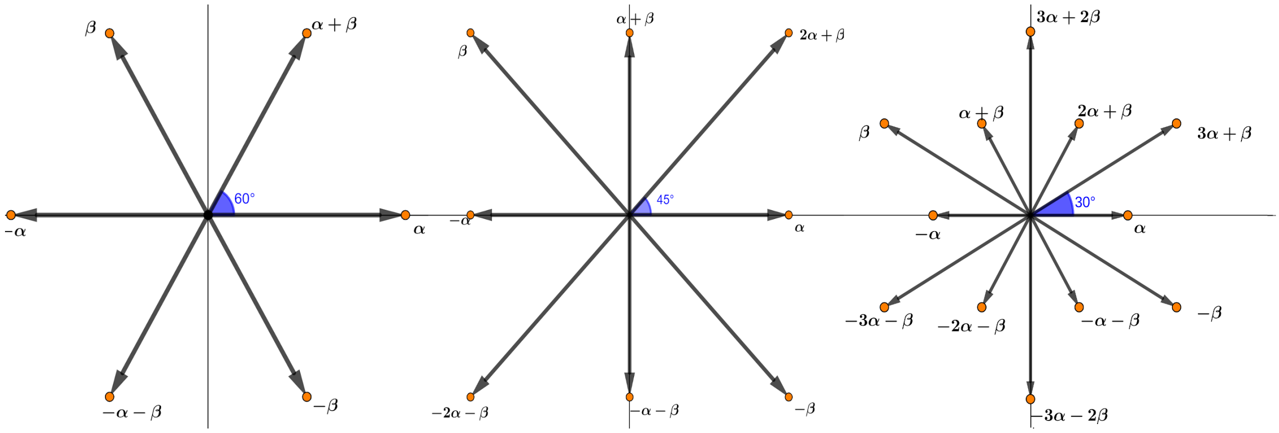
(c) $[H_i, X_j] = \langle \alpha_j, \alpha_i{}^\vee \rangle X_j$ *and* $[H_i, Y_j] = -\langle \alpha_j, \alpha_i{}^\vee \rangle Y_j$;

(d) $ad_{X_i}^{\langle \alpha_j, \alpha_i{}^\vee \rangle + 1}(X_j) = 0$;

(e) $ad_{Y_i}^{\langle \alpha_j, \alpha_i{}^\vee \rangle + 1}(Y_j) = 0$.

If $\mathfrak{g}$ is any semi-simple Lie algebra, and its root system Φ is irreducible then $\mathfrak{g}$ is simple. Moreover, if two Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ have isomorphic root systems then they are isomorphic.

Figure 3.1.: Example of Irreducible Root Systems



This motivates the classification of irreducible root systems, which will be done through their rank (number of elements in its base) and the respective Cartan product. Given a root system Φ with basis $\Delta = \{\alpha_1, \cdots, \alpha_\ell\}$, its **Cartan matrix** is defined by the matrix:

$$C_\Phi = (\langle \alpha_i, {\alpha_j}^\vee \rangle)_{ij} \in M_{\ell \times \ell}(\mathbb{Z}).$$

For example the Cartan matrices for the root systems presented in Figure 3.1 are given by:

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$$

Proposition 3.3.5 *If (Φ, V) and (Ψ, W) are root systems with same Cartan matrix C then they are isomorphic. That is, the Cartan matrix defines the root system up to isomorphism.*

Proof: Let $\Delta_\Phi = \{\alpha_1, \cdots, \alpha_\ell\}$ and $\Delta_\Psi = \{\beta_1, \cdots, \beta_\ell\}$ be the basis from which the Cartan matrix is defined, then the linear morphism:

$$\begin{aligned} T : V &\to W \\ \alpha_i &\mapsto \beta_i \end{aligned}$$

is an isomorphism. It is trivially a linear isomorphism as it sends a basis of V onto a basis of W, now letting $\alpha \in \Delta$ be arbitrary, one can see that the diagram given by:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow{\scriptstyle s_\alpha} & & \downarrow{\scriptstyle s_{T(\alpha)}} \\ V & \xrightarrow{T} & W \end{array}$$

commutes through the relation:

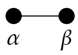
$$s_{T(\alpha)}(T(\gamma)) = T(\gamma) - \langle T(\gamma), T(\alpha)^\vee \rangle T(\alpha) = T(\gamma) - \langle \gamma, \alpha^\vee \rangle T(\alpha) = T(s_\alpha(\gamma))$$

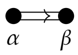
for any $\gamma \in \Delta$. This in turn implies that the Weyl groups are isomorphic as they are generated by $\{s_\alpha, \alpha \in \Delta\}$ and $\{s_{T(\alpha)}, \alpha \in \Delta\}$ respectively, through the mapping $s \mapsto T \circ s \circ T^{-1}$.
Since the Weyl group acting on the basis generates the whole root system we find $T(\Phi) = \Psi$. Moreover T preserves the Cartan product as it is the unique coefficient $\langle T(\alpha), T(\beta)^\vee \rangle T(\alpha) = T(\beta) - s_{T(\alpha)}(T(\beta))$. □

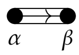
This result shows us that rather than classify all root systems, it is sufficient to classify Cartan matrices. This classification is usually done through Dynkin diagrams.

Definition 3.3.6 *The **Dynkin diagram** of Φ with respect to a basis Δ with an ordering $\{\alpha_1, \cdots, \alpha_\ell\}$ is a graph having ℓ-vertices, with the i-th vertex joined to the j-th vertex with $\langle \alpha_i, \alpha_j{}^\vee \rangle$ edges. If $\|\alpha_i\| \neq \|\alpha_j\|$, all of these edges are directed as to point to the larger of the two roots (with respect to $\|\cdot\|$).*

Example 3.3.7 *The root systems in Figure 3.1 have the respective Dynkin diagrams, with respect to the basis $\{\alpha, \beta\}$:*

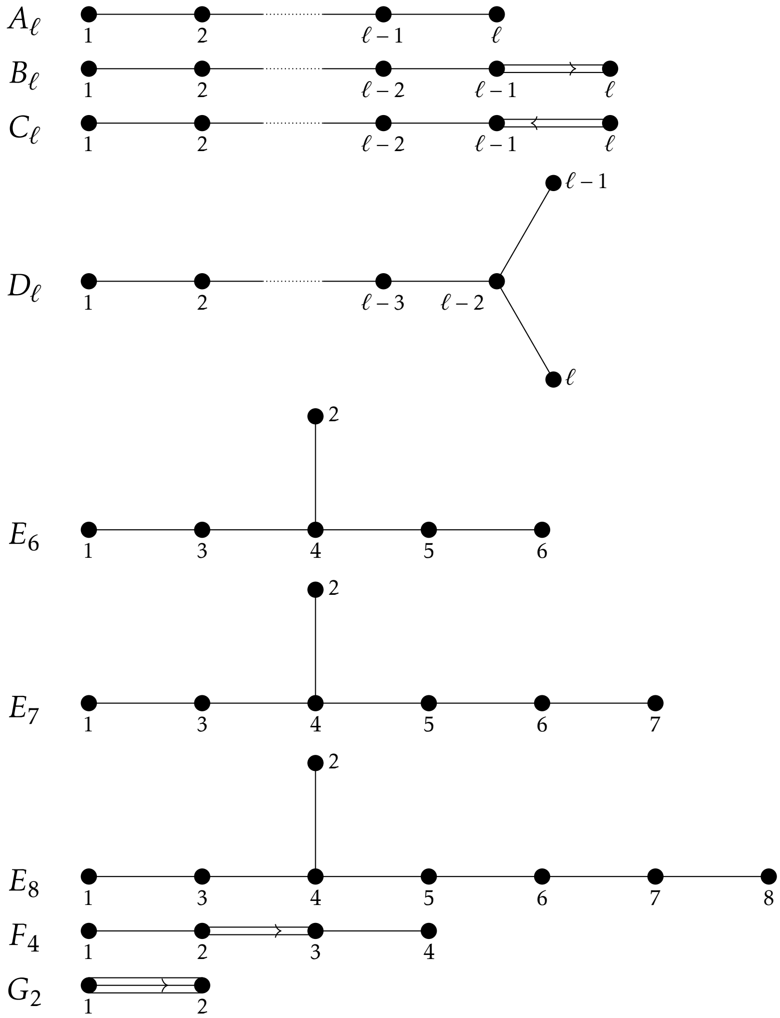


, 

, *and* 

It is clear that the Dynkin diagram will be connected if and only if the underlying root system is irreducible.
It is also clear that a Dynkin diagram uniquely defines the Cartan matrix for the underlying root system given the restrictions of the Cartan product (Table 3.1).
All of these Dynkin diagrams have been classified into 4 infinite families of root systems and 5 exceptional root systems. One can see [Hum72, p.57] for a full proof of this classification.

Table 3.2.: All possible connected Dynkin diagrams



4 Universal Enveloping Algebras

The universal enveloping algebra of a Lie algebra is one of the more important tools in representation theory, it is an associative algebra that contains an isomorphic copy of $\mathfrak{g}$ and extends all its representations. This algebra provides us many tools to the study of representations of Lie algebras, for example, a general construction for Casimir operators (Proposition 1.4.11), visualizing representations as modules over rings, and commutation identities through the Poincaré-Birkhoff-Witt theorem.

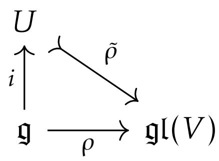
4.1. Universal Enveloping Algebras

Only in this section, we let $\mathfrak{g}$ be an arbitrary Lie algebra over any field. This section was based on [Mar09, sec 10.1].

Definition 4.1.1 (Universal Enveloping Algebra) *A pair (U, i), U being an associative algebra with unity and $i : \mathfrak{g} \to U$ an algebra morphism is an universal enveloping pair of a Lie algebra $\mathfrak{g}$ if:*

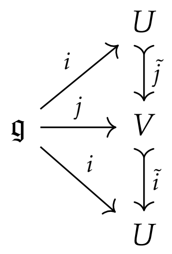
1. *It is enveloping, meaning that $i : \mathfrak{g} \to U$ is injective and its image $i(\mathfrak{g})$ generates U as an algebra.*
2. *It is universal. If $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation, then there exists a unique morphism of associative algebras $\tilde{\rho} : U \to \mathfrak{gl}(V)$ satisfying:*

$$\tilde{\rho}(i(X)) = \rho(X) \text{ for all } X \in \mathfrak{g}.$$



The idea of an universal object implies uniqueness, so before constructing the enveloping algebra explicitly, let us check uniqueness.

Remark 4.1.2 *If there are two pairs (U, i) and (V, j) that are universal and enveloping then there is bijective morphism between then:*



The idea to construct the universal enveloping algebra for any Lie algebra comes from the fact that the tensor algebra is an universal associative algebra for a vector space. So we just need to reduce its structure to contain the Lie algebra bracket in some way.

Proposition 4.1.3 *If $\mathfrak{g}$ is a Lie algebra and I is the ideal generated by*

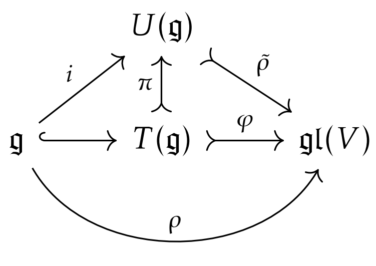
$$\{[X,Y] - X \otimes Y + Y \otimes X \mid X, Y \in \mathfrak{g}\}$$

in the tensor algebra $\mathcal{T}(\mathfrak{g})$, then $U(\mathfrak{g}) = \mathcal{T}(\mathfrak{g})/\mathcal{I}$ is an universal enveloping algebra of $\mathfrak{g}$.

Proof: Notice that $U(\mathfrak{g})$ contains at least the scalars since the ideal only contains elements of order higher than 1.
If $\pi : \mathcal{T}(\mathfrak{g}) \to U(\mathfrak{g})$ is the canonical projection then we define i to be the restriction of π to $\mathfrak{g} \subset \mathcal{T}(\mathfrak{g})$. If $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is any representation, then let φ be the morphism from $\mathcal{T}(\mathfrak{g}) \to \mathfrak{gl}(V)$ extending ρ. Now since ρ is a representation then for an element in $\mathcal{I}$ we get

$$\varphi(X \otimes Y - Y \otimes X - [X,Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X) - \rho([X,Y]) = 0,$$

therefore $I \subset \ker\varphi$ and φ induces a morphism $\tilde{\rho} : U(\mathfrak{g}) \to \mathfrak{gl}(V)$ satisfying $\tilde{\rho} \circ i = \rho$.



The injectivity of i is not trivial, but is a direct consequence of the PBW theorem presented in the next section. □

Poincaré-Birkhoff-Witt Theorem

We now focus our attention in a natural basis of $U(\mathfrak{g})$ and finish to prove that it actually is the universal enveloping algebra.
For this, let $B = \{X_i \mid i \in I\}$ be a well-ordered basis of $\mathfrak{g}$ by the ordering $(\leq)$ in the set of indices.

Theorem 4.1.4 *The set of monomials*

$$\{X_{i_1} \cdots X_{i_k} \mid k \in \mathbb{Z}_{\geq 0} \text{ and } i_j \leq i_{j+1} \text{ for } j \text{ in } 1, \cdots, k-1\}$$

is a basis of $U(\mathfrak{g})$. In particular if we have a finite basis $\{X_1, \cdots, X_n\}$, then the monomials:

$$X_1^{m_1} X_2^{m_2} \cdots X_n^{m_n} \quad \text{with } m_i \geq 0$$

form a basis of $U(\mathfrak{g})$.

Proof: See Theorem B.0.3 or [Mar09, p.272-275]. □

Corollary 4.1.5 *$i : \mathfrak{g} \to U(\mathfrak{g})$ is injective.*

In the case that $\mathfrak{g}$ is a semi-simple Lie algebra over $\mathbb{C}$, we can consider a maximal toral subalgebra $\mathfrak{h}$ and roots Φ such that:

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

If Δ is a basis of Φ, we can partition the root system as $\Phi = \Phi^- \cup \Phi^+$. Let

$$\mathfrak{n}_- = \bigoplus_{\alpha \in \Phi^-} \mathfrak{g}_\alpha \text{ and } \mathfrak{n}_+ = \bigoplus_{\alpha \in \Phi^+} \mathfrak{g}_\alpha.$$

It is clear that $\mathfrak{n}_-$ and $\mathfrak{n}_+$ are subalgebras of $\mathfrak{g}$ through additivity ($[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$).
The decomposition $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$ is called a **Cartan decomposition**.
By the PBW theorem, suitably ordering a basis of $\mathfrak{g}$:

Corollary 4.1.6 $U(\mathfrak{g}) = U(\mathfrak{n}_-) \otimes_{\mathbb{C}} U(\mathfrak{h}) \otimes_{\mathbb{C}} U(\mathfrak{n}_+)$.

Remark 4.1.7 *If V is a $\mathfrak{g}$-module (as defined in 1.4.1), then the induced algebra morphism gives V a natural $U(\mathfrak{g})$-module structure.*
For any $U(\mathfrak{g})$-module V, we can restrict the action to $\mathfrak{g}$ in order to obtain a representation of $\mathfrak{g}$. These two approaches are totally equivalent.

4.2. Representations

Let $\mathfrak{g}$ be a finite-dimensional semi-simple Lie algebra over an algebraically closed field $\mathbb{F}$ of characteristic 0, with maximal toral (equivalently, cartan) subalgebra $\mathfrak{h}$ and V any finite-dimensional $\mathfrak{g}$-module. We have seen that $\mathfrak{h}$ acts semisimply on V, this allows us to decompose V into eigenspaces with respect to $\mathfrak{h}$.

$$V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda,$$

where $V_\lambda = \{v \in V | H \cdot v = \lambda(H)v \text{ for all } H \in h\}$, we will call it a **weight space** whenever $V_\lambda \neq 0$ and we will call λ a **weight** of V.

Example 4.2.1 *Let $V = \mathfrak{g}$ with the adjoint action. The weights λ are the roots α, along with 0, where $\mathfrak{h}$ is the 0-weight space.*
Letting $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{F})$, $\mathfrak{h}^*$ is one-dimensional, and the weights of any particular finite-dimensional representation are subsets of the lattice of integers.

If $\dim V = \infty$, there is no assurance that such decomposition holds, when it does we call V a **weight module**. Weight modules are one of, if not the most, importan class of modules to study in representation theory.
To formalize the concept of weights on arbitrary representations of semi-simple Lie algebra, we will generalize the results of additivity under the $\mathfrak{g}$-action of weights already seen in the case of $\mathfrak{sl}(2, \mathbb{F})$ and the adjoint action.

Lemma 4.2.2 *If V is an arbitrary $\mathfrak{g}$-module. Then:*

(a) $\mathfrak{g}_\alpha$ maps V_λ to $V_{\lambda+\alpha}$($\alpha \in \Phi$ and $\lambda \in \mathfrak{h}^*$).

(b) The sum $V' = \sum_{\lambda \in \mathfrak{h}^*} V_\lambda$ is direct, and V' is a submodule of V.

(c) If V is finite-dimensional, then $V = V'$.

Proof: Let $X \in \mathfrak{g}_\alpha$, then $H(Xv) = X(Hv) + [H, X]v = \lambda(H)Xv + \alpha(H)Xv = (\lambda + \alpha)(H)Xv$. Since $\mathfrak{g}$ is the sum of root-spaces, then V' is closed under the action of $\mathfrak{g}$ by the previous note, as it is also a sum of vector spaces it is a vector space. As eigenspaces of distinct eigenvalues, the sum is direct.
Finally, if V is finite-dimensional, since we are taking an algebraically closed field, it decomposes as a direct sum of eigenspaces. □

As we did with $\mathfrak{sl}(2,\mathbb{F})$, we shall define a maximal vector, let $\Delta \subset \Phi$ be a basis of the root system, and Φ^+, Φ^- be the positive and negative roots with respect to Δ.

Given V a $\mathfrak{g}$-module, we call v^+ a **maximal vector** of weight λ if $\mathfrak{g}_\alpha v = 0$ for any $\alpha \in \Phi^+$ and $v \in V_\lambda$. If $\dim V = \infty$, there is no reason for a maximal vector to exist. On finite-dimensional representation, maximal vectors have to exist because the algebra $\mathfrak{h} \oplus n_+$ is solvable, so the existence of a common eigenvector is necessary by Lie's Theorem, and the eigenvalues of n_+ as nilpotent elements has to be 0.

An essential sub-class of weight modules are those generated by a single maximal vector. We call V a **highest weight module** if $V = \mathfrak{g}v^+$ with v^+ being a maximal vector. To justify this terminology, let us look at the following results:

Theorem 4.2.3 *Let V be a highest weight $\mathfrak{g}$-module with maximal vector $v^+ \in V_\lambda$. Denote $\Phi^+ = \{\beta_1, \cdots, \beta_m\}$, $\Delta = \{\alpha_1, \cdots, \alpha_\ell\}$, X_k the vector that spans $\mathfrak{g}_{\beta_k}$, and Y_k as the one that spans $\mathfrak{g}_{-\beta_k}$. Then:*

(a) V is spanned by the vectors $Y_1^{i_1} \cdots Y_m^{i_m} v^+$ $(i_k \in \mathbb{Z}_{\geq 0})$, in particular V is a weight module.

(b) The weights of V are of the form $\mu = \lambda - \sum_{i=1}^{\ell} k_i \alpha_i$, with $k_i \in \mathbb{Z}_{\geq 0}$.

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

(d) Each submodule of V is a direct sum of its weight spaces.

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

(f) Every non-zero homomorphic image of V is also a highest weight module of weight λ.

Proof:

(a) By Corollary 4.1.6, we know that $U(\mathfrak{g}) = U(\mathfrak{n}_-) \otimes_{\mathbb{C}} U(\mathfrak{h}) \otimes_{\mathbb{C}}$. $U(\mathfrak{n}_+)$ acts as 0 on v^+ and $U(\mathfrak{h})$ acts as a scalar, therefore $U(\mathfrak{g})v^+ = U(\mathfrak{n}_-)v^+$. Applying the PBW theorem to the algebra $U(\mathfrak{n}_-)$ we know that a basis for $U(\mathfrak{n}_-)$ is given by the ordered monomials $Y_1^{i_1} \cdots Y_m^{i_m}$. Since v^+ generates V, the result follows.

(b) Since every weight β_i is positive, they are written as a sum of simple roots, and by (a) every element of V is the image of the action of a negative root vector on v^+.

(c) Only a finite number of combinations of the roots $-\beta_m$ centered at λ can give rise to a particular weight $\mu = \lambda - \sum_{i=1}^{\ell} k_i \alpha_l$. In view of (a) they span this weight space and therefore $\dim V_\mu < \infty$. Since there is only one way to get to the weight λ, through $i_k = 0$ for all k, $\dim V_\lambda = 1$.

(d) Since V is a weight module by (a), it rests to prove that if a submodule contains a sum, then they must contain the weight vectors.
Let $W \subset V$ be a submodule and $w = \sum_{i=1}^{r} v_i$ for v_i weight vectors of weight μ_i. If some v_i lie in W, then $w - v_i \in W$ and we may assume $w = v_1 + \cdots + v_k$ with no v_i in W minimal on k. Choose $H \in \mathfrak{h}$ such that $\mu_1(H) \neq \mu_2(H)$, then $Hw \in W$ and so is $\mu_1(H)w$. Therefore:

$$(H - \mu_1(H)I)w = (\mu_2(H) - \mu_1(H))v_2 + \cdots + (\mu_k(H) - \mu_1(H))v_k \neq 0,$$

which means one of those elements lie in W, contradicting minimality.

(e) Since v^+ generates V, no proper submodules contains v^+. Therefore the sum of all proper submodules is still a proper submodule and trivially maximal, let us call this sum W. Every proper submodule is contained in W, proving that V cannot be decomposed.
Moreover, since W is unique and maximal, there is a unique irreducible quotient V/W.

(f) Let $\phi(V)$ be such non-zero homomorphic image, then $\phi(v^+)$ generates the image, and by preservation it is also maximal of weight λ.

□

Corollary 4.2.4 *Let V be an irreducible highest weight $\mathfrak{g}$-module with maximal vector v^+ of weight λ, then there is no other maximal vector (up to scalar multiples) in V.*

Proof: If w^+ is another maximal vector, then $U(\mathfrak{g})w^+ = V$ since V is irreducible. If μ is the weight of w^+ then Theorem 4.2.3(b) applies to both μ and λ, and therefore $\mu = \lambda$. By part (c) they must be proportional. □

Based on the previous corollary, one might naturally think about how to describe all the irreducible modules of that form by their highest weight.
Given $\lambda \in \mathfrak{h}^*$ there always exists an irreducible module with λ as the highest weight (see Section 5.2), furthermore, they are always isomorphic.

Theorem 4.2.5 *If V, W are irreducible highest weight of highest weight λ, then they are isomorphic.*

Proof: Let $M = V \oplus W$ and v^+, w^+ be their respective maximal vectors, then $m^+ = (v^+, w^+)$ is also a maximal vector by how we extend the $\mathfrak{g}$-action on M.
Let $N = U(\mathfrak{g})m^+$, it is then a highest weight sub $\mathfrak{g}$-module of M, then the projections from N to the initial modules form irreducible quotients of N and are therefore isomorphic by Theorem 4.2.3(e). □

5 Category $\mathcal{O}$

We begin the study of category $\mathcal{O}$ by reminding ourselves of the equivalence between $\mathfrak{g}$ and $U(\mathfrak{g})$ modules, every $\mathfrak{g}$ module induces an $U(\mathfrak{g})$ module by the universal property and every $U(\mathfrak{g})$ module gives rise to a $\mathfrak{g}$-module through the action of the length one elements. The results here presented are based on [Hum08, chpt 0&1].

The category $U(\mathfrak{g})$-mod of representations is definitely too big for any practical purposes. BGG category $\mathcal{O}$ was originally constructed by Joseph Bernstein, Sergei Gelfand, and Israel Gelfand with the intent to extend on results for finite-dimensional representations as well as understand problems raised by Daya-Nand Verma in his Ph.D. thesis in 1968.

There are many reasons to study Category $\mathcal{O}$:

- It is an abelian category, meaning that it is a category in which morphisms and objects can be added, kernels and cokernels exist and have desirable properties.
- It contains imporant objects to the study of representation theory, such as the finite-dimensional modules and highest weight modules.
- Its irreducible objects have been completely classified as quotients of highest weight modules.
- Its results are relatively accesible when compared to other sub-categories of $U(\mathfrak{g})$-mod.

On this chapter fix a finite-dimensional semi-simple Lie algebra $\mathfrak{g}$, a maximal toral subalgebra $\mathfrak{h}$ and a basis Δ for its root system Φ. Let $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$ be the Cartan decomposition of $\mathfrak{g}$ with respect to Δ.

5.1. Axiomatic

Definition 5.1.1 *We will say that a $U(\mathfrak{g})$-module M is on BGG category $\mathcal{O}$ if it satisfies the following conditions:*

($\mathcal{O}$1) M is finitely generated as a $U(\mathfrak{g})$-module.

($\mathcal{O}$2) M is a weight module.

($\mathcal{O}$3) For each $v \in M$ the subspace generated by the action of $\mathfrak{n}_+$ on v is finite.

As a category, we define as the objects of $\mathcal{O}$ the modules satisfying conditions ($\mathcal{O}$1), ($\mathcal{O}$2) and ($\mathcal{O}$3). The morphisms are the ones inherited from $U(\mathfrak{g})$-mod, and since we define it this way we say $\mathcal{O}$ is a **full subcategory** of $U(\mathfrak{g})$-mod, as all arrows betweeen two objects of $\mathcal{O}$ in the category $U(\mathfrak{g})$-mod are also arrows in $\mathcal{O}$.

Proposition 5.1.2 *As a direct consequence of the axioms, an $M \in \mathcal{O}$ satisfies:*

($\mathcal{O}$4) M is finitely generated by weight vectors.

($\mathcal{O}$5) Every weight space of M is finite-dimensional.

($\mathcal{O}$6) There exists a maximal vector $v^+ \in M$.

Proof:

($\mathcal{O}$4) Since M is a finitely generated weight module, we can take the weight vectors that make up our set of generators (every generator is a linear combination of weight vectors).

($\mathcal{O}$5) We can assume M to be generated by a single weight vector v. Since $U(\mathfrak{n}_+)v$ is finite dimensional and there are only a finite number of combinations of roots that go from a certain weight to a lower one, the result follows.

($\mathcal{O}$6) Otherwise $U(\mathfrak{n}_+)v$ would not be finite dimensional for any non-zero weight vector v.

□

Since every finite-dimensional module is a weight module then they are trivially modules in $\mathcal{O}$. It is also easily seen that highest weight modules are also in $\mathcal{O}$.

Theorem 5.1.3 *As a category $\mathcal{O}$ satisfies:*

(a) *$\mathcal{O}$ is a noetherian category, meaning all $M \in \mathcal{O}$ are noetherian modules.*

(b) *$\mathcal{O}$ is closed under submodules, quotients and finite direct sums.*

(c) *$\mathcal{O}$ is an abelian category.*

Proof:

(a) *Since $U(\mathfrak{g})$ is Noetherian [MR01, p.31] and all $M \in \mathcal{O}$ are finitely generated, the result follows.*

(b) *A submodule of a module in $\mathcal{O}$ is finitely generated by (a). It is a weight module since for finite sum of weight vectors v_i of distinct weights λ_i we can find $H \in \mathfrak{h}$ such that $\lambda_i(H) = 0$ for all but one i. Finally, condition $(\mathcal{O}3)$ is trivial.*
For quotients, a quotient of a Noetherian module satisfies $(\mathcal{O}1)$ and $(\mathcal{O}3)$ trivially. For $(\mathcal{O}2)$, letting π denote the canonical projection, note that it will either send a weight vector v_λ to 0 or to a weight vector whose weight is λ through:

$$H\pi(v_\lambda) = \pi(Hv_\lambda) = \pi(\lambda(H)v_\lambda) = \lambda(H)\pi(v_\lambda).$$

The case for finite direct sums is trivial.

(c) *Since $U(\mathfrak{g})$-mod is an abelian category, it is sufficient to check that $\mathcal{O}$ is closed under finite direct sums, kernels and cokernels. All of which follow from (b).*

□

We shall now prove that the highest-weight modules introduced in Section 4.2 are, in a loose sense, building blocks for all non-zero modules of $\mathcal{O}$.

Proposition 5.1.4 *Let M be any nonzero module in $\mathcal{O}$. Then M has a finite filtration $0 \subset M_1 \subset M_2 \subset \cdots \subset M_n = M$ with nonzero quotients each of which is a highest weight module.*

Proof: M is finitely generated by weight vectors. Let W be one such set of generators, then $V = \mathfrak{n}_+ W$ is a finite-dimensional $\mathfrak{n}_+$-module. If $\dim V = 1$ then W is itself a highest weight module. Otherwise proceed by induction.
Let $v \in V$ be a maximal vector of M and $\overline{M} := M/M_1 \in \mathcal{O}$, where $M_1 = vM$. Then $\overline{M}$ is generated by the image $\overline{V}$ of V through the canonical projection, and $\dim \overline{V} < \dim V$. The induction hypothesis can be applied to $\overline{M}$ yielding a chain of modules in $\overline{M}$, whose pre-images in M are the desired modules. □

5.2. Verma Modules

We are able to construct a large family of highest weight modules by inducing an easily constructed family of highest modules for a subalgebra of $\mathfrak{g}$ to the whole of $\mathfrak{g}$. This will be done through the use of the **Borel subalgebra** $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}_+$.

The quotient algebra $\mathfrak{b}/\mathfrak{n}_+$ is isomorphic to $\mathfrak{h}$, therefore any $\lambda \in \mathfrak{h}^*$ defines a 1-dimensional $\mathfrak{b}$-module with trivial $\mathfrak{n}_+$-action, denote this by $\mathbb{C}_\lambda$. This means that an element $(H + X) \in \mathfrak{b}$ acts as $\lambda(H)$ on $\mathbb{C}_\lambda$.

Definition 5.2.1 *The $U(\mathfrak{b})$-module $M(\lambda) := U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} \mathbb{C}_\lambda$ has a natural structure of $U(\mathfrak{g})$-module and, as a $U(\mathfrak{g})$-module is called the **Verma module**.*

The Verma module is a highest weight module with maximal vector $v^+ = 1 \otimes 1$ of weight λ, in fact, it is the **universal highest weight module** of weight λ.

Having this constructed, we shall denote $N(\lambda)$ as the unique maximal submodule of $M(\lambda)$ and $L(\lambda)$ as the unique simple quotient.

Theorem 5.2.2 *For every highest weight module M of weight λ, there is a surjective module morphism $\varphi : M(\lambda) \to M$.*

Proof: Since M is a highest weight module, there exists an $U(\mathfrak{g})$-module morphism $\rho : U(\mathfrak{g}) \to M$, with $\rho(X) = Xv^+$.

M is highest weight of weight λ, therefore $\mathfrak{n}_+$ and $\{H - \lambda(H) : H \in \mathfrak{h}\}$ are contained in the kernel, let I denote the ideal generated by these two sets.

By the definition of $M(\lambda)$, the $U(\mathfrak{g})$-module morphism $\psi : U(\mathfrak{g}) \to M(\lambda)$ has as kernel exactly I, therefore $M(\lambda)$ is isomorphic to $U(\mathfrak{g})/I$. Finally we can induce a surjective morphism from $M(\lambda)$ to M:

$$M(\lambda) \xrightarrow{\tilde{\psi}} U(\mathfrak{g})/I \longrightarrow U(\mathfrak{g})/\ker\rho \xrightarrow{\tilde{\rho}^{-1}} M$$

□

Theorem 5.2.3 *Every irreducible module $M \in \mathcal{O}$ is isomorphic to $L(\lambda)$ for some $\lambda \in \mathfrak{h}^*$.*

Proof: Since every module $M \in \mathcal{O}$ has a maximal vector v^+ of weight λ, by irreducibility $M = U(\mathfrak{g})v^+$.

But then M is a highest module of weight λ, and so is $L(\lambda)$, therefore, they are isomorphic by Theorem 4.2.5. □

A General Algebra

Definition A.0.1 *A pair* $(\mathcal{A}, \cdot)$ *where* $\mathcal{A}$ *is a vector space over a field* $\mathbb{F}$ *and* $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ *is bilinear is called an* $\mathbb{F}$*-algebra.*

$\mathbb{F}$-algebras are everywhere in mathematical research, from matrix spaces to polynomials and are the main topic of this dissertation.
Note that this is a convention used in the studies of linear properties of algebraic structures, and in another context an F-algebra can denote something different.
If some additional structure is present in the algebra, we shall denote so by the properties of its product, for example, the following:

- If for all $x, y \in \mathcal{A}$, $x \cdot y = y \cdot x$ it is called **commutative**.
- If for all $x, y, z \in \mathcal{A}$, $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ it is called **associative**.
- If there exists an element $e \in \mathcal{A}$ such that $e \cdot x = x \cdot e = x$ for all $x \in \mathcal{A}$ then it is called **unital**.

Definition A.0.2 *A subspace* B *of an algebra* A *which satisfies* $x, y \in B \Rightarrow x \cdot y \in B$ *is called a subalgebra which will be represented by* $B \leq A$*. When it satisfies the further property that* $x \in A, y \in B \Rightarrow x \cdot y \in B$ $(y \cdot x \in B)$ *it will be called a left-ideal (right-ideal), being simply called an ideal if it satisfies both, which will be denoted by* $A \trianglelefteq B$*.*
An algebra morphism is defined as a linear transformation that preserves the product, that is, given two algebras $(A, \cdot), (B, \times)$ *then a linear transformation* $\varphi : A \to B$ *is an algebra morphism if:*

$$\varphi(x \cdot y) = \varphi(x) \times \varphi(y).$$

It is generally useful on proving properties of algebras to consider different induced algebras from known ones, such as a quotient or direct sum. There is a natural way to expand the product by composing algebras in those cases.

Proposition A.0.3 *If B is an ideal of $(A,\cdot)$ then the quotient space A/B with the product $\times$ defined as $[x]\times[y]=[x\cdot y]$ is an algebra and the projection $[\ ]: A \to A/B$ is an algebra morphism.*
If $(A,\cdot)$ and $(B,\times)$ are algebras, then the vector space $A\oplus B$ is an algebra with $(a,b)(a',b')=(a\cdot a', b\times b')$.

Proof: It just needs to be shown that the product is well defined, but given $x+B$ and $y+B$, the product $(x+B)\cdot(y+B)=x\cdot y+x\cdot B+B\cdot y+B\cdot B=x\cdot y+B$ is in the same coset independently of the representing vector used. As for the direct sum, bilinearity follows directly from the vector space structure from $A\oplus B$ and the bilinearity of the product. $\square$
We will state some theorems that will be useful later for the specific cases of Lie algebras but are valid in the general case of algebras:

Proposition A.0.4 *Given algebras $(A,\cdot)$ and $(B,\times)$, and a morphism $\varphi: A\to B$:*

(a) $\ker(\varphi)\trianglelefteq A$.

(b) $\varphi(A)\leq B$.

(c) $A/\ker(\varphi)\simeq im(\varphi)$ *where $(\simeq)$ denotes the existence of a bijective morphism (isomorphism).*

Proof: Given $X\in\ker(\varphi)$ and $Y\in A$, $\varphi(X\cdot Y)=\varphi(X)\times\varphi(Y)=0$ and therefore $X\cdot Y\in A$. That is, $\ker(\varphi)$ is an ideal of A.
Given $\varphi(X),\varphi(Y)\in\varphi(A)$, $\varphi(X)\times\varphi(Y)=\varphi(X\cdot Y)$, and since $X\cdot Y\in A$, then $\varphi(X)\times\varphi(Y)\in\mathrm{im}(\varphi)$. That is, $\varphi(A)$ is a subalgebra of B
Finally consider the morphism $\psi: X+\ker(\varphi)\mapsto\varphi(X)$, ψ is well defined since

$$\psi(X+\ker(\varphi))=\psi(Y+\ker(\varphi))\Rightarrow\varphi(X)=\varphi(Y)\Rightarrow\varphi(X-Y)=0\Rightarrow X-Y\in\ker\varphi.$$

It is a morphism by the way we define the product in the quotient:

$$\begin{aligned}\psi([X]\cdot[Y])=\psi([X\cdot Y])=\varphi(X\cdot Y)&=\varphi(X)\times\varphi(Y)\\&=\psi([X])\times\psi([Y]).\end{aligned}$$

Finally, it is bijective because if $\psi([X])=0$ then $\varphi(X)=0$ which implies that $[X]=[0]$. Moreover, for all $\varphi(X)\in\varphi(A)$, $\varphi(X)=\psi([X])$, with $[X]\in A/\ker(\varphi)$. $\square$

This section is a compilation of results that did not fit the main text, their proof do not connect well with the topics at hand, but are necessary to the study.

Proposition B.0.1 (Jordan-Chevalley Decomposition) *If V is a finite-dimensional vector space over an algebraically closed field $\mathbb{F}$, then every linear transformation $T : V \to V$ satisfies:*

(a) There exists unique $S, N : V \to V$ satisfying $T = S + N$ with S semi-simple and N nilpotent.

(b) There exist polynomials $p(x)$ and $q(x)$ without constant term such that $S = p(T)$ and $N = q(T)$, in particular, S and N commute with every endomorphism commuting with T.

(c) If $A \subset B \subset V$ are subspaces, and T maps B into A, then S and N also map B into A.

Proof: Let $\Pi(x-a_i)^{m_i}$ be the characteristic polynomial of T, then $V = \oplus\ker(T-a_iI)^{m_i}$. By the Chinese Remainder Theorem, we can find a polynomial $p(x)$ such that:

$$\begin{cases} p(x) \equiv a_i \pmod{(x-a_i)^{m_i}}, \\ p(x) \equiv 0 \pmod{x}. \end{cases}$$

Now let $q(x) = x - p(x)$ and set $S = p(T)$ and $N = q(T)$. Since they are both polynomials in T, then they commute with every endomorphism commuting with T. Furthermore if T sends $B \supset A$ into A, then these also do so in accordance with (c).
To show that S is semi-simple, notice that $S - a_iI$ restricted to $\ker(T - a_iI)^{m_i}$ is 0 and therefore it acts diagonally on each of those spaces, but then it acts diagonally on V as a direct sum. Now N is clearly nilpotent as it has no non-zero eigenvalues.
For uniqueness, since $S_1 + N_1 = S_2 + N_2 = T$, then $S_1 - S_2 = N_2 - N_1$. The sum of commuting semi-simple operators is semi-simple and the same is valid for nilpotent operators, therefore $S_1 - S_2$ is both semi-simple and nilpotent, that is $S_1 - S_2 = 0$.

This implies that $S_1 = S_2$ and $N_1 = N_2$. □

Lemma B.0.2 *If V is a vector space over an infinite field $\mathbb{F}$, then a union of finitely many proper subspaces V_i cannot equal V.*

Proof: If $V = \bigcup_{i=1}^{n} V_i$, and let $x \in V_1$ non-zero, since V_1 is proper there exists an $y \in V \backslash V_1$, there are infinitely many vectors of the form $x + \alpha y$ for $\alpha \in \mathbb{F}$ that are not in V_1, but therefore there exists some V_i for which there are infinitely many of these vectors, in particular there are 2 of those vectors in this V_i and therefore $(x + \alpha_1 y) - (x - \alpha_2 y) = (\alpha_1 - \alpha_2)y \in V_i \Rightarrow y \in V_i$, but therefore V_i contains x, since the choice of x was arbitrary, we find: $V_1 \subseteq \bigcup_{i=2}^{n} V_i$, repeating this process we find $V = V_n$, a contradiction.
Note that the field has to be infinite, otherwise the pigeonhole principle does not apply for the union of more subspaces than distinct elements of the field. □

Theorem B.0.3 *If $\{X_i \mid i \in I\}$ is a well-ordered basis of a Lie algebra $\mathfrak{g}$ through $\leq$, the set of monomials*

$$\{X_{i_1} \cdots X_{i_k} \mid k \in \mathbb{Z}_{\geq 0} \text{ and } i_j \leq i_{j+1} \text{ for } j \text{ in } 1, \cdots, k-1\}$$

is a basis of $U(\mathfrak{g})$. In particular if we have a finite basis $\{X_1, \cdots, X_n\}$, the monomials:

$$X_1^{m_1} X_2^{m_2} \cdots X_n^{m_n} \quad \text{with } m_i \geq 0$$

form a basis of $U(\mathfrak{g})$.

Proof: To show that these ordered monomials form a basis of $U(\mathfrak{g})$, we show that any monomial $m = X_{i_1} \cdots X_{i_k}$ can be written as a linear combination of elements in the ordered set, but this can be trivially seen from the relation:

$$M(X_i X_j)N = M(X_j X_i + [X_j, X_i])N \text{ for any } M, N \in U(\mathfrak{g}).$$

If there are m which have $i_j > i_l$ we can swap them and add a single element, reducing the length of the monomial and, by induction, since monomials of length 1 are always ordered the result follows.
The biggest hurdle lies in proving linear independence, one approach to this is filtering $U(\mathfrak{g})$ through the filtration in the tensor algebra and proving that the graded algebra is isomorphic to the symmetric algebra [Hum72].

The approach we are going to use is to construct an endomorphism σ on the tensor algebra in a way that $\mathcal{I}$ is mapped to 0 and $\sigma(m) = m$ if m is an ordered monomial. Showing that the span of ordered monomials does not intersect I and therefore it is linearly independent in $\mathcal{T}/\mathcal{I}$, since it is independent in $\mathcal{T}$.
Fix $m = X_{i_1} \cdots X_{i_k}$ any monomial. If m is ordered, set $\sigma(m) = m$. Otherwise, it is possible to find an index i_s such that $i_s > i_{s+1}$, setting the number of these as $d(m)$ and defining σ inductively on $d(m)$.

$$\sigma(m) = \sigma(X_{i_1} \cdots X_{i_{s+1}} X_{i_s} \cdots X_{i_k}) + \sigma(X_{i_1} \cdots [X_{i_s}, X_{i_{s+1}}] \cdots X_{i_k}),$$

where the right-hand side is defined by induction.
Since we are defining σ only in a basis of T, we extend it naturally to an endomorphism. Now to finish the proof, it rests to prove that the endomorphism is null in $\mathcal{I}$ and that it is well defined, that is, the above recursion does not depend on the choice of s.
To show that it is null on $\mathcal{I}$ pick a general element $x = M(X_i X_j - X_j X_i - [X_i, X_j])N \in \mathcal{I}$. If $i = j$ then $x = 0$, if $i \neq j$ assume without loss of generality that $i > j$, then by the definition of σ, if it is well defined, and we get:

$$\sigma(MX_j X_i N) = \sigma(MX_i X_j N) + \sigma(M[X_i, X_j]N) \Rightarrow \sigma(x) = 0.$$

To show that σ well defined, consider the two possible cases for more than one unordered pair on m.

1. $$X_{i_1} \cdots X_{i_r} X_{i_{r+1}} \cdots X_{i_s} X_{i_{s+1}} \cdots X_{i_k},$$

   with $i_r > i_{r+1}$ and $i_s > i_{s+1}$, we need to show that applying σ for s then r, is the same as applying σ for r then s:
   Let $A = X_{i_1} \cdots X_{i_{r-1}}$, $B = X_{i_{r+2}} \cdots X_{i_{s-1}}$ and $C = X_{i_{s+2}} \cdots X_{i_k}$ and $X_{i_k} = X_k$ to simplify notation.

   $$\begin{aligned} \sigma(AX_r X_{r+1} BX_s X_{s+1} C) = {} & \sigma(AX_{r+1} X_r BX_{s+1} X_s C) \\ & + \sigma(A[X_r, X_{r+1}]BX_{s+1} X_s C) \\ & + \sigma(AX_{r+1} X_r B[X_s, X_{s+1}]C) \\ & + \sigma(A[X_r, X_{r+1}]B[X_s, X_{s+1}]C). \end{aligned}$$

   Therefore σ does not depend on the choice between two indices and is well defined in this case.

2.

$$AX_rX_{r+1}X_{r+2}B,$$

with $i_r > i_{r+1} > i_{r+2}$, first calculating the permutation of i_{r+1} and i_r we find:

$$\begin{aligned}
\sigma(AX_rX_{r+1}X_{r+2}B) &= \sigma(AX_{r+1}X_rX_{r+2}B) + \sigma(A[X_r, X_{r+1}]X_{r+2}B) \\
&= \sigma(AX_{r+1}X_{r+2}X_rB) + \sigma(AX_{r+1}[X_r, X_{r+2}]B) + \sigma(A[X_r, X_{r+1}]X_{r+2}B) \\
&= \sigma(AX_{r+2}X_{r+1}X_r) \\
&+ \sigma(A[X_{r+1}, X_{r+2}]X_rB) \\
&+ \sigma(AX_{r+1}[X_r, X_{r+2}]B) \\
&+ \sigma(A[X_r, X_{r+1}]X_{r+2}B),
\end{aligned}$$

where the terms to permute after the first step are uniquely determined. Similarly for the other permutation

$$\begin{aligned}
\sigma(AX_rX_{r+1}X_{r+2}B) &= \sigma(AX_{r+2}X_{r+1}X_r) \\
&+ \sigma(A[X_{r+1}, X_{r+2}]X_rB) \\
&+ \sigma(AX_{r+1}[X_r, X_{r+2}]) \\
&+ \sigma(A[X_r, X_{r+1}]X_{r+2}B).
\end{aligned}$$

The difference between the two expressions, evaluated using the induction hypothesis on the length of monomials is given by:

$$\sigma(A([X_r, [X_{r+1}, X_{r+2}]] + [X_{r+1}, [X_{r+2}, X_r]] + [X_{r+2}, [X_r, X_{r+1}]])B),$$

which is equal to 0 by the Jacobi identity.
Finally it follows that the span of ordered monomials does not intersect I and therefore they are linearly independent in $U(\mathfrak{g}) \cong \mathcal{T}/\mathcal{I}$.

□

C Lie Algebras from Matrix Lie Groups

In this appendix, we give a brief construction of Lie algebras from Matrix Lie groups. This approach can be found in [Hal04]. A more classical and general construction can be found in [Kir08] and a more algebraic one can be found in [Kac10, lec 2].

Lie groups are manifolds with a group structure, we will focus our attention on matrix Lie groups as they have a simpler extension to Lie algebras.

Given a field $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, define $GL(n;\mathbb{F})$ as the group of invertible matrices, known as the **general linear group**. The metric used on $GL(n;\mathbb{F})$ is the **operator norm** with respect to the euclidean one:

$$\|A\| = \sup_{\|v\|_2=1} \{\|Av\|_2\} \text{ where } \|v\|_2 = \sum_{i=1}^{n} \sqrt{v_i}.$$

A sequence of matrices $A_n = (a_{ij}^n)$ converges to $A = (a_{ij})$ in this norm if it converges entry-wise $a_{ij}^n \to a_{ij}$.

Definition C.0.1 *A matrix Lie group G is a closed subgroup of $GL(n;\mathbb{F})$, meaning that it is a group and whenever a sequence of matrices in $A_m \in G$ converges to a matrix $A \in GL(n;\mathbb{F})$, we have $A \in G$.*

Example C.0.2 *$GL(n;\mathbb{F})$ is a matrix Lie group since it is a closed group.*

Example C.0.3 *Moreover, since the determinant is a continuous function, the group $SL(n;\mathbb{F})$ of matrices with determinant 1 is a matrix Lie group ($\{1\} \subseteq \mathbb{R}$ is closed).*

Example C.0.4 *Not every subgroup of $GL(2;\mathbb{C})$ is closed, in fact, given $a \in \mathbb{R}/\mathbb{Q}$ irrational and:*

$$G = \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{ita} \end{pmatrix} \middle| \, t \in \mathbb{R} \right\}.$$

$-I$ is not in G because t has to be an odd multiple of π in which case ta cannot be an odd multiple of π since a is irrational. On the other hand we can take $t = (2n+1)\pi$ for carefully chosen integers to make ta arbitrarily close to an odd multiple of π.

We will briefly remind some of the properties of matrix exponentials necessary to show properties of the Lie algebra as the “logarithm” of a Lie group in some sense.

Proposition C.0.5 *The exponential matrix of a complex or real $X \in M_{n\times n}$, defined as*

$$e^X = I + \sum_{m=1}^{\infty} \frac{X^m}{m!},$$

satisfies the following properties:

(a) It is a continuous map on the corresponding matrix spaces.

(b) $(e^X)^* = e^{X^*}$.

(c) If $XY = YX$ then $e^{X+Y} = e^X e^Y$. As a consequence, e^X is invertible with inverse e^{-X} since $e^0 = I$.

(d) If C is any invertible matrix, then $e^{CXC^{-1}} = Ce^X C^{-1}$.

(e) $\|e^X\| \le e^{\|X\|}$.

(f) $\frac{d}{dt} e^{tX} = X e^{tX}$.

(g) ***(Lie's Product Formula)*** $e^{X+Y} = \lim_{m\to\infty} (e^{\frac{X}{m}} e^{\frac{Y}{m}})^m$.

Proof:

(a) The sum is well defined, given that the norm converges:

$$\sum_{m=1}^{\infty} \left\| \frac{X^m}{m!} \right\| = \sum_{m=1}^{\infty} \frac{\|X^m\|}{m!} \le \sum_{m=1}^{\infty} \frac{\|X\|^m}{m!} = e^{\|X\|} - 1. \tag{C.1}$$

Now given $X, Y \in M_{n\times n}$ it follows that (abusing notation as to say $0^0_{n\times n} = I$):

$$\begin{aligned} \|e^{X+Y} - e^X\| &= \left\| \sum_{m\ge 0} \frac{(X+Y)^m - X^m}{m!} \right\| \\ &\le \sum_{m\ge 0} \frac{(\|X\| + \|Y\|)^m - \|X\|^m}{m!} \\ &= e^{\|X\|+\|Y\|} - e^{\|X\|} \\ &= e^{\|X\|}(e^{\|Y\|} - 1) \le \|Y\| e^{\|X\|} e^{\|Y\|}. \end{aligned}$$

Continuity follows directly by choosing Y in the neighborhood of a chosen X.

(b) Follows directly from $(X^m)^* = (X^*)^m$ and continuity of the transpose operator.

(c) As X and Y commute, then $(X+Y)^m = \sum_{k=0}^{\infty} \binom{n}{k} X^k Y^{m-k}$ now since the exponential converges in absolute value:

$$e^X e^Y = \sum_{m=0}^{\infty} \sum_{k=0}^{m} \frac{X^k}{k!} \frac{Y^{m-k}}{(m-k)!} = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{k=0}^{m} \binom{n}{k} X^k Y^{m-k} = \sum_{m=0}^{\infty} \frac{(X+Y)^m}{m!} = e^{X+Y}.$$

(d)

$$e^{CXC^{-1}} = \sum_{m=0}^{\infty} \frac{(CXC^{-1})^m}{m!} = \sum_{m=0}^{\infty} \frac{CX^m C^{-1}}{m!} = Ce^X C^{-1}.$$

(e) The same result presented in (1.1).

(f) Again we differentiate term by term since the power series converges uniformly.

$$\frac{d}{dt} e^{tX} = \frac{d}{dt} 1 + \sum_{m=1}^{\infty} \frac{d}{dt} \frac{(tX)^m}{m!} = \sum_{m \geq 1} \frac{t^{m-1} X^m}{(m-1)!} = Xe^{tX}.$$

(g) Define $A = e^{(X+Y)/k}$ and $B = e^{X/k} e^{Y/k}$, then by the norm inequality from (1.1) and the triangle inequality imply:

$$\|A\|, \|B\| \leq (e^{\|A\|+\|B\|})^{1/k}.$$

On the other hand, reordering terms for B in terms of the power of k by absolute convergence of the exponential:

$$B = \sum_{i=0}^{\infty} \frac{(X/k)^i}{i!} \cdot \sum_{j=0}^{\infty} \frac{(Y/k)^j}{j!} = \sum_{m=0}^{\infty} k^{-m} \sum_{i=0}^{m} \frac{A^i}{i!} \cdot \frac{B^{m-i}}{(m-i)!}.$$

Which allows us to bound the norm of the difference by:

$$\begin{aligned}
\|A-B\| &= \left\| \sum_{i=0}^{\infty} \frac{([A+B]/k)^i}{i!} - \sum_{m=0}^{\infty} k^{-m} \sum_{j=0}^{m} \frac{A^i}{i!} \frac{B^{m-i}}{(m-i)!} \right\| \\
&= \left\| \sum_{i=2}^{\infty} k^{-i} \frac{(A+B)^i}{i!} - \sum_{m=2}^{\infty} k^{-m} \sum_{j=0}^{m} \frac{A^i}{i!} \frac{B^{m-i}}{(m-i)!} \right\| \\
&\leq \frac{1}{k^2} \left[e^{\|A\|+\|B\|} + \sum_{m=2}^{\infty} \frac{1}{m!} \sum_{i=0}^{m} \frac{m}{i} \|A^i\| \|B^{m-i}\| \right] \\
&= \frac{1}{k^2} \left[e^{\|A\|+\|B\|} + \sum_{m=2}^{\infty} \frac{(\|A\|+\|B\|)^m}{m!} \right] \\
&\leq \frac{2}{k^2} e^{\|A\|+\|B\|}.
\end{aligned}$$

□

Now we are ready to define and prove some properties of the Lie algebra of these matrix Lie groups.

Definition C.0.6 *Given a matrix Lie Group $G \subset M_{n\times n}(\mathbb{F})$, its Lie algebra is the set $\mathfrak{g} = \{X \in M_{n\times n}(\mathbb{F}) \mid e^{tX} \in G$ for all $t \in \mathbb{R}\}$.*

Considering this set instead of the Lie group itself is very useful as they have very nice algebraic properties and an underlying structure that is quite rich and unique. One of the nicest properties this set possesses is that it is a vector subspace of $M_{n\times n}$ and is closed under a special commutator operator, in fact:

Proposition C.0.7 *Given any two elements X, Y in a Lie algebra $\mathfrak{g}$ of a Lie matrix group G, then:*

(a) $sX \in \mathfrak{g}$ *for all* $s \in \mathbb{R}$,

(b) $X + Y \in \mathfrak{g}$,

(c) $XY - YX \in \mathfrak{g}$.

Proof:

(a) $e^{t(sX)} = e^{(ts)X} \in G$ for all t since $ts \in \mathbb{R}$

(b) We will use Lie's product formula (Proposition C.0.5g): $e^{t(X+Y)} = \lim_{m\to\infty}(e^{tX/m}e^{tY/m})^m$, since G is a group then $(e^{tX/m}e^{tY/m})^m$ is in G for all $m \in \mathbb{N}$ and since it converges, by definition its limit is in the matrix Lie group G, proving that $X + Y \in \mathfrak{g}$.

(c) Since $\mathfrak{g}$ is a vector subspace of $GL(n;\mathbb{F})$, it is a closed set.
Since $e^{tX} \in G$ is invertible and $Y \in \mathfrak{g}$, Proposition C.0.5(d) implies that

$$e^{e^{tX}Ye^{-tX}} = e^{tX}e^{Y}e^{-tX} \in G$$

which by definition means $e^{tX}Ye^{-tX} \in \mathfrak{g}$. Therefore

$$\frac{e^{tX}Ye^{-tX} - Y}{t} \in \mathfrak{g} \text{ for all } t.$$

But $\mathfrak{g}$ is closed:

$$\begin{aligned}
\lim_{t\to 0} \frac{e^{tX}Ye^{-tX} - Y}{t} &\in \mathfrak{g} \\
\lim_{t\to 0} \frac{e^{tX}Ye^{-tX} - Y}{t} &= \frac{d}{dt} e^{tX}Ye^{-tX}\Big|_{t=0} \\
&= (Xe^{tX}Ye^{-tX} + e^{tX}Y(-X)e^{-tX})|_{t=0} \\
&= XY - YX \in \mathfrak{g}
\end{aligned}$$

□

Bibliography

[Hal04] B.C. Hall. *Lie Algebras and Representations: An Elementary Introduction*. Springer, 2004.

[Hum72] J.E. Humphreys. *Introduction to Lie algebras and representation theory*. Graduate texts in mathematics 9. Springer-Verlag, 1972.

[Hum08] J.E. Humphreys. *Representations of Semisimple Lie Algebras in the BGG Category O*. Graduate Studies in Mathematics. AMS, 2008.

[JGBS66] B. Kolman J. G. Belinfante and H. A. Smith. An Introduction to Lie Groups and Lie Algebras, with Applications. *SIAM Review*, 8(1), 1966.

[Kac10] V. Kac. Lecture Notes: Introduction to Lie Algebras. *MIT*, 2010. `https://math.mit.edu/classes/18.745/classnotes.html`.

[Kir08] A. A. Kirillov. *Introduction to Lie Groups and Lie Algebras*. Cambridge University Press, Department of Mathematics, Stony Brook, NY 11794, 2008. `https://www.math.stonybrook.edu/~kirillov/mat552/liegroups.pdf`.

[Mar09] L.A.B.S. Martin. *Algebras de Lie*. Coleção Livro-Texto. Unicamp, 2009.

[MR01] J. C. McConnell and J. C. Robson. *Noncommutative Noetherian rings*. Graduate studies in mathematics 30. American Mathematical Society, 2 edition, 2001.