**Delft University of Technology**
**Faculty of Electrical Engineering, Mathematics and Computer Science**
**Delft Institute of Applied Mathematics**

# Classification of finite-dimensional modules over semisimple Lie algebras

**(Dutch Title: Classificatie van eindig dimensionale modulen over semisimpele Lie algebras)**

Thesis submitted to
Delft Institute of Applied Mathematics
as partial fulfilment of the requirements

for the degree of

**BACHELOR OF SCIENCE**
**in**
**APPLIED MATHEMATICS**

**by**

**Jin Jun Liu**

**Delft, Netherlands**
**July 2023**

**Bsc THESIS APPLIED MATHEMATICS**

**"Classification of finite-dimensional modules over semisimple Lie algebras"**

**(Dutch title: "Classificatie van eindig dimensionale modulen over semisimpele Lie algebras")**

Jin Jun Liu

**Delft University of Technology**

**Supervisors**

Dr. ir. W.G.M. Groenevelt
ir. C.C.M.L. Wagenaar

**Other thesis committee members**

Dr. R. Versendaal

July, 2023 Delft

# Contents

# 1 Introduction

Sophus Lie (1842-1899) known as the founder of the theory of transformation groups, originally aimed to study solutions of differential equations via their symmetries. Over the decades this theory has evolved into the theory of Lie groups. These Lie groups are of an analytic and geometric nature, but Sophus Lie's principal discovery was that these groups can be studied by their "infinitessimal generators" leading to a linearization of the group. The group structure endows this linearized space with a special bracket operation, $[x, y] = xy - yx$, which gives rise to Lie algebras.

The main applications for Lie algebras stem from physics, notably in quantum mechanics and particle physics. It turns out that representations of Lie algebras are the way to describe symmetries of physical systems. So, it becomes an important task to figure out what all the possible representations are. Thus, our main goal for this thesis is to classify all finite-dimensional semisimple Lie algebra representations.

We start by discussing the basic definitions of Lie algebras in Section 2 and develop the theory necessary to classify all the representations, using the language of module theory. This is done with a particular emphasis on the prime example $\mathfrak{sl}(2, \mathbb{C})$, which consists of $2 \times 2$ traceless matrices. Then in Section 3, we focus our attention on the roots of Lie algebras, which are a generalization of eigenvalues of what is called the adjoint representation. They reveal an important part of the underlying structure that semisimple Lie algebras have. The section finishes by studying the symmetries of such roots via the Weyl group. In Section 4, we need to introduce the notion of a universal enveloping algebra. This allows us to imbed a Lie algebra $\mathfrak{g}$ into a more familiar space, so that modules over $\mathfrak{g}$ can be studied through the modules over its universal enveloping algebra. Additionally, an important theorem regarding the universal enveloping algebra, the Poincaré-Birkhoff-Witt theorem, and its consequences will be proven. Finally, in Section 5 the classification of general modules will be discussed, culminating in a correspondence between finite-dimensional modules and dominant integral weights, which concludes our classification problem.

Our treatment of Lie algebras and their representations mainly follows Humphrey's [8] book. Major parts are also inspired by Brian Hall's [7] exposition on Lie algebras. Lastly, a particularly helpful source for Section 3 has been Erdmann and Wildon's [5] book. All the proofs in this thesis were learned from the sources listed in the Bibliography.

*Notation:* The set of nonnegative integers is written as $\mathbb{Z}_{\geq 0}$.

# 2 Lie algebras

Originally coming from the especially analytic and geometric theory of Lie groups, the Lie algebras arise as a linear approximation of the Lie groups together with the operation $[x, y] = xy - yx$. This brings forth an algebraic structure that can be studied separately from its analytic origins, which is exactly the approach we will take.

First, in Section 2.1 we start by introducing the basic notions for the study of Lie algebras. Then we proceed similarly in Section 2.2 with representations and modules. It ends with Weyl's Theorem, which states under what conditions we can decompose the representations into the smallest type of building blocks. Following this, Section 2.3 will treat some necessary linear algebra background. Finally, in Section 2.4 we classify all the finite-dimensional representations of $\mathfrak{sl}(2, \mathbb{C})$.

## 2.1 Basics

**Definition 2.1.** A **Lie algebra** is a vector space $\mathfrak{g}$ over $\mathbb{C}$ with an operation called the **Lie bracket** $[-,-] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the following properties:

(L1) $[-,-]$ is *bilinear*, so linear in the two arguments,

(L2) $[x, x] = 0$ for all $x \in \mathfrak{g}$,

(L3) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in \mathfrak{g}$. (Jacobi identity)

*Remark.* The notion of a Lie algebra can be generalized as a vector space over arbitrary fields $\mathbb{F}$. We will only concern ourselves with *complex Lie algebras* over $\mathbb{C}$. However, all the theorems discussed in this thesis will still hold for other fields, so long as they are algebraically closed and of characteristic zero.

*Remark.* Throughout the thesis, we will only work with *finite-dimensional* Lie algebras, which will be assumed unless otherwise stated.

*Remark.* Bilinearity and property (L2) imply that the bracket is also *anticommutative*:

$$0 = [x + y, x + y] = [x, x] + [x, y] + [y, x] + [y, y] = [x, y] + [y, x],$$

thus we have $[x, y] = -[y, x]$ for all $x, y \in \mathfrak{g}$.

We can view the Lie bracket as a strange type of product, which is in general not commutative. In addition, it also fails to be associative in general. The associativity is instead replaced by the Jacobi identity. If we rewrite $[-,-]$ with $\cdot$ for a moment, we see that the Jacobi identity says the following

$$x \cdot (y \cdot z) + y \cdot (z \cdot x) + z \cdot (x \cdot y) = 0,$$

using anticommutativity we obtain

$$x \cdot (y \cdot z) - (x \cdot y) \cdot z = -y \cdot (z \cdot x),$$

from which we can see how associativity is obstructed by the extra term on the right-hand side.

*Example.* Any vector space $V$ can be turned into a Lie algebra by trivially defining the Lie bracket as $[x, y] := 0$ for all $x, y \in V$. When the bracket of two elements is zero, we say that they *commute* and in the same vein a Lie algebra where all elements commute, as in this example, is called *abelian.*

*Example.* The first example of a Lie algebra that the reader might already be familiar with, is the vector space $\mathbb{R}^3$ with the cross product $\times$ as the bracket operation.

A more important example arises when $V$ is a complex vector space, then the set of all linear maps $V \to V$ is denoted $\mathfrak{gl}(V)$ or $\mathrm{End}(V)$; we will sometimes call these linear maps *endomorphisms.* This new space is again a complex vector space, and here we define the Lie bracket as follows $[x, y] := x \circ y - y \circ x$ for all $x, y \in \mathfrak{gl}(V)$, making it a Lie algebra called the *general linear algebra*: the conditions (L1) and (L2) follow quite directly, and the Jacobi identity is also satisfied as we show below. Let $x, y, z \in \mathfrak{gl}(V)$, then leaving the $\circ$-notation out, we obtain

$$\begin{aligned}[x,[y,z]]+[y,[z,x]]+[z,[x,y]] &= x(yz)-x(zy)-(yz)x+(zy)x \\ &\quad + y(zx)-y(xz)-(zx)y+(xz)y \\ &\quad + z(xy)-z(yx)-(xy)z+(yx)z,\end{aligned}$$

where all the terms cancel out. To show the Jacobi identity, associativity inherent to the map composition within $\mathfrak{gl}(V)$ was essential.

**Definition 2.2.** A vector subspace $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is called a **subalgebra**, if it is closed under the Lie bracket,

$$[x, y] \in \mathfrak{h} \quad \text{for all } x, y \in \mathfrak{h}.$$

*Remark.* A subalgebra is itself again a Lie algebra, where the properties (L1)-(L3) are inherited from the parent Lie algebra.

*Example.* Examining subspaces of $\mathfrak{gl}(V)$, we find that the traceless linear maps form the subspace $\mathfrak{sl}(V)$ called the *special linear algebra.* This is indeed a subalgebra of $\mathfrak{gl}(V)$, because of the trace identity $\mathrm{tr}(xy) = \mathrm{tr}(yx)$ for all $x, y \in \mathfrak{gl}(V)$ and linearity of the trace, which imply that the bracket $[x, y]$ of zero trace maps returns a linear map with $\mathrm{tr}([x, y]) = 0$.

We can also look at the matrix variant of these linear algebras. As we then encounter a fundamental example of a Lie algebra, which will play an immense role in the rest of the thesis. Namely, the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$, which is the set of traceless $2 \times 2$ complex matrices and has a basis given by the elements

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

This Lie bracket is still taken as $[x, y] = xy - yx$, which yields the following bracket relations: $[h, x] = 2x$, $[h, y] = -2y$ and $[x, y] = h$. These bracket relations completely determine the Lie algebra structure of $\mathfrak{sl}(2,\mathbb{C})$. So, we can study $\mathfrak{sl}(2,\mathbb{C})$ separately from its original definition as a subspace consisting of traceless matrices, where the bracket now functions as the only "product", forgetting the matrix product.

Next, we define a special type of map between Lie algebras, which in some sense tells us that two Lie algebras are similar in terms of their structure.

**Definition 2.3.** Given two Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, we say that a map $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a **Lie algebra homomorphism**, if it is linear and

$$\phi([x,y]) = [\phi(x), \phi(y)],$$

where the left bracket is in $\mathfrak{g}_1$ and the right is in $\mathfrak{g}_2$.

We also define the *kernel* and *image* of a Lie algebra homomorphism $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ as follows,

$$\ker\phi = \{x \in \mathfrak{g}_1 : \phi(x) = 0\}, \quad \phi[\mathfrak{g}_1] = \{\phi(x) \in \mathfrak{g}_2 : x \in \mathfrak{g}_1\}.$$

*Example.* Let $\mathfrak{g}$ be a Lie algebra and take an $x \in \mathfrak{g}$, now define $\mathrm{ad}(x) := [x, -]$, then $\mathrm{ad}(x)$ is a map from $\mathfrak{g} \to \mathfrak{g}$. It is also linear by the bilinear property of the bracket, so we have that $\mathrm{ad}(x) \in \mathrm{End}(\mathfrak{g})$. Then the map

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}) \text{ with } x \mapsto \mathrm{ad}(x),$$

called the *adjoint representation* of $\mathfrak{g}$, can be shown to be a Lie algebra homomorphism. This map also obtains its linearity from the bilinear property of the bracket, so it remains to show that the bracket is preserved under the map ad. Let $x, y, z \in \mathfrak{g}$, then the Jacobi identity and anticommutativity yield

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

Hence, the adjoint representation is a Lie algebra homomorphism.

**Definition 2.4.** A subspace $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is said to be an **ideal**, when

$$[x,y] \in \mathfrak{h} \quad \text{for all } x \in \mathfrak{h},\, y \in \mathfrak{g}.$$

*Example.* Any Lie algebra $\mathfrak{g}$ contains the *trivial ideals*, namely $\{0\}$ and $\mathfrak{g}$.

*Example.* The kernel of a Lie algebra homomorphism $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is an ideal of $\mathfrak{g}_1$. This is seen as follows: let $x \in \ker\phi$ and $x' \in \mathfrak{g}_1$, then $\phi([x,x']) = [0, \phi(x')] = 0$.

**Definition 2.5.** A Lie algebra $\mathfrak{g}$ is called **simple** when it is non-abelian and the only ideals it contains are trivial.

*Example.* We will show that $\mathfrak{sl}(2,\mathbb{C})$ is an example of a simple Lie algebra. Its bracket relations are given by $[h,x]=2x$, $[h,y]=-2y$ and $[x,y]=h$. Now, suppose $\mathfrak{sl}(2,\mathbb{C})$ does contain a nonzero proper ideal $\mathfrak{h}$, then we take an arbitrary nonzero element $g := ax+by+ch \in \mathfrak{h}$ with $a,b,c \in \mathbb{C}$. We split this into two cases: firstly, if $a=b=0$ then $c \neq 0$ and $ch \in \mathfrak{h}$, which implies $h \in \mathfrak{h}$, and from the bracket relations we can then obtain both $x$ and $y$ resulting in $\mathfrak{h}=\mathfrak{g}$. Secondly, if $a \neq 0$ or $b \neq 0$ then we have

$$[x,g]=[x,ax+by+ch]=bh-2cx \in \mathfrak{h}, \text{ so } [x,bh-2cx]=-2bx \in \mathfrak{h}$$

and

$$[y,g]=[y,ax+by+ch]=-ah+2cy \in \mathfrak{h}, \text{ so } [y,-ah+2cy]=-2ay \in \mathfrak{h},$$

implying that either $x$ or $y$ has to be in the ideal $\mathfrak{h}$. The bracket relations would in both cases yield $h \in \mathfrak{h}$ and then $\mathfrak{h}=\mathfrak{g}$. Hence, a proper nonzero ideal of $\mathfrak{sl}(2,\mathbb{C})$ cannot exist, so that it is a simple Lie algebra.

From the definition of simple Lie algebras, they seem to be like atoms, in the sense that they cannot be broken up further. However, we might ask if they are also the building blocks for all Lie algebras. This is almost true, which we will discuss after defining a new type of Lie algebra.

Let us recall the direct sum of vector spaces. There are two notions of a direct sum, one external and one internal. Let $V$ and $W$ be arbitrary vector spaces, then the external direct sum yields a vector space $V \oplus W$ with elements $(v,w)$ where $v \in V$ and $w \in W$, and having componentwise operations. The internal direct sum is a property of a vector space $U$, where $U=\{v+w : v \in V, w \in W\}$ and $V \cap W = 0$, as a result each $u \in U$ can be written uniquely as a sum $u=v+w$ where $v \in V$ and $w \in W$. These constructions are interchangeable as they lead to isomorphic vector spaces. With this behind us, we define how a direct sum of Lie algebras can be made.

**Definition 2.6.** The **direct sum** of two Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ is defined as the vector space $\mathfrak{g}=\mathfrak{g}_1 \oplus \mathfrak{g}_2$ with the bracket operation defined as

$$[(x,x'),(y,y')]_{\mathfrak{g}_1\oplus\mathfrak{g}_2} = ([x,x']_{\mathfrak{g}_1},[y,y']_{\mathfrak{g}_2}) \text{ for all } (x,x'),(y,y') \in \mathfrak{g},$$

so that $[(x,0),(0,x')]=0$ for all $x \in \mathfrak{g}_1$, $x' \in \mathfrak{g}_2$.

**Definition 2.7.** If a Lie algebra $\mathfrak{g}$ is a direct sum of simple Lie algebras $\mathfrak{g}_1, \dots \mathfrak{g}_n$, so that $\mathfrak{g}=\mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$, then we call $\mathfrak{g}$ a **semisimple** Lie algebra.

*Remark.* So, any simple Lie algebra is also a semisimple Lie algebra.

In all the following sections we will be focusing only on semisimple Lie algebras. This is motivated by a result of the *Levi-Malcev Theorem* (see Elduque [4], p27), which tells us that all Lie algebras can be divided into a semisimple part and in a so called solvable part. Thus, understanding the semisimple Lie algebras will bring us halfway to understanding Lie algebras in general.

## 2.2 Lie algebra modules and representations

In this section we discuss representations and the equivalent notion of modules. Where we finish by looking at the building blocks of such representations and modules, culminating in Weyl's Theorem.

*In this section* $\mathfrak{g}$ *will denote a complex semisimple Lie algebra, unless otherwise stated.*

**Definition 2.8.** Given a (finite-dimensional) vector space $V$ and a Lie algebra $\mathfrak{g}$. A (finite-dimensional) **representation** of $\mathfrak{g}$ is a Lie algebra homomorphism $\rho : \mathfrak{g} \to \mathrm{End}(V)$, so the map is linear and satisfies:

$$\rho([x,y]_{\mathfrak{g}}) = [\rho(x), \rho(y)]_{\mathrm{End}(V)} = \rho(x)\rho(y) - \rho(y)\rho(x).$$

*Remark.* When the homomorphism $\rho$ is clear, the vector space $V$, usually named the representation space in this context, is sometimes also called the representation.

*Example.* For subalgebras of $\mathrm{End}(V)$ the inclusion map, is called the *natural representation* and immediately satisfies the linearity and bracket condition. This is essentially how we view $\mathfrak{sl}(2,\mathbb{C})$, via the representation $i : \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V)$.

*Example.* As the name suggests, the adjoint representation of $\mathfrak{g}$ from our example of a Lie algebra homomorphism is indeed a representation. Taking the adjoint representation of $\mathfrak{sl}(2,\mathbb{C})$ we can find the matrices of $\mathrm{ad}(x), \mathrm{ad}(y)$ and $\mathrm{ad}(h)$ by examining the bracket relations. The Lie algebra has an ordered basis given by $(x, y, h)$, so

$$\mathrm{ad}(x) = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathrm{ad}(y) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & -1 & 0 \end{bmatrix}, \quad \mathrm{ad}(h) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

In Section 3 we will see particularly fruitful results from studying the adjoint representation. Which show numerous structural properties of semisimple Lie algebras.

An equivalent way, as we will show, to talk about representations is through the language of modules and actions.

**Definition 2.9.** A **$\mathfrak{g}$-module** is a vector space $V$ with an operation $\mathfrak{g} \times V \to V$ called the **$\mathfrak{g}$-action**, denoted $(x, v) \mapsto x \cdot v$ or just $xv$, satisfying the following conditions: Given $x, y \in \mathfrak{g}$; $v, w \in V$; $a, b \in \mathbb{C}$,

(M1) $(ax + by) \cdot v = a(x \cdot v) + b(y \cdot v)$,

(M2) $x \cdot (av + bw) = a(x \cdot v) + b(x \cdot w)$,

(M3) $[x, y] \cdot v = x \cdot (y \cdot v) - y \cdot (x \cdot v)$.

*Remark.* This is in particular a *left-module* over the Lie algebra $\mathfrak{g}$, as the multiplication is performed from the left. Similarly, a right-module or a bimodule can be defined.

The condition (M1) and (M2) give two types of linearities, one in the Lie algebra elements and one in the vectors. The final condition (M3) encodes the structure of $\mathfrak{g}$. A particularly helpful interpretation of the action is to view it as *scalar multiplication*, where $\mathfrak{g}$ functions as the scalar set on the vector space $V$. The concept of a module came forth from a generalization of vector spaces, so in light of its origins the interpretation of scalar multiplication is justified.

The equivalence between the representation and modules formulations can be seen by translating the representation to the action and vice versa. Indeed, if we have a representation $\rho : \mathfrak{g} \to \operatorname{End}(V)$ with $V$ a vector space, then we can define the action as

$$x \cdot v := \rho(x)v, \text{ for all } x \in \mathfrak{g}, v \in V.$$

This action satisfies (M1) and (M2) by the linearity of the representation, then (M3) follows from $\rho$ being a Lie algebra homomorphism and from the bracket relations of endomorphisms on $V$. Conversely, if $V$ is a $\mathfrak{g}$-module, then we can find the corresponding representation by defining the linear map $\rho(X) : v \mapsto x \cdot v$ for all $x \in \mathfrak{g}$.

*Remark.* We use a shorthand notation for repeated application of the action: $x^k \cdot v$ will mean $\rho(x)^k v$, and not $\rho(x^k)v$.

*Example.* Let $V$ and $W$ be $\mathfrak{g}$-modules, then given $v \in V$ and $w \in W$, we can define the action on the direct sum $V \oplus W$ by $x \cdot (v, w) = (x \cdot v, x \cdot w)$ for all $x \in \mathfrak{g}$. Conditions (M1) and (M2) are straightforward, the third we work out below. For any $x, y \in \mathfrak{g}$ and $(v, w) \in V \oplus W$,

$$\begin{aligned}[x, y] \cdot (v, w) &= \big(x \cdot (y \cdot v) - y \cdot (x \cdot v),\, x \cdot (y \cdot w) - y \cdot (x \cdot w)\big) \\ &= \big(x \cdot (y \cdot v),\, x \cdot (y \cdot w)\big) - \big(y \cdot (x \cdot v),\, y \cdot (x \cdot w)\big) \\ &= x \cdot (y \cdot (v, w)) - y \cdot (x \cdot (v, w)).\end{aligned}$$

Thus, $V \oplus W$ is also a $\mathfrak{g}$-module.

### 2.2.1 Module homomorphisms

As in any other algebraic theory, we can introduce the notion of a homomorphism for modules and hope for the important homomorphism results to apply as well.

**Definition 2.10.** A **homomorphism of $\mathfrak{g}$-modules** is a linear map $\phi : V \to W$ between two $\mathfrak{g}$-modules that preserves the action, so for any $x \in \mathfrak{g}$ and $v \in V$ we have

$$\phi(x \cdot v) = x \cdot \phi(v).$$

This map is called an **isomorphism** of $\mathfrak{g}$-modules, if it is bijective and we write $V \cong W$.

*Example.* Given a direct sum of $\mathfrak{g}$-modules $V \oplus W$, we can define the corresponding *projections* $p_1 : V \oplus W \to V$ and $p_2 : V \oplus W \to W$ onto the first and second argument respectively. We show that $p_1$ is a module homomorphism, the argument for $p_2$ is analogous. Let $x \in \mathfrak{g}$ and $(v, w) \in V \oplus W$, then

$$p_1(x \cdot (v, w)) = p_1((x \cdot v, x \cdot w)) = x \cdot v = x \cdot p_1((v, w)).$$

A homomorphism is always accompanied with its *kernel* and *image*, which are defined as usual. For a $\mathfrak{g}$-module homomorphism $\phi : V \to W$,

$$\ker \phi = \{v \in V : \phi(v) = 0\}, \quad \phi[V] = \{\phi(v) \in W : v \in V\}.$$

These two sets already say a great deal about the homomorphism $\phi$, namely $\ker \phi = 0$ is equivalent to $\phi$ being injective and $\phi[V] = W$ is equivalent to $\phi$ being surjective. Having a homomorphism and isomorphism theorem, as in the group theory or vector space setting, at our disposal will be especially helpful once we work on the classification of the Lie algebra modules. Although, before we can call to such theorems, we need to introduce two new types of modules.

**Definition 2.11.** Let $V$ be a $\mathfrak{g}$-module, then a subspace $W$ satsifying

$$x \cdot w \in W \text{ for all } x \in \mathfrak{g} \text{ and } w \in W,$$

is called a **$\mathfrak{g}$-submodule**, or just a submodule when it is clear what Lie algebra is acting on $W$. We also call a $\mathfrak{g}$-submodule a **$\mathfrak{g}$-invariant subspace**, as the action does not move the space out of $W$.

*Remark.* We introduce the notation $\mathfrak{g} \cdot W := \{x \cdot w : x \in \mathfrak{g}, w \in W\}$, which translates the submodule condition to $\mathfrak{g} \cdot W \subset W$.

*Example.* For a $\mathfrak{g}$-module $V$, there are of course the trivial examples of submodules, namely 0 and $V$ itself.

*Example.* Given a $\mathfrak{g}$-module homomorphism $\phi : V \to W$, its kernel and image are both $\mathfrak{g}$-submodules of $V$ and $W$ respectively. Indeed,

$$\phi(\mathfrak{g} \cdot \ker \phi) = \mathfrak{g} \cdot \phi(\ker \phi) = \mathfrak{g} \cdot 0 = 0 \text{ and } \mathfrak{g} \cdot \phi[V] = \phi[\mathfrak{g} \cdot V] \subset \phi[V].$$

This immediately implies that the $\mathfrak{g}$-modules $V$ and $W$ are submodules of $V \oplus W$, as images of their corresponding projections.

Very similar to the vector space quotient, we can construct a $\mathfrak{g}$-module quotient. So, let us briefly recall what a *vector space quotient* is. Given a vector space $V$ with subspace $W$, the quotient space $V/W$ is the vector space consisting of the equivalence classes $[v] = v + W := \{v + w : w \in W\}$, where the equivalence relation is defined as $v \sim v'$ if and only if $v - v' \in W$. The operations in the quotient space are defined as $(v + W) + (v' + W) = (v + v') + W$ and $\lambda(v + W) = \lambda v + W$, where $v, v' \in V$ and $\lambda \in \mathbb{C}$. These operations were also well-defined.

**Definition 2.12.** Given a $\mathfrak{g}$-module $V$ with submodule $W$, the **quotient module** $V/W$ is the quotient vector space $V/W$ with a natural $\mathfrak{g}$-module structure given by

$$x \cdot (v + W) = x \cdot v + W \in V/W \text{ for all } x \in \mathfrak{g}.$$

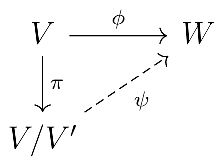
The *natural* or *canonical map* is given by $\pi : V \to V/W$ with $v \mapsto v + W$.

The $\mathfrak{g}$-module structure of the quotient module $V/W$ can be shown to be well-defined as follows. Take an arbitrary $v \in V$, if $v' + W = v + W$ then $v' = v + w$ for some $w \in W$. Now, as $W$ is a submodule we obtain the following for all $x \in \mathfrak{g}$,

$$\begin{aligned} x \cdot (v' + W) &= x \cdot v' + W = x \cdot (v + w) + W \\ &= x \cdot v + x \cdot w + W = x \cdot v + W = x \cdot (v + W). \end{aligned}$$

Finally, we arrive at the homomorphism and isomorphism theorems.

**Theorem 2.13** (Homomorphism theorem)**.** *If $\phi : V \to W$ is a $\mathfrak{g}$-module homomorphism and $V'$ is any $\mathfrak{g}$-submodule of $V$ such that $V' \subset \ker \phi$, then there exists a unique homomorphism $\psi : V/V' \to W$ such that $\phi = \psi \circ \pi$, thus making the following diagram commute:*




**Theorem 2.14** (Isomorphism theorem)**.** *If $\phi : V \to W$ is a $\mathfrak{g}$-module homomorphism, then $V/\ker \phi \cong \phi[V]$.*

*Remark.* The proof for these theorems are nearly identical as in the group theoretic or vector space variant. The reader can find the proof in any standard algebra textbook.

In general, the purpose of homomorphisms is to preserve the relevant structure. We observe this in modules through the preservation of the action, in Lie algebras through preservation of the Lie bracket, and even for vector spaces this same notion applies where its homomorphisms are the linear maps. All the results and constructions regarding homomorphism can be passed over in an analogous manner to any algebraic theory, which we will be doing again in Section 4 for objects called *associative algebras*.

### 2.2.2 Complete reducibility

Returning to more module-specific concepts, we naturally want to reduce the study of modules to possible building blocks. So, we introduce the following definitions to help us attain such modules.

**Definition 2.15.** A $\mathfrak{g}$-module $V$ is called **irreducible**, when the only submodules it contains are trivial, so 0 and $V$. Otherwise, we call the module **reducible**.

**Definition 2.16.** When a $\mathfrak{g}$-module $V$ can be written as a direct sum of proper nonzero submodules, so $V = V_1 \oplus \cdots \oplus V_n$ where $V_i \neq 0$ are proper submodules of $V$, then $V$ is called **decomposable**. Otherwise, we say that $V$ is **indecomposable**.

These definitions feel very similar, but they should not be confused with each other because in general they are not equivalent. The definition of decomposability directly implies reduciblity, as a decomposable module must have a proper nonzero submodule in its sum. So, irreducible implies indecomposable, but the other way does not hold in general, this fails in particular for certain infinite-dimensional modules which we will encounter in Section 5. Thus, irreducibility is a stricter condition than indecomposability, therefore the irreducible modules will be chosen as the building blocks for our modules.

**Definition 2.17.** A $\mathfrak{g}$-module $V$ is called **completely reducible** if $V$ is a direct sum of irreducible $\mathfrak{g}$-submodules.

*Remark.* With this definition, irreducible modules are also completely reducible.

Fortunately, for finite-dimensional modules we do not have to dwell on the differences between decomposable and reducible as they are equivalent in this context, as a result of Weyl's theorem. We will not prove Weyl's theorem, since the proof will not add much to the focus of our discussion. For the proof we refer the reader to (Humphreys [8], Theorem 6.3).

**Theorem 2.18** (Weyl's Theorem)**.** *If $V$ is a finite-dimensional module over a semisimple Lie algebra, then $V$ is completely reducible.*

The important takeaway of Weyl's theorem is that we can restrict our study of finite-dimensional modules to purely the irreducible modules. Knowing how they behave will yield enough information to push their properties further to reducible finite-dimensional modules via the direct sums.

## 2.3 Linear algebra background

*All the vector spaces in this section will be finite-dimensional, unless otherwise stated.*

**Definition 2.19.** Let $V$ be a vector space over $\mathbb{C}$, then we can define the **(algebraic) dual** of $V$, denoted $V^*$, as the collection of all **linear functionals** $f : V \to \mathbb{C}$.

*Example.* Let $V$ be an inner product space with inner product $\langle -, - \rangle$. Then, given a vector $v \in V$, the map $f_v := \langle v, - \rangle$ is an example of a linear functional.

*Remark.* The dual of a vector space $V$ becomes a vector space under the following operations, for all $f, g \in V^*$, $v \in V$ and $a \in \mathbb{C}$,

$$(f + g)(v) := f(v) + g(v) \text{ and } (af)(v) := a(f(v)).$$

We call a linear map $A : V \to V$ *diagonalizable*, when there exists a basis of the vector space $V$ consisting of eigenvectors of the map $A$. This allows us to represent $A$ by a diagonal matrix, where the entries are the eigenvalues of $A$. We can rephrase this into the existence of an eigenspace decomposition of $V$, where we write $V$ as a direct sum of the $A$-eigenspaces

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n},$$

where the $\lambda_i$ are distinct eigenvalues of $A$. It can happen that a finite collection of linear maps are all diagonalizable with the same set of eigenvectors (with possibly different eigenvalues), then we say that they are *simultaneously diagonalizable*. There is a beautiful and powerful result regarding simultaneous diagonalizability, which we will be using quite a few times in later sections. A proof can be found in (Erdmann and Wildon [5], Lemma 16.7).

**Theorem 2.20.** *Let $A_1, \ldots, A_m : V \to V$ be a family of diagonalizable linear maps, then they are simultaneously diagonalizable if and only if they commute.*

Admittedly, not all linear maps can be diagonalized. However, we are working over the algebraically closed field $\mathbb{C}$, therefore all the characteristic polynomials will split into linear factors. So, every eigenvalue is contained in the field we are working in, albeit some with different multiplicities than others. This leads to the well known *Jordan canonical form*:

$$J = \begin{bmatrix} J_{t_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{t_r}(\lambda_r) \end{bmatrix},$$

where the empty entries are zero and the $J_{t_i}(\lambda_i)$ are Jordan blocks for some $t_i \in \mathbb{Z}_{>0}$ and $\lambda_i \in \mathbb{C}$ of the form

$$J_{t_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}.$$

Every linear map $A \in \mathrm{End}(V)$ can be written in such a matrix form via an appropriate basis. Leaving out the details, this form shows that it is possible to decompose $A$ into a diagonalizable and nilpotent part such that $A = D + N$ where $D$ and $N$ commute, and is uniquely determined. This is called the **Jordan decomposition**.

It turns out, that it is also possible to introduce an *abstract Jordan decomposition* for elements in any semisimple Lie algebra $\mathfrak{g}$. Each $x \in \mathfrak{g}$ determines unique elements $d, n \in \mathfrak{g}$ such that $x = d + n$ is the usual Jordan decomposition of $\mathrm{ad}(x) \in \mathrm{End}(\mathfrak{g})$. This then yields the abstract Jordan decomposition $x = d + n$, where $\mathrm{ad}(d)$ is diagonalizable, $\mathrm{ad}(n)$ is nilpotent and $[d, n] = 0$. We finish this section with a useful theorem.

**Theorem 2.21.** *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\rho : \mathfrak{g} \to End(V)$ a (finite-dimensional) representation of $\mathfrak{g}$. If $x = d + n$ is the abstract Jordan decomposition of $x \in \mathfrak{g}$, then $\rho(x) = \rho(d) + \rho(n)$ is the usual Jordan decomposition of $\rho(d)$.*

The consequence that we will use in later sections is that elements $d \in \mathfrak{g}$ where $\mathrm{ad}(d)$ is diagonalizable, return a diagonalizable endomorphism $\rho(d)$ for any representation $\rho : \mathfrak{g} \to \mathrm{End}(V)$.

## 2.4 Representations of $\mathfrak{sl}(2, \mathbb{C})$

Before we study the representations of arbitrary semisimple Lie algebras we first need to focus on $\mathfrak{sl}(2, \mathbb{C})$, as it will play a key role in determining the representations of other Lie algebras. We will see hints of how fundamental $\mathfrak{sl}(2, \mathbb{C})$ really is in all of this in Section 3.

Recall that the basis of this Lie algebra is given by

$$x = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

inducing the relations $[h, x] = 2x$, $[h, y] = -2y$, $[x, y] = h$. Let $\rho : \mathfrak{g} \to \mathrm{End}(V)$ be any representation. Since we have seen that $\mathrm{ad}(h)$ is represented by a diagonal matrix, the endomorphism $\rho(h) \in \mathrm{End}(V)$ is diagonalizable by Theorem 2.21. Hence we can decompose $V$ into a direct sum of eigenspaces $V_\lambda = \{v \in V : \rho(h)v = \lambda v\}$ where $\lambda \in \mathbb{C}$. Rephrasing this in the language of modules, we say that the action of $h$ on $V$ is diagonalizable and satisfies $h \cdot v = \lambda v$ for $v \in V_\lambda$. We introduce the following terminology.

**Definition 2.22.** Let $V$ be a module over $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$. We call $\lambda \in \mathbb{C}$ a **weight** of $V$ when there exists a nonzero vector $v \in V$ such that

$$h \cdot v = \lambda v.$$

The vector $v$ is called a **weight vector** and the corresponding space $V_\lambda$ a **weight space**.

It might seem strange to define new terminology, while we can simply use the already existing language of eigenvalues and eigenvectors. Though once we look at arbitrary semisimple Lie algebras, we will encounter more elements like $h$ that have diagonalizable actions on $V$, which necessitates us to work with a particular generalization of eigenvalues and eigenvectors, which we will first encounter in Section 3.

Naturally we would also like to know how $x$ and $y$ act on the weight spaces of $V$, as this will allow us to understand how $\mathfrak{sl}(2, \mathbb{C})$ acts on $V$ as a whole. Given a vector $v$ in the weight space $V_\lambda$ we find that

$$\begin{aligned} h \cdot (x \cdot v) &= [h, x] \cdot v + x \cdot (h \cdot v) \\ &= 2x \cdot v + x \cdot (\lambda v) = (\lambda + 2) x \cdot v, \end{aligned}$$

which shows that for a weight vector $v \in V_\lambda$ the vector $x \cdot v$ is again a weight vector, but then in $V_{\lambda+2}$. We perform a similar calculation for $y$,

$$\begin{aligned} h \cdot (y \cdot v) &= [h, y] \cdot v + y \cdot (h \cdot v) \\ &= -2y \cdot v + y \cdot (\lambda v) = (\lambda - 2) y \cdot v, \end{aligned} \tag{2.1}$$

showing that $v \in V_\lambda$ under the action of $y$ gets mapped into $V_{\lambda-2}$. We summarize this in the following lemma.

**Lemma 2.23.** *If $v \in V_\lambda$, then $x \cdot v \in V_{\lambda+2}$ and $y \cdot v \in V_{\lambda-2}$.*

So we can find new weights and weight spaces through the actions of $x$ and $y$. But since $V = \bigoplus_\lambda V_\lambda$ has finite dimension, we cannot do this indefinitely. So, there must be a weight space $V_m \neq 0$ such that $V_{m+2} = 0$. Meaning that any nonzero vector in $V_m$ is *annihilated* by $x$, so $x \cdot v = 0$ for $v \in V_m$. Such vectors are said to be **maximal vectors** and the associated weight $m$ is called the **highest weight** of $V$. Now we are set to state the fundamental results regarding finite-dimensional irreducible modules of $\mathfrak{sl}(2, \mathbb{C})$.

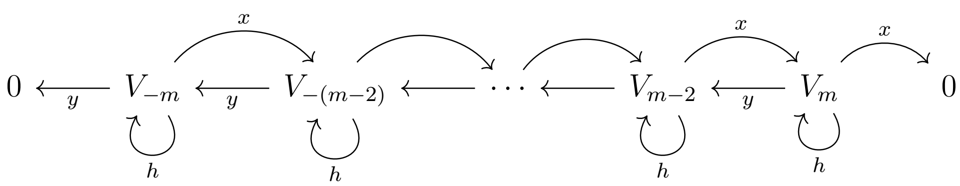
**Theorem 2.24.** *Let $V$ be a finite-dimensional irreducible module over $\mathfrak{sl}(2, \mathbb{C})$.*

*(1) Relative to $h$, the module $V$ is a direct sum of weight spaces $V_\lambda$ with $\lambda = m, m-2, \ldots, -(m-2), -m$, where $m = \dim V - 1$.*

*(2) All the weight spaces of $V$ are one-dimensional and in particular $V$ has a unique maximal vector up to scalar multiples, which has highest weight $m$.*

*(3) For every $m \in \mathbb{Z}_{\geq 0}$, there exists a unique irreducible module $V(m)$ of $\mathfrak{sl}(2, \mathbb{C})$, up to isomorphism.*

From the last statement we can say that there exists a bijection from $\mathbb{Z}_{\geq 0}$ to the collection of finite-dimensional irreducible modules of $\mathfrak{sl}(2, \mathbb{C})$ via the map $m \mapsto V(m)$ where $m$ is the highest weight. We can find an analogous bijection for such modules of arbitrary semisimple Lie algebras. This will be proven in Section 5. Moreover, the structure of irreducible modules $V$ can now be visualized as follows,




To prove the theorem, we first need to understand the Lie algebra actions on irreducible modules. Luckily we can find explicit relations for the actions as shown in the following lemma.

**Lemma 2.25.** *Let $V$ be a finite-dimensional irreducible module over $\mathfrak{sl}(2, \mathbb{C})$ and choose a maximal vector $v_0 \in V_m$, now set $v_{-1} = 0$ and $v_k = y^k \cdot v_0$ for $k \geq 0$. Then we find that for $k \geq 0$,*

*(a)* $h \cdot v_k = (m - 2k)v_k$,

*(b)* $y \cdot v_k = v_{k+1}$,

*(c)* $x \cdot v_k = k(m - (k-1))v_{k-1}$.

*Remark.* These relations also hold for infinite-dimensional modules. As seen in the proof, no part of the arguments rely on the finite-dimensionality of $V$.

*Proof.* The first condition $(a)$ is obtained by repeated application of the calculation in (2.1) and $(b)$ comes from how we have defined $v_k$. Lastly $(c)$ requires some more work. To prove this we apply induction on $k$. So for $k = 0$ we have $x \cdot v_0 = 0$ since $v_0$ is a maximal vector. This aligns with the condition in $(c)$. Now suppose that the equality holds for $k$, then

$$\begin{aligned} x \cdot v_{k+1} &= x \cdot (y \cdot v_k) = [x, y] \cdot v_k + y \cdot (x \cdot v_k) \\ &= h \cdot v_k + y \cdot (k(m - (k-1))v_{k-1}) \\ &= (m - 2k)v_k + k(m - (k-1))y \cdot v_{k-1} \\ &= ((m - 2k) + k(m - (k-1)))v_k = (k+1)(m-k)v_k. \end{aligned}$$

□

If we define the subspace $W \subset V$ as the span of the vectors in $\{v_k : k \geq 0\}$, then Lemma 2.25 tells us that every $\mathfrak{sl}(2, \mathbb{C})$-action sends elements in $W$ back into $W$. So the subspace is a submodule of $V$. But since $V$ is irreducible, we must have that $W$ is all of $V$ or only 0. Because $W$ contains a nonzero element, namely the maximal vector $v_0$, the collection $\{v_k : k \geq 0\}$ must span $V$. Additionaly, property $(a)$ of Lemma 2.25 implies that each $v_k$ is linearly independent from the other vectors, since they each have different eigenvalues. So, we can conclude that $(v_0, v_1, v_2, \dots)$ actually forms an ordered basis of $V$. However, remember that our module is finite-dimensional, so there must exist some nonnegative integer $d = \dim V - 1$ such that $v_d \neq 0$ and $v_{d+1} = 0$. Hence,

$$0 = x \cdot v_{d+1} = (d+1)(m-d)v_d,$$

from which it follows that $m = d$. So we have obtained the following,

**Lemma 2.26.** *Given a maximal vector $v_0$ with weight $m$ of an irreducible $\mathfrak{sl}(2, \mathbb{C})$-module $V$, an ordered basis of $V$ is given by $(v_0,\ y \cdot v_0,\ \dots,\ y^m \cdot v_0)$.*

*Example.* Now that we have a basis for $V$ with representation $\rho : \mathfrak{sl}(2, \mathbb{C}) \to \mathrm{End}(V)$, we can find the corresponding matrices of the endomorphisms $\rho(x), \rho(y)$ and $\rho(h)$. Let us look at the case $\dim V = 4$ (assuming such a representation exists, which the theorem justifies), so with the ordered basis $(v_0, v_1, v_2, v_3)$. From Lemma 2.25 we can determine what the matrices look like with this basis,

$$\rho(x) = \begin{bmatrix} 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \rho(y) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \rho(h) = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$

Notice that $\rho(x)$, $\rho(y)$ are nilpotent and $\rho(h)$ is diagonal, exactly as $x, y$ and $h$ are in their standard matrix representations, in accordance to what Theorem 2.21 implies.

Now we are ready to prove Theorem 2.24.

*Proof of Theorem 2.24.* Notice that now the highest weight of $V$ is fully determined to be $m = \dim V - 1$ and is therefore also a nonnegative integer. Property $(a)$ of Lemma 2.25 also shows that all the weights of $h$ in $V$ are $\lambda = m, m-2, \ldots, -(m-2), -m$, which completes the first result *(1)*. In addition, $(a)$ of Lemma 2.25 also implies that the weight vectors in the basis for $V$ as in Lemma 2.26, have distinct weights. So the weights $\lambda$ each occur only for one vector, therefore the weight spaces $V_\lambda$ are each spanned by one weight vector and thus they all have dimension one. Using that $V$ uniquely determines the highest weight, we can conclude that the maximal vector $v_0$ with weight $m$ is the only one up to scalar multiplication. This proves the second result *(2)*.

For the third result we first note that there exists *at most one* irreducible $\mathfrak{sl}(2,\mathbb{C})$-module $V$ for each possible dimension $m+1$, with $m \in \mathbb{Z}_{\geq 0}$. Indeed, Lemma 2.25 fully determines the action of the Lie algebra on an irreducible module of a given dimension, allowing for only one module $V$. Secondly, to show *existence* for every highest weight $m \in \mathbb{Z}_{\geq 0}$, we can define a module $V(m)$ with basis as in Lemma 2.26 and the actions defined as in Lemma 2.25. Then for any $v_k$,

$$\begin{aligned}[h,x]\cdot v_k &= h\cdot(x\cdot v_k) - x\cdot(h\cdot v_k) = k(m-(k-1))h\cdot v_{k-1} - (m-2k)x\cdot v_k \\ &= (k(m-(k-1))(m-2(k-1)) - (m-2k)k(m-(k-1)))v_{k-1} \\ &= 2k(m-(k-1))v_{k-1} = 2x\cdot v_k,\end{aligned}$$

similarly we can check the other relations to find that they are satisfied by the actions defined in this manner. To prove *irreducibility* of $V$, we take $W$ as a nonzero invariant subspace of $V(m)$. Let $w \in W$ be nonzero, then we can write the vector as a linear combination of the basis elements over $\mathbb{C}$,

$$w = a_0v_0 + \cdots + a_{m-1}v_{m-1}, \text{ where at least one coefficient } a_i \neq 0.$$

Since $W$ is an invariant subspace, all actions on the elements in $W$ will end up in $W$ again. Exploiting this property, we use that the action of $y$ on $w$ increases the indices of the basis vectors by one, so we obtain

$$y^{m-i-1}\cdot w = a_0v_{m-i-1} + \cdots + a_iv_{m-1} \in W,$$

where the vectors with higher indices than $v_i$ are annihilated. Repeating a similar calculation with an action of $x$ yields,

$$x^{m-i-1}\cdot(y^{m-i-1}\cdot w) = ca_iv_i \in W,$$

where $c \in \mathbb{Z}$ comes from the factors $k(m-(k-1))$ that the action of $x$ brings. Thus $v_i \in W$ and now we can obtain any other basis vector via the actions of $x$ or $y$. So this would imply that $W = V$, thus $V$ only contains trivial submodules and hence $V$ is irreducible. Proving that we can find a unique irreducible module for every $m \in \mathbb{Z}_{\geq 0}$. □

Finally we arrive at arbitrary modules. Since Weyl's theorem (Theorem 2.18) applies here, as $\mathfrak{sl}(2,\mathbb{C})$ is semisimple, we can decompose any module into a direct sum of irreducible modules. This allows us to extract a great amount from Theorem 2.24 and apply those results to the irreducible components.

**Theorem 2.27.** *Let $V$ be any finite-dimensional module over $\mathfrak{sl}(2, \mathbb{C})$. Then the following statements hold.*

*(1) The weights of $h$ in $V$ are all integers. Furthermore, when a weight vector $v$ has weight $\lambda$ and $x \cdot v = 0$, then $\lambda$ is nonnegative.*

*(2) If $k \in \mathbb{Z}$ is a weight, then so are $-|k|, -(|k| - 2), \ldots, |k| - 2, |k|$.*

*(3) In the decomposition of $V$ into a direct sum of irreducible submodules, $V = \bigoplus_{j=1}^{n} W_j$, the number of summands is $n = \dim V_0 + \dim V_1$.*

*Proof.* We can assume $V \neq 0$ and write the module as a direct sum of irreducible submodules, $V = \bigoplus_{j=1}^{n} W_j$. All the weight spaces of $V$ are each weight spaces of a particular $W_j$. From this we can conclude due to Theorem 2.24 that all weights are integers. In particular, when a weight vector is annihilated by $x$ then the weight itself is a highest weight in the irreducible submodule that contains the weight vector, hence the weight is also nonnegative. Using Theorem 2.24 again, *(2)* now follows directly from how all the weights in the irreducible case are determined by the highest weights.

For the last result, we note that every irreducible module will always contain either the weight 0 or 1 but not both. Indeed, every weight in such an irreducible module has the form $m - 2n$ for some $n = 0, 1, \ldots m$. So their weights are all odd or all even, hence only 0 or 1 will occur and not both. This allows us to count how many $W_j$'s the module $V$ is composed of, namely by counting the number of one-dimensional weight spaces for the weight 0 and also for 1, which is equal to $\dim V_0 + \dim V_1$. □

# 3 Roots of Lie algebras

In Section 2.4, we have seen the important role that the element $h$ has played in first determining the action of $\mathfrak{sl}(2, \mathbb{C})$ on its modules as in Lemma 2.23, and secondly in decomposing the modules in their weight spaces as in Theorem 2.24. When we broaden our focus to more general semisimple Lie algebras, we encounter multiple elements that have the same properties as $h$ in $\mathfrak{sl}(2, \mathbb{C})$.

This leads us to define a collection of such elements, namely the Cartan subalgebra, which we discuss in Section 3.1. This opens up the path to study roots of Lie algebras, which reveal numerous important structural properties of semisimple Lie algebras that are crucial to our main classification problem. Then, we will see that the roots themselves form a peculiar geometric structure, called a root system. So, we focus on these systems in Section 3.2 and in particular on their symmetries under the Weyl group. The properties arising from these symmetries will return once we are in Section 5.3.

A handful of proofs and other details have been left out, if the reader desires to read them, we refer them to (Humphreys [8], Sec. 8-10).

*In this section* $\mathfrak{g}$ *denotes a complex semisimple Lie algebra, unless otherwise stated.*

## 3.1 Cartan subalgebras

**Definition 3.1.** For a complex semisimple Lie algebra $\mathfrak{g}$ we call a subalgebra $\mathfrak{h} \subset \mathfrak{g}$ a **Cartan subalgebra** if it has the following properties:

(C1) $\mathfrak{h}$ is abelian,

(C2) for all $h \in \mathfrak{h}$, the endomorphism $\mathrm{ad}(h)$ is diagonalizable,

(C3) $\mathfrak{h}$ is a maximal Lie subalgebra satisfying the two above conditions.

*Remark.* The maximality in (C3), means that any subalgebra $\mathfrak{h}' \subset \mathfrak{g}$ satisfying (C1) and (C2) are contained in a corresponding Cartan subalgebra $\mathfrak{h}$.

Properties (C1) and (C2) will allow us to use Theorem 2.20 to find an eigendecomposition for semisimple Lie algebras. The importance of the third property will become clear after Theorem 3.7.

With the given definition of a Cartan subalgebra, we cannot be sure that one even exists. Fortunately, this is not an issue for semisimple Lie algebras, though in general a different definition for a Cartan subalgebra is taken to ensure every Lie algebra contains such a subalgebra. A proof for the following theorem is found in (Hall [7], Prop. 7.11).

**Theorem 3.2.** *Every semisimple Lie algebra contains a Cartan subalgebra as defined in Definition 3.1.*

Since the eigenvalues of $\mathrm{ad}(h)$ played a major role for the $\mathfrak{sl}(2, \mathbb{C})$-modules, we want to approach in a similar manner for more general semisimple Lie algebra. However, since all the elements of the Cartan subalgebras assume the role of $h$, we need to make a small modification for the notion of eigenvalues. This is achieved by having our "eigenvalues" be functions where the input are elements of the Cartan subalgebra, which is done in the following definition.

**Definition 3.3.** Given $\lambda \in \mathfrak{h}^*$, we define

$$\mathfrak{g}_\lambda := \{x \in \mathfrak{g} : [h, x] = \lambda(h)x, \text{ for all } h \in \mathfrak{h}\}$$

If for a nonzero $\alpha \in \mathfrak{h}^*$, we have $\mathfrak{g}_\alpha \neq 0$, then we call $\alpha$ a **root** of $\mathfrak{g}$ and $\mathfrak{g}_\alpha$ the corresponding **root space**. The collection of all the roots of $\mathfrak{g}$ is denoted by $\Phi$.

*Example.* In $\mathfrak{sl}(2, \mathbb{C})$ the roots were given by 2 and -2, and we had the following root spaces: $\mathfrak{g}_2 = \mathbb{C}x$, $\mathfrak{g}_{-2} = \mathbb{C}y$.

With the roots we will be able to reveal numerous structural properties of semisimple Lie algebras. The two most important results regarding the roots that we will encounter, are the root space decomposition and Theorem 3.8 which links $\mathfrak{sl}(2, \mathbb{C})$ to all semisimple Lie algebras.

Before we can achieve those results, we have to introduce a symmetric bilinear form that allows us to advance our study of the roots. This bilinear form also plays an important role in opening up a geometric approach to roots, as we will see in Section 3.2.

**Definition 3.4.** For a Lie algebra $\mathfrak{g}$, the **Killing form** $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ is defined as the symmetric bilinear form

$$\kappa(x, y) = \operatorname{tr}(\operatorname{ad}(x) \circ \operatorname{ad}(y)) \quad \text{for all } x, y \in \mathfrak{g}.$$

The Killing form is indeed symmetric, because of the identity $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for all endomorphisms $A, B$. The bilinearity is a consequence of the adjoint representation being linear, we show the linearity in the first argument. Given $a, b \in \mathbb{C}$ and $x, x', y \in \mathfrak{g}$ we have

$$\begin{aligned}\kappa(ax + bx', y) &= \operatorname{tr}(\operatorname{ad}(ax + bx') \circ \operatorname{ad}(y)) \\ &= \operatorname{tr}\big((a \cdot \operatorname{ad}(x) + b \cdot \operatorname{ad}(x')) \circ \operatorname{ad}(y)\big) \\ &= a \cdot \operatorname{tr}(\operatorname{ad}(x) \circ \operatorname{ad}(y)) + b \cdot \operatorname{tr}(\operatorname{ad}(x') \circ \operatorname{ad}(y)) = a \cdot \kappa(x, y) + b \cdot \kappa(x', y).\end{aligned}$$

The linearity in the second argument is shown similarly. Another important property of the Killing form is its *associativity*, for all $x, y, z \in \mathfrak{g}$

$$\kappa([x, y], z) = \kappa(x, [y, z]),$$

which follows from the identity $\operatorname{tr}(B(AC)) = \operatorname{tr}((AC)B)$ leading to $\operatorname{tr}([A, B]C) = \operatorname{tr}(A[B, C])$ for all $A, B, C \in \operatorname{End}(V)$ on some vector space $V$.

We introduce two concepts for the Killing form. First, we say that for $x, y \in \mathfrak{g}$, they are *orthogonal* relative to $\kappa$, when $\kappa(x, y) = 0$. Next, we say the form is *nondegenerate* when the set $S = \{x \in \mathfrak{g} : \kappa(x,y) = 0 \text{ for all } y \in \mathfrak{g}\}$ only contains 0. So, when $\kappa$ is nondegenerate, the only vector that is orthogonal to all other vectors is 0. With this set, we state the following theorem, without proof, which shows an equivalent definition for semisimplicity of Lie algebras.

**Theorem 3.5** (Cartan's criterion)**.** *A Lie algebra $\mathfrak{g}$ is semisimple if and only if its Killing form $\kappa$ is nondegenerate.*

The following theorem shows how the spaces $\mathfrak{g}_\lambda$, which we defined earlier, interact with each other.

**Theorem 3.6.** *Let $\mathfrak{h} \subset \mathfrak{g}$ be a Cartan subalgebra and $\kappa$ the Killing form of $\mathfrak{g}$. Then the following statements hold.*

*(1) There exist an eigenspace decomposition of $\mathfrak{g}$ as follows:*

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

*(2) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$, for all $\alpha, \beta \in \mathfrak{h}^*$.*

*(3) For all $\alpha, \beta \in \mathfrak{h}^*$, if $\alpha + \beta \neq 0$, then $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$, so we say $\mathfrak{g}_\alpha$ is orthogonal to $\mathfrak{g}_\beta$.*

*(4) The restriction of the Killing form to $\mathfrak{g}_0$ is nondegenerate.*

*Example.* In the case of $\mathfrak{sl}(2,\mathbb{C})$, the decomposition is given by $\mathfrak{sl}(2,\mathbb{C}) = \mathbb{C}y \oplus \mathbb{C}h \oplus \mathbb{C}x$.

*Proof.* For *(1)*, we use the diagonalizability of the $\mathrm{ad}(h)$ maps. Since all $h \in \mathfrak{h}$ commute with each other, this must also be true for each $\mathrm{ad}(h)$, as representations preserve the bracket relations. So, now we can apply Theorem 2.20. From this we conclude that they are all simultaneously diagonalizable. Hence, there exists an eigenbasis that diagonalizes all the maps $\mathrm{ad}(h)$ simultaneously and this yields an eigenspace decomposition in the desired form.

The second part is proven with the Jacobi identity. For $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_\beta$ and $h \in \mathfrak{h}$,

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

so $[x,y] \in \mathfrak{g}_{\alpha+\beta}$.

For *(3)*, we can take an $h \in \mathfrak{h}$ such that $(\alpha+\beta)(h) \neq 0$. Let $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_\beta$, then by associativity of the Killing form we obtain

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

bringing all terms to the left-hand side yields the equality

$$(\alpha+\beta)(h)\kappa(x,y) = 0.$$

So, by our assumption, it follows that $\kappa(x,y) = 0$.

To prove *(4)*, we suppose that the restriction is degenerate, then we can choose an $x \in \mathfrak{g}_0$ such that $\kappa(x, \mathfrak{g}_0) = 0$. According to *(3)*, $\mathfrak{g}_0$ is orthogonal to all $\mathfrak{g}_\alpha$ with $\alpha \in \mathfrak{h}^*$ and $\alpha \neq 0$, thus via the decomposition in *(1)* we find that

$$\kappa(x, \mathfrak{g}) = \sum_{\alpha \in \mathfrak{h}^*} \kappa(x, \mathfrak{g}_\alpha) = 0.$$

So, $\kappa$ is degenerate over $\mathfrak{g}$, but this contradicts Cartan's criterion (Theorem 3.5) for semisimplicity. Hence, $\kappa|_{\mathfrak{g}_0 \times \mathfrak{g}_0}$ is nondegenerate. □

Since all elements of a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ already commute with each other, we know that $\mathfrak{h} \subset \mathfrak{g}_0$. It turns out that the other inclusion holds as well, which is stated in the following theorem that we show without proof. The proof can be found in Humphreys [8].

**Theorem 3.7.** *Let $\mathfrak{h} \subset \mathfrak{g}$ be a Cartan subalgebra of the Lie algebra $\mathfrak{g}$, then $\mathfrak{h} = \mathfrak{g}_0$.*

This has two important consequences: that the eigenspace decomposition of $\mathfrak{g}$ now becomes what is called the *root space decomposition* $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$, and the Killing form restricted to $\mathfrak{h}$ is nondegenerate. The first consequence is quite important; after we look more into the root spaces, we will see how much of the structure for the $\mathfrak{g}$-modules can be gathered out of this decomposition. The second consequence allows us to make a correspondence between $\mathfrak{h}$ and its dual space $\mathfrak{h}^*$, as for any $\alpha \in \mathfrak{h}^*$ there is a unique element $t_\alpha \in \mathfrak{h}$ satisfying, $\alpha(h) = \kappa(t_\alpha, h)$ for all $h \in \mathfrak{h}$. Indeed, if we had another $r_\alpha \in \mathfrak{h}$ with the same property, then

$$\kappa(t_\alpha - r_\alpha, h) = \kappa(t_\alpha, h) - \kappa(r_\alpha, h) = \alpha(h) - \alpha(h) = 0 \quad \text{for all } h \in \mathfrak{h}.$$

But nondegeneracy over $\mathfrak{h}$ forces $t_\alpha - r_\alpha = 0$, which locks in the uniqueness.

Now we arrive at a key element for the classification of finite-dimensional modules. The following result gives us a way to transfer results obtained for $\mathfrak{sl}(2, \mathbb{C})$-modules to general semisimple Lie algebra modules.

**Theorem 3.8.** *Let $\mathfrak{g}$ be a semisimple Lie algebra. For any root $\alpha \in \Phi$, if $x_\alpha$ is a nonzero element of $\mathfrak{g}_\alpha$, then there exists $y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $x_\alpha, y_\alpha$ and $h_\alpha = [x_\alpha, y_\alpha]$ together span a subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}(2, \mathbb{C})$.*

We give a sketch of the proof, for a complete proof see Humphreys [8] or (Erdmann and Wildon [5], Lemma 10.5).

*Sketch of proof.* The proof relies heavily on the nondegeneracy of the Killing form for semisimple Lie algebras to prove nonzero conditions. It starts by showing there exists a $y \in \mathfrak{g}_{-\alpha}$ such that $\kappa(x_\alpha, y) \neq 0$ and $[x_\alpha, y] \neq 0$, where the third part of Theorem 3.6 and the associativity of the Killing form are applied. Then it has to be shown that

$\kappa(t_\alpha, t_\alpha) \neq 0$ holds, where $t_\alpha$ is as defined in the discussion above, as this allows us to find a specific $y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $\kappa(x_\alpha, y_\alpha) = \dfrac{2}{\kappa(t_\alpha, t_\alpha)}$. Then, setting

$$h_\alpha := \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)} \tag{3.1}$$

we obtain all the $\mathfrak{sl}(2, \mathbb{C})$ bracket relations. The elements $x_\alpha, y_\alpha$ and $h_\alpha$ now span a three dimensional subalgebra of $\mathfrak{g}$ which has to be isomorphic to $\mathfrak{sl}(2, \mathbb{C})$. □

We end this section by making a link to the geometry of the roots using the Killing form. The geometric approach will allow us to focus on the important properties without diluting our vision with other baggage. In addition, we will be able to visualize a lot of concepts better.

Because of the correspondence made between $\mathfrak{h}$ and $\mathfrak{h}^*$, we can introduce the dual form of $\kappa$ over $\mathfrak{h}^*$ as follows,

$$(\lambda, \mu) := \kappa(t_\lambda, t_\mu) \quad \text{for all } \lambda, \mu \in \mathfrak{h}^*.$$

We state the following theorem without proof, see Humphreys [8] or (Erdmann and WIldon [5], Lemma 10.11) for full proof.

**Lemma 3.9.** *The set $\Phi$ spans $\mathfrak{h}^*$.*

The lemma allows us to take a basis of roots $\{\alpha_1, \ldots, \alpha_n\}$ for $\mathfrak{h}^*$. It turns out that all the roots in $\Phi$ can be expressed as a linear combination with rational coefficients of the $\alpha_i$'s. Hence, $\mathrm{span}(\{\alpha_1, \ldots, \alpha_n\})$ seen as a real subspace of $\mathfrak{h}^*$ will contain all the roots, and we call this space $E$. It can be shown that the dual form $(-,-)$ is actually a real-valued inner product over $E$. With this we can view $E$ as a *Euclidean space*, a finite-dimensional real vector space endowed with an inner product, allowing us to approach the study of roots geometrically. The details have been left out, but can be found in Humphreys [8].

We now list properties of $\Phi$ without proof (see [8]), which will lay the foundations for how the geometric approach will be carried out.

**Theorem 3.10.** *Let $\mathfrak{g}$ be a Lie algebra, $\mathfrak{h} \subset \mathfrak{g}$ a Cartan subalgebra, $\Phi$ the set of roots of $\mathfrak{g}$ and $E$ the Euclidean space as above. Then:*

*(1) $\Phi$ spans $E$, and 0 does not belong to $\Phi$.*

*(2) If $\alpha \in \Phi$, the only scalar multiples of $\alpha \in \Phi$ are $\pm\alpha$.*

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

*(4) If $\alpha, \beta \in \Phi$, then $\dfrac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z}$.*

## 3.2 Root systems

In this section we focus purely on what structure arises from the properties listed in Theorem 3.10, temporarily leaving the theory we have built up from Lie algebras. Because these structures, called *root systems*, also arise in different areas of mathematics, they are of interest in their own right. Surprisingly, this short break from Lie algebras is not a deviation at all, as by some miracle every root system is isomorphic to a set of roots of a complex semisimple Lie algebra. Though, slightly unfortunate, we will only discuss a small part of this theory on root systems.

We fix an Euclidean space $E$ with an inner product $(-,-)$. Since Theorem 3.10 serves as the foundation for our approach here, we take a closer look at

$$\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha =: \sigma_\alpha(\beta)$$

of part *(3)* but in this more stripped down context. The expression $\sigma_\alpha(\beta)$ can be interpreted quite simple geometrically, namely as the *reflection* of the vector $\beta \in E$ in the hyperplane $P_\alpha = \{\beta \in E : (\beta,\alpha) = 0\}$ orthogonal to $\alpha$. Indeed, for $\beta \in P_\alpha$ the

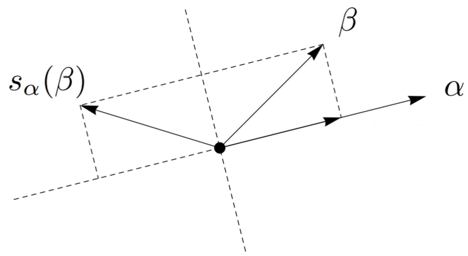



Figure 1: The reflection of $\beta$ in the hyperplane orthogonal to $\alpha$ (Schnell [11])

reflection yields $\sigma_\alpha(\beta) = \beta$, for $\alpha$ itself we have $\sigma_\alpha(\alpha) = -\alpha$, and the reflection preserves the inner product:

$$\begin{aligned}
(\sigma_\alpha(x), \sigma_\alpha(y)) &= \left(x - \frac{2(x,\alpha)}{(\alpha,\alpha)}\alpha,\, y - \frac{2(y,\alpha)}{(\alpha,\alpha)}\alpha\right) \\
&= (x,y) - \frac{2(x,\alpha)}{(\alpha,\alpha)}(\alpha,y) - \frac{2(y,\alpha)}{(\alpha,\alpha)}(x,\alpha) + \frac{4(x,\alpha)(y,\alpha)}{(\alpha,\alpha)^2}(\alpha,\alpha) \\
&= (x,y) - \frac{4(x,\alpha)(y,\alpha)}{(\alpha,\alpha)} + \frac{4(x,\alpha)(y,\alpha)}{(\alpha,\alpha)} = (x,y), \quad \text{for all } x,y \in E.
\end{aligned}$$

It should be noted that any transformation that preserves the inner product, so is orthogonal, is also a linear transformation. Hence, the reflections $\sigma_\alpha$ are linear as well.

We abbreviate $2(\beta,\alpha)/(\alpha,\alpha)$ by $\langle\beta,\alpha\rangle$ as it comes up quite often. Note that $\langle\beta,\alpha\rangle$ is only linear in the first argument.

**Definition 3.11.** A subset $\Phi \subset E$ is called a **root system** in $E$ if the following axioms are satisfied.

(R1) $\Phi$ is finite, spans $E$, and does not contain 0.

(R2) If $\alpha \in \Phi$, the only scalar multiples of $\alpha$ in $\Phi$ are $\pm\alpha$.

(R3) If $\alpha \in \Phi$, the reflection $\sigma_\alpha$ permutes the elements in $\Phi$, so $\sigma_\alpha(\Phi) \subset \Phi$.

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

Compare this definition with Theorem 3.10, and we see that the roots of a semisimple Lie algebra form a root system per construction.

The axioms are already quite restrictive; for lower dimensions of $E$ there are only very few possibilities for root systems. We call $l = \dim E$ the **rank** of the root system $\Phi$. When $l = 1$, the only possibility is determined by (R2) and corresponds to the roots of $\mathfrak{sl}(2, \mathbb{C})$:

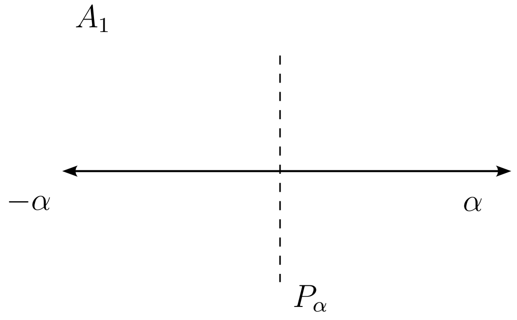



Figure 2: Root system $A_1$.

For rank 2 root systems, the inner product can be exploited to determine all the angles that the roots can have with each other, which is done via axiom (R4). This, in combination with (R3), ultimately determines the following four possibilities shown in Figure 3 and Figure 4, see (Erdmann and Wildon [5], Example 11.6) for the proofs:

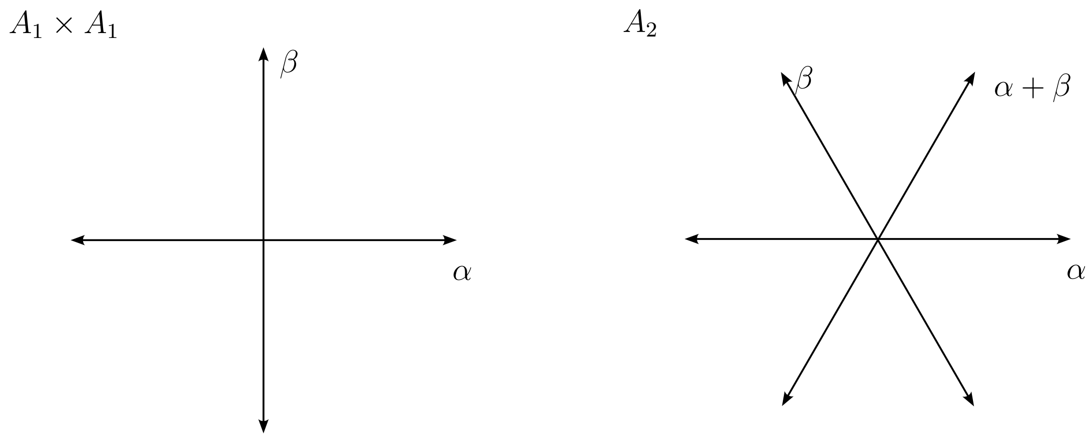



Figure 3: Root system $A_1 \times A_1$ and $A_2$.

In Figure 3 and Figure 4 we see that $\alpha$ and $\beta$ are the only roots that have to be specified to determine the whole root system, in particular a partition between positive

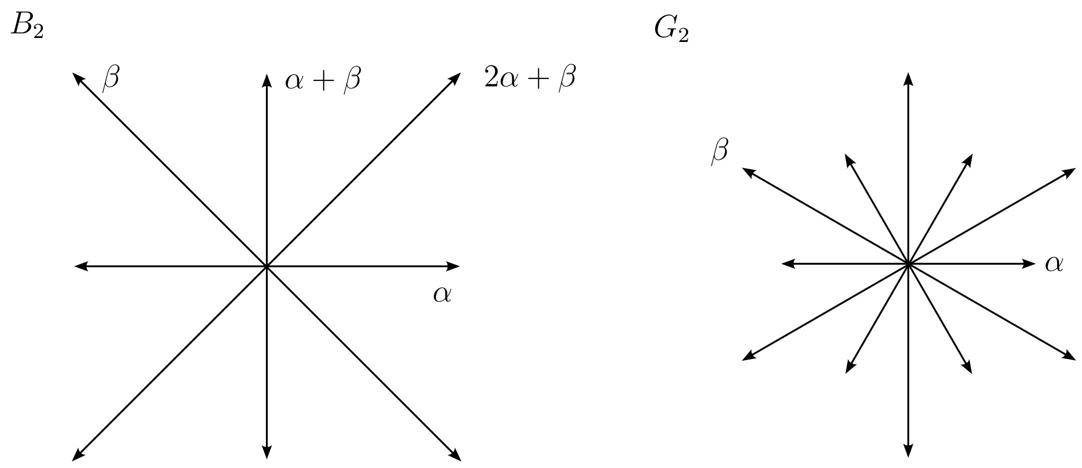



Figure 4: Root system $B_2$ and $G_2$.

and negative linear combinations of $\alpha$ and $\beta$ can be seen. This motivates us to define the following.

**Definition 3.12.** A subset $\Delta$ of $\Phi$ is called a **base** if

(B1) $\Delta$ is a vector basis of $E$,

(B2) each root $\beta$ can be written uniquely as $\beta = \sum_{\alpha\in\Delta} k_\alpha \alpha$ with $k_\alpha \in \mathbb{Z}$, where all the nonzero coefficients $k_\alpha$ are either all positive or negative.

The roots in a base $\Delta$ are called **simple**.

If for a root $\beta$ all the coefficients satisfy $k_\alpha \geq 0$, then we call $\beta$ *positive* and write $\beta \succ 0$, this also means that all simple roots are positive per definition. Analogously, if the coefficients all satisfy $k_\alpha \leq 0$, then we call $\beta$ *negative* and write $\beta \prec 0$. The collection of positive roots is denoted by $\Phi^+$ and respectively for negative roots the set is denoted by $\Phi^-$, so $\Phi$ can be written as a disjoint union of $\Phi^+$ and $\Phi^-$. Note that given a base $\Delta$ of $\Phi$, already the set $-\Delta$ can be a base of $\Phi$ as well, and in general there are additional possibilities for bases to exist, so there is not a unique base for a root system. This means that the positive and negative labels of the roots are relative to the chosen base $\Delta$.

The definition does not make it clear if a base even exists. Fortunately, they can always be constructed in a particular manner such that the desired labeling for a positive and negative side are obtained. The proof of the following theorem can be found in Humphreys [8].

**Theorem 3.13.** *Every root system $\Phi$ of $E$ has a base $\Delta$.*

Since axiom (R3) tells us that any root system $\Phi$ is symmetric with respect to all reflections $\sigma_\alpha$ where $\alpha \in \Phi$, we will examine these symmetries more deeply through the group the reflections generate, called the **Weyl group** $\mathcal{W}$. So, we can write $\mathcal{W} = \langle \sigma_{\alpha_1}, \ldots \sigma_{\alpha_n} \rangle$ when $\#\Phi = n$. We remind ourselves that the reflections $\sigma_\alpha$ are all *linear* and even *orthogonal* (hence invertible), so the Weyl group $\mathcal{W}$ is a subgroup of $GL(E)$ the general linear group of $E$, consisting of all invertible linear transformations of $E$.

**Lemma 3.14.** *The Weyl group of any root system $\Phi$ is finite.*

*Proof.* Let $\Phi$ be a root system of $E$. Because axiom (R3) tells us that the Weyl group $\mathcal{W}$ of $\Phi$ permutes all the roots, there exists a group homomorphism $f : \mathcal{W} \to S(\Phi)$, where $S(\Phi)$ denotes the group of all permutations of $\Phi$. As the root system is finite, $S(\Phi)$ is finite as well. Now, once we prove that the homomorphism $f$ is injective, we can conclude that the Weyl group is finite. Indeed, given $\sigma \in \ker f$ we know that $\sigma$ fixes all roots, but by (R1) $\Phi$ spans $E$ so by linearity $\sigma$ must be the identity in $\mathcal{W}$. □

Even though the Weyl group is finite, there is still a way to simplify working with the Weyl group. First, we define a following subgroup of $\mathcal{W}$, namely $\mathcal{W}_0 := \langle \sigma_\alpha : \alpha \in \Delta \rangle$ and we call such reflections along simple roots the *simple reflections*. Our aim is to prove $\mathcal{W} = \mathcal{W}_0$, which will be needed in the classification of finite-dimensional modules in Section 5.3.

**Lemma 3.15.** *Let $\alpha \in \Delta$, then the simple reflection $\sigma_\alpha$ permutes $\Phi^+ \setminus \{\alpha\}$.*

*Proof.* Given any $\beta \in \Phi^+ \setminus \{\alpha\}$, we can write the root as $\beta = \sum_{\gamma \in \Delta} k_\gamma \gamma$. Since $\beta$ cannot be any multiple of $\alpha$, per assumption, there must exist a simple root $\gamma$ that contributes a difference between $\beta$ and $\alpha$. So, for such a $\gamma$ the coefficient is $k_\gamma \neq 0$ and is also greater than 0 because $\beta$ is positive. Now, the simple reflection $\sigma_\alpha$ of $\beta$ yields

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha = \sum_{\gamma \in \Delta} k_\gamma \gamma - \langle \beta, \alpha \rangle \alpha = (k_\alpha - \langle \beta, \alpha \rangle)\alpha + \sum_{\gamma \in \Delta \setminus \{\alpha\}} k_\gamma \gamma. \tag{3.2}$$

This reflection of $\beta$ is now written as a linear combination of the simple roots as well. However, the coefficients from roots unequal to $\alpha$ have not changed, so the same coefficient $k_\gamma > 0$ is still present. The positive coefficient $k_\gamma$ forces all the coefficients of $\sigma_\alpha(\beta)$ to be positive as well, so $\sigma_\alpha(\beta) \in \Phi^+$. Furthermore, $\sigma_\alpha(\beta) \neq \alpha$ as only $-\alpha$ is mapped to $\alpha$ under this reflection. Hence, the lemma is proven. □

Now, we can turn our focus to the following theorem which states that any root can be seen as a permutation of a simple root.

**Theorem 3.16.** *For every root $\beta \in \Phi$ there exist $\tau \in \mathcal{W}_0$ and $\alpha \in \Delta$ such that $\beta = \tau(\alpha)$.*

*Proof.* We can assume $\beta \in \Phi^+$, otherwise $-\beta$ is positive and when $-\beta = \tau(\alpha)$ for appropriate $\tau \in \mathcal{W}_0$ and $\alpha \in \Delta$, then $\beta = \tau(-\alpha) = \tau(\sigma_\alpha(\alpha))$ where $\tau\sigma_\alpha \in \mathcal{W}_0$.

First, we define the *height* of a root $\beta = \sum_{\gamma \in \Delta} k_\gamma \gamma$ to be the sum of its coefficients,

$$\mathrm{ht}(\beta) = \sum_{\gamma \in \Delta} k_\gamma.$$

Then, we proceed by induction on the height of $\beta$. For $\mathrm{ht}(\beta) = 1$, we can say that $\beta$ has to be simple and there is nothing to prove. Now, assume that $\mathrm{ht}(\beta) \geq 2$ and the corresponding induction hypothesis for positive roots holds. We claim that there exists a $\gamma \in \Delta$ such that $(\beta, \gamma) > 0$, so that there exists a simple root which forms an acute angle with $\beta$. Indeed, if this were not the case then $(\beta, \gamma) \leq 0$ for all simple roots $\gamma$. However, this implies

$$\|\beta\|^2 = (\beta, \beta) = \sum_{\gamma \in \Delta} k_\gamma (\beta, \gamma) \leq 0,$$

because all the coefficients $k_\gamma \geq 0$. Since $\beta \neq 0$, the norm cannot be zero or less, so we have a contradiction. Thus we can take a $\gamma$ as in the claim, then using (3.2) the height of $\sigma_\gamma(\beta)$ is found to be

$$\mathrm{ht}(\sigma_\gamma(\beta)) = \sum_{\delta \in \Delta} k_\delta - \langle \beta, \gamma \rangle = \mathrm{ht}(\beta) - \langle \beta, \gamma \rangle.$$

Because $\langle \beta, \gamma \rangle > 0$, as $\gamma$ was taken to be acute to $\beta$, we have $\mathrm{ht}(\sigma_\gamma(\beta)) < \mathrm{ht}(\beta)$. Since Lemma 3.15 preserves $\sigma_\gamma(\beta)$ as a positive root, we are allowed to apply the induction hypothesis to obtain a $\tau \in \mathcal{W}_0$ and $\alpha \in \Delta$ such that $\sigma_\gamma(\beta) = \tau(\alpha)$. Hence, $\beta = \sigma_\gamma \tau(\alpha)$ and the induction step is proven, which finishes the proof. □

Finally, we can show that the Weyl group is indeed generated by the simple reflections.

**Theorem 3.17.** *The Weyl group is generated by all simple reflections, $\sigma_\alpha$ where $\alpha \in \Delta$, so $\mathcal{W} = \mathcal{W}_0$.*

*Proof.* The Weyl group is per definition generated by all the reflections $\sigma_\beta$ where $\beta \in \Phi$. So, we need to prove that these reflections can be written as a composition of simple reflections. To do so, we first prove that for any $\beta \in \Phi$ and $\tau \in \mathcal{W}$ the following holds, $\tau \sigma_\beta \tau^{-1} = \sigma_{\tau(\beta)}$. Let $\gamma \in \Phi$, then using linearity and orthogonality we obtain

$$\tau\sigma_\beta(\gamma) = \tau(\gamma - \langle \gamma, \beta \rangle \beta) = \tau(\gamma) - \langle \gamma, \beta \rangle \tau(\beta) = \tau(\gamma) - \langle \tau(\gamma), \tau(\beta) \rangle \tau(\beta) = \sigma_{\tau(\beta)} \tau(\gamma)$$

So, the identity indeed holds. From Theorem 3.16 we can find, for any root $\beta$, a simple reflection $\tau \in \mathcal{W}_0$ and $\alpha \in \Delta$ such that $\beta = \tau(\alpha)$. Then the identity we just proved, yields $\sigma_\beta = \tau \sigma_\alpha \tau^{-1}$, which is a composition of only simple reflections, thus we obtain the desired result. □

We move on to another aspect of the root systems, which can only be studied because we are taking this geometric approach. The hyperplanes $P_\alpha$, where $\alpha \in \Phi$, partition the space $E$ into finitely many regions. These regions, without the hyperplane borders, are called **Weyl chambers** (see Figure 5). The example shown in Figure 5 is a special type of Weyl chamber, namely the fundamental Weyl chamber which we define as follows.

**Definition 3.18.** Given a base $\Delta$ for a root system $\Phi$, we define the **fundamental Weyl chamber** relative to $\Delta$ as $\mathfrak{C}(\Delta) := \{x \in E : (x, \alpha) > 0 \text{ for all } \alpha \in \Delta\}$. So, it contains all the vectors whose angles between each simple root is acute.

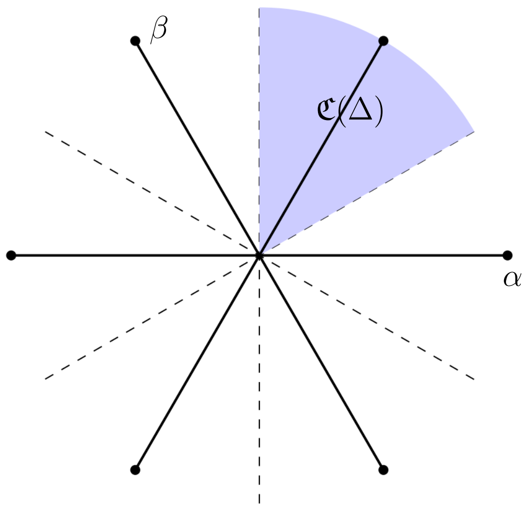



Figure 5: Fundamental Weyl chamber relative to the base $\Delta = \{\alpha, \beta\}$ in blue (Schnell [11]).

We end this section by stating a theorem that will be necessary for the classification of finite-dimensional modules in Section 5.3. The proof will be omitted but can be found in Humphreys [8].

**Theorem 3.19.** *For each vector $x \in E$, there is exactly one vector $y$ in the closure of the Fundamental Weyl chamber $\overline{\mathfrak{C}(\Delta)}$, such that $\sigma(x) = y$ for some $\sigma \in \mathcal{W}$.*

In Figure 6 we show the $\mathcal{W}$-*orbit* of a vector $x \in E$, which is defined by the set $\mathcal{W}x := \{\sigma(x) \in E : \sigma \in \mathcal{W}\}$ that contains all the vectors the Weyl group can permute $x$ to. From the figure we can also see that there is indeed only one vector in the $\mathcal{W}$-orbit of $x$. It also becomes clear that the only vectors that the Weyl group sends to this $y$ are the vectors in the orbit $\mathcal{W}x$.

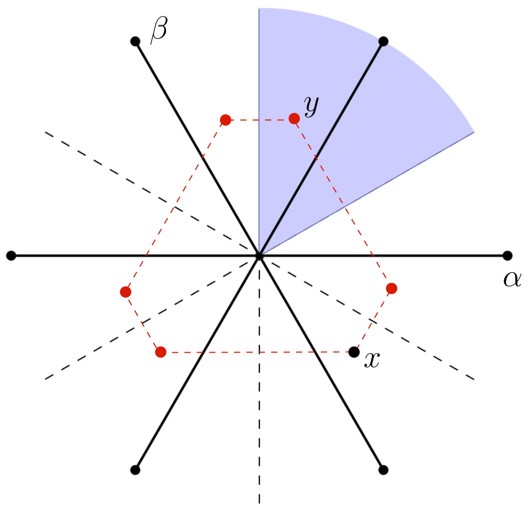



Figure 6: $\mathcal{W}$-orbit of a vector $x \in E$.

# 4 Universal enveloping algebras

We have seen that End($V$) forms a Lie algebra with the bracket $[x, y] = xy - yx$, defined via its own associative product. This is an example of an associative algebra and they can all be formed into a Lie algebra by defining the bracket in this way. Now come the universal enveloping algebras, which are associative algebras to which every Lie algebra can be injected to. So, even though Lie algebras are not associative on their own, they can be viewed as a Lie subalgebra of larger associative algebras. The main reason we are interested in the universal enveloping algebra, as it turns out, is that all the representation theory of Lie algebras amounts to the representation theory for their universal enveloping algebra (Dixmier [3], Cor 2.2.2).

Since the formal construction of the universal enveloping algebras uses tensor products, we give a brief description of them in Section 4.1. Of course, we also need to consider a few standard properties of associative algebras, which is done in Section 4.2. Then, we arrive at Section 4.3, where we define and construct universal enveloping algebras. Additionaly, we treat the Poincaré-Birkhoff-Witt (PBW) Theorem and its consequences, which allows us to view the universal enveloping algebra as a polynomial algebra and have a complete description of its basis. Lastly, in Section 4.4 we prove the PBW Theorem, though the proof is very technical and does not add a lot to the understanding of the theorem, so it can be skipped without any issues.

## 4.1 Tensor products

We briefly discuss tensor products as in [1], because we only need to know a few properties for our purposes. For a more in depth treatment of tensor products, we refer the reader to (Conrad [2]).

The given definition on its own might not clarify how to think of tensor products and what they are supposed to represent. In general, the use of tensor products come from their connection with multilinear maps, as they have what is called a universal property which allows multilinear maps to be identified with linear maps. Though, we will only need to see the tensor product as a formal product without their link to multilinear algebra.

**Definition 4.1.** Let $V$ and $W$ be complex vector spaces. To define their **tensor product** $V \otimes W$ we consider the vector space $U$ which has a basis given by $V \times W = \{(v, w) : v \in V, w \in W\}$. Now, define the subspace $T \subset U$ that is generated by elements of the form:

$$\begin{aligned} &(v_1 + v_2, w) - (v_1, w) - (v_2, w), \\ &(v, w_1 + w_2) - (v, w_1) - (v, w_2), \\ &(\lambda v, w) - (v, \lambda w), \end{aligned} \tag{4.1}$$

where $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $\lambda \in \mathbb{C}$. Then the tensor product $V \otimes W$ is defined to be the quotient space $U/T$. The image of $(v, w)$ in $V \otimes W$ is denoted by $v \otimes w$ and called a **tensor** or *the tensor product of $v$ and $w$*.

By factoring $T$ out, we force the elements in $U$ of the form in (4.1) to be equal to zero in the quotient. Hence, the tensors in $V \otimes W$ have, per construction, the following properties:

$$\begin{aligned}
(v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w, \\
v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2, \\
\lambda v \otimes w &= v \otimes \lambda w.
\end{aligned}$$

So, the *tensor product of two vectors* shares the bilinear properties of regular multiplication, we can therefore view $\otimes$ as a formal product of vectors. The bilinearity also implies that the *canonical map* $\phi : V \times W \to V \otimes W$ with $\phi(x, y) = x \otimes y$ is bilinear.

**Lemma 4.2.** *In $V \otimes W$, we have $v \otimes 0 = 0 \otimes w = 0$ for all $v \in V$ and $w \in W$.*

*Proof.* Let $v \in V$ and $w \in W$, then

$$v \otimes 0 = v \otimes (0 + 0) = v \otimes 0 + v \otimes 0 \implies v \otimes 0 = 0.$$

We can argue similarly for $0 \otimes w$ to see that it is also 0. □

Our motivation to work with tensor products for universal enveloping algebras is partially due to its associativity.

**Property 4.3.** *Let $U, V, W$ be vector spaces, then there is a canonical isomorphism such that $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$ and maps $(u \otimes v) \otimes w$ to $u \otimes (v \otimes w)$.*

We add that the tensor product of two vectors is in general *noncommutative*.

## 4.2 Algebras

**Definition 4.4.** We say that a vector space $\mathcal{A}$ over $\mathbb{C}$ is an **algebra**, when it has a bilinear map $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ with $(x, y) \mapsto xy$, which we call the *product*. If the product in $\mathcal{A}$ satisfies

$$x(yz) = (xy)z \text{ for all } x, y, z \in \mathcal{A},$$

then $\mathcal{A}$ is said to be an *associative algebra*. Moreover, if $\mathcal{A}$ has an element $1_{\mathcal{A}}$ such that

$$1_{\mathcal{A}} x = x 1_{\mathcal{A}} = x \text{ for all } x \in \mathcal{A},$$

we call it *unital*.

*Remark.* In this thesis, beside Lie algebras, we will only encounter the associative algebras with unit. So, for brevity, when we say "*algebra*" we really mean "*unital associative algebra*".

*Remark.* An algebra (with unit) always contains the underlying field, in this case $\mathbb{C} \cong \text{span}\{1_{\mathcal{A}}\}$.

*Example.* The vector space $\mathrm{End}(V)$ consisting of linear maps on $V$ is an algebra with map composition as the product and the identity map as the unit.

*Example.* All complex-valued polynomials in the indeterminate $X$ form a vector space $\mathbb{C}[X]$, and this is an algebra over $\mathbb{C}$. The product is defined as the regular product between polynomials.

**Definition 4.5.** Let $\mathcal{A}, \mathcal{B}$ be algebras. An **algebra homomorphism** is a $\mathbb{C}$-linear map $\phi : \mathcal{A} \to \mathcal{B}$ between two algebras such that for all $x, y \in \mathcal{A}$ we have

$$\phi(xy) = \phi(x)\phi(y).$$

This is an **isomorphism** if it is bijective and we write $\mathcal{A} \cong \mathcal{B}$.

The kernel and image of an algebra homomorphism $\phi : \mathcal{A} \to \mathcal{B}$ are defined in the usual manner. Again, the injectivity of $\phi$ is also equivalent to the condition $\ker \phi = 0$ and surjectivity holds if and only if $\phi[\mathcal{A}] = \mathcal{B}$.

**Definition 4.6.** A subspace $I$ of an algebra $\mathcal{A}$ is said to be a **left-ideal**, when

$$xy \in I \quad \text{for all } x \in I,\, y \in \mathcal{A},$$

a **right-ideal** when the product is reversed to $yx$ with $y \in \mathcal{A}$. Moreover, $I$ is said to be **two-sided** if it is both a left- and right-ideal.

*Remark.* This is essentially the same definition as for Lie algebra ideals. Though, we will only encounter algebra ideals in the proceeding discussions, so from here on "ideal" purely refers to algebra ideals, unless otherwise stated.

*Remark.* In case the algebra is commutative, then all the ideals are two-sided.

*Example.* The kernel of a homomorphism $\phi : \mathcal{A} \to \mathcal{B}$ is also a two-sided ideal: let $x \in \ker \phi$ and $y \in \mathcal{A}$, then

$$\phi(xy) = \phi(x)\phi(y) = 0 = \phi(y)\phi(x) = \phi(yx).$$

We define a two-sided ideal *generated* by an element $x \in \mathcal{A}$ as follows,

$$(x) = \mathcal{A}x\mathcal{A} = \{axc : a, c \in \mathcal{A}\}.$$

We define a two-sided ideal generated by a finite subset $S \subset \mathcal{A}$ as follows,

$$(S) = \mathcal{A}S\mathcal{A} = \{\textstyle\sum_{s\in S} a_s s c_s : a_s, c_s \in \mathcal{A} \text{ for all } s\}.$$

The definitions of $(x)$ and $(S)$ are defined in a way such that the generated two-sided ideals are the smallest two-sided ideals containing $x$ or $S$. For the ideal generated by a subset, the sum appears because we need the ideal to be closed under linear combinations for it to be a subspace. There are corresponding definitions for the left- and right-ideals, but we will only be working with two-sided ideals, because they are the only ideals that allow us to define a quotient algebra.

*Example.* Let $\mathbb{C}[X]$ be the polynomial algebra with variable $X$. Then, the ideal generated by $X^2+1$ contains all polynomials divisible by this factor. So, $5X^2+5$ and $X^4-1$ are contained in the ideal, but $1+X$ is not.

**Definition 4.7.** Let $\mathcal{A}$ be an algebra and $I \subset \mathcal{A}$ a two-sided ideal. Then we define $\mathcal{A}/I$ as the quotient vector space, with a product defined as

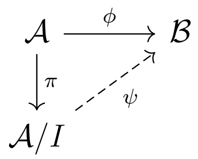
$$(x+I)(y+I) = xy + I.$$

This product yields an algebra structure and we call $\mathcal{A}/I$ the **quotient algebra**. It has a *canonical map* $\pi : \mathcal{A} \to \mathcal{A}/I$ with $x \mapsto x+I$ which is an algebra homomorphism.

*Remark.* The kernel of the canonical map $\pi$ as in the definition is $I$.

*Example.* We look at the following quotient algebra, $\mathcal{A} := \mathbb{C}[X]/(X^2+1)$. Its canonical map satisfies $\pi(X^2+1) = 0$ so $\pi(X)^2 = -1$. Therefore, the following holds in the quotient algebra $X^2 \equiv -1$ and we can rewrite any polynomial in $\mathcal{A}$. In this manner $X^3 + iX^2 + 5X - 1$ becomes $4X - (1+i)$ in $\mathcal{A}$.

**Theorem 4.8** (Homomorphism theorem)**.** *If $\phi : \mathcal{A} \to \mathcal{B}$ is an algebra homomorphism and $I$ is any ideal of $\mathcal{A}$ such that $I \subset \ker\phi$, then there exists a unique homomorphism $\psi : \mathcal{A}/I \to \mathcal{B}$ such that $\phi = \psi \circ \pi$, thus making the following diagram commute:*




**Theorem 4.9** (Isomorphism theorem)**.** *If $\phi : \mathcal{A} \to \mathcal{B}$ is an algebra homomorphism, then $\mathcal{A}/\ker\phi \cong \phi[\mathcal{A}]$.*

Lastly, we mention that there is also a notion of an *algebra module* instead of a Lie algebra module. The only difference is in the bracket condition (M3) for the Lie algebra modules. Instead the following conditions have to hold, such that associativity and the unit of the algebra are compatible with the action:

$$\text{(M3')}\ \ (xy)\cdot v = x\cdot(y\cdot v),$$
$$\text{(M4')}\ \ 1\cdot v = v.$$

*Warning.* At this point we have tools for vector spaces, Lie algebras, Lie algebra modules, algebras and algebra modules at our disposal, and for every object the tools are quite similar in their definition and behaviour. So, it is important to keep a sharp eye out on what types of objects we are working with, to not confuse ourselves when we, for example, only need vector space quotients instead of quotient algebras. We try to make the distinction clear when necessary, so that confusion can be avoided. However, to preserve readability we leave the specifications out when the context should be clear enough.

## 4.3 Poincaré-Birkhoff-Witt Theorem

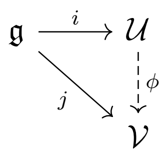
Now that we have discussed the preliminaries for the universal enveloping algebra, we will discuss its definition and construction. Afterwards we set up quite a handful of algebras and maps between them to eventually state the Poincaré-Birkhoff-Witt Theorem and its consequences.

**Definition 4.10.** For an arbitrary Lie algebra $\mathfrak{g}$, which can be infinite-dimensional, we define an **enveloping algebra** of $\mathfrak{g}$ as a pair $(\mathcal{U}, i)$, where $\mathcal{U}$ is a unital associative algebra over the underlying field $\mathbb{C}$ and $i : \mathfrak{g} \to \mathcal{U}$ is a linear map satisfying

$$i([x, y]) = i(x)i(y) - i(y)i(x), \quad x, y \in \mathfrak{g}. \tag{4.2}$$

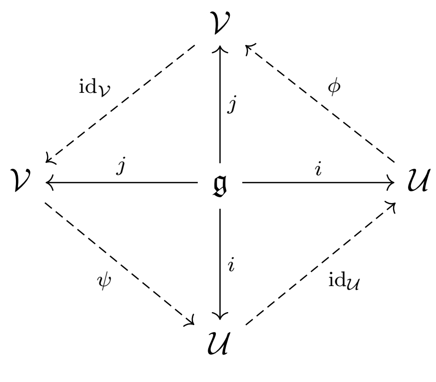
We call the pair $(\mathcal{U}, i)$ a **universal enveloping algebra** when it has the following *universal property*: given any other pair $(\mathcal{V}, j)$ over $\mathfrak{g}$ with the above properties, there exists a unique algebra homomorphism $\phi : \mathcal{U} \to \mathcal{V}$ such that $\phi \circ i = j$.

The universal property of the universal enveloping algebra can be formulated in terms of the following commuting diagram.



**Theorem 4.11.** *For every Lie algebra $\mathfrak{g}$ there exists, up to algebra isomorphism, a unique universal enveloping algebra.*

*Proof.* For *uniqueness* we take two universal enveloping algebras of $\mathfrak{g}$, namely $(\mathcal{U}, i)$ and $(\mathcal{V}, j)$. The universal property yields four unique homomorphisms between each of the algebras and therefore the following diagram commutes.



Now because the diagram commutes we find $\phi \circ \psi = \mathrm{id}_{\mathcal{V}}$ and $\psi \circ \phi = \mathrm{id}_{\mathcal{U}}$. This shows that $\phi$ and $\psi$ are inverses, therefore they are bijective and hence they are also isomorphisms, so $\mathcal{U} \cong \mathcal{V}$. Therefore the pairs are isomorphic to eachother.

To show *existence* we will construct the universal enveloping algebra via tensor products. Given the Lie algebra $\mathfrak{g}$ over the field $\mathbb{C}$, we first define the **tensor algebra** on $\mathfrak{g}$

as $\mathcal{T}(\mathfrak{g}) = \mathbb{C} \oplus \mathfrak{g} \oplus (\mathfrak{g} \otimes \mathfrak{g}) \oplus (\mathfrak{g} \otimes \mathfrak{g} \otimes \mathfrak{g}) \cdots = \bigoplus_{m \geq 0} T^m \mathfrak{g}$, where $T^m\mathfrak{g}$ is the $m$-fold tensor product of $\mathfrak{g}$. The direct sum only contains elements that are finite linear combinations of the tensors, and it preserves the vector space structure of the Lie algebra. To make this into an algebra we define a product by

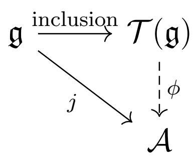
$$(v_1 \otimes \cdots \otimes v_m) \cdot (w_1 \otimes \cdots \otimes w_n) = v_1 \otimes \cdots \otimes v_m \otimes w_1 \otimes \cdots \otimes w_n \in T^{m+n}\mathfrak{g},$$

which we extend linearly over the field $\mathbb{C}$. This product is associative because of the associativity of the tensor product. Now $1 \in \mathbb{C}$ is the multiplicative identity for $\mathcal{T}(\mathfrak{g})$, making the tensor algebra an associative algebra with unit.

We claim to have a universal property for the tensor algebra due to $\mathcal{T}(\mathfrak{g})$ being generated as an algebra by 1 and the basis elements of the vector space $\mathfrak{g}$. This implies that given any linear map $j : \mathfrak{g} \to \mathcal{A}$ into an associative algebra $\mathcal{A}$, there exists a unique homomorphism $\phi : \mathcal{T}(\mathfrak{g}) \to \mathcal{A}$, which is the extension of $j$ into $\mathcal{T}(\mathfrak{g})$. Indeed, we define $\phi$ by $\phi|_{\mathfrak{g}} = j$, then what remains is fully determined by the generators of the tensor algebra as follows,

$$\phi(x_1 \otimes \cdots \otimes x_m) = \phi(x_1 \cdots x_m) = \phi(x_1) \cdots \phi(x_m) = j(x_1) \cdots j(x_m),$$

where the first equality is due to the definition of the product in the tensor algebra. This makes the following diagram commute, yielding the universal property as claimed.




We proceed by introducing the two-sided ideal $J$ in $\mathcal{T}(\mathfrak{g})$ generated by the elements $x \otimes y - y \otimes x - [x, y]$ for $x, y \in \mathfrak{g}$ and claim that $\mathcal{U}(\mathfrak{g}) = \mathcal{T}(\mathfrak{g})/J$ is the universal enveloping algebra of $\mathfrak{g}$. Let $\pi : \mathcal{T}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})$ be the canonical homomorphism, which then automatically has to satisfy (4.2) as $\pi(J) = 0$, and take $i : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ as the restriction of $\pi$ to $\mathfrak{g}$. Then given any algebra $\mathcal{V}$ and a linear map $j : \mathfrak{g} \to \mathcal{V}$ satisfying $j([x, y]) = j(x)j(y) - j(y)j(x)$ we obtain a unique homomorphic extension $\phi' : \mathcal{T}(\mathfrak{g}) \to \mathcal{V}$ by the universal property of the tensor algebra. So the elements $x \otimes y - y \otimes x - [x, y]$ are contained in $\ker \phi'$, thus $J \subset \ker \phi'$. By the homomorphism theorem (Theorem 4.8) we find that $\phi'$ induces a unique homomorphism $\phi : \mathcal{U}(\mathfrak{g}) \to \mathcal{V}$ with $\phi \circ i = j$ as desired. Hence $\mathcal{U}(\mathfrak{g})$ is the universal enveloping algebra of $\mathfrak{g}$. □

*Example.* We take a look at the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$. Recall that the defining relations are $[h, x] = 2x$, $[h, y] = -2y$, $[x, y] = h$. Then the map $i : \mathfrak{sl}(2, \mathbb{C}) \to \mathcal{U}(\mathfrak{sl}(2, \mathbb{C}))$ is the unique linear map such that $i(x) = X$, $i(y) = Y$, $i(h) = H$, where the capital letters are elements in the universal enveloping algebra of $\mathfrak{sl}(2, \mathbb{C})$. Then elements of the tensor and enveloping algebra might for example look like $X \otimes Y \otimes X$, though we will leave out the tensor notation for readability, so we write $XYX$ instead. Beware that this is not equal to $X^2Y$. Because of the bracket relations that are carried over from the Lie algebra into the enveloping algebra we can for example rewrite $YXH + HY^2$ using amongst other things $YX = XY - [X,Y] = XY - H$,

$$\begin{aligned} YXH + HY^2 &= (XY - [X,Y])H + HY^2 \\ &= (XYH - H^2) + ([H,Y] + YH)Y \\ &= XYH - H^2 - 4Y^2 + Y^2H. \end{aligned}$$

So, even though the product in the universal enveloping algebra is not commutative, we can still reorder $YX$ to $XY$ but with an extra term $[X,Y]$ to compensate the factor swapping. While such a reordering is not possible in the tensor algebra.

Next we will construct yet another algebra on $\mathfrak{g}$, namely the symmetric algebra. This new algebra contains elements that resemble polynomials in Lie algebra elements. The goal is to connect this algebra to the universal enveloping algebra via the PBW Theorem, so we can better understand the structure of the universal enveloping algebra. We start with the two-sided ideal $I$ in $\mathcal{T}(\mathfrak{g})$ generated by all $x \otimes y - y \otimes x$ for $x, y \in \mathfrak{g}$. Define the **symmetric algebra** on $\mathfrak{g}$ as

$$\mathcal{S}(\mathfrak{g}) = \mathcal{T}(\mathfrak{g})/I \text{ with } \sigma : \mathcal{T}(\mathfrak{g}) \to \mathcal{S}(\mathfrak{g}) \text{ the canonical map.}$$

Notice that by factoring out $x \otimes y - y \otimes x$ we are demanding the elements in the symmetric algebra to satisfy $x \otimes y = y \otimes x$, hence forcing all the elements to be commutative. Moreover, we can write $\mathcal{S}(\mathfrak{g}) = \bigoplus_{m \geq 0} S^m\mathfrak{g}$ where $S^m\mathfrak{g} = T^m\mathfrak{g}/(I \cap T^m\mathfrak{g})$.

The reason to construct objects like the tensor and symmetric algebras is because they, in some sense, mimic polynomial algebras. This is made clearer by the following definition.

**Definition 4.12.** An algebra $\mathcal{A}$ is **graded** when it can be decomposed as vector spaces into $\mathcal{A} = \bigoplus_{m \geq 0} A^m$ and $A^m \cdot A^n \subset A^{m+n}$ for all $m, n \geq 0$. If $a \in A^m$ then we call $m$ the **degree** of $a$ and call $a$ **homogeneous** of degree $m$.

*Example.* The fundamental example of a graded algebra is a polynomial algebra, say $\mathbb{C}[z_1, z_2, z_3]$ which is the collection of complex valued polynomials in $z_1, z_2$ and $z_3$. This algebra can be written as a direct sum of its homogeneous components of degree $m$, denoted by $\mathbb{C}[z_1, z_2, z_3]_m$, with $m \in \mathbb{Z}_{\geq 0}$, making it a graded algebra.

In this way the tensor and symmetric algebra look like polynomial algebras, since they both have a graded structure endowed by their products. Though the tensor algebra does not share the commutativity, whilst the symmetric algebra does and it can therefore be identified with a polynomial algebra. This identification with the polynomial algebra makes working with the symmetric algebra easier, which we will see in the proof of the PBW theorem.

Now, because we do not have much information on the structure of the universal enveloping algebra, we will construct a graded algebra from it. This allows us to find a suitable mapping from the symmetric algebra to the associated graded algebra of the universal enveloping algebra. To this end, we have to introduce a few new spaces. For brevity we leave out the $\mathfrak{g}$ in the notation, so for example we write $\mathcal{T}$ instead of $\mathcal{T}(\mathfrak{g})$. Now we define a *filtration* on $\mathcal{T}$ by $T_m = T^0 \oplus \cdots \oplus T^m$, and let $U_m = \pi[T_m] \in \mathcal{U}$ and $U_{-1} = 0$, where $\pi : \mathcal{T} \to \mathcal{U}$ is the canonical homomorphism as in the construction of $\mathcal{U}$ above. So, intuitively we take noncommutative polynomials of degree $m$ in Lie algebra elements and map them to the universal enveloping algebra. Since $T_m \cdot T_n \subset T_{m+n}$ by the definition of the product in the tensor algebra, we find that

$$U_m \cdot U_n = \pi[T_m] \cdot \pi[T_n] = \pi[T_m \cdot T_n] \subset \pi[T_{m+n}] = U_{m+n}, \text{ but also } U_m \subset U_{m+1},$$

which means we cannot immediately decompose the universal enveloping algebra into a direct sum of the subspaces $U_m$. Hence, we define $G^m = U_m/U_{m-1}$ as a quotient of vector spaces, this yields $G^m \cap G^n = \{0\}$ for any $m, n \geq 0$, which allows us to set up a direct sum as $\mathcal{G} = \bigoplus_{m\geq 0} G^m$. Additionally, the multiplication in $\mathcal{U}$ defines a bilinear map $G^m \times G^n \to G^{m+n}$ which extends to a multiplication $\mathcal{G} \times \mathcal{G} \to \mathcal{G}$, making $\mathcal{G}$ a graded associative algebra with unit. To show that the bilinear map is actually well-defined, we take arbitrary elements in $G^m$ and $G^n$. We write the elements as $u + u'$ and $v + v'$ where $u, v$ are in $U_m, U_n$ and $u', v' \in U_{m-1}, U_{n-1}$ respectively. Now,

$$\begin{aligned}(u + u') \cdot (v + v') &= uv + \underbrace{u'v + uv'}_{\text{in } U_{m+n-1}} + \overbrace{u'v'}^{\text{in } U_{m+n-2}} \\ &= uv + w \quad (\text{where } w \in U_{m+n-1}) \\ &\in U_{m+n}/U_{m+n-1} = G^{m+n}.\end{aligned}$$

Note that for $T^m \subset T_m$ the canonical algebra homomorphism $\pi$ maps the subspace to $U_m$. So composing $\pi$ with the natural linear map $f_m : U_m \to G^m$, that arose from the vector space quotient, makes sense and we obtain the linear map $\varphi_m = f_m \circ \pi : T^m \to G^m$. These maps combine to form the linear map $\varphi : \mathcal{T} \to \mathcal{G}$.

We have quite a handful of sets and functions to keep track of now, each in relation to different types of objects. So we list all of those that we will use for the upcoming theorems and proofs:

| | |
|---|---|
| $\mathcal{T}$ | tensor algebra of $\mathfrak{g}$; $\mathcal{T} = \bigoplus_{m\geq 0} T^m\mathfrak{g}$ |
| $T_m$ | filtration on $\mathcal{T}$ of degree $m$; $T_m = T^1 \oplus \cdots \oplus T^m$ |
| $I$ | (algebra) ideal generated by elements $x \otimes y - y \otimes x$ |
| $\mathcal{S}$ | symmetric algebra of $\mathfrak{g}$; $\mathcal{S} = \mathcal{T}/I$ |
| $\sigma$ | canonical (algebra) homomorphism $\mathcal{T} \to \mathcal{S}$ |

| | |
|---|---|
| $J$ | (algebra) ideal generated by elements $x \otimes y - y \otimes x - [x, y]$ |
| $\mathcal{U}$ | universal enveloping algebra of $\mathfrak{g}$; $\mathcal{U} = \mathcal{T}/J$ |
| $\pi$ | canonical (algebra) homomorphism $\mathcal{T} \to \mathcal{U}$ |
| $U_m$ | image of $T_m$ into $\mathcal{U}$ under $\pi$; $U_m = \pi[T_m]$ |
| $G^m$ | a quotient of vector spaces; $G^m = U_m/U_{m-1}$ |
| $\mathcal{G}$ | graded algebra of the universal enveloping algebra; $\mathcal{G} = \bigoplus_{m \geq 0} G^m$ |
| $f_m$ | natural linear map $U_m \to G^m$ induced by the quotient space $G^m$ |
| $\varphi$ | a linear map $\mathcal{T} \to \mathcal{G}$ build from the maps $\varphi_m : T^m \xrightarrow{\pi} U_m \xrightarrow{f_m} G^m$ |

**Lemma 4.13.** *The linear map $\varphi : \mathcal{T} \to \mathcal{G}$ is a surjective algebra homomorphism. Moreover, $\varphi(I) = 0$ so we obtain an induced homomorphism $\omega : \mathcal{S} = \mathcal{T}/I \to \mathcal{G}$ that is surjective.*

*Proof.* For $\varphi$ to be an algebra homomorphism, we only have to check that it is multiplicative on $\mathcal{T}$ as it is already a linear map. To this end we take arbitrary homogeneous tensors $x \in T^m$ and $y \in T^n$. Since the product in $\mathcal{G}$ is really a carried over product from $\mathcal{T}$ through $\mathcal{U}$ we have

$$\varphi(xy) = f_{m+n}(\pi(xy)) = f_{m+n}(\pi(x)\pi(y)) = f_m(\pi(x))f_n(\pi(y)) = \varphi(x)\varphi(y),$$

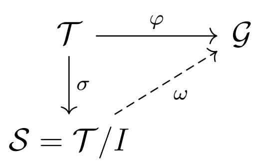
where we used the well-definedness of the bilinear map $G^m \times G^n \to G^{m+n}$ associated with the multiplication in $\mathcal{G}$ in the third equality. So this makes $\varphi$ an algebra homomorphism.

The surjectiveness of $\varphi$ comes from surjectiveness of the canonical maps. To show this, let $m \geq 0$ then the filtration $T_m$ is mapped as follows,

$$\begin{aligned}\varphi(T_m) &= \varphi_1(T^1) \oplus \cdots \oplus \varphi_m(T^m) = f_1(\pi(T^1)) \oplus \cdots \oplus f_m(\pi(T^m)) \\ &= f_1(U_1) \oplus \cdots \oplus f_m(U_m) = G^1 \oplus \cdots \oplus G^m.\end{aligned}$$

From this we can conclude that $\varphi(\mathcal{T}) = \mathcal{G}$, hence the map is surjective.

To show $\varphi(I) = 0$, see that for a given generator $x \otimes y - y \otimes x \in T^2$ with $x, y \in \mathfrak{g}$ we have $\pi(x \otimes y - y \otimes x) \in U_2$, this means that $\varphi$ maps this generator into $U_2/U_1$. But the universal enveloping algebra factors out $x \otimes y - y \otimes x - [x, y]$, such that also $\pi(x \otimes y - y \otimes x) = \pi([x, y]) \in U_1$ after use of linearity. Now $\varphi(x \otimes y - y \otimes x) \in U_1/U_1 = 0$, thus $I \subset \ker \varphi$. Finally the homomorphism theorem (Theorem 4.8) yields a unique homomorphism $\omega$ as desired, shown in the commutative diagram below.




□

**Theorem 4.14** (PBW Theorem)**.** *The homomorphism $\omega : \mathcal{S} \to \mathcal{G}$ is an algebra isomorphism.*

*Remark.* Note that by Lemma 4.13 we are left to show injectivity of the map $\omega$.

We defer the proof of the Poincaré-Birkhoff-Witt Theorem to Section 4.4 and first discuss the corollaries of the theorem, which will be of more importance for our purposes than the theorem itself.

**Corollary 4.14.1.** *If a subspace $W \subset T^m$ is sent isomorphically onto $S^m$ by the canonical map $T^m \to S^m$, then $\pi[W]$ is a complement to $U_{m-1}$ in the subspace $U_m$, so $U_m = U_{m-1} \oplus \pi[W]$.*

The corollary is drenched in algebraic language, so let us examine what it really says. Since, $T^m$ essentially contains noncommutative homogeneous polynomials of degree $m$ and $S^m$ their commutative variants, not any subspace $W \subset T^m$ can be sent isomorphically onto $S^m$. For each polynomial in $S^m$, there are numerous noncommutative variants in $T^m$. So, to have $W$ map isomorphically onto $S^m$ it must contain exactly one of all those noncommutative variants. The corollary then tells us that after factoring out the ideal $J$, the elements in $W$ remain homogeneous of degree $m$.

*Proof.* The following diagram commutes

$$\begin{array}{ccccc} & & U_m & & \\ & \overset{\pi}{\nearrow} & & \overset{f_m}{\searrow} & \\ T^m & & & & G^m \\ & \underset{\sigma}{\searrow} & & \underset{\omega}{\nearrow} & \\ & & S^m & & \end{array}$$

This is because the top half from $T^m$ to $G^m$ is given by the map $\varphi_m$ and we see from the commutative diagram in the proof of Lemma 4.13 that $\varphi = \omega \circ \sigma$. So, by restricting to degree $m$ we can conclude that the above diagram is commutative. From the assumption and the PBW theorem we know that $W \cong S^m \cong G^m$ via the bottom half of the diagram. If we examine the top half, we see that then $\pi[W] \subset U_m$ must map bijectively onto $G^m = U_m/U_{m-1}$, meaning that $U_m = U_{m-1} \oplus \pi[W]$ □

**Corollary 4.14.2.** *The canonical map $i : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ is injective. Allowing us to identify $\mathfrak{g}$ with $i(\mathfrak{g})$ in the universal enveloping algebra.*

Now the term "enveloping" makes sense as we can imbed any Lie algebra with any bracket operation into its universal enveloping algebra, where the bracket is always given by $[x, y] = xy - yx$.

*Proof.* This follows as a special case of Corollary 4.14.1 with $W = \mathfrak{g} = T^1$. The Lie algebra is sent isomorphically onto $S^1$ because the ideal $I$ that is factored out is generated by $T^2$ elements. This means that the canonical map $T^1 \to S^1$ is the identity, hence the isomorphism holds. Now the corollary yields $i(\mathfrak{g}) = \pi[\mathfrak{g}] \cong U_1/U_0 \subset \mathcal{U}(\mathfrak{g})$, so $i$ maps injectively to a subspace of the universal enveloping algebra. □

**Corollary 4.14.3.** *Let* $(x_1, x_2, \dots)$ *be any ordered basis of* $\mathfrak{g}$*. Then the elements* $x_{j_1} \cdots x_{j_m} = \pi(x_{j_1} \otimes \cdots \otimes x_{j_m})$*, for* $m \in \mathbb{Z}_{>0}$ *and* $j_1 \leq \cdots \leq j_m$*, along with* $1$ *form a basis of* $\mathcal{U}(\mathfrak{g})$ *over* $\mathbb{C}$*.*

Usually Corollary 4.14.3 is referred to as the PBW Theorem and the given basis is called a **PBW basis**.

*Proof.* To prove this we use induction on $m$ and show that the given elements certainly do form a basis for each $U_m$. For $m = 0$ the set $\{1\}$ forms a basis for $\mathbb{C} = U_0$. Suppose now that the corollary holds for $U_{m-1}$, then we will aim to use Corollary 4.14.1, so that we can apply the induction hypothesis on the $U_{m-1}$ component.

Let $W$ be the subspace of $T^m$ spanned by all elements of the form $x_{j_1} \otimes \cdots \otimes x_{j_m}$ with $j_1 \leq \cdots \leq j_m$. Note that this is not all of $T^m$ yet because of the noncommutativity in $T^m$, as we specifically take the span of elements with an increasing order of indices only. Since we can identify $S^m$ with the collection of homogeneous polynomials of degree $m$, it has a basis described exactly by degree $m$ monomials in $\{x_{j_1} \cdots x_{j_m} : \textit{increasing indices}\}$. So, we can conclude that $W$ maps isomorphically onto $S^m$, in accordance with the discussion below Corollary 4.14.1. Now Corollary 4.14.1 implies that $U_m = U_{m-1} \oplus \pi[W]$ and with the induction hypothesis we see that Corollary 4.14.3 also holds for $U_m$. Hence by induction we can form a basis for $\mathcal{U}(\mathfrak{g})$ in this manner. □

**Corollary 4.14.4.** *Let* $\mathfrak{h}$ *be a subalgebra of* $\mathfrak{g}$*, and extend an ordered basis* $(h_1, h_2, \dots)$ *of* $\mathfrak{h}$ *to an ordered basis* $(h_1, \dots, x_1, \dots)$ *of* $\mathfrak{g}$*. Then the homomorphism* $\mathcal{U}(\mathfrak{h}) \to \mathcal{U}(\mathfrak{g})$ *induced by the imbedding* $\mathfrak{h} \to \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ *is itself an injection, and* $\mathcal{U}(\mathfrak{g})$ *is a* $\mathcal{U}(\mathfrak{h})$*-module with basis consisting of all* $x_{j_1} \cdots x_{j_m}$ *with* $j_1 \leq \cdots \leq j_m$*, along with 1.*

This corollary will be fundamental for constructing what are called Verma modules, in Section 5.

*Proof.* We use the injectivity of $i$ to see that $\mathfrak{h} \to \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ is injective. The universal property of $\mathcal{U}(\mathfrak{h})$ induces a homomorphism $\mathcal{U}(\mathfrak{h}) \to \mathcal{U}(\mathfrak{g})$ that maps the basis elements $h_{j_1} \cdots h_{j_m}$ to itself. So the map imbeds $\mathcal{U}(\mathfrak{h})$ into $\mathcal{U}(\mathfrak{g})$ making it injective. Furthermore, the $\mathcal{U}(\mathfrak{h})$-action is given by repeated multiplication of elements in $\mathfrak{h}$. Then, in light of the bases that both the universal enveloping algebras have over the field $\mathbb{C}$, we notice that the $\mathcal{U}(\mathfrak{h})$-basis must be given by the indicated basis of the corollary, which consists of all the finite products of the basis elements from the extension to $\mathfrak{g}$. □

## 4.4 Proof of PBW Theorem

*The proof will be especially technical, without being necessary to understand the PBW theorem itself better, so the reader does not lose out on anything by skipping the proof.*

We start the proof by showing that there is a $\mathfrak{g}$-module structure on the symmetric algebra $\mathcal{S}$, which is defined in such a way that we have a description of how to map tensors in the tensor algebra to the symmetric algebra. This will require most of the work for the proof. Then we proceed with Lemma 4.21, which yields some information on what tensors lie in the ideal $I$, the ideal generated by elements of the form $x\otimes y - y\otimes x$. This leads us to the final part, where we use the previous lemma to show the map $\omega : \mathcal{T}/I \to \mathcal{G}$ is injective.

Let us now set up the necessary notations for the proof. Fix an ordered basis $(x_\lambda : \lambda \in \Omega)$ of $\mathfrak{g}$, then we identify $\mathcal{T}$ with the noncommutative polynomial algebra $\mathbb{C}\langle x_\lambda : \lambda \in \Omega\rangle$ and the symmetric algebra $\mathcal{S}$ with the polynomial algebra $\mathbb{C}[z_\lambda : \lambda \in \Omega]$, where the factors do commute. As for the tensor algebra, we also define the filtration on $\mathcal{S}$ of degree $m$ by $S_m := S^1 \oplus \cdots \oplus S^m$.

To shorten notation we define for each sequence $\Sigma = (\lambda_1, \ldots, \lambda_m)$ of indices of length $m$, the following notation, $x_\Sigma = x_{\lambda_1} \otimes \cdots \otimes x_{\lambda_m} \in T^m$ and $z_\Sigma = z_{\lambda_1} \cdots z_{\lambda_m} \in S^m$. Furthermore, we call the sequence $\Sigma$ *increasing* if $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_m$ and for the (increasing) sequence $\varnothing$ we define $z_\varnothing = 1$.

Having set up the start of the proof, we can take $\{z_\Sigma : \Sigma \text{ increasing}\}$ as a basis for the symmetric algebra. This is because commutativity allows us to reorder the factors of each monomial in $z_\lambda$'s in our desired increasing order. We begin the proof by defining a module structure on the symmetric algebra.

**Proposition 4.15.** *There exists a $\mathfrak{g}$-module structure on $\mathcal{S}$ which is uniquely determined by the following conditions: (for brevity the action is written without the dot $\cdot$) given $x_\lambda \in \mathfrak{g}$ and $m \in \mathbb{Z}_{\ge 0}$,*

$(A_m)$ $x_\lambda z_\Sigma = z_\lambda z_\Sigma$ *for* $\lambda \le \Sigma$, $z_\Sigma \in S_m$,

$(B_m)$ $x_\lambda z_\Sigma \equiv z_\lambda z_\Sigma$ *(mod* $S_k$*) when* $z_\Sigma \in S_k$ *and* $k \le m$.

*Remark.* Note that once a condition holds for some $m \in \mathbb{Z}_{\ge 0}$, then it also holds for all the $k \le m$. This is because the filtration $S_m$ contains all the lower degree filtrations as well.

The conditions $(A_m)$ give us a way to inductively map the tensors in $T_m$ to mononomials in the indeterminates $z_\lambda$. Conditions $(B_m)$ can be rewritten to state

$(B_m)$ $x_\lambda z_\Sigma = z_\lambda z_\Sigma + y$ for some unique $y \in S_k$, when $z_\Sigma \in S_k$ and $k \le m$.

So, the conditions $(B_m)$ capture the idea that we can reorder the factors but with added error terms of lower degree that come from the bracket structure.

*Proof of Proposition 4.15.* We need to prove that such an action exists and that it is uniquely defined as well. To do so, besides $(A_m)$ and $(B_m)$, we also need the following condition to hold for each $m \in \mathbb{Z}_{\ge 0}$:

$(C_m)$ $x_\lambda(x_\mu z_T) = x_\mu(x_\lambda z_T) + [x_\lambda, x_\mu] z_T$, for all $x_\lambda, x_\mu \in \mathfrak{g}$ and $z_T \in S_{m-1}$,

where $(C_m)$ appears from the $\mathfrak{g}$-module structure. To define an action that satisfies these conditions, we will apply induction on the length of $\Sigma$.

**Lemma 4.16** (Base case)**.** *If the sequence $\Sigma$ has length $m = 0$, then there is a unique fitting $\mathfrak{g}$-action on the $S_0$ elements.*

*Proof.* For $m = 0$ we only have $z_\Sigma = z_\varnothing = 1$, so we can only define the action in one way such that it satisfies $(A_0)$, namely as $x_\lambda \cdot 1 = z_\lambda$, and this immediately satisfies the other conditions. Then we extend the definition to satisfy the two types of linearities as follows, for $\alpha, \beta \in \mathbb{C}$

$$x_\lambda(\alpha + \beta) = \alpha z_\lambda + \beta z_\lambda, \text{ and } (x_\lambda + x_\mu)\alpha = \alpha(z_\lambda + z_\mu). \tag{4.3}$$

Hence, we have successfully defined the desired action on $S_0$ elements. □

Now, for the *induction hypothesis* suppose that there is a uniquely defined action that satisfies all the conditions $(A_{m-1}), (B_{m-1})$ and $(C_{m-1})$ for sequences of length $m-1$. We aim to extend this to sequences of length $m$. It suffices to define the action $x_\lambda z_\Sigma$ where $\Sigma$ is increasing and of length $m$. Because the commutativity in $\mathcal{S}$ allows us to permute the factors into an increasing order without problem. For the linearity, the extension can be done as in (4.3) but with $\alpha z_\Sigma$ instead of $\alpha$ and similarly for $\beta$.

For the definition of the extension for length $m$ sequences, we fix an arbitrary sequence $\Sigma$ of length $m$. Then we split the definition of the extension into two cases, first when $\lambda \leq \Sigma$ and second when $\lambda \not\leq \Sigma$. The lemmas 4.17 and 4.18 will yield the definitions for the desired extension of the $\mathfrak{g}$-action.

**Lemma 4.17** (First extension)**.** *Let $x_\lambda \in \mathfrak{g}$. When $\lambda \leq \Sigma$, the extension of the $\mathfrak{g}$-action to $S_m$ elements is uniquely defined as $x_\lambda z_\Sigma = z_\lambda z_\Sigma$ and satisfy $(A_m)$.*

*Proof.* We define the extension such that $(A_m)$ is forced to hold. This can only be done by defining the action $x_\lambda z_\Sigma = z_\lambda z_\Sigma$, because the indices are already increasing. □

When $\lambda \not\leq \Sigma$, we cannot use the same definition because we want to preserve the increasing order of the indices to ensure uniqueness, otherwise multiple elements will be sent to monomials in $S_m$ that are the same after reordering the factors. Our next best bet is to use $(C_m)$ for the definition, as $(B_m)$ does not give an explicit form.

**Lemma 4.18** (Second extension)**.** *Let $x_\lambda \in \mathfrak{g}$. When $\lambda \not\leq \Sigma$, write $\Sigma = (\mu, T)$ where $\mu$ is the first index of $\Sigma$, $\mu \leq T$ and $T$ has length $m - 1$. The extension of the action to $S_m$ elements is uniquely defined as*

$$x_\lambda z_\Sigma = z_\mu z_\lambda z_T + x_\mu y + [x_\lambda, x_\mu] z_T \textit{ for a unique } y \in S_{m-1}$$

*and satisfies $(C_m)$ in the case of $\mu < \lambda$ and $\mu \leq T$.*

*Proof.* Since $\lambda \not\leq \Sigma$, the index $\mu$ is strictly less than $\lambda$. By the induction hypothesis $(A_{m-1})$ holds and tells us that $z_\Sigma = z_\mu z_T = x_\mu z_T$ for some $x_\mu \in \mathfrak{g}$. Condition $(B_{m-1})$ also holds by assumption, so $x_\lambda z_T \equiv z_\lambda z_T \pmod{S_{m-1}}$, which means there is a unique

$y \in S_{m-1}$ such that $x_\lambda z_T = z_\lambda z_T + y$. Now, $(C_m)$ can only be satisfied when the following holds,

$$\begin{aligned} x_\lambda z_\Sigma = x_\lambda(x_\mu z_T) &= x_\mu(x_\lambda z_T) + [x_\lambda, x_\mu]z_T \\ &= x_\mu(z_\lambda z_T + y) + [x_\lambda, x_\mu]z_T \\ &= z_\mu z_\lambda z_T + x_\mu y + [x_\lambda, x_\mu]z_T. \end{aligned} \tag{4.4}$$

The second equality is how $(C_m)$ is forced. The last equality comes from applying linearity and $(A_m)$. This allows us to properly define the action $x_\lambda z_\Sigma$ in a unique way. □

Now we are left to show that these two extensions from lemmas 4.17 and 4.18 combined satisfy all the conditions, instead of only the forced ones under their specific cases. Lemma 4.17 already forces condition $(A_m)$ to be true, so this has been cleared. We have not yet shown if any of the two extensions satisfy $(B_m)$.

**Lemma 4.19** (Satisfaction of $(B_m)$)**.** *The combined extension of lemmas 4.17 and 4.18 satisfies condition $(B_m)$:*

$(B_m)$ $x_\lambda z_\Sigma \equiv z_\lambda z_\Sigma$ *(mod $S_k$) when $z_\Sigma \in S_k$ and $k \leq m$.*

*Proof.* Fortunately, when the sequences are of length $k \leq m-1$, the condition is equivalent to $(B_{m-1})$, which holds by the induction hypothesis. When the sequence $\Sigma$ is of length $k = m$ and $\lambda \leq \Sigma$, then the condition is implied by $(A_m)$.

It remains to show that $(B_m)$ holds when $k = m$ and $\lambda \not\leq \Sigma$. Since we do not have an increasing order with $\lambda$, we use the definition in Lemma 4.18 attained from forcing $(C_m)$. As in Lemma 4.18 we write $\Sigma = (\mu, T)$ where $\mu \leq T$ and $T$ has length $m-1$. So, the extension is defined as

$$x_\lambda z_\Sigma = z_\mu z_\lambda z_T + x_\mu y + [x_\lambda, x_\mu]z_T \text{ for a unique } y \in S_{m-1}.$$

Because $y, z_T \in S_{m-1}$, we must have $x_\mu y$ and $[x_\lambda, x_\mu]z_T$ in $S_m$. This leads us to

$$x_\lambda z_\Sigma = z_\mu z_\lambda z_T + \underbrace{x_\mu y + [x_\lambda, x_\mu]z_T}_{\text{in } S_m} \equiv z_\lambda z_\Sigma \pmod{S_m}. \tag{4.5}$$

Proving that $(B_m)$ is also true in this final case. □

Lastly, we have to show that $(C_m)$ is satisfied in all cases by the extensions.

**Lemma 4.20** (Satisfaction of $(C_m)$)**.** *The combined extension of lemmas 4.17 and 4.18 satisfies condition $(C_m)$:*

$(C_m)$ $x_\lambda(x_\mu z_T) = x_\mu(x_\lambda z_T) + [x_\lambda, x_\mu]z_T$, *for all $x_\lambda, x_\mu \in \mathfrak{g}$ and $z_T \in S_{m-1}$,* *Proof.* We have to split the approach into four cases, each determined by the indices as in the order they are denoted in $(C_m)$.

(I) $\mu < \lambda$ and $\mu \leq T$;

This case coincides with the case of Lemma 4.18 in which we defined the action $x_\lambda z_\Sigma$ through $(C_m)$. Hence, $(C_m)$ is certainly true per construction.

(II) $\lambda < \mu$ and $\lambda \leq T$;

Here the situation is almost identical, but we have to be careful. In Lemma 4.18 the indices had the relation $\lambda > \mu$ and $\mu \leq T$, but in our situation we are dealing with indices where $\lambda < \mu$ and $\lambda \leq T$. So, Lemma 4.18 does yield an expression as in $(C_m)$, but with the indices swapped, so that

$$x_\mu(x_\lambda z_T) = x_\lambda(x_\mu z_T) + [x_\mu, x_\lambda]z_T.$$

This is not yet in the desired form, but using that $[x_\lambda, x_\mu] = -[x_\mu, x_\lambda]$, we find

$$x_\lambda(x_\mu z_T) = x_\mu(x_\lambda z_T) - [x_\mu, x_\lambda]z_T = x_\mu(x_\lambda z_T) + [x_\lambda, x_\mu]z_T.$$

Showing that $(C_m)$ is true in this case as well.

(III) $\lambda = \mu$;

In this case, the bracket is always equal to zero and $x_\lambda = x_\mu$, thus $(C_m)$ is satisfied.

(IV) $\lambda \not\leq T$ and $\mu \not\leq T$;

The final case will be fairly more involved and we will have to use the Jacobi identity. We write $T = (\nu, \Psi)$, where $\nu \leq \Psi$ and $\nu < \lambda, \mu$. Since $T$ has length $m-1$ and $\Psi$ has length $m-2$, the induction hypothesis allows us to apply $(A_{m-1})$ and $(C_{m-1})$ to obtain,

$$x_\mu z_T = x_\mu(x_\nu z_\Psi) = x_\nu(x_\mu z_\Psi) + [x_\mu, x_\nu]z_\Psi.$$

But also $(B_{m-1})$ can be applied such that $x_\mu z_\Psi = z_\mu z_\Psi + w$ with $w \in S_{m-2}$. Now since $\nu \leq \Psi$ and $\nu < \lambda$, by case (I) we can use $(C_m)$ on $x_\lambda(x_\nu(z_\mu z_\Psi))$ as $z_\mu z_\Psi \in S_{m-1}$, and without issues also on $x_\lambda(x_\nu w)$. Hence, $(C_m)$ applies to $x_\lambda(x_\nu(x_\mu z_\Psi))$ by linearity of the action. Therefore,

$$\begin{aligned} x_\lambda(x_\mu z_T) &= x_\lambda(x_\nu(x_\mu z_\Psi)) + x_\lambda([x_\mu, x_\nu]z_\Psi) \\ &= (x_\nu(x_\lambda(x_\mu z_\Psi)) + [x_\lambda x_\nu](x_\mu z_\Psi)) + ([x_\mu, x_\nu](x_\lambda z_\Psi) + [x_\lambda, [x_\mu, x_\nu]]z_\Psi). \end{aligned} \tag{4.6}$$

Since both $\lambda$ and $\mu$ have been indistinguishable throughout this argument for (IV), except for utilizing case (I), we can swap their roles and instead use case (II) to obtain the same expression as in (4.6) but with $\lambda$ and $\mu$ interchanged. Subtracting both expressions then yields,

$$\begin{aligned} x_\lambda(x_\mu z_T) - x_\mu(x_\lambda z_T) = x_\nu(x_\lambda(x_\mu z_\Psi)) + [x_\lambda, x_\nu](x_\mu z_\Psi) + [x_\mu x_\nu](x_\lambda z_\Psi) + [x_\lambda, [x_\mu, x_\nu]]z_\Psi \\ - x_\nu(x_\mu(x_\lambda z_\Psi)) - [x_\mu, x_\nu](x_\lambda z_\Psi) - [x_\lambda, x_\nu](x_\mu z_\Psi) - [x_\mu, [x_\lambda, x_\nu]]z_\Psi. \end{aligned}$$

Cancelling the gray-colored terms yields

$$= x_\nu(x_\lambda(x_\mu z_\Psi)) + [x_\lambda, [x_\mu, x_\nu]]z_\Psi - x_\nu(x_\mu(x_\lambda z_\Psi)) - [x_\mu, [x_\lambda, x_\nu]]z_\Psi.$$

Using linearity of the action we can gather the terms with $x_\nu$ acting on the left

$$= x_\nu(x_\lambda(x_\mu z_\Psi) - x_\mu(x_\lambda z_\Psi)) + [x_\lambda, [x_\mu, x_\nu]]z_\Psi - [x_\mu, [x_\lambda, x_\nu]]z_\Psi.$$

Now we can apply $(C_{m-1})$ and anticommutativity of the bracket to see

$$= x_\nu([x_\lambda, x_\mu]z_\Psi) + [x_\lambda, [x_\mu, x_\nu]]z_\Psi + [x_\mu, [x_\nu, x_\lambda]]z_\Psi.$$

This time applying $(C_m)$ which is possible as $\nu \leq \Psi$, a situation covered by the previous cases, yields

$$\begin{aligned} &= [x_\lambda, x_\mu](x_\nu z_\Psi) + ([x_\nu, [x_\lambda, x_\mu]] + [x_\lambda, [x_\mu, x_\nu]] + [x_\mu, [x_\nu, x_\lambda]])z_\Psi \\ &= [x_\lambda, x_\mu](x_\nu z_\Psi). \end{aligned}$$

In the last equality the Jacobi identity finishes the calculation, which proves $(C_m)$ for the final case. □

Through induction we have obtained an action that satisfies all the conditions $(A_m)$, $(B_m)$ and $(C_m)$ for all $m$, so we have a $\mathfrak{g}$-module structure on $\mathcal{S}$. Hence Proposition 4.15 is proven. □

Since Proposition 4.15 shows that there is a $\mathfrak{g}$-module structure on $\mathcal{S}$, an equivalent formulation for representations is possible, which says that there exists a representation $\rho : \mathfrak{g} \to \mathrm{End}(\mathcal{S})$ satisfying all the conditions $(A_m)$ and $(B_m)$, where the action of $x_\lambda$ is replaced by a mapping of $\rho(x_\lambda)$. A representation is itself a linear map preserving the Lie bracket operation, so for any $x, y \in \mathfrak{g}$ the map $\rho$ satisfies

$$\rho([x, y]_\mathfrak{g}) = [\rho(x), \rho(y)]_{\mathrm{End}(\mathcal{S})} = \rho(x)\rho(y) - \rho(y)\rho(x).$$

Hence, by the universal property of the universal enveloping algebra, $\rho$ can be extended to a unique algebra homomorphism $\tilde{\rho} : \mathcal{U}(\mathfrak{g}) \to \mathrm{End}(\mathcal{S})$. This way we obtain a unique algebra homomorphism $\rho' : \mathcal{T} \xrightarrow{\pi} \mathcal{U} \xrightarrow{\tilde{\rho}} \mathrm{End}(\mathcal{S})$, where $J = \ker \pi \subset \ker \rho'$ (see table above Lemma 4.13). This map behaves as we would expect, so for example for the element $1 - 5x_1 \otimes x_2 \otimes x_3 \in \mathcal{T}$ and a $z_\Sigma \in S_m \subset \mathcal{S}$ with $3 \leq \Sigma$, we have that

$$\begin{aligned} \rho'(1 - 5x_1 \otimes x_2 \otimes x_3)(z_\Sigma) &= \rho'(1)(z_\Sigma) - 5\rho'(x_1)\rho'(x_2)\rho'(x_3)(z_\Sigma) \\ &= z_\Sigma - 5z_1 z_2 z_3 z_\Sigma, \end{aligned}$$

and for $x_2 \otimes x_1$, we have

$$\rho'(x_2 \otimes x_1)(z_\Sigma) = \rho'(x_2)\rho'(x_1)(z_\Sigma) \equiv z_2 z_1 z_\Sigma \pmod{S_{m+1}}.$$

**Lemma 4.21.** *Let $t \in T_m \cap J$, then the homogeneous component $t_m$ of $t$ of degree $m$ lies in the ideal $I$ (see table above Lemma 4.13 to recall the definitions).*

*Proof.* Now, let $t \in T_m \cap J$, then it is also contained in the kernel of $\rho'$, so $\rho'(t) = 0$. Write the homogeneous component $t_m$ as $\sum_{j=1}^{r} c_j x_{\Sigma_j}$, a linear combination of basis elements $x_{\Sigma_j}$ with $1 \leq j \leq r$, where each sequence $\Sigma_j$ (not necessarily increasing) is of length $m$ and $c_j \in \mathbb{C}$ scalars. We let $t' := t - t_m \in T_{m-1}$ denote the part of $t$ that is not of degree $m$. Then, for $1 \in \mathcal{S}$ we have that

$$0 = \rho'(t)(1) = \rho'(t')(1) + \rho'(t_m)(1),$$

where the right-hand side is a polynomial of variables expressed in the monomials $z_{\Sigma_1}, \ldots, z_{\Sigma_r}$. However, components of the same degree on the left- and right-hand side need to equal each other. As $\rho'(t')(1)$ has no degree $m$ component, this means that the degree $m$ component of $\rho'(t_m)(1)$ is zero in $\mathcal{S}$ because cancellations with terms of this degree in $\rho'(t')(1)$ will not occur. The properties $(B_m)$ imply that

$$\rho'(t_m)(1) \equiv \sum_{j=1}^{r} c_j z_{\Sigma_j} \pmod{S_{m-1}},$$

where the degree $m$ component $\sum_{j=1}^{r} c_j z_{\Sigma_j} = 0$ in $\mathcal{S}$. So the canonical map $\sigma : \mathcal{T} \to \mathcal{T}/I = \mathcal{S}$ sends $t_m = \sum_{j=1}^{r} c_j x_{\Sigma_j}$ to 0:

$$\sigma(t_m) = \sigma\left(\sum_{j=1}^{r} c_j x_{\Sigma_j}\right) = \sum_{j=1}^{r} c_j \sigma(x_{\Sigma_j}) = \sum_{j=1}^{r} c_j z_{\Sigma_j} = 0.$$

From this we conclude that $t_m \in I$, as desired. □

*proof of PBW Theorem.* It remains to prove that the map $\omega : \mathcal{S} \to \mathcal{G}$ is injective. This can be done by showing the kernel of $\omega$ is zero. But we can trace this back to the condition $\ker \varphi \subset I$, because $\varphi = \omega \circ \sigma$ implies that $\omega$ only maps $\sigma(I)$ to 0, which is zero in $\mathcal{S}$.

To show $\ker \varphi \subset I$, we let $m \in \mathbb{Z}_{>0}$ and $t \in T^m \cap \ker \varphi$, and we claim that $\pi(t) \in U_{m-1}$. This would mean that $t$ as a homogeneous noncommutative polynomial of degree $m$ has factors that lie in the ideal $J$, resulting in the degree of the polynomial being lowered to $m-1$ once it is sent to the universal enveloping algebra $\mathcal{U}$. Indeed, $t$ is in particular an element of $\ker \varphi_m$ and as $\varphi_m = f_m \circ \pi$ we must have $\pi(\ker \varphi_m) \subset \ker f_m$. Now, $\ker f_m = U_{m-1}$ and therefore $\pi(t) \in U_{m-1}$.

On the other hand, because per definition $U_{m-1} = \pi[T_{m-1}]$, there must be a $t'$ in the filtration $T_{m-1}$ such that $\pi(t') = \pi(t) \in U_{m-1}$. But now $t - t' \in \ker \pi = J$, so we can apply Lemma 4.21 to $t - t' \in T_m \cap J$. As the homogeneous component of degree $m$ is $t$, we obtain $t \in I$. □

# 5 Classification of semisimple Lie algebra modules

Finally we have gathered all the necessary knowledge to explore how all the *finite-dimensional irreducible* modules of semisimple Lie algebras over $\mathbb{C}$ can be classified. This whole section will contain a substantial amount of results that are analogous to the $\mathfrak{sl}(2,\mathbb{C})$ case. Besides the analogies, we even have to use some results from Section 2.4 on $\mathfrak{sl}(2,\mathbb{C})$-modules to make the classification possible for these semisimple Lie algebra modules.

We start in Section 5.1 by introducing standard cyclic modules as $\mathfrak{g}$-modules generated by their maximal vector, which leads us to construct what are called Verma modules for every highest weight in Section 5.2. These Verma modules are generally reducible, but by taking a particular quotient module from it, we end up with an irreducible module. Unfortunately, Verma module are always infinite-dimensional, even when the Lie algebra is finite-dimensional. Though, in Section 5.3 we will encounter the necessary and sufficient conditions to have a finite-dimensional standard cyclic module. This culminates in a bijection between the so called dominant integral weights and the isomorphism classes of finite-dimensional irreducible modules.

*Throughout the section we fix a complex semisimple Lie algebra* $\mathfrak{g}$*, with Cartan subalgebra* $\mathfrak{h}$*. The corresponding root system is denoted by* $\Phi$ *with a base* $\Delta$ *and Weyl group* $\mathcal{W}$*.*

## 5.1 Standard cyclic modules

Since $\mathfrak{h}$ is a Cartan subalgebra, all the linear maps in $\operatorname{ad}(\mathfrak{h})$ are diagonalizable, hence by Theorem 2.21 the corresponding endomorphisms $\rho(h)$ of $h \in \mathfrak{h}$ are diagonalizable in the *finite-dimensional* $\mathfrak{g}$-module $V$. But as all the elements in $\mathfrak{h}$ commute, so do the corresponding endomorphisms over $V$. Hence they are all simultaneously diagonalizable by Theorem 2.20, resulting in the decomposition

$$V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda \text{ with } V_\lambda = \{v \in V : h \cdot v = \lambda(h)v, \text{ for all } h \in \mathfrak{h}\},$$

similarly as in the case of $\mathfrak{sl}(2,\mathbb{C})$ where the $V_\lambda$ were eigenspaces. In this more general situation we say $V_\lambda$ is a **weight space** if $V_\lambda \neq 0$ and the corresponding $\lambda \in \mathfrak{h}^*$ a **weight** of $\mathfrak{h}$ on $V$, analogous to Definition 2.22.

*Remark.* If $V$ is infinite-dimensional, then it is possible that $V$ is not necessarily equal to the sum of its weight spaces.

Now we would like to know how the rest of $\mathfrak{g}$ acts on weight spaces of $V$. Since $\mathfrak{g} = \mathfrak{h} + \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$, we will examine the action of such $\mathfrak{g}_\alpha$ on $V$. Let $\alpha \in \Phi$ and $x \in \mathfrak{g}_\alpha$, then given $h \in \mathfrak{h}$ and $v \in V_\lambda$,

$$h \cdot (x \cdot v) = [h,x] \cdot v + x \cdot (h \cdot v) = \alpha(h)x \cdot v + x \cdot (\lambda(h)v) = (\lambda + \alpha)(h)x \cdot v.$$

So elements in $\mathfrak{g}_\alpha$ send weight vectors in $V_\lambda$ to $V_{\lambda+\alpha}$, thus we can conclude the following.

**Lemma 5.1.** *Let $V$ be any representation of $\mathfrak{g}$, then given $\lambda \in \mathfrak{h}^*$ and $\alpha \in \Phi$ we have that $\mathfrak{g}_\alpha \cdot V_\lambda \subset V_{\lambda+\alpha}$.*

*Remark.* From the lemma we see that weight spaces on their own are not submodules, but are only linear subspaces.

In the $\mathfrak{sl}(2, \mathbb{C})$ case, we could fully determine the irreducible modules based on what the highest weight was. We want to push this idea further to general semisimple Lie algebras. But before we can do that, we need to extend the definition of what a "highest weight" means in this context, as now the weights are not just numbers but linear functionals.

Recall from Lemma 3.9 that the roots in $\Phi$ span $\mathfrak{h}^*$. So, we can define a *partial order* on the weights: we say $\lambda \succ \mu$ if and only if $\lambda - \mu$ can be written as $\sum_{\alpha\in\Phi^+} k_\alpha \alpha$ where all $k_\alpha \geq 0$. When $\lambda \succ \mu$, we say that $\lambda$ is *higher* than $\mu$ and correspondingly we say that $\mu$ is *lower* than $\lambda$. This is visualized in Figure 7.

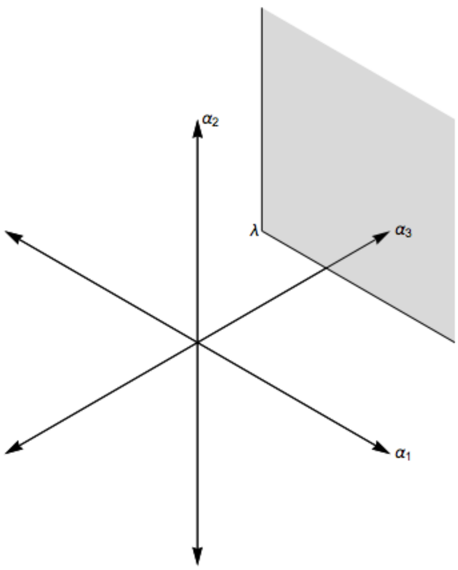



Figure 7: Base given by $\alpha_1$ and $\alpha_2$, the points higher than $\lambda$ are in grey (Wikipedia [13]).

**Definition 5.2.** A **highest weight** of $V$ is a weight $\lambda \in \mathfrak{h}^*$ that is higher than all the other weights, so $\lambda \succ \mu$ for all weights $\mu$ of $V$.

*Remark.* From this definition it follows that a highest weight is necessarily unique.

**Definition 5.3.** A **maximal vector** $v^+ \in V_\lambda$ is a nonzero element satisfying

$$x_\alpha \cdot v^+ = 0,$$

for all $x_\alpha \in \mathfrak{g}_\alpha$ where $\alpha \in \Phi^+$. So $v^+$ gets annihilated by all the elements of the positive root spaces.

*Remark.* We will see later, that for the modules $V$ we are interested in, their maximal vectors are of highest weight.

*Remark.* If the module is finite-dimensional, then there will always exist a maximal vector. But for infinite-dimensional modules, there is no guarantee that such a vector exists, so we have to be more careful there.

We have seen in the proof of Theorem 2.24 part *(3)* that irreducible $\mathfrak{sl}(2,\mathbb{C})$-modules can be constructed by knowing the action of the Lie algebra on their maximal vector. This motivates us to study $\mathfrak{g}$-modules that are *generated* by a maximal vector, meaning that the elements in this module are the result of repeated $\mathfrak{g}$-action on the maximal vector.

**Definition 5.4.** Let $V$ be a $\mathfrak{g}$-module, then $V$ is called a **standard cyclic module** of weight $\lambda$, if there is a maximal vector $v^+ \in V$ with weight $\lambda$ such that $V = \mathcal{U}(\mathfrak{g}) \cdot v^+$.

The definition says that a standard cyclic module $V$ is spanned by elements of the form $a_1 a_2 \cdots a_m \cdot v^+$ with $a_i \in \mathfrak{g}$, so in this manner is $V$ generated by a maximal vector $v^+$.

Now we prove a list of important properties for these standard cyclic modules. A few interesting things to note here are that we will see that the weight of maximal vectors of standard cyclic modules are in fact the highest weight of the module. Furthermore, standard cyclic modules are in general reducible, but we do find that they are always indecomposable. To obtain an irreducible variant, we will have to take appropriate quotients.

**Theorem 5.5.** *Let $V$ be a standard cyclic $\mathfrak{g}$-module, with maximal vector $v^+ \in V_\lambda$ of weight $\lambda \in \mathfrak{h}^*$. Let $\Phi^+ = \{\beta_1, \ldots, \beta_m\}$. Then,*

*(1) $V$ is spanned by the vectors $y_{\beta_1}^{i_1} \cdots y_{\beta_m}^{i_m} \cdot v^+$ where $i_j \in \mathbb{Z}_{\geq 0}$ and $y_{\beta_j} \in \mathfrak{g}_{-\beta_j}$. In particular $V$ is the direct sum of its weight spaces.*

*(2) The weights of $V$ are of the form $\mu = \lambda - \sum_{j=1}^{l} k_j \alpha_j$, where $k_j \in \mathbb{Z}_{\geq 0}$ and $\alpha_j \in \Delta$. So all the weights satisfy $\mu \prec \lambda$, making $\lambda$ a highest weight.*

*(3) For each weight $\mu \in \mathfrak{h}^*$, the weight space $V_\mu$ is finite-dimensional and* $\dim V_\lambda = 1$.

*(4) Each submodule of $V$ is a direct sum of its weight spaces.*

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

*(6) If $f : V \to W$ is a nonzero homomorphism between $\mathfrak{g}$-modules, then the image $f[V]$ is also a standard cyclic $\mathfrak{g}$-module of weight $\lambda$.*

*Proof.* We begin by writing $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{b}$ where $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+$, called the *Borel subalgebra*, and $\mathfrak{n}^\pm = \bigoplus_{\alpha \succ 0} \mathfrak{g}_{\pm\alpha}$. Now we choose a basis $(y_1, \ldots, y_m, z_1, \ldots, z_n)$ for $\mathfrak{g}$ where $(y_1, \ldots, y_m)$

forms a basis for $\mathfrak{n}^-$ and $(z_1, \ldots, z_n)$ a basis for $\mathfrak{b}$. Applying Corollary 4.14.4 of the PBW theorem leads us to the following,

$$\mathcal{U}(\mathfrak{g}) \cdot v^+ = \mathcal{U}(\mathfrak{n}^-) \cdot (\mathcal{U}(\mathfrak{b}) \cdot v^+),$$

This tells us that given, for example, $y_1 y_2 z_1 \in \mathcal{U}(\mathfrak{g})$, we have $y_1 y_2 z_1 \cdot v^+ = y_1 y_2 \cdot (z_1 \cdot v^+)$. Furthermore, we observe that

$$\mathcal{U}(\mathfrak{b}) \cdot v^+ = \mathbb{C}v^+. \tag{5.1}$$

This is because each element $z \in \mathfrak{b}$ can be written as $z = h + \bar{x}$ with $h \in \mathfrak{h}$ and $\bar{x}$ a sum of elements in the positive root spaces, while all the terms in $\bar{x}$ annihilate $v^+$ per definition and $h \cdot v^+ = \lambda(h) v^+$ with $\lambda(h) \in \mathbb{C}$. Since $\mathcal{U}(\mathfrak{b}) \cdot v^+$ constitutes repeated action of $\mathfrak{b}$-elements on $v^+$, we indeed find that it is equal to the span of $v^+$. So in the end, we conclude

$$\mathcal{U}(\mathfrak{g}) \cdot v^+ = \mathbb{C}(\mathcal{U}(\mathfrak{n}^-) \cdot v^+). \tag{5.2}$$

Now Corollary 4.14.3 of the PBW theorem says that $\mathcal{U}(\mathfrak{n}^-)$ has a basis consisting of the monomials $y_{\beta_1}^{i_1} \cdots y_{\beta_m}^{i_m}$ with $i_j \in \mathbb{Z}_{\geq 0}$. So interpreting Equation (5.2): the module $V$ is spanned by the vectors $y_{\beta_1}^{i_1} \cdots y_{\beta_m}^{i_m} \cdot v^+$, which is exactly the first part of statement *(1)*. To show that $V$ is indeed equal to the direct sum of its weight spaces, we note that the direct sum of the weight spaces is a subspace of $V$. Because the vectors $y_{\beta_1}^{i_1} \cdots y_{\beta_m}^{i_m} \cdot v^+$ span $V$, the sum of all the weight spaces corresponding to such vectors must also span $V$.

To prove statement *(2)* we use Lemma 5.1 to obtain $y_{\beta_j}^{i_j} \in \mathfrak{g}_{-\beta_j}^{i_j}(V_\lambda) \subset V_{\lambda - i_j \beta_j}$, where the vector receives the weight $\lambda - i_j \beta_j$. Repeating this calculation yields,

$$\lambda - \sum_{j=1}^{m} i_j \beta_j \tag{5.3}$$

as the weight of the vector $y_{\beta_1}^{i_1} \cdots y_{\beta_m}^{i_m} \cdot v^+$. Since all $\beta_j$ are positive roots, we can rewrite them as nonnegative integer linear combinations of the simple roots in the base $\Delta$, hence obtaining the desired form.

For *(3)*, we note that there are only a finite number of vectors $y_{\beta_1}^{i_1} \cdots y_{\beta_m}^{i_m} \cdot v^+$ of which their weight as in (5.3) can be equal to $\mu = \lambda - \sum_{j=1}^{l} k_j \alpha_j$. Certainly, both sums are finite, so there are only finitely many ways to have different sums be equal to each other. As $V$ is the direct sum of its weight spaces, the finite set of vectors described above must span the weight space $V_\mu$. Moreover, the only vector which has weight $\mu = \lambda$ is $v^+$ itself, hence the dimension of $V_\lambda$ is 1.

Proceeding with *(4)*, we let $W$ be a submodule of $V$. We write elements in $W$ as a unique sum with nonzero coefficients of vectors $v_i \in V_{\mu_i}$, each belonging to different weights, which is possible by part *(1)* of Theorem 5.5. The spaces $V_{\mu_i}$ are not necessarily known to be weight spaces of $W$, though this is exactly what we want to prove, since the directness of the sums already holds for weight spaces in general. Suppose the $V_{\mu_i}$ are not all weight spaces of $W$, then we can choose weight vectors

$$v_i \notin W \text{ such that } v_1 + \cdots + v_n \in W \text{ where } n > 1. \tag{5.4}$$

Now we can always choose a minimal length sum of $v_i$'s that satisfy (5.4), in the sense that $n$ is minimal. This is because a sum as in (5.4) must exist by our assumption, this will have a length of say $m > 1$ terms. If no shorter sums satisfy (5.4), then the minimal amount is $n = m$, otherwise we take the shorter sum of length $1 < n < m$. Take one such shortest sum and call it $w$. Since the weights are distinct we take an $h \in \mathfrak{h}$ for which $\mu_1(h) \neq \mu_2(h)$. Then,

$$h \cdot w = \sum_{i=1}^{n} \mu_i(h) v_i \in W,$$

because $W$ is a submodule, hence invariant under Lie algebra actions. But then also,

$$(h - \mu_1(h) \cdot 1) \cdot w = \underbrace{(\mu_2(h) - \mu_1(h))v_2}_{\neq 0} + \cdots + (\mu_n(h) - \mu_1(h))v_n \in W.$$

So we have found an element in $W$ that is a sum with $n - 1$ weight vectors not in $W$, which is in contradiction with the minimality of $n$. Hence no shortest sum satisfying (5.4) exists. But then $n = 0$ or $n = 1$ so all $v_i \in V_{\mu_i}$ for all $i$ are elements of $W$. Thus all these $V_{\mu_i}$ are weight spaces of $W$.

To prove *(5)*, we first note that any submodule containing $V_\lambda$ must be equal to $V$, as the maximal vector $v^+ \in V_\lambda$ generates the whole space. Combining this with *(4)* of the theorem, we can say that any proper submodule of $V$ lies in the sum of all weight spaces excluding $V_\lambda$. We call this sum $W$, which is still a submodule and proper, hence it is the unique maximal submodule of $V$. This implies that $V$ cannot be decomposed into a direct sum of two proper submodules, as they and hence also the sum would both be contained in $W$, thus $V$ is indeed indecomposable. This maximal submodule comes with a unique irreducible quotient module $V/W$.

*Irreducibility* is due to the maximality of $W$, indeed if there was a non-trivial submodule $U$ of the quotient, then $U + W$ would be a proper submodule of $V$. But $U + W$ should then per definition be contained in $W$, though this implies that $U = 0$ in the quotient module $V/W$. Hence the quotient module is irreducible.

*Uniqueness* also follows from the maximality: take a different irreducible quotient $V/W'$, where $W'$ has to be a submodule distinct from $W$. Then $W' \subset W$, because $W$ is maximal and now $W/W'$ is a nonzero submodule of $V/W'$, contradicting the irreducibility. Thus we do end up with a unique irreducible quotient $V/W$ from the maximal submodule $W$.

Lastly *(6)* follows by taking the image of the maximal vector $v^+ \in V$ to see that $f(v^+) \in f[V]$ is again a maximal vector of the same weight $\lambda$. Indeed, for any $\alpha \in \Phi^+$, $x_\alpha \in \mathfrak{g}_\alpha$ and $h \in \mathfrak{h}$,

$$\begin{aligned} x_\alpha \cdot f(v^+) &= f(x_\alpha \cdot v^+) = 0, \\ h \cdot f(v^+) &= f(h \cdot v^+) = f(\lambda(h)v^+) = \lambda(h) f(v^+). \end{aligned}$$

Since $v^+$ generates $V$, the element $f(v^+)$ must therefore also generate $f[V]$ because $f$ is a homomorphism of $\mathfrak{g}$-modules. □

## 5.2 Irreducible standard cyclic modules

Having proven the necessary properties of general standard cyclic modules, we are all set to move on with their irreducible variants. We will encounter two uniqueness properties, first of the maximal vector and second of the standard cyclic modules themselves. Then we construct a type of standard cyclic modules that are called Verma modules to show the existence of such modules. We finish by showing that irreducible standard cyclic modules of any weight can be obtained from Verma modules.

**Theorem 5.6.** *Let $V$ be an irreducible standard cyclic $\mathfrak{g}$-module, with maximal vector $v^+ \in V_\lambda$ of weight $\lambda$. Then $v^+$ is the unique maximal vector in $V$, up to nonzero scalar multiples.*

*Proof.* Given a maximal vector $w^+$ in $V$ of weight $\lambda'$, because of the irreducibility, we know that the only submodule containing $w^+$ must be $V$ itself. So we can say that $w^+$ generates $V$ as a $\mathfrak{g}$-module, or $V = \mathcal{U}(\mathfrak{g}) \cdot w^+$. This allows us to use the results from Theorem 5.5 with $w^+$ as the maximal vector instead, yielding $\lambda \prec \lambda'$ because $\lambda'$ is now the highest weight by part *(2)*, but at the same time $\lambda' \prec \lambda$ holds via the same argument as $V = \mathcal{U}(\mathfrak{g}) \cdot v^+$. Hence $\lambda = \lambda'$, but the weight space of the highest weight is just the span of the maximal vector as is said in statement *(3)*, which means $w^+$ and $v^+$ are scalar multiples of each other. □

The uniqueness of maximal vectors will help us prove that, up to isomorphism, at most one irreducible standard cyclic module of weight $\lambda \in \mathfrak{h}^*$ can exist. This tremendously simplifies the classification of the modules.

**Theorem 5.7.** *Let $V$ and $W$ be irreducible standard cyclic $\mathfrak{g}$-module with the same highest weight $\lambda$. Then $V$ and $W$ are isomorphic modules.*

*Proof.* First we define the $\mathfrak{g}$-module $X = V \oplus W$. Suppose the maximal vectors of $V$ and $W$ are $v^+$ and $w^+$ respectively, then we claim that $x^+ = (v^+, w^+)$ is in fact a maximal vector of $X$. Indeed, let $\alpha \in \Phi^+$ and $x_\alpha \in \mathfrak{g}_\alpha$, then

$$x_\alpha \cdot x^+ = x_\alpha \cdot (v^+, w^+) = (x_\alpha \cdot v^+, x_\alpha \cdot w^+) = (0, 0).$$

Now we look at the module $Y = \mathcal{U}(\mathfrak{g}) \cdot x^+$ and introduce the projections $p : Y \to V$ on the first coordinate and $p' : Y \to W$ on the second coordinate. These projections are also module homomorphism, so we can conclude that they send maximal vectors to maximal vectors by the argument made in the proof of *(6)* of Theorem 5.5. Hence by the uniqueness of maximal vectors (up to scalar multiples, which would not affect the rest of the argument), we have that $p(x^+) = v^+$ and $p'(x^+) = w^+$. Because the projections are homomorphisms we find that

$$p[Y] = \mathcal{U}(\mathfrak{g}) \cdot p(x^+) = \mathcal{U}(\mathfrak{g}) \cdot v^+ = V \quad \text{and} \quad p'[Y] = \mathcal{U}(\mathfrak{g}) \cdot p'(x^+) = \mathcal{U}(\mathfrak{g}) \cdot w^+ = W,$$

so $p[Y] = V$ and $p'[Y] = W$. Now applying the isomorphism theorem yields,

$$Y/\ker(p) \cong p[Y] = V \quad \text{and} \quad Y/\ker(p') \cong p'[Y] = W.$$

This allows us to interpret $V$ and $W$ as quotients of the module $Y$. But since both $V$ and $W$ were assumed to be irreducible, the uniqueness of irreducible quotient modules as stated in *(5)* of Theorem 5.5 forces them to be isomorphic to each other, which proves the theorem. □

All these theorems on standard cyclic modules have been proven, but we have not determined the existence of these modules yet. Problems might arise as infinite-dimensional modules are not guaranteed to have a maximal vector. So, we want to construct a module that contains a nonzero maximal vector, which in turn will be a standard cyclic module. Achieving this is not so trivial and can be done in a few ways. We will follow the *induced module* or *extension of scalars* construction.

This construction is motivated by what we have seen should hold for maximal vectors of standard cyclic modules in Equation (5.1),

$$\mathcal{U}(\mathfrak{b}) \cdot v^+ = \mathbb{C}v^+.$$

Showing that $\mathfrak{b}$-actions on the maximal vector only scale the vector, so the new vector remains in its span. So the one-dimensional space spanned by the given maximal vector is a $\mathfrak{b}$-module. So, to construct a standard cyclic $\mathfrak{g}$-module we start with a one-dimensional vector space $D_\lambda$ which has a basis given by a vector $v^+$. This vector will not be the maximal vector of our standard cyclic $\mathfrak{g}$-module, but we will only make this vector into a maximal vector for $D_\lambda$ seen as a $\mathfrak{b}$-module. To do so, we define the action of the Borel subalgebra $\mathfrak{b} \subset \mathfrak{g}$ on $D_\lambda$ as follows,

$(a)$ for all $h \in \mathfrak{h},\ h \cdot v^+ = \lambda(h)v^+$,
$(b)$ for all $x \in \mathfrak{n}^+,\ x \cdot v^+ = 0$,

which makes our vector space $D_\lambda$ into a $\mathfrak{b}$-module. This can be shown as follows, given $b_1, b_2 \in \mathfrak{b}$, we write $b_i = h_i + x_i$ with $i = 1, 2$, where $h_i \in \mathfrak{h}$ and $x_i \in \mathfrak{n}^+$,

$$\begin{aligned}[b_1, b_2] \cdot v^+ &= [h_1 + x_1, h_2 + x_2] \cdot v^+ \\ &= ([h_1, h_2] + [h_1, x_2] + [x_1, h_2] + [x_1, x_2]) \cdot v^+ \\ &= (0 + \alpha_2(h_1)x_2 - \alpha_1(h_2)x_1 + [x_1, x_2]) \cdot v^+ = 0,\end{aligned}$$

and

$$\begin{aligned}b_1 \cdot (b_2 \cdot v^+) - b_2 \cdot (b_1 \cdot v^+) &= (h_1 + x_1) \cdot \lambda(h_2)v^+ - (h_2 + x_2) \cdot \lambda(h_1)v^+ \\ &= \lambda(h_1)\lambda(h_2)v^+ - \lambda(h_2)\lambda(h_1)v^+ = 0.\end{aligned}$$

So the action satisfies condition (M3) and hence $D_\lambda$ is a $\mathfrak{b}$-module. Thus we can also view $D_\lambda$ as an $\mathcal{U}(\mathfrak{b})$-module, by defining the action as repeated $\mathfrak{b}$-action. Via Corollary

4.14.4 we know that $\mathcal{U}(\mathfrak{g})$ is also a $\mathcal{U}(\mathfrak{b})$-module, as the Borel subalgebra $\mathfrak{b}$ is certainly a subalgebra of $\mathfrak{g}$. This allows us to define the following tensor product of $\mathcal{U}(\mathfrak{b})$-modules called the **Verma module** of weight $\lambda$,

$$Z(\lambda) := \mathcal{U}(\mathfrak{g}) \otimes_{\mathcal{U}(\mathfrak{b})} D_\lambda.$$

The name "extension of scalars" is given to this construction because we start by viewing $D_\lambda$ as a module with scalars in $\mathcal{U}(\mathfrak{b})$, but through the tensor product we can interpret it with scalars in $\mathcal{U}(\mathfrak{g})$ instead. We first make $Z(\lambda)$ into a $\mathcal{U}(\mathfrak{g})$-module (generated by $1 \otimes v^+$), which is done by defining the action of $x \in \mathfrak{g}$ as the natural left action,

$$x \cdot (1 \otimes v^+) = x \otimes v^+. \tag{5.5}$$

Since we can move scalars of $\mathcal{U}(\mathfrak{g})$ and $D_\lambda$ viewed as $\mathcal{U}(\mathfrak{b})$-modules through the tensors, we have

$$x \otimes v^+ = \begin{cases} x \otimes v^+, & \text{if } x \in \mathfrak{n}^-, \\ 1 \otimes (x \cdot v^+), & \text{if } x \in \mathfrak{b}. \end{cases} \tag{5.6}$$

So we can see $Z(\lambda)$ as a $\mathcal{U}(\mathfrak{g})$-module via repeated $\mathfrak{g}$-action. Now, the scalars of $D_\lambda$ have in some fashion been extended to $\mathcal{U}(\mathfrak{g})$, with a particular element being $xyz \otimes v^+$. Then viewing the tensor as a formal product allows the notation $xyzv^+$, as if $xyz$ is a scalar of the vector $v^+$.

We have defined $D_\lambda$ such that $\mathcal{U}(\mathfrak{b}) \cdot v^+ = D_\lambda$ holds, so via the definition in (5.6) we can obtain $1 \otimes D_\lambda$ from $1 \otimes v^+$. Now applying the definition in (5.5) to let $\mathcal{U}(\mathfrak{g})$ act on $1 \otimes D_\lambda$ yields $\mathcal{U}(\mathfrak{g}) \otimes D_\lambda$. Hence $1 \otimes v^+$ generates $Z(\lambda)$. The tensor is nonzero and for any $x \in \mathfrak{n}^+ \subset \mathfrak{b}$,

$$x \cdot (1 \otimes v^+) = 1 \otimes (x \cdot v^+) = 1 \otimes 0 = 0,$$

thus $1 \otimes v^+$ is a maximal vector of weight $\lambda$ and therefore the Verma module $Z(\lambda)$ is a standard cyclic $\mathfrak{g}$-module of weight $\lambda$.

So for every weight a Verma module, $Z(\lambda)$, can be constructed, but they are not always irreducible. Fortunately, Theorem 5.5 solves this issue.

**Theorem 5.8.** *For every weight $\lambda \in \mathfrak{h}^*$ there exists an irreducible standard cyclic $\mathfrak{g}$-module $V(\lambda)$ of weight $\lambda$.*

*Proof.* Given any weight $\lambda \in \mathfrak{h}^*$, we can construct the Verma module $Z(\lambda) = \mathcal{U}(\mathfrak{g}) \otimes_{\mathcal{U}(\mathfrak{b})} D_\lambda$ of weight $\lambda$. Because the Verma module is a standard cyclic $\mathfrak{g}$-module, we apply part *(5)* Theorem 5.5 to it, resulting in the unique maximal submodule $Y(\lambda)$. Then the same statement says that the quotient

$$V(\lambda) := Z(\lambda)/Y(\lambda)$$

is an irreducible $\mathfrak{g}$-module. The canonical homomorphism $f : Z(\lambda) \to Z(\lambda)/Y(\lambda)$ has image $f[Z(\lambda)] = V(\lambda)$, so by part *(6)* of Theorem 5.5 the quotient module $V(\lambda)$ must be standard cyclic of weight $\lambda$ as well. □

Combining the results from Theorem 5.8 with Theorem 5.7, we can conclude that, there exists a unique irreducible standard cyclic $\mathfrak{g}$-module for every weight $\lambda \in \mathfrak{h}^*$, which has to be isomorphic to the quotient $V(\lambda)$.

## 5.3 Finite-dimensional modules

Our goal is to classify all finite-dimensional modules, though because of Weyl's theorem we can restrict ourselves to the irreducible modules of finite dimension. Thus far we have only studied standard cyclic modules, which can be infinite-dimensional. In particular, the Verma modules are always infinite-dimensional because of the factor $\mathcal{U}(\mathfrak{g})$ in the tensor product. First, we will show that all finite-dimensional irreducible modules are also standard cyclic. In fact, they are isomorphic to the Verma module quotient $V(\lambda)$, if the module has highest weight $\lambda$. What remains, is to find out under what conditions $V(\lambda)$ is finite-dimensional, which will be the focus of the rest of the section.

**Theorem 5.9.** *If $V$ is a finite-dimensional irreducible $\mathfrak{g}$-module of highest weight $\lambda \in \mathfrak{h}^*$, then $V \cong V(\lambda)$.*

*Proof.* If $V$ is a finite-dimensional irreducible $\mathfrak{g}$-module, then $V$ has at least one maximal vector $v^+ \in V$. If this were not the case, we could obtain infinitely many distinct weight spaces by repeated application of Lemma 5.1. But a finite-dimensional module is a direct sum of its weight spaces as we have seen in the discussion of Section 5.1 , making it impossible for infinitely many weight spaces of $V$ to exist. So there must exist a maximal vector in $V$, which is killed by the actions of all the positive root spaces $\mathfrak{g}_\alpha$ where $\alpha \in \Phi^+$, thus respecting the finite-dimensionality. This maximal vector has a uniquely determined weight $\lambda \in \mathfrak{h}^*$, and the vector generates a submodule $\mathcal{U}(\mathfrak{g}) \cdot v^+$ of $V$ which has to be equal to $V$ by irreducibility. As $V$ is now a standard cyclic $\mathfrak{g}$-module of weight $\lambda$, the uniqueness from Theorem 5.7 implies that $V \cong V(\lambda)$ where $V(\lambda)$ is the quotient module as in Theorem 5.8. □

In the following discussions, we will make considerable use of Theorem 3.8 to extract $\mathfrak{sl}(2,\mathbb{C})$-properties for the modules $V(\lambda)$. We will denote such an $\mathfrak{sl}(2,\mathbb{C})$-copy in $\mathfrak{g}$, for the simple roots $\alpha \in \Delta$, by $\mathfrak{s}_\alpha$, where $x_\alpha, y_\alpha$, and $h_\alpha$ are the corresponding generators. With this, we can already establish the necessary condition for finite-dimensionality, which amounts to a condition on the highest weight.

**Theorem 5.10.** *If $V$ is a finite-dimensional irreducible $\mathfrak{g}$-module of highest weight $\lambda$, then $\lambda(h_\alpha)$ is a nonnegative integer for all $\alpha \in \Delta$.*

*Proof.* Let $\alpha \in \Delta$. Since $V \cong V(\lambda)$ is a finite-dimensional irreducible $\mathfrak{g}$-module, and $\mathfrak{s}_\alpha$ is a subalgebra of $\mathfrak{g}$, it follows that $V$ is also a finite-dimensional irreducible $\mathfrak{s}_\alpha$-module. Additionally, a maximal vector in $V$ as a $\mathfrak{g}$-module is killed by all elements in $\mathfrak{n}^+$, but $x_\alpha \in \mathfrak{n}^+$, so it remains as a maximal vector when we view $V$ as an $\mathfrak{s}_\alpha$-module. Hence, $\lambda(h_\alpha)$ is the highest weight of $V$ as an $\mathfrak{s}_\alpha$-module. Now we can apply Theorem 2.24, to conclude that the highest weight $\lambda(h_\alpha)$ is a nonnegative integer for every simple root $\alpha \in \Delta$. □

Theorem 2.24 also implies that all the weights of an $\mathfrak{s}_\alpha$-module satisfy $\mu(h_\alpha) \in \mathbb{Z}$, combining this with the expression for $h_\alpha$ in (3.1) and using the definitions of the discussions from Section 3.1, we obtain

$$\begin{aligned} \mu(h_\alpha) &= \kappa(t_\mu, h_\alpha) = \kappa\left(t_\mu, \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}\right) \\ &= 2\frac{\kappa(t_\mu, t_\alpha)}{\kappa(t_\alpha, t_\alpha)} = 2\frac{(\mu, \alpha)}{(\alpha, \alpha)} = \langle \mu, \alpha \rangle. \end{aligned} \tag{5.7}$$

So, $\langle \mu, \alpha \rangle \in \mathbb{Z}$ for each $\alpha \in \Delta$. Leaving out the details (see Humphreys [8], Sec. 13) this yields a *lattice* as in the following example for $\mathfrak{sl}(3, \mathbb{C})$ shown in Figure 8.

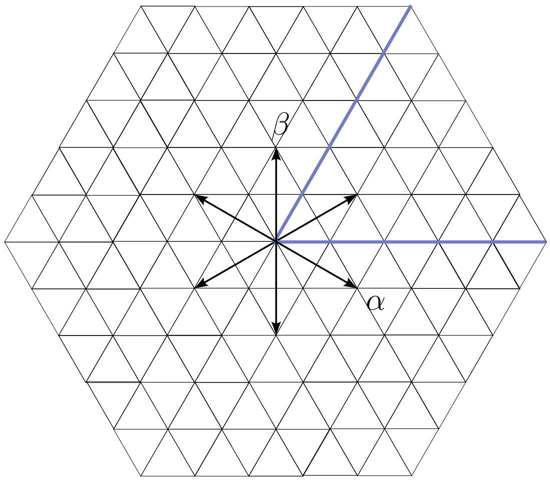



Figure 8: Weight lattice $\mathfrak{sl}(3, \mathbb{C})$, where the weights are given by the corners of the triangles.

In Theorem 5.10 we have a weight $\lambda$ where $\lambda(h_\alpha)$ is a nonnegative integer for all $\alpha \in \Delta$, we will call such weights **dominant integral**. In light of (5.7), for any $\alpha \in \Delta$, the dominant integral weights $\lambda \in \mathfrak{h}^*$ yield $\langle \lambda, \alpha \rangle \geq 0$, but then $(\lambda, \alpha) \geq 0$. Remember that the closure of the Fundamental Weyl chamber was given by

$$\overline{\mathfrak{C}(\Delta)} = \{x \in E : (x, \alpha) \geq 0 \text{ for all } \alpha \in \Delta\}.$$

So, all the weights in $\overline{\mathfrak{C}(\Delta)}$ are in particular dominant integral weights. This will return in the proof of the following theorem, to prove the collection of weights of $V(\lambda)$, which we denote by $\Pi(\lambda)$, is finite.

**Theorem 5.11.** *If $\lambda \in \mathfrak{h}^*$ is dominant integral, then the irreducible $\mathfrak{g}$-module $V(\lambda)$ is finite-dimensional.*

The proof for this direction will be a lot more involved than the converse, so it has been split smaller parts. Because the highest weight of the $V(\lambda)$ we are considering is dominant integral, we can call to the geometry established for root systems as in the above discussion. This motivates the main idea for the proof, which is to use the Weyl group to show the weights of $V(\lambda)$ are of finite amount.

Before we reach that point, we prove two lemmas that help in ensuring any vector in $V(\lambda)$ is contained in a sum of finite-dimensional $\mathfrak{s}_\alpha$-invariant subspaces. This allows us to use a few results of $\mathfrak{sl}(2,\mathbb{C})$-modules in this context, which are necessary for proving the invariance of the weights when applying Weyl group permutations on them.

*Proof of Theorem 5.11.* Let $v_0 \in V(\lambda)$ be the maximal vector of highest weight $\lambda \in \mathfrak{h}^*$ and let $m$ denote the highest weight in $\mathfrak{s}_\alpha$ for each simple root $\alpha$, so $m := \lambda(h_\alpha)$. Per assumption $m \in \mathbb{Z}_{\geq 0}$ as $\lambda$ is dominant integral.

**Lemma 5.12.** *If $\lambda$ is dominant integral and $v_0$ a maximal vector of $V(\lambda)$, then the subspace $W$ with ordered basis $(v_0, y_\alpha \cdot v_0, \ldots, y_\alpha^m \cdot v_0)$ is an $\mathfrak{s}_\alpha$-invariant subspace.*

*Proof.* We will consider the vectors

$$v_k := y_\alpha^k \cdot v_0, \quad \text{for } k \geq 0,$$

as in Lemma 2.25 for the $\mathfrak{sl}(2,\mathbb{C})$ case work towards showing their span is an $\mathfrak{s}_\alpha$-invariant subspace. So, first we examine the vector $v_{m+1} = y_\alpha^{m+1} \cdot v_0$. For positive roots $\beta \in \Phi^+$ with $\beta \neq \alpha$, we will show that their root spaces annihilate $v_{m+1}$. Let $x_\beta \in \mathfrak{g}_\beta$, if $x_\beta \cdot v_{m+1} \neq 0$ the vector would be a weight vector of weight $\mu = \lambda - (m+1)\alpha + \beta$ , by Lemma 5.1. But then $\lambda - \mu = (m+1)\alpha - \beta$ where $\alpha, \beta \in \Phi^+$, while now $\lambda - \mu$ is not a linear combination of simple roots with solely nonnegative coefficients. So $\lambda - \mu \not\succ 0$ and therefore $\lambda \not\succ \mu$, but this contradicts with $\lambda$ being the highest weight of $V(\lambda)$. Thus $x_\beta \cdot v_{m+1} = 0$ and all positive root spaces, excluding $\mathfrak{g}_\alpha$, annihilate the vector $v_{m+1}$. For the root $\alpha$ we need a different approach, since the weight would otherwise be given by $\mu = \lambda - m\alpha$, which does not cause the same problem necessary for the previous argument. Now, because we are working within an $\mathfrak{sl}(2,\mathbb{C})$-isomorphic submodule for $x_\alpha$, the results in Lemma 2.25 can be used here,

$$x_\alpha \cdot v_k = k(m - (k-1))v_{k-1}.$$

So $x_\alpha$ kills the vector $v_{m+1}$, as

$$x_\alpha \cdot v_{m+1} = (m+1)(m-m)v_m = 0.$$

Hence all positive root spaces annihilate $v_{m+1}$. From this we can conclude that $v_{m+1}$ has to either be zero or a maximal vector of $V(\lambda)$. Suppose it were a maximal vector, then $v_{m+1}$ has weight $\lambda - (m+1)\alpha \neq \lambda$ which goes against the property of $V(\lambda)$ having a unique maximal vector, thus $v_{m+1} = 0$.

Subsequently, the subspace $W$ spanned by the vectors $v_0, v_1, \ldots, v_m$ is a nonzero $\mathfrak{s}_\alpha$-invariant subspace. Since $x_\alpha \cdot v_0 = 0$ and $y_\alpha \cdot v_m = 0$, as shown above, and the rest is due to the relations in Lemma 2.25 still being applicable here, which tell us that every $\mathfrak{s}_\alpha$-action sends the vectors in $W$ back into $W$. □

It has now been established that there exists a nonzero $\mathfrak{s}_\alpha$-invariant subspace. Remarkable is that it is done via a subspace similar to one we have seen for $\mathfrak{sl}(2,\mathbb{C})$. The next step is to show that $V(\lambda)$ is a sum of finite dimensional $\mathfrak{s}_\alpha$-invariant subspaces.

**Lemma 5.13.** *If $\lambda$ is dominant integral, then $V(\lambda)$ is equal to the sum of finite-dimensional $\mathfrak{s}_\alpha$-invariant subspaces for any fixed $\alpha \in \Delta$.*

*Proof.* Let $T_\alpha$ be the sum of all finite-dimensional $\mathfrak{s}_\alpha$-invariant subspaces, which by Lemma 5.12 at least contains $W \neq 0$ as defined in the lemma. We will show that $T_\alpha$ is invariant under the action of $\mathfrak{g}$, making it a nonzero $\mathfrak{g}$-submodule of $V(\lambda)$ and thus by irreducibility $T_\alpha = V(\lambda)$, proving the lemma. Fix a vector $v \in T_\alpha$ and an arbitrary Lie algebra element $x \in \mathfrak{g}$; we have to prove that $x \cdot v$ is again in $T_\alpha$. Let $S \subset T_\alpha$ be a finite-dimensional $\mathfrak{s}_\alpha$-invariant subspace containing $v$. We then define $S'$ to be the span of all vectors having the form $z \cdot w$, with $z \in \mathfrak{g}$ and $w \in S$. This subspace is finite-dimensional as $\dim S' \leq (\dim \mathfrak{g}) \cdot (\dim S)$. Now for $s \in \mathfrak{s}_\alpha$ we find that

$$s \cdot (z \cdot w) = z \cdot (s \cdot w) + [s, z] \cdot w \in S',$$

where we used that $s \cdot w \in S$ and $[s, z] \in \mathfrak{g}$ , to conclude that both terms of the sum are in the desired form for $S'$. So $S'$ is also a finite-dimensional $\mathfrak{s}_\alpha$-invariant subspace and therefore contained in $T_\alpha$. But as $x \cdot v \in S'$, it has to be in $T_\alpha$, which completes the proof. □

Now, the role of the Weyl group will become clear in the following lemma. For this we have to define what the action of the Weyl group on the weights $\mu \in \Pi(\lambda)$ is. Since the Weyl group is defined as the group generated by all reflections along the roots, we only have to specify how these reflections act on $\mu$. Let $\sigma_\alpha$ be a reflection along $\alpha \in \Phi$, then we define the action as $\sigma_\alpha(\mu) := \mu - \mu(h_\alpha)\alpha$.

**Lemma 5.14.** *If $\lambda$ is dominant integral, then the set of weights $\Pi(\lambda)$ of $V(\lambda)$ is invariant under the action of the Weyl group. So, for any $\mu \in \Pi(\lambda)$ and $\sigma \in \mathcal{W}$, we have $\sigma(\mu) \in \Pi(\lambda)$.*

This proof will follow Elduque's ([4], Theorem 4.1) treatment, in contrast to the standard proof that requires us to work with exponentials of $x_\alpha$ and $y_\alpha$ in some representation.

*Proof.* The Weyl group is generated by the simple reflections $\sigma_\alpha$ where $\alpha \in \Delta$, by Theorem 3.17. So, we only have to show that the set of weights $\Pi(\lambda)$ is invariant under simple reflections. To this end, we let $\alpha \in \Delta$ and let $\mu \in \Pi(\lambda)$ be a weight with a nonzero weight vector $v \in V(\lambda)$ and show that we can obtain $\sigma_\alpha(\mu)$ as a weight.

Lemma 5.13 ensures that there is a sum of finite-dimensional $\mathfrak{s}_\alpha$-invariant subspaces containing the vector $v$, then we apply Weyl's theorem to decompose the sum further into a direct sum of (irreducible) $\mathfrak{s}_\alpha$-invariant subspaces. So, $v \in W_1 \oplus \cdots \oplus W_n$ where each $W_j$ is a finite-dimensional $\mathfrak{s}_\alpha$-invariant subspace. Now we write

$$v = w_1 + \cdots + w_n, \text{ with } w_j \in W_j \text{ for all } j,$$

and assume that $w_1 \neq 0$ as $v$ itself is nonzero, so there must be one term that is also nonzero. Since $h_\alpha \cdot v = \mu(h_\alpha)v$, we have

$$h_\alpha \cdot w_1 + \cdots + h_\alpha \cdot w_n = \mu(h_\alpha)w_1 + \cdots + \mu(h_\alpha)w_n,$$

in particular this means that $h_\alpha \cdot w_1 = \mu(h_\alpha)w_1$, so $w_1$ is a weight vector with weight $\mu(h_\alpha)$. By Theorem 2.27 for finite $\mathfrak{sl}(2,\mathbb{C})$-modules, $-\mu(h_\alpha)$ is also a weight of $h_\alpha$ on $W_1$.

If we can somehow lower the weight $\mu$ of the vector $v$ to $\mu - \mu(h_\alpha)\alpha$ by applying Lemma 5.1 without killing $v$, then we have succesfully shown that $\sigma_\alpha(\mu) = \mu - \mu(h_\alpha)\alpha$ is a weight as well. Fortunately, this is possible through $w_1$, though we need to do a case analysis for this. Firstly, if $\mu(h_\alpha) \geq 0$, then by Lemma 2.23 the vector $y_\alpha^{\mu(h_\alpha)} \cdot w_1$ is a weight vector, hence nonzero, with weight $-\mu(h_\alpha)$. So, now applying Lemma 5.1 we obtain the following, $y_\alpha^{\mu(h_\alpha)} \cdot v \in V(\lambda)_{\mu-\mu(h_\alpha)\alpha} = V(\lambda)_{\sigma_\alpha(\mu)}$ as a nonzero weight space. On the other hand, if $\mu(h_\alpha) < 0$, then we let $x_\alpha^{-\mu(h_\alpha)}$ act on $v$, via the same reasoning as for $y_\alpha^{\mu(h_\alpha)}$ to obtain $0 \neq x_\alpha^{-\mu(h_\alpha)} \cdot v \in V(\lambda)_{\sigma_\alpha(\mu)}$. Thus $\sigma_\alpha(\mu)$ is also a weight of $V(\lambda)$ and we can conclude that the set of weights $\Pi(\lambda)$ is invariant under all simple reflections $\sigma_\alpha$, hence also under the action of the whole Weyl group. □

There is still one more step remaining until we can prove the finite-dimensionality of $V(\lambda)$, namely to prove that there are only a finite amount of weights in $\Pi(\lambda)$.

**Lemma 5.15.** *If $\lambda$ is dominant integral, then the collection of weights $\Pi(\lambda)$ of the module $V(\lambda)$ is finite.*

*Proof.* We can restrict our view solely to weights in the closure of the fundamental Weyl chamber. Because, Theorem 3.19 states that for each weight $\mu \in \Pi(\lambda)$ there exists exactly one vector $y \in \overline{\mathfrak{C}(\Delta)}$ to which $\mu$ can be sent to by $\mathcal{W}$, and Lemma 5.14 ensures that this $y$ must be a weight in $\Pi(\lambda)$ as well, so in particular $y \prec \lambda$ as $\lambda$ is the highest weight. Hence, by the finiteness of the Weyl group there are only a finite amount of weights that are sent to this particular weight $y$. So, if the weights in $C^+ := \overline{\mathfrak{C}(\Delta)} \cap \Pi(\lambda)$ are of finite amount, then $\Pi(\lambda)$ is finite.

Now, the weights that are in the closure of the fundamental Weyl chamber are exactly the dominant integral weights. We claim that there are only finitely many dominant integral weights $\mu \in C^+$.

Indeed, let $\mu \in C^+$ so that $\mu \prec \lambda$. Then, as both $\mu$ and $\lambda$ are dominant integral, their sum $\lambda + \mu$ will satisfy $(\lambda+\mu)(h) \geq 0$ for all $h \in \mathfrak{h}$ and is hence also dominant integral. Per definition of the relation $\mu \prec \lambda$, we have that $\lambda - \mu = \sum_{\alpha\in\Delta} k_\alpha \alpha$ where each $k_\alpha \geq 0$. Now, as $\lambda + \mu$ is in $C^+ \subset \overline{\mathfrak{C}(\Delta)}$, their angles between all the simple roots will be acute or straight, so

$$(\lambda,\lambda) - (\mu,\mu) = (\lambda+\mu, \lambda-\mu) = \sum_{\alpha\in\Delta} k_\alpha(\lambda+\mu, \alpha) \geq 0.$$

However, now $\mu$ is an element of the compact set $\{x \in E : (x,x) \leq (\lambda,\lambda)\}$ and also of $C^+$ which is discrete, while the intersection of a discrete set and a compact set is finite, proving the claim. This implies that $C^+$ is finite and hence $\Pi(\lambda)$ is also finite. □

Finally, as part *(3)* of Theorem 5.5 implies that all weight spaces of the standard cyclic module $V(\lambda)$ are finite-dimensional, we can combine it with finiteness of $\Pi(\lambda)$ to obtain our desired conclusion. The module $V(\lambda)$ is a finite sum of finite-dimensional weight spaces, therefore it is a finite-dimensional module. □

Now that both Theorem 5.10 and Theorem 5.11 have been proven, we have established for irreducible modules that finite-dimensional implies that the highest weight is dominant integral and conversely that a dominant integral highest weight implies the modules is finite-dimensional. Hence, we have the following one to one correspondence (up to module isomorphisms), where we denote $\Lambda^+$ for the collection of all dominant integral weights,

$$\begin{aligned}\Lambda^+ &\to \{\text{finite-dimensional irreducible modules}\}\\ \lambda &\mapsto V(\lambda).\end{aligned}$$

Finally, Weyl's Theorem (Theorem 2.18) allows us to decompose any finite-dimensional module of semisimple Lie algebras to direct sums of irreducible modules. Thus, this bijection fully concludes the classification problem for finite-dimensional modules over semisimple Lie algebras.

# References

[1] Nicolas Bourbaki. *Elements of Mathematics, Algebra I.* Springer, 1998.

[2] Keith Conrad. Tensor products. https://kconrad.math.uconn.edu/blurbs/linmultialg/tensorprod.pdf.

[3] Jacques Dixmier. *Enveloping algebras.* North-Holland mathematical library ; v. 14 (GSM/11). North-Holland (Elsevier) 1977 /AMS (1996), 1977 /1996.

[4] Alberto Elduque. Lie algebras course notes. https://personal.unizar.es/elduque/files/LAElduque.pdf, 2005-2021. Universidad de Zaragoza.

[5] Karin Erdmann and Mark J. Wildon. *Introduction to Lie Algebras.* SUMS. Springer London, 2006.

[6] William Fulton and Joe Harris. *Representation Theory*, volume 129 of *GTM.* Springer New York, 2004.

[7] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations*, volume 222 of *GTM.* Springer International Publishing, 2 edition, 2015.

[8] James E. Humphreys. *Introduction to Lie Algebras and Representation Theory*, volume 9 of *GTM.* Springer New York, 1972.

[9] Kyoshi Igusa. Math 223a notes 2011 lie algebras. https://people.brandeis.edu/~igusa/Math223aF11/Notes223a.pdf, 2011. Brandeis University.

[10] Jr. Kirillov, Alexander. *Introduction to Lie Groups and Lie Algebras*, volume 113 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, 2010.

[11] Christian Schnell. Simple roots and the weyl group. https://www.math.stonybrook.edu/~cschnell/mat552/lecture-april-22.pdf, April 2020. Lecture slides.

[12] Josh Swanson. The pbw theorem. https://www.jpswanson.org/talks/2015_Humphreys_PBW.pdf, July 2015.

[13] Wikipedia, the free encyclopedia. Weight (representation theory). https://commons.wikimedia.org/wiki/File:Illustration_of_notion_of_%22higher%22_for_root_systems.png, 2017. [Online; accessed July 7, 2023].