The Convex Hull of the Highest Weight Orbit and the Carathéodory Orbitope

Nigel Redding

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the degree of
Master of Science in Mathematics[1]

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

[1]The M.Sc. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

Abstract

In this thesis, we study the polynomial equations that describe the highest weight orbit of an irreducible finite dimensional highest weight module under a semisimple Lie group. We also study the connection of the convex hull of this orbit and the Carathéodory orbitope.

Dedications

To my parents.

Acknowledgement

I would like to thank my supervisor Dr. Hadi Salmasian for all his guidance during the writing of this thesis. I have learned a great deal under his supervision. I would also like to thank my parents for being supportive when I was writing my thesis.

Contents

Introduction

Since its birth, Lie theory has always had a strong connection to geometry. One of the geometric structures of particular interest in Lie theory is the construction of a homogeneous space. The study of homogeneous spaces (or, equivalently, orbits of group actions) is important because they naturally arise in applications of Lie theory such as invariant theory and quantization. In this thesis, we consider questions related to the structure of the orbit of the highest weight vector of an irreducible representation of a real semisimple Lie algebra from the view points of algebraic and convex geometry.

Roughly speaking, the setting of this thesis is as follows. Let G be a real split semisimple Lie group, and let V_λ be a finite dimensional irreducible representation of G of highest weight λ. Let $v_\lambda \in V_\lambda$ be a highest weight vector. Set $\mathcal{O} = G \cdot v_\lambda$ and let $\mathcal{O}_\mathbb{C}$ be the orbit of v_λ under the action of the complexification $G_\mathbb{C}$ of G inside the corresponding complex highest weight module. Then $\mathcal{O}_\mathbb{C} \cup \{0\}$ is indeed a complex algebraic variety, and therefore $\mathcal{O} \cup \{0\}$ is a semialgebraic set. From the Tarski-Seidenberg Theorem (see Chapter 1), it follows that $\mathrm{conv}(\mathcal{O} \cup \{0\})$ is a semialgebraic set, and therefore it is described by a system of polynomial inequalities. While there is a constructive proof of the Tarski-Seidenberg theorem which provides an algorithm to describe the constraints of $\mathrm{conv}(\mathcal{O} \cup \{0\})$, this algorithm turns out to be very inefficient, and it is rather difficult to gain insight into the structure of this set from this algorithm alone. Instead, we focus on techniques from representation theory. Our main goal in this thesis, which is achieved in Chapter 5, is to obtain a concrete description of the latter semi-algebraic set.

We now elaborate on the content of every chapter. The first chapter outlines the algebraic objects used in this thesis, some results from real algebraic geometry, and finally some basic notions about Lie algebras. In Section 1.1, we define the tensor algebra, symmetric algebra, and the space of symmetric tensors. These are used extensively in this thesis as representation spaces for Lie groups and Lie algebras. Following this, in Section 1.2, we discuss some basic notions from real algebraic geometry. We define the notion of a semialgebraic set, and prove that the convex hull of a semialgebraic set is semialgebraic. In this proof, we use the celebrated Tarski-Seidenberg theorem, which states that the projection of a semialgebraic set is semialgebraic. This result is the starting point of our understanding of the highest weight orbitope, which appears in Chapter 5. Finally, in Section 1.3, we discuss some basic Lie theory, such as the notion of the universal enveloping algebra, the Casimir operator, and the Killing form, all of which play an important role in later chapters.

In the second chapter, we discuss some results about the structure theory of real semisimple Lie algebras and Lie groups. In Section 2.1, we discuss the structure of real semisimple Lie algebras. The most important notions from this section include the Iwasawa decomposition for Lie algebras and the notion of a Cartan subalgebra. In Section 2.2, we consider the Cartan and Iwasawa decompositions in the group setting. In Section 2.3, we illustrate the theory from the previous two sections using the examples of $\mathfrak{sl}_2(\mathbb{R})$ and $\mathrm{SL}_2(\mathbb{R})$.

In the third chapter, we discuss the representation theory of Lie groups and Lie algebras. Section 3.1 is devoted to some generalities and elementary facts. In Section 3.2, we classify the finite dimensional representations of $\mathrm{SL}_2(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{R})$. There are some subtleties associated with classifying the real representations of $\mathfrak{sl}_2(\mathbb{R})$, as opposed to the complex representations, which arise due to the fact that $\mathbb{R}$ is not algebraically closed. Standard references such as Humphrey's book [10] only consider the classification problem over algebraically closed base fields, which is why we include the details for the real case. Finally, in Section 3.3, we classify the real finite dimensional representations of $\mathrm{SO}_2(\mathbb{R})$.

In the fourth chapter, we concentrate on the highest weight orbit $\mathcal{O}$. In Section 4.1, we restrict our attention to the complex orbit $\mathcal{O}_{\mathbb{C}}$. We show that $\mathcal{O}_{\mathbb{C}} \cup \{0\}$ is a complex algebraic variety. We also show that $\mathcal{O}_{\mathbb{C}}$ is the only orbit of $G_{\mathbb{C}}$ acting on V_λ with the latter property. In Section 4.2, we consider $X_\lambda := \mathcal{O} \cup \{0\}$. We prove a modified version of *Kostant's theorem*. Kostant originally proved that for a complex simple Lie group G, the orbit $\mathcal{O}_{\mathbb{C}} \cup \{0\}$ is described by a set of quadratic equations. In this thesis, we prove that for $G = \mathrm{SL}_n(\mathbb{R})$, X_λ is a semialgebraic set, given by the intersection of variety determined by quadratic equations and a specific semialgebraic set E_λ. We focus on $\mathrm{SL}_n(\mathbb{R})$, but our proof works (with some modification) for any real split semisimple Lie group. The statment of this theorem (Theorem 4.2.3) explicitly describes the constraints on X_λ, and we explicitly give the equations for the 5-dimensional representation of $\mathrm{SL}_2(\mathbb{R})$ (Example 4.2.14). Our version of Kostant's theorem is original.

In the fifth chapter, we introduce convex hulls of orbits. In Chapter 4, we proved that X_λ is semialgebraic and we described its constraints. It is thus natural to ask whether we can do the same for $\mathrm{conv}(X_\lambda)$, or at least describe $\mathrm{conv}(X_\lambda)$ in a more explicit way. For this, we need to focus on the theory of convex hulls of orbits of compact groups, known as *orbitopes*. In [20], there is an extensive discussion of of $\mathrm{SO}_2(\mathbb{R})$ orbitopes. In this paper, the authors give an explicit description of the orbitopes of $\mathrm{SO}_2(\mathbb{R})$ as a *spectrahedra*, i.e. sets in $\mathbb{R}^n$ which are described by a linear matrix inequality. In this chapter we obtain results about G-orbits similar to those in [20], where $G = \mathrm{SL}_2(\mathbb{R})$ (the difference between the work here and that in [20] is that G is non-compact). To this end, we first observe that the G-orbit $\mathcal{O}$ contains a K-orbit in a natural way, where K is a maximal compact subgroup of G. In Section 5.1, we describe the convex hull of X_λ in terms of this K-orbit. Finally, in Section 5.2, we prove that in the case of $G = \mathrm{SL}_2(\mathbb{R})$, the convex hull of X_λ can be described as a cone over a certain orbitope, which is known as the Carathéodory orbitope (Theorem 5.2.15). This work is original.

There are a number of questions which naturally arise after the work done in this thesis. In particular, the following questions are interesting. Does a version of Kostant's theorem hold when G is compact? Do the combinatorial results of [20] extend to groups which are *not* compact? Given a general semialgebraic set, is there an efficient way to construct the constraints for the cone over that set? More specifically, is it possible to efficiently construct the constraints for the cone $\mathrm{conv}(X_\lambda) = \mathbb{R}^+ \mathrm{conv}(K \cdot v_\lambda)$?

Chapter 1

Preliminaries

1.1 Basic Algebra

In this section, we review some of the basic algebraic objects used in this thesis. Unless stated otherwise, the ground field is a field $\mathbb{F}$ of characteristic 0.

Definition 1.1.1. Let V be a vector space over a field $\mathbb{F}$. For each $k \geq 0$, we define

$$\mathcal{T}^k(V) := \underbrace{V \otimes \cdots \otimes V}_{k\text{-times}}.$$

Note that $\mathcal{T}^0(V) = \mathbb{F}$ and $\mathcal{T}^1(V) = V$. Then we define the *tensor algebra* of V to be the algebra $\mathcal{T}(V)$ defined by

$$\mathcal{T}(V) := \bigoplus_{k=0}^{\infty} \mathcal{T}^k(V)$$

where multiplication is defined by

$$(v_1 \otimes \cdots \otimes v_r) \cdot (w_1 \otimes \cdots \otimes w_s) := v_1 \otimes \cdots \otimes v_r \otimes w_1 \otimes \cdots \otimes w_s$$

for every $r, s \geq 1$, $v_1 \otimes \cdots \otimes v_r \in \mathcal{T}^r(V)$, and $w_1 \otimes \cdots \otimes w_s \in \mathcal{T}^s(V)$. We extend this multiplication linearly to all of $\mathcal{T}(V)$.

Definition 1.1.2. Let V be a vector space over a field $\mathbb{F}$. Let $\mathcal{I}$ be the ideal of $\mathcal{T}(V)$ generated by the set $\{v \otimes w - w \otimes v : v, w \in V\}$. Then we define the *symmetric algebra* $\mathcal{S}(V)$ to be the quotient algebra

$$\mathcal{S}(V) := \mathcal{T}(V)/\mathcal{I}.$$

Note that $\mathcal{I}$ is a homogeneous ideal, and therefore

$$\mathcal{T}(V)/\mathcal{I} = \bigoplus_{k=0}^{\infty} \mathcal{T}^k(V)/(\mathcal{I} \cap \mathcal{T}^k(V)).$$

We denote the k-th graded component of $\mathcal{S}(V)$ by $\mathcal{S}^k(V)$. We remark that

$$\mathcal{S}^k(V) \cong \mathcal{T}^k(V)/(\mathcal{I} \cap \mathcal{T}^k(V)).$$

For an element $v_1 \otimes \cdots \otimes v_k \in \mathcal{T}^k(V)$, we denote its image in $\mathcal{S}(V)$ by

$$v_1 \cdots v_k. \tag{1.1.1}$$

Thus, all elements of $\mathcal{S}(V)$ are linear combinations of elements of the form (1.1.1). We observe that the image of $\mathcal{T}^k(V)$ in $\mathcal{S}(V)$ is $\mathcal{S}^k(V)$.

Let $k \geq 0$. Let S_k denote the permutation group on $\{1, 2, \ldots, k\}$. Then there is an action of S_k on $\mathcal{T}^k(V)$ defined by

$$\sigma \cdot v_1 \otimes \cdots \otimes v_k = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)} \qquad \sigma \in S_k, v_1 \otimes \cdots \otimes v_k \in \mathcal{T}^k(V).$$

We extend this action linearly to all of $\mathcal{T}(V)$.

Definition 1.1.3. Let $k \geq 0$. We say a tensor $t \in \mathcal{T}^k(V)$ is a *symmetric k-tensor* if $\sigma \cdot t = t$ for all $\sigma \in S_k$. We let $\mathrm{Sym}^k(V)$ denote the vector space of symmetric k-tensors. We define the vector space $\mathrm{Sym}(V)$ by

$$\mathrm{Sym}(V) := \bigoplus_{k=0}^{\infty} \mathrm{Sym}^k(V).$$

Proposition 1.1.4. *For each $k \geq 0$, there is a vector space isomorphism*

$$\mathcal{S}^k(V) \to \mathrm{Sym}^k(V) \qquad v_1 \cdots v_k \mapsto \frac{1}{k!} \sum_{\sigma \in S_k} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}.$$

These isomorphisms induce a vector space isomorphism

$$\mathcal{S}(V) \to \mathrm{Sym}(V).$$

Proof: See [5, §11.5, Proposition 40]. ∎

1.2 Basic Semi-algebraic Geometry

In this section, we review the basic theory of semi-algebraic sets needed for this thesis. We follow [1, Chapter 2]. The most important result of this section is that the convex hull of a semi-algebraic set is semi-algebraic. This is obtained as a corollary of the Tarski-Seidenberg theorem.

Definition 1.2.1. We say $S \subseteq \mathbb{R}^n$ is a *basic semi-algebraic set* if S has the form

$$S = \bigcap_{i=1}^{k} \{x \in \mathbb{R}^n : f_i(x) *_i 0\}$$

where $f_i \in \mathbb{R}[x_1, \dots, x_n]$ and $*_i$ is either the symbol $=$ or $>$, depending on i. We call $\{(f_1, *_1), \dots, (f_k, *_k)\}$ the *constraints* of S. A *semi-algebraic set* is defined as a finite union of basic semi-algebraic sets.

The following proposition is the celebrated *Tarski-Seidenberg theorem*.

Proposition 1.2.2. *Let S be a semi-algebraic subset of $\mathbb{R}^{n+1}$, and let $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ be the projection*

$$\pi(x_1, \dots, x_n, x_{n+1}) = (x_1, \dots, x_n).$$

Then $\pi(S)$ is a semi-algebraic subset of $\mathbb{R}^n$.

Proof: See [1, Theorem 2.2.1] ∎

Corollary 1.2.3. *Let S be a semi-algebraic subset of $\mathbb{R}^n$. If $1 \leq i_1 < \cdots < i_k \leq n$ are integers, and $\pi : \mathbb{R}^n \to \mathbb{R}^k$ is the projection*

$$\pi(x_1, \dots, x_n) = (x_{i_1}, \dots, x_{i_k}),$$

then $\pi(S)$ is semi-algebraic.

Definition 1.2.4. For a subset S of a real vector space V, we define the *convex hull* of S by

$$\operatorname{conv}(S) := \left\{ t_1 v_1 + \cdots + t_k v_k : k \in \mathbb{N}, t_i \geq 0, \sum t_i = 1, v_i \in S \right\}.$$

Equivalently, $\operatorname{conv}(S)$ is the smallest convex subset of V containing S.

Theorem 1.2.5. *(Carathéodory, 1911) For any set $S \subseteq \mathbb{R}^d$, and any point $x \in \operatorname{conv}(S)$, x is a convex combination of at most $d + 1$ points in S.*

Proof: See [18, §17]. ∎

Recall that for $n \geq 1$, the *n-simplex* Δ^n is the subset of $\mathbb{R}^{n+1}$ defined by

$$\Delta^n := \left\{(t_0, t_1, \ldots, t_n) : t_i \geq 0, \sum t_i = 1\right\}.$$

It is straightforward to verify that Δ^n is semi-algebraic.

Proposition 1.2.6. *Let S be a semi-algebraic subset of $\mathbb{R}^n$. Then the convex hull* $\operatorname{conv}(S)$ *is semi-algebraic.*

Proof: Write $S = S_1 \cup \cdots \cup S_k$ where each S_i is a basic semi-algebraic set. Now define

$$A = \left\{(s_1, \ldots, s_{n+1}, t_1, \ldots, t_{n+1}, \sum_{i=1}^{n+1} t_i s_i) : s_i \in S, (t_1, \ldots, t_{n+1}) \in \Delta^n\right\}.$$

By Corollary 1.2.3 and Theorem 1.2.5, it suffices to show that A is a semi-algebraic subset of $\mathbb{R}^{n^2+3n+1}$.

As Δ^n is semi-algebraic as mentioned, one can write $\Delta^n = D_1 \cup \cdots \cup D_\ell$ where each D_i is a basic semi-algebraic set. Now for each $(n+1)$-tuple $\alpha = (\alpha_1, \ldots, \alpha_{n+1})$ in $\{1, 2, \ldots, k\}^{n+1}$ and each $\beta \in \{1, 2, \ldots, \ell\}$, define

$$A_{\alpha\beta} = \left\{(s_1, \ldots, s_{n+1}, t_1, \ldots, t_{n+1}, \sum_{i=1}^{n+1} t_i s_i) : s_i \in S_{\alpha_i}, (t_1, \ldots, t_{n+1}) \in D_\beta\right\}.$$

Then clearly we have

$$A = \bigcup_\alpha \bigcup_\beta A_{\alpha\beta}.$$

Now we only need to show each $A_{\alpha\beta}$ is semi-algebraic. Let $\{(f^i_\gamma) *^i_\gamma\}$ be the constraints for S_{α_i} and let $\{(g_\mu, *_\mu)\}$ be the constraints for D_β. Then for $v \in \mathbb{R}^{n^2+3n+1}$, denote v by $v = (x_1, \ldots, x_{n+1}, y, z)$ where $x_i \in \mathbb{R}^n$, $y \in \mathbb{R}^{n+1}$ and $z \in \mathbb{R}^n$. Then the constraint polynomials for $A_{\alpha\beta}$ are given by

$$\begin{aligned} h^i_\gamma(v) &:= f^i_\gamma(x_i) \\ k_\mu(v) &:= g_\mu(t_1, \ldots, t_{n+1}) \\ l(v) &:= z - y_1x_1 - y_2x_2 - \cdots - y_{n+1}x_{n+1} \end{aligned}$$

and the constraints are given by $\{(h^i_\gamma, *^i_\gamma)\}$, $\{(k_\mu), *_\mu\}$ and $\{(l, =)\}$. ∎

1.3 Some Facts about Lie Algebras

In this section, we review some facts concerning semisimple Lie algebras which will be used in this thesis. We assume the reader is familiar with the basic theory of Lie algebras. In this section, $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.3.1. A Lie algebra $\mathfrak{g}$ is *semisimple* if $\mathfrak{g}$ has no nonzero solvable ideals.

Remark 1.3.2. Equivalently, one sees that $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g}$ has no nonzero abelian ideals.

Definition 1.3.3. Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{F}$, and let $X \in \mathfrak{g}$. We define the map $\mathrm{ad}(X) : \mathfrak{g} \to \mathrm{End}_{\mathbb{F}}(\mathfrak{g})$ by

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

Alternatively, we sometimes use the notation

$$\mathrm{ad}_X = \mathrm{ad}(X).$$

Definition 1.3.4. Let $\mathfrak{g}$ be a Lie algebra. We define the *Killing form* on $\mathfrak{g}$ by

$$B(X,Y) = \operatorname{tr}(\operatorname{ad}(X) \cdot \operatorname{ad}(Y)) \qquad X, Y \in \mathfrak{g}.$$

Proposition 1.3.5. *The Killing form B satisfies the following properties.*

(1) B is bilinear and symmetric.

(2) B is associative*, i.e. $B([X,Y],Z) = B(X,[Y,Z])$ for all $X, Y, Z \in \mathfrak{g}$.*

(3) B is nondegenerate if and only if $\mathfrak{g}$ is semisimple.

Proof:

(1) The fact that B is bilinear is a straightforward calculation. The fact that B is symmetric follows from the identity $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ for any two linear endomorphisms X and Y.

(2) This is a straightforward calculation.

(3) See [12, Theorem 1.45].

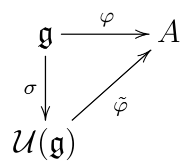
∎

Recall that for an associative $\mathbb{F}$-algebra $\mathcal{A}$, there is an associated Lie algebra structure which has $\mathcal{A}$ as the underlying set and with the Lie bracket defined by

$$[X,Y] = XY - YX \qquad \text{for } X, Y \in \mathcal{A}.$$

Definition 1.3.6. Let $\mathfrak{g}$ be a Lie algebra. A pair $(\mathcal{U}(\mathfrak{g}), \sigma)$, where $\mathcal{U}(\mathfrak{g})$ is a unital associative algebra and $\sigma : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ is a Lie algebra homomorphism is called a *universal enveloping algebra* if for every unital associative algebra $\mathcal{A}$, and for every

Lie algebra homomorphism $\varphi : \mathfrak{g} \to \mathcal{A}$, there exists a unique homomorphism of unital associative algebras $\tilde{\varphi} : \mathcal{U}(\mathfrak{g}) \to \mathcal{A}$ such that $\varphi = \tilde{\varphi} \circ \sigma$. This is expressed by the following commutative diagram.



Remark 1.3.7. (Uniqueness of universal enveloping algebra) Let $(\mathcal{U}(\mathfrak{g}), \sigma)$ and $(\tilde{\mathcal{U}}(\mathfrak{g}), \tilde{\sigma})$ be two universal enveloping algebras of $\mathfrak{g}$. Then there exists an isomorphism $\varphi : \mathcal{U}(\mathfrak{g}) \to \tilde{\mathcal{U}}(\mathfrak{g})$ of unital associative algebras satisfying $\varphi \circ \sigma = \tilde{\sigma}$. See [9, Lemma 7.1.2].

We now provide a construction of the universal enveloping algebra $(\mathcal{U}(\mathfrak{g}), \sigma)$. Let $\mathcal{T}(\mathfrak{g})$ be the tensor algebra of $\mathfrak{g}$), i.e.

$$\mathcal{T}(\mathfrak{g}) = \mathbb{F} \oplus \mathfrak{g} \oplus (\mathfrak{g} \otimes \mathfrak{g}) \oplus \cdots .$$

Then let $\mathcal{J}$ be the ideal generated by the set

$$\{X \otimes Y - Y \otimes X - [X, Y] : X, Y \in \mathfrak{g}\}.$$

Then define $\mathcal{U}(\mathfrak{g}) = \mathcal{T}(\mathfrak{g})/\mathcal{J}$, and define $\sigma : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ by $\sigma(X) = X + \mathcal{J}$. See [9, Proposition 7.1.3] for the proof that this is indeed a universal enveloping algebra of $\mathfrak{g}$.

Theorem 1.3.8. *(Poincare-Birkhoff-Witt Theorem (PBW)) Let $\mathfrak{g}$ be a Lie algebra and let $\{X_1, \ldots, X_n\}$ be a basis of $\mathfrak{g}$. Then the set*

$$\{X_1^{\mu_1} \cdots X_n^{\mu_n} : \mu_i \geq 0\}$$

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

Proof: See [9, Theorem 7.1.9]. ∎

Definition 1.3.9. Let $\mathfrak{g}$ be a semisimple Lie algebra, and let $X_1, \ldots, X_n$ be a basis of $\mathfrak{g}$. Let $X^1, \ldots, X^n$ be another basis of $\mathfrak{g}$ which satisfies

$$B(X_i, X^j) = \delta_{ij}.$$

Then define the element

$$\mathbf{C} := \sum_{i=1}^{n} X_i X^i \in \mathcal{U}(\mathfrak{g}).$$

We call this the *Casimir element* of $\mathfrak{g}$.

Remark 1.3.10. The Casimir element is unique. That is to say, it does not depend on the choice of basis $\{X_1, \ldots, X_n\}$. See [21, Chapter 6, §3].

Chapter 2

Real Semisimple Structure Theory

2.1 Advanced Theory of Lie Algebras

In this section, we compile some of the more advanced definitions and results from the theory of Lie algebras which will be used in this thesis. In this section, we assume that the base field is $\mathbb{F} = \mathbb{R}$. Our main reference is [12].

Definition 2.1.1. Let $\mathfrak{g}$ be a real Lie algebra, and let $\theta : \mathfrak{g} \to \mathfrak{g}$ be a Lie algebra automorphism. Assume θ satisfies

(1) $\theta^2 = \mathrm{id}$, and

(2) the bilinear form B_θ on $\mathfrak{g}$ given by

$$B_\theta(X, Y) = -B(X, \theta(Y)) \tag{2.1.1}$$

is positive definite.

Then we say θ is a *Cartan involution* of $\mathfrak{g}$.

Remark 2.1.2. The bilinear from B_θ on $\mathfrak{g}$ is symmetric. Moreover, if $\mathfrak{g}$ is semisimple, then the non-degeneracy of B (from Cartan's criterion) implies that B_θ is nondegenerate as well.

Let $\mathfrak{g}$ be a real Lie algebra and let θ be a Cartan involution of $\mathfrak{g}$. Since $\theta^2 = \text{id}$, the eigenvalues of θ are exactly $+1$ and -1. Therefore, $\mathfrak{g}$ decomposes as

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$$

where

$$\mathfrak{k} = \{X \in \mathfrak{g} : \theta(X) = X\} \text{ and } \mathfrak{p} = \{X \in \mathfrak{g} : \theta(X) = -X\}.$$

We call this the *Cartan decomposition* of $\mathfrak{g}$. We sometimes refer to this as the *polar decomposition*.

The following result has significant importance in the theory of real semisimple Lie algebras.

Proposition 2.1.3. *Let $\mathfrak{g}$ be a real semisimple Lie algebra. Then $\mathfrak{g}$ has a Cartan involution θ.*

Proof: See [12, Corollary 6.18] ∎

Corollary 2.1.4. *Every real semisimple Lie algebra admits a Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$.*

Remark 2.1.5. Let $\mathfrak{g}$ be a real semisimple Lie algebra with a Cartan involution θ and Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Let B_θ be the bilinear form on $\mathfrak{g}$ given as above in (2.1.1). By Definition 2.1.1 we see that B_θ is an inner product on $\mathfrak{g}$. Henceforth, all references to the notion of orthogonality and adjunction are with respect to B_θ.

Lemma 2.1.6. *Let* $\mathfrak{g}$ *be a real semisimple Lie algebra, and let* θ *and* B_θ *be as above. Then*

$$\operatorname{ad}(X)^* = -\operatorname{ad}(\theta(X)) \text{ for all } X \in \mathfrak{g}.$$

Proof: The following proof is essentially from [12, Lemma 6.27]. Let $X, Y, Z \in \mathfrak{g}$. Then

$$\begin{aligned}
B_\theta(\operatorname{ad}(X)^*Y, Z) = B_\theta(Y, \operatorname{ad}(X)Z) &= -B(Y, \theta(\operatorname{ad}(X)Z)) \\
&= -B(Y, \theta([X, Z]) \\
&= -B(Y, [\theta(X), \theta(Z)]) \\
&= -B([Y, \theta(X)], \theta(Z)) \\
&= B([\theta(X), Y], \theta(Z)) \\
&= B(\operatorname{ad}(\theta(X))Y, \theta(Z)) \\
&= -B_\theta(\operatorname{ad}(\theta(X))Y, Z)
\end{aligned}$$

Since $\mathfrak{g}$ is semisimple, B is nondegenerate and therefore B_θ is nondegenerate as well. Thus $\operatorname{ad}(X)^* = -\operatorname{ad}(\theta(X))$. ∎

Let $\mathfrak{a}$ be a maximal abelian subspace of $\mathfrak{p}$. We know this exists since $\mathfrak{p}$ is finite dimensional. By the above lemma, the set

$$\mathcal{F} = \{\operatorname{ad}(H) : H \in \mathfrak{a}\}$$

is a commuting family of self-adjoint transformations on $\mathfrak{g}$. From linear algebra, we know that this family is simultaneously diagonalizable (since the family commutes) with real eigenvalues (since each transformation is self-adjoint).

Let $X \in \mathfrak{g}$ be an eigenvector of the family $\mathcal{F}$. Then let $H, H' \in \mathfrak{a}$ and suppose

H and H' have eigenvalues λ_H and $\lambda_{H'}$, respectively. Also let $\alpha \in \mathbb{R}$. Then we have

$$\mathrm{ad}(\alpha H + H')X = \alpha\,\mathrm{ad}(H)X + \mathrm{ad}(H')X = \alpha\lambda_H X + \lambda_{H'} X$$

so

$$\lambda_{\alpha H + H'} = \alpha\lambda_H + \lambda_{H'}.$$

Hence, our simultaneous eigenvalues are members of the dual space $\mathfrak{a}^*$. For $\lambda \in \mathfrak{a}^*$, we write

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} : \mathrm{ad}(H)X = \lambda(H)X \text{ for all } H \in \mathfrak{a}\}.$$

If $\mathfrak{g}_\lambda \neq 0$ and $\lambda \neq 0$, we call λ a *restricted root* of the pair $(\mathfrak{g}, \mathfrak{a})$. We denote the set of restricted roots by Σ. For $\lambda \in \Sigma$, we call $\mathfrak{g}_\lambda$ a *restricted root space.*

Proposition 2.1.7. *The restricted roots and the restricted root spaces have the following properties:*

(a) With respect to B_θ, $\mathfrak{g}$ is the orthogonal direct sum

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\lambda \in \Sigma} \mathfrak{g}_\lambda.$$

(b) For $\lambda, \mu \in \mathfrak{a}^*$, $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$.

(c) $\theta\mathfrak{g}_\lambda = \mathfrak{g}_{-\lambda}$. Thus, $\lambda \in \Sigma$ implies $-\lambda \in \Sigma$.

(d) $\mathfrak{g}_0 = \mathfrak{a} \oplus \mathfrak{m}$, orthogonally, where $\mathfrak{m} = Z_\mathfrak{k}(\mathfrak{a})$.

Proof:

(a) This follows directly from the above discussion.

(b) Let $\lambda, \mu \in \mathfrak{a}^*, X \in \mathfrak{g}_\lambda$, $Y \in \mathfrak{g}_\mu$ and $H \in \mathfrak{a}$. Then given the Jacobi identity

$$[H, [X, Y]] = -[X, [Y, H]] - [Y, [H, X]]$$

we get

$$\begin{aligned}\mathrm{ad}(H)([X,Y]) &= [H,[X,Y]] \\ &= -[X,[Y,H]] - [Y,[H,X]] \qquad = [X,[H,Y]] + [[H,X],Y] \\ &= [X,\mu(H)Y] + [\lambda HX,Y] \\ &= \mu(H)[X,Y] + \lambda(H)[X,Y] \\ &= (\lambda+\mu)(H)[X,Y]\end{aligned}$$

(c) Let $X \in \mathfrak{g}_\lambda$ and $H \in \mathfrak{a}$. Then

$$\begin{aligned}\mathrm{ad}(H)(\theta(X)) &= [H,\theta(X)] \\ &= \theta[\theta(H),X] \\ &= -\theta[H,X] \text{ since } H \in \mathfrak{a} \subseteq \mathfrak{p} \\ &= -\lambda(H)\theta(X)\end{aligned}$$

(d) First note that $\mathfrak{g}_0 = (\mathfrak{k}\cap\mathfrak{g}_0)\oplus(\mathfrak{p}\cap\mathfrak{g}_0)$. It then suffices to show that $\mathfrak{k}\cap\mathfrak{g}_0 = Z_{\mathfrak{k}}(\mathfrak{a})$ and $\mathfrak{p}\cap\mathfrak{g}_0 = \mathfrak{a}$.

Showing $\mathfrak{k}\cap\mathfrak{g}_0 = Z_{\mathfrak{k}}(\mathfrak{a})$ is trivial. To show $\mathfrak{p}\cap\mathfrak{g}_0 = \mathfrak{a}$ first note that $\mathfrak{a} \subseteq \mathfrak{p}\cap\mathfrak{g}_0$. Now suppose $\mathfrak{a} \subsetneq \mathfrak{p}\cap\mathfrak{g}_0$. Then there exists $X \in (\mathfrak{p}\cap\mathfrak{g}_0)\setminus\mathfrak{a}$. Then $\mathfrak{a}\oplus\mathbb{F}X$ is a larger abelian subalgebra of $\mathfrak{p}$ than $\mathfrak{a}$, a contradiction.

∎

As in the case of any root system, we choose an ordering on Σ and let Σ^+ be the set of positive roots in Σ. Define

$$\mathfrak{n} = \bigoplus_{\lambda\in\Sigma^+} \mathfrak{g}_\lambda.$$

By Proposition 2.1.7, we see that $\mathfrak{n}$ is indeed a nilpotent subalgebra of $\mathfrak{g}$.

Theorem 2.1.8. *(Iwasawa decomposition for Lie algebras) With the notation as above, the real semisimple Lie algebra $\mathfrak{g}$ admits a vector space decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$.*

Proof: See [12, Proposition 6.43]. ∎

Definition 2.1.9. Let $\mathfrak{g}$ be a Lie algebra. We say a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is a *Cartan subalgebra* of $\mathfrak{g}$ if (1) $\mathfrak{h}$ is nilpotent and (2) $\mathfrak{h} = \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$.

Proposition 2.1.10. *If $\mathfrak{t}$ is a maximal abelian subspace of $\mathfrak{m} = Z_{\mathfrak{k}}(\mathfrak{a})$, then $\mathfrak{h} = \mathfrak{a} \oplus \mathfrak{t}$ is a Cartan subalgebra of $\mathfrak{g}$.*

Proof: See [12, Proposition 6.47]. ∎

Definition 2.1.11. We say a real semisimple Lie algebra $\mathfrak{g}$ is *split* if $\mathfrak{t} = 0$. Alternatively, we say $\mathfrak{h}$ is a *split Cartan subalgebra*.

Definition 2.1.12. Let $\mathfrak{g}$ be a real semisimple Lie algebra, and let $\mathfrak{h}$ be a split Cartan subalgebra. Chose a set of positive roots Σ^+ for $(\mathfrak{g}, \mathfrak{h})$. The *fundamental Weyl chamber* in $\mathfrak{h}$ is defined to be the set

$$\mathfrak{h}^+ := \{X \in \mathfrak{h} : \alpha(X) > 0 \text{ for all } \alpha \in \Sigma^+\}.$$

2.2 Facts about Lie Groups

We assume the reader is familiar with the basic theory of Lie groups. We use this section to review some more advanced definitions and propositions from the theory of real semisimple Lie groups.

Definition 2.2.1. We say a Lie group G is *semisimple* if it is connected and its Lie algebra is semisimple.

Theorem 2.2.2. *Let G be a semisimple Lie group, and let θ be a Cartan involution of its Lie algebra $\mathfrak{g}$. Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the corresponding Cartan decomposition. Let K be the analytic subgroup of G with Lie algebra $\mathfrak{k}$. Then*

(a) There exists a Lie group automorphism Θ of G with differential θ and $\Theta^2 = id$.

(b) The subgroup of G fixed by Θ is K.

(c) The map $K \times \mathfrak{p} \to G$ given by $(k, X) \mapsto k exp\ X$ is a diffeomorphism.

(d) K is closed.

(e) K contains the center Z of G.

(f) K is compact if and only if Z is finite.

(g) When Z is finite, K is a maximal compact subgroup of G.

We refer to the decomposition in (c) as the *polar decomposition* of G.

Proof: See [12, Theorem 6.31] ∎

Example 2.2.3. An example of a split semisimple Lie group with infinite center is the universal cover $\widetilde{\mathrm{SL}_2(\mathbb{R})}$. See [15, Example 1.4.13] for more details.

Theorem 2.2.4. *(Iwasawa decomposition) Let G be a semisimple Lie group, and let $\mathfrak{g}$ be its Lie algebra. Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ be its Iwasawa decomposition. Let K, A and N be the analytic subgroups of G with Lie algebras $\mathfrak{k}$, $\mathfrak{a}$ and $\mathfrak{n}$, respectively. Then the* *map*

$$K \times A \times N \to G$$
$$(k, a, n) \mapsto kan$$

is a diffeomorphism. Moreover, the groups A and N are simply connected.

Proof: See [12, Theorem 6.46] ∎

Our next goal is to describe another important decomposition of a semisimple Lie group known as the *Bruhat decomposition*. We only describe this decomposition in the case of the group $G = \mathrm{SL}_n(\mathbb{R})$. Recall that $\mathrm{SL}_n(\mathbb{R})$ is the group of $n \times n$ real matrices of determinant 1. Let H be the standard Cartan subgroup consisting of diagonal matrices in G. Let B be the *Borel subgroup* of G consisting of upper triangular matrices in G.

Recall that a matrix M over a field $\mathbb{F}$ is a called a *monomial matrix* if there exists exactly one nonzero entry in each row and exactly one nonzero entry in each column.

Lemma 2.2.5. *The normalizer $N_G(H)$ is equal to the set of monomial matrices in G.*

Proof: Let $g = (a_{ij}) \in N_G(H)$, and let $h = \mathrm{diag}(1, 2, \ldots, n)$. Then $x := ghg^{-1} \in H$. Since h and x have the same eigenvalues, there exists some $\sigma \in S_n$ such that $x = \mathrm{diag}(\sigma(1), \ldots, \sigma(n))$. Since $x = ghg^{-1}$, we also have the equation $xg = gh$. A simple calculation shows that this implies that

$$a_{ij}(\sigma(i) - j) = 0$$

for all $1 \leq i, j \leq n$. Fix some $1 \leq j \leq n$. Then the above equation implies that for each i such that $\sigma(i) \neq j$, we have $a_{ij} = 0$. It's not hard to see that g cannot have any entirely zero columns. This proves our claim. ∎

Recall that the Weyl group of a semisimple Lie group G with a chosen maximal torus T is defined as $W := N_G(T)/T$. Notice that H is a maximal torus in $\mathrm{SL}_n(\mathbb{R})$.

Corollary 2.2.6. *The Weyl group for* $\mathrm{SL}_n(\mathbb{R})$ *is isomorphic to* S_n*.*

Theorem 2.2.7. *(Bruhat decomposition) We have the decomposition*

$$G = \coprod_{w \in W} BwB.$$

Proof: Let $g \in G$. Then there exists some $b \in B$ such that every row of $b \cdot g$ contains a different number of 0's. Therefore, there exists $n_\sigma \in N_G(H)$ such that $n_\sigma \cdot b \cdot g$ is in upper triangular form, i.e. $b' = n_\sigma \cdot b \cdot g \in B$. So $g = b^{-1} \cdot n_\sigma^{-1} \cdot b' \in BWB$.

Recall that the double cosets in the above union are disjoint since they are the equivalence classes of the relation $\sim$ on G defined by $x \sim y$ iff there exists $b, b' \in B$ such that $bxb' = y$. ∎

2.3 Structure of $\mathrm{SL}_2(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{R})$

Recall that the group $\mathrm{SL}_n(\mathbb{R})$ is defined to be the subgroup of $\mathrm{GL}_n(\mathbb{R})$ consisting of matrices of determinant 1. As in Section 2.2, we define the standard *Cartan subgroup* of $\mathrm{SL}_n(\mathbb{R})$ to be the subgroup H of $\mathrm{SL}_n(\mathbb{R})$ consisting of diagonal matrices, and the standard *Borel subgroup* B of $\mathrm{SL}_n(\mathbb{R})$ to be the subgroup of $\mathrm{SL}_n(\mathbb{R})$ consisting of upper triangular matrices. Clearly H is a subgroup of B.

Also recall that the Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ is defined to be the subalgebra of $\mathfrak{gl}_n(\mathbb{R})$ consisting of matrices with trace 0. We let $\mathfrak{h}$ be the standard Cartan subalgebra of $\mathfrak{g}$ and $\mathfrak{b}$ the standard Borel subalgebra of $\mathfrak{g}$. We know that $\mathfrak{h}$ consists of all diagonal matrices in $\mathfrak{g}$ and $\mathfrak{b}$ consists of upper triangular matrices in $\mathfrak{g}$.

In particular,

$$\mathfrak{sl}_2(\mathbb{R}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R} \text{ and } a + d = 0 \right\}$$

and we have a basis of $\mathfrak{sl}_2(\mathbb{R})$ given by

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which satisfies the relations

$$[H, E] = 2E, \qquad [H, F] = -2F, \qquad [E, F] = H.$$

Concretely, the standard Cartan subalgebra of $\mathfrak{sl}_2(\mathbb{R})$ is

$$\mathfrak{h} = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} : a, b \in \mathbb{R} \text{ and } a + b = 0 \right\}$$

and its standard Borel subalgebra is

$$\mathfrak{b} := \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : a, b, c \in \mathbb{R} \text{ and } a + c = 0 \right\}$$

which is indeed a subalgebra of $\mathfrak{sl}_2(\mathbb{R})$.

Example 2.3.1. A Cartan involution on $\mathfrak{sl}_2(\mathbb{R})$ is given by $\theta(X) = -X^T$. A standard Iwasawa decomposition for $\mathfrak{sl}_2(\mathbb{R})$ is given by $\mathfrak{sl}_2(\mathbb{R}) = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ where $\mathfrak{k}$ is the set of matrices of the form

$$\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix},$$

$\mathfrak{a}$ is the set of all matrices of the form

$$\begin{pmatrix} x & 0 \\ 0 & -x \end{pmatrix},$$

and $\mathfrak{n}$ is the set of all matrices of the form

$$\begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}$$

where x is an arbitrary element of $\mathbb{R}$.

Proposition 2.3.2. *The Casimir element of* $\mathfrak{sl}_2(\mathbb{R})$ *is given by*

$$\mathbf{C} = \frac{1}{4}EF + \frac{1}{4}FE + \frac{1}{8}H^2.$$

Proof: Using the relations above, we see that the matrices for $\mathrm{ad}(E), \mathrm{ad}(F)$ and $\mathrm{ad}(H)$ are given as follows

$$\mathrm{ad}(E) = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathrm{ad}(F) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \mathrm{ad}(H) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

An easy calculation then confirms that

$$E' = \frac{1}{4}F, \qquad F' = \frac{1}{4}E, \qquad H' = \frac{1}{8}H$$

is a basis of $\mathfrak{sl}(2,\mathbb{R})$ which satisfies $\kappa(E,E') = 1, \kappa(E,F') = \kappa(E,H') = 0$, and $\kappa(F,E') = 0, \kappa(F,F') = 1, \kappa(F,H') = 0$ and $\kappa(H,E') = 0, \kappa(H,F') = 0, \kappa(H,H') = 1$. Therefore, our Casimir element is given by

$$C = EE' + FF' + HH' = \frac{1}{4}EF + \frac{1}{4}FE + \frac{1}{8}H^2.$$

∎

Remark 2.3.3. Note that $EF+FE = EF-FE+2FE = [E,F]+2FE = H+2FE$. Therefore, one can also write

$$C = \frac{1}{2}FE + \frac{1}{4}H + \frac{1}{8}H^2.$$

Chapter 3

Real Versus Complex Representations

In this chapter, we outline some basic results on the representation theory of Lie groups and Lie algebras. Following this, we classify the real finite dimensional representations of $\mathrm{SL}_2(\mathbb{R})$ and $\mathrm{SO}_2(\mathbb{R})$.

3.1 Basic Representation Theory

In this section, we outline some basic definitions and results in the representation theory of Lie groups and Lie algebras. We assume that the reader is familiar with some basic results, such as Schur's Lemma. We take $\mathbb{F}$ to be either $\mathbb{R}$ or $\mathbb{C}$.

3.1.1 Representations of Lie Groups

Definition 3.1.1. Let G be a Lie group. An $\mathbb{F}$-*representation* of G is a pair (π, V) where π is a smooth group homomorphism

$$\pi : G \to \mathrm{GL}(V)$$

and V is a finite dimensional topological vector space over $\mathbb{F}$.

Remark 3.1.2. If the underlying field $\mathbb{F}$ of V is $\mathbb{R}$, we say (π, V) is a *real representation* of G. Similarly, if $\mathbb{F} = \mathbb{C}$, we say (π, V) is a *complex representation* of G.

Example 3.1.3. Let G be a Lie group, and let $\mathfrak{g}$ be its Lie algebra. Let e be the identity element of G. Thus, $\mathfrak{g}$ is the tangent space T_eG. For each $g \in G$, define the inner automorphism $\Psi_g : G \to G$ by $\Psi_g(h) = ghg^{-1}$. Define $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ to be the differential of Ψ_g at e. If we realize G as a matrix Lie group, then Ad_g acts on $\mathfrak{g}$ by conjugation, i.e.

$$\mathrm{Ad}_g(X) = gXg^{-1} \qquad g \in G, X \in \mathfrak{g}.$$

We thus have the *adjoint representation* Ad given by

$$\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g}) \qquad g \mapsto \mathrm{Ad}_g.$$

Definition 3.1.4. Let G be a Lie group and let (π, V) be a representation of G. We define the *dual representation* of G to be the representation (π^*, V^*) given by

$$\pi^*(g) := \pi(g^{-1})^T$$

where $X \mapsto X^T$ denotes the transpose operation and V^* denotes the linear dual of V. Explicitly, for $g \in G$, $\lambda \in V^*$, and $v \in V$, we have

$$\pi^*(g)(\lambda)(v) = \lambda(\pi(g^{-1})v).$$

Proposition 3.1.5. *Let G be a compact Lie group, and let (π, V) be a finite dimensional $\mathbb{F}$-representation of G. Suppose $\langle \cdot, \cdot \rangle$ is an inner product on V. Then there is* *an inner product $(\cdot,\cdot)$ on V that is G-invariant, i.e. satisfies*

$$(\pi(g)v, \pi(g)w) = (v, w) \qquad \textit{for all } g \in G, v, w \in V.$$

Proof: We give a sketch of the proof. Define $(\cdot,\cdot) : V \times V \to \mathbb{F}$ by

$$(v, w) := \int_G \langle \pi(g)v, \pi(g)w \rangle dg$$

for all $v, w \in G$ and where dg is the Haar measure on G. Then $(\cdot,\cdot)$ is G-invariant. ∎

Remark 3.1.6. Typically, we would say that $(\cdot,\cdot)$ makes (π, V) into a *unitary representation*, but the notion of unitary representations will not be used in this thesis.

3.1.2 Representations of Lie Algebras

Definition 3.1.7. Let $\mathfrak{g}$ be a Lie algebra. A *representation* of $\mathfrak{g}$ is a pair (π, V) where V is a vector space and

$$\pi : \mathfrak{g} \to \mathfrak{gl}(V)$$

is a Lie algebra homomorphism.

Remark 3.1.8. If V is a real (respectively, complex) vector space, we say (π, V) is a *real (complex) representation* of $\mathfrak{g}$.

Example 3.1.9. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Recall that we have the adjoint representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ from Example 3.1.3. Define $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ to be the differential of Ad at e, i.e. $\mathrm{ad} := d(\mathrm{Ad})(e)$. Then one sees that $\mathrm{ad}(X)(Y) = [X, Y]$ for $X, Y \in \mathfrak{g}$.

Definition 3.1.10. Let $\mathfrak{g}$ be a Lie algebra and let (π, V) be a representation of $\mathfrak{g}$.

We define the *dual representation* of (π, V) to be the representation (π^*, V^*) given by

$$\pi^*(X) := -\pi(X)^T \qquad X \in \mathfrak{g}.$$

Definition 3.1.11. Suppose $\mathfrak{g}$ is a real semisimple Lie algebra with Iwasawa decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$. Let $\lambda : \mathfrak{a} \to \mathbb{R}$ be a linear functional, and let (π, V) be a real representation of $\mathfrak{g}$. We say a nonzero vector $v_\lambda \in V$ is a *highest weight vector of weight λ with respect to $\mathfrak{a}$ and $\mathfrak{n}$* if

$$\pi(H)v_\lambda = \lambda(H)v_\lambda, \pi(X)v_\lambda = 0 \qquad \text{for all } H \in \mathfrak{a}, X \in \mathfrak{n}.$$

We say V is a *highest weight module of weight λ* if V is generated by a highest weight vector v_λ of weight λ.

Theorem 3.1.12. *Let $\mathfrak{g}$ be a real split semisimple Lie algebra (in the sense of Definition 2.1.11) and let V be an irreducible finite dimensional representation. Then V is a highest weight representation.*

Proof: See [9, Theorem 7.3.15]. ∎

Remark 3.1.13. There is a parallel theory in the complex case which is more standard. We will not review the complex case, and refer the reader to references such as [10].

Theorem 3.1.14. *Let G be a Lie group, and let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. Let (π, V) be an $\mathbb{F}$-representation of G. Then for every $X \in \mathrm{Lie}(G)$, the map*

$$t \mapsto \pi(\exp(tX)) \qquad t \in \mathbb{R}$$

is smooth. Set

$$\pi(X) := \frac{d}{dt}|_{t=0}\pi(\exp(tX)).$$

Then $\pi(X) \in \mathrm{End}_{\mathbb{F}}(V)$ *and the map*

$$\mathrm{Lie}(G) \to \mathrm{End}_{\mathbb{F}}(V) \qquad X \mapsto \pi(X)$$

is a Lie algebra homomorphism.

Note that this notation is abusive, and several authors use the notation $d\pi$ for the representation on $\mathfrak{g}$.

Proof: See [8, Chapter 2]. ∎

3.2 Representations of $\mathrm{SL}_2(\mathbb{R})$, $\mathfrak{sl}_2(\mathbb{R})$ and $\mathrm{SL}_n(\mathbb{R})$

The goal of this section is to classify the real finite dimensional irreducible representations of $\mathrm{SL}_2(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{R})$ up to isomorphism. The classification is, in principle, the same for the complex finite dimensional representations of $\mathrm{SL}_2(\mathbb{C})$, but some subtleties occur in the real case which need to be addressed. The main issue is that the action of the Cartan subalgebra of $\mathfrak{sl}_2(\mathbb{R})$ is not obviously diagonalizable over $\mathbb{R}$, and therefore we need to modify the argument from the complex case to obtain weight spaces which are *defined* over $\mathbb{R}$.

For each $d \geq 0$, there is exactly one $(d+1)$-dimensional real irreducible representation of $\mathrm{SL}_2(\mathbb{R})$, which we denote by (π_d, V_d). We begin by constructing this family. We then show that this is an exhaustive family of representations of $\mathrm{SL}_2(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{R})$. All representations in this section will be real. We end with a word on representations of $\mathrm{SL}_n(\mathbb{R})$.

3.2.1 Representations of $\mathrm{SL}_2(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{R})$

Fix an integer $d \geq 0$ and let V_d denote the space of homogeneous polynomials in x and y of degree d with coefficients in $\mathbb{R}$. Recall that we have the usual action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{R}^2$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.(x,y) := (ax+by, cx+dy).$$

We then define the representation

$$\pi_d : \mathrm{SL}_2(\mathbb{R}) \to \mathrm{GL}(V_d)$$

as follows. For $P(x,y) \in V_d$ and $g \in \mathrm{SL}_2(\mathbb{R})$, we set

$$\pi_d(g)P(x,y) = P(g^{-1}(x,y)).$$

Note that if

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$$

then

$$g^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

so

$$\pi_d(g)P(x,y) = P(dx - by, -cx + ay).$$

Proposition 3.2.1. *Let $d \geq 1$. Then (π_d, V_d) is a representation of* $\mathrm{SL}_2(\mathbb{R})$.

Proof: The fact that π_d is a homomorphism is a straightforward calculation. The smoothness of π_d follows from the fact that the the map

$$\mathrm{SL}_2(\mathbb{R}) \times \mathbb{R}^2 \to \mathbb{R}^2 \qquad (g,v) \mapsto g^{-1}v$$

is smooth. ∎

We now differentiate (π_d, V_d) to obtain a representation of $\mathfrak{sl}_2(\mathbb{R})$ which we still denote by (π_d, V_d) (by abuse of notation).
Let $t \in \mathbb{R}$. Then

$$\exp tE = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad \exp tF = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}, \qquad \exp tH = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}.$$

So

$$(\exp tE)^{-1} = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}, \qquad (\exp tF)^{-1} = \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}, \qquad (\exp tH)^{-1} = \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix}.$$

Now let $f \in V_d$. Then

$$\begin{aligned} \frac{d}{dt}|_{t=0}\pi(\exp tE)f(x,y) &= \frac{d}{dt}|_{t=0}\pi \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} f(x,y) \\ &= \frac{d}{dt}|_{t=0} f(\begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix} (x,y)) \\ &= \frac{d}{dt}|_{t=0} f(x-ty,y) \\ &= -y\frac{\partial f}{\partial x} \end{aligned}$$

and

$$\begin{aligned}\frac{d}{dt}_{t=0}\pi(\exp tF)f(x,y) &= \frac{d}{dt}_{t=0}f(\begin{pmatrix}1 & 0\\ -t & 1\end{pmatrix}(x,y))\\ &= \frac{d}{dt}_{t=0}f(x,-tx+y)\\ &= -x\frac{\partial f}{\partial y}\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dt}_{t=0}\pi(\exp tH)f(x,y) &= \frac{d}{dt}_{t=0}f(\begin{pmatrix}e^{-t} & 0\\ 0 & e^{t}\end{pmatrix}(x,y))\\ &= \frac{d}{dt}_{t=0}f(e^{-t}x,e^{t}y)\\ &= -x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}\end{aligned}$$

So we have the following proposition.

Proposition 3.2.2. *Let $d \geq 1$. Let $\{v_0, \ldots, v_d\}$ be a basis of V_d, where $v_i = x^{d-i}y^i$ for $0 \leq i \leq d$. Then (π_d, V_d) is given by the following relations:*

(a) $\pi_d(H)(v_i) = (2i-d)v_i$

(b) $\pi_d(E)(v_i) = (i-d)v_{i+1}$

(c) $\pi_d(F)(v_i) = -iv_{i-1}$

Here we assume $v_{-1} = 0$ and $v_{d+1} = 0$.

Proof: This immediately follows from the above calculations. See [23] for the details. ∎

Remark 3.2.3. Recall the Iwasawa decomposition $\mathfrak{sl}_2(\mathbb{R}) = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ from Example 2.3.1. Clearly $\mathfrak{a} = \mathbb{R}H$ and $\mathfrak{n} = \mathbb{R}E$. In the module V_d, the vector v_d satisfies

$$\pi_d(H)v_d = dv_d \qquad \pi_d(E)v_d = 0.$$

Thus, if we define $\lambda : \mathfrak{a} \to \mathbb{R}$ by $\lambda(H) = d$ then we see that v_d is a highest weight vector of V_d with weight λ. Moreover, V_d is generated by v_d, so V_d is a highest weight module.

Remark 3.2.4. This is a specific case of a very general principle on the classification of representations of split semisimple Lie algebras.

Lemma 3.2.5. *Each module (π_d, V_d) is irreducible.*

The proof of the above lemma is identical to the complex case, and may be found in standard books on Lie algebras (e.g. [6, Theorem 8.2]).

Lemma 3.2.6. *Let $n \geq 0$. Then*

1. $[H, F^n] = -2nF^n$.
2. $[E, F^n] = n(H + (n-1)\mathrm{id})F^{n-1}$.
3. $[F, E^n] = -n(H - (n-1)\mathrm{id})E^{n-1}$.

Proof: We only prove (a) for $n = 2$. See [9, Lemma 6.2.2] for a full proof. By the

PBW theorem (Theorem 1.3.8) we obtain

$$\begin{aligned}[H, F^2] &= HF^2 - F^2H \\ &= (HF - FH)F + FHF - F^2H \\ &= [H, F]F + FHF - F^2H \\ &= -2F^2 + FHF - F^2H \\ &= F(-2F + HF) - F^2H \\ &= F(-4F + FH) - F^2H \\ &= -4F^2\end{aligned}$$

∎

Lemma 3.2.7. *Let (ρ, V) be a finite dimensional $\mathfrak{sl}_2(\mathbb{R})$ representation. Suppose there exists $v \in V$ such that $\rho(E)v = 0$ and $\rho(H)v = \lambda v$ for some $\lambda \in \mathbb{R}$. Then*

(i) λ is a non-negative integer.

(ii) v generates a submodule of V isomorphic to V_λ.

Proof:

(i) Let $n \geq 0$. By Lemma 3.2.6, we have

$$\rho(H)\rho(F)^n v = ([\rho(H), \rho(F)^n] + \rho(F)^n \rho(H))v = (\lambda - 2n)\rho(F)^n v.$$

Similarly, we have

$$\begin{aligned}\rho(E)\rho(F)^n v &= ([\rho(E), \rho(F)^n] + \rho(F)^n \rho(E))v \\ &= n\rho(F)^{n-1}(\rho(H) - n + 1)v = n(\lambda - n + 1)\rho(F)^{n-1}v.\end{aligned}$$

This shows that the submodule W generated by v is

$$W = \text{span}_{\mathbb{R}}\{\rho(F)^n v : n \geq 0\}.$$

Since V is finite dimensional, $\rho(H)$ only has finitely many eigenvalues on V. Hence, there is a minimal $N \geq 0$ with $\rho(F)^{N+1}v = 0$. From the fact that $\rho(E)\rho(F)^{N+1}v = 0$, we obtain $\lambda = N$.

(ii) For each $0 \leq k \leq \lambda - 1$, define

$$v_k := \frac{\rho(F)^{\lambda-k}v}{\lambda(\lambda-1)\cdots(k+1)}.$$

Then $W = \text{span}_{\mathbb{R}}\{v_0, \ldots, v_\lambda\}$. A simple computation shows that $W \cong V_N$.

∎

Definition 3.2.8. Let V be a vector space over $\mathbb{R}$. Consider the real vector space $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$. We define the action of $\mathbb{C}$ on $V_{\mathbb{C}}$ by

$$\alpha.(v \otimes \beta) = v \otimes (\alpha\beta) \qquad v \in V, \alpha, \beta \in \mathbb{C}.$$

We then view $V_{\mathbb{C}}$ as a complex vector space with this scalar multiplication.

Remark 3.2.9. Any time we use the notation $V_{\mathbb{C}}$ in this thesis, we are viewing $V_{\mathbb{C}}$ as a complex vector space.

Definition 3.2.10. Let $\mathfrak{g}$ be a real Lie algebra. Then we define its *complexification* $\mathfrak{g}_{\mathbb{C}}$ to be the Lie algebra $\mathfrak{g}_{\mathbb{C}} := \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$ where the Lie bracket on $\mathfrak{g}_{\mathbb{C}}$ is the unique extension to $\mathfrak{g}_{\mathbb{C}}$ of the Lie bracket on $\mathfrak{g}$.

Proposition 3.2.11. *Let* (ρ, V) *be an irreducible* $\mathfrak{sl}_2(\mathbb{R})$*-module of dimension* $d+1$ *for some* $d \geq 0$*. Then*

$$(\rho, V) \cong (\pi_d, V_d).$$

Proof: Denote the Lie subalgebra of $\mathfrak{gl}(V)$ generated by $\rho(E)$ and $\frac{1}{2}\rho(H)$ by $\mathfrak{b}$. Note that $\mathfrak{b}$ is solvable. It follows that the complexification $\mathfrak{b}_{\mathbb{C}}$ is a solvable subalgebra of $\mathfrak{gl}(V_{\mathbb{C}})$. Thus, by Lie's theorem, there exists a basis of $V_{\mathbb{C}}$ such that both $\pi(E)$ and $\frac{1}{2}\rho(H)$ are upper triangular. But note that

$$[\frac{1}{2}\rho(H), \rho(E)] = \frac{1}{2}\rho([H, E]) = \rho(E).$$

The commutator of two upper triangular matrices is strictly upper triangular. It thus follows that $\rho(E)$ is nilpotent. Let $d \geq 1$ be the minimal positive integer such that $\rho(E)^d = 0$. By Lemma 3.2.6 we have

$$0 = [\rho(F), \rho(E)^d] = -d(\rho(H) - (d-1)\mathrm{id})\rho(E)^{d-1}.$$

Thus, any $v_0 \in \rho(E)^{d-1}V$ is an eigenvector for $\rho(H)$ with eigenvalue $d-1$. By Lemma 3.2.5 and Lemma 3.2.7, $V \cong V_d$. ∎

Theorem 3.2.12. *The Casimir operator* $\mathbf{C}$ *of* $\mathfrak{sl}_2(\mathbb{R})$ *acts on a highest weight module* V_λ *by the scalar* $\frac{1}{4}\lambda + \frac{1}{8}\lambda^2$*.*

Proof: Let $v \in V_\lambda$ be the highest weight vector. Note that since $\mathbf{C}$ is in the center of $\mathcal{U}(\mathfrak{sl}_2(\mathbb{R}))$ and $V_\lambda = \mathcal{U}(\mathfrak{sl}_2(\mathbb{R}))v$, $\mathbf{C}$ must act by a scalar on all of V_λ. The scalar may be computed by the action of $\mathbf{C}$ on v. We obtain

$$\begin{aligned}\mathbf{C}.v &= \frac{1}{2}(FE).v + \frac{1}{4}H.v + \frac{1}{8}(H^2).v \\ &= 0 + \frac{1}{4}\lambda v + \frac{1}{8}\lambda^2 v \\ &= (\frac{1}{4}\lambda + \frac{1}{8}\lambda^2)v.\end{aligned}$$

∎

3.2.2 Representations of $\mathrm{SL}_n(\mathbb{R})$

Let $G = \mathrm{SL}_n(\mathbb{R})$ and let $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{R})$ be the Lie algebra of G. We let B be the standard Borel subgroup of G, and H be the standard Cartan subgroup of G. Similarly, we let $\mathfrak{b}$ be the standard Borel subalgebra of $\mathfrak{g}$ and $\mathfrak{h}$ be the standard Cartan subalgebra of $\mathfrak{g}$. The reader may revisit the end of Section 2.2 for the definitions of G, H, B, $\mathfrak{g}$ and $\mathfrak{b}$. We have the triangular decomposition

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$$

where $\mathfrak{n}^-$ and $\mathfrak{n}^+$ are the subalgebras consisting of strictly lower triangular and strictly upper triangular matrices, respectively.

Before ending this section, we take a quick look at the highest weight modules for G and $\mathfrak{g}$, which will be used in the proof of Kostant's theorem for G, in Section 4.2.1.

By 3.1.12, all irreducible finite dimensional representations of G and $\mathfrak{g}$ are highest weight representations. Let (π, V_λ) be a highest weight representation of G with highest weight $\lambda \in \mathfrak{h}^*$ (we also denote the representation of $\mathfrak{g}$ by (π, V_λ)). Let Γ_λ be the set of weights of (π, V_λ), and decompose V_λ into $\mathfrak{h}$-weight spaces

$$V_\lambda = \bigoplus_{\mu \in \Gamma_\lambda} V_\lambda(\mu).$$

Moreover, if $X \in \mathfrak{n}^+$ and $x \in B$, then $X.v = 0$ and $x.v \in (\mathbb{R} \setminus \{0\})v_\lambda$.

Assume that V_λ is not the trivial G-module. Then $\lambda \neq 0$. Now, define $\varepsilon_i \in \mathfrak{h}^*$ by

$$\varepsilon_i(\mathrm{diag}(t_1, \ldots, t_n)) = t_i.$$

Then one can write $\lambda = \lambda_1\varepsilon_1 + \cdots + \lambda_n\varepsilon_n$ where $\lambda_i - \lambda_{i+1} \in \mathbb{N} \cup \{0\}$ for $1 \leq i \leq n-1$. Note that this representation of λ is not unique, as one can also write

$$\lambda = \lambda_1\varepsilon_1 + \cdots + \lambda_n\varepsilon_n + \alpha(\varepsilon_1 + \cdots + \varepsilon_n)$$

where $\alpha \neq 0$. In particular, we can assume that $\lambda_i \in \mathbb{Z}$ for $1 \leq i \leq n$. Since we have $\varepsilon_1 + \cdots + \varepsilon_n = 0$, and we know that $\lambda \neq 0$, we must have $\lambda_i > \lambda_j$ for some $i < j$.

The action of H on V_λ is given as follows. Suppose $x = \mathrm{diag}(t_1, \ldots, t_n) \in G$. Then for $v \in V_\lambda(\lambda)$,

$$x.v = {t_1}^{\lambda_1} \cdots {t_n}^{\lambda_n} v.$$

3.3 Representations of $\mathrm{SO}_2(\mathbb{R})$

In this section, we classify the real finite dimensional representations of $\mathrm{SO}_2(\mathbb{R})$. All representations will be over real and finite dimensional vector spaces.

First recall that the group $\mathrm{SO}_n(\mathbb{R})$ is defined to be the group of $n \times n$ orthogonal matrices with determinant 1. When $n = 2$, we know that

$$\mathrm{SO}_2(\mathbb{R}) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in [0, 2\pi) \right\}.$$

We let $(\rho_0, \mathbb{R})$ be the trivial representation of $\mathrm{SO}_2(\mathbb{R})$ on $\mathbb{R}$, i.e.

$$\rho_0(g)x = x \qquad \text{for all } g \in \mathrm{SO}_2(\mathbb{R}), x \in \mathbb{R}.$$

For $k \in \mathbb{Z} \setminus \{0\}$, we define the representation $(\rho_k, \mathbb{R}^2)$ by

$$\rho_k\left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\right) := \begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix} \qquad \text{for all } \theta \in [0, 2\pi).$$

Proposition 3.3.1. *The representations* $(\rho_0, \mathbb{R})$ *and* $(\rho_k, \mathbb{R}^2)$*,* $k \in \mathbb{Z} \setminus \{0\}$*, are irreducible.*

Proof: The fact that $(\rho_0, \mathbb{R})$ is irreducible is trivial. Now let $k \in \mathbb{Z} \setminus \{0\}$. To show that $(\rho_k, \mathbb{R}^2)$ is irreducible, suppose there exists an invariant 1-dimensional subspace $U \subseteq \mathbb{R}^2$. Say U is spanned by some vector

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Then

$$\rho_k\left(\begin{pmatrix} \cos\pi/2k & -\sin\pi/2k \\ \sin\pi/2k & \cos\pi/2k \end{pmatrix}\right) v = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -v_2 \\ v_1 \end{pmatrix}.$$

But the right hand side is not a member of U, because

$$\begin{pmatrix} -v_2 \\ v_1 \end{pmatrix} \text{ and } \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

always span $\mathbb{R}^2$. This is a contradiction. ∎

Remark 3.3.2. Let X be an $n \times n$ matrix with real entries and an eigenvalue $\lambda \in \mathbb{C}$, with a corresponding eigenvector $v \in \mathbb{C}^n$. If λ is non-real and $v = u + iw$ where $u, w \in \mathbb{R}^n$, then $w \neq 0$.

Lemma 3.3.3. *Let $n \geq 2$ and let $X \in \mathrm{SO}_n(\mathbb{R})$. Suppose $\lambda = a + ib \in \mathbb{C}$ is a non-real (i.e. $b \neq 0$) eigenvalue of X with eigenvector $u = v + iw \in \mathbb{C}^n$, where $v, w \in \mathbb{R}^n$. Then the space $U := span_{\mathbb{R}}\{v, w\}$ has real dimension 2 and is invariant under X.*

Proof: We first prove that U has real dimension 2. Indeed, suppose for a contradiction that U has real dimension 1. Then there exists some scalar $\alpha \in \mathbb{R}$ such that $v = \alpha w$. But notice that $\lambda' = a - ib$ is also an eigenvalue of X with eigenvector $u' = v - iw$. Since $\lambda \neq \lambda'$, and eigenvectors with distinct eigenvalues are linearly independent, we know that $\{u, u'\}$ is a linearly independent set over $\mathbb{C}$. But the fact that $v = \alpha w$ forces $u = \alpha w + iw = (\alpha + i)w$ and $u' = \alpha w - iw = (\alpha - i)w$. This is a contradiction. So U has real dimension 2.

We now show that U is invariant under X. Notice that $Xu = Xv + iXw$ so $\lambda u = (a + ib)(v + iw) = (av - bw) + i(av + bu)$ thus

$$Xv = av - bw \qquad \text{and} \qquad Xw = aw + bv.$$

Since Xv and Xw are members of U, we obtain $XU \subseteq U$. Hence U is invariant under X. ∎

Proposition 3.3.4. *Let $V = \mathbb{R}^n$, and let $\mathcal{F}$ be a commuting subset of $\mathrm{SO}_n(\mathbb{R})$. Then one can write*

$$V = U_1 \oplus \cdots \oplus U_k$$

where each U_i is either a 1-dimensional common eigenspace of $\mathcal{F}$ or a 2-dimensional $\mathcal{F}$-invariant subspace of V. Moreover, each U_i is irreducible, that is, it does not have a proper $\mathcal{F}$-invariant subspace.

Proof: Note that we may assume that $\mathcal{F}$ is a linearly independent finite set $\{X_1, \ldots, X_d\}$ of $\mathrm{SO}_n(\mathbb{R})$ by taking a basis of the span of $\mathcal{F}$ over $\mathbb{R}$. Moreover, since each X_i is an orthogonal matrix, if U is $\mathcal{F}$ invariant, then so is $U^{\perp}$. Thus, it suffices to show that we may find a subspace U such that either (a) U is a 1-dimensional common eigenspace of $\mathcal{F}$ or (b) U is a 2-dimensional $\mathcal{F}$-invariant subspace of V. Indeed, one can then write

$$V = U \oplus U^{\perp}$$

and notice that we may then restrict $\mathcal{F}$ to $U^{\perp}$, and apply induction.

We proceed by induction on $n = \dim_{\mathbb{R}} V$. For the base case, note that if $n = 1$, then the claim is trivial.

Now suppose $n \geq 2$. Recall that members of $\mathrm{SO}_n(\mathbb{R})$ are diagonalizable over $\mathbb{C}$, so $\{X_1, \ldots, X_d\}$ is a commuting family of matrices which are diagonalizable over $\mathbb{C}$. Thus, one can write

$$V_{\mathbb{C}} = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

where each V_i is a common eigenspace of $\{X_1, \ldots, X_d\}$. Note that there exist $\lambda_1, \ldots, \lambda_d \in \mathbb{C}$ such that

$$X_i v = \lambda_i v \qquad \text{for all } v \in V_1.$$

There are two cases. In the first case, each λ_i is real. In the second case, at least one λ_i is complex.

Case 1. Suppose each λ_i is real. Then there is an common eigenvector $w \in V_1$ of $\{X_1, \ldots, X_d\}$ such that $w \in \mathbb{R}^n$. Set $U := \mathrm{span}_{\mathbb{R}}\{w_1\}$. It is clear that U is $\mathcal{F}$-invariant and has no proper $\mathcal{F}$-invariant subspace.

Case 2. Suppose at least one of the λ_i's is non-real. Without loss of generality, we may assume that λ_1 is non-real. Then by Remark 3.3.2, there is an eigenvector $w \in V_1$ such that $w \in \mathbb{C}^n \setminus \mathbb{R}^n$. There exist $u, v \in \mathbb{R}^n$ such that $w = u + \sqrt{-1}v$. Let $U := \text{span}_{\mathbb{R}}\{u, v\}$. For $1 \le i \le d$, if λ_i is complex, then Lemma 3.3.3 implies that U is invariant under X_i and is 2-dimensional. If λ_i is real, then note that $\overline{w} = u - \sqrt{-1}v$ is also an eigenvector of X_i with eigenvalue λ_i. Thus, we obtain

$$X_i(u) = X_i(\frac{1}{2}(w + \overline{w})) = \lambda_i u$$

and

$$X_i(v) = X_i(\frac{1}{2i}(w - \overline{w}) = \lambda_i v.$$

So u and v are both eigenvectors of X_i with eigenvalue λ_i. So U is invariant under X_i. So U is $\mathcal{F}$ invariant. Moreover, U has no proper $\mathcal{F}$-invariant real subspace, as both u and v are complex eigenvectors of X_1. This concludes our proof. ∎

Lemma 3.3.5. *Let (ρ, V) be an n-dimensional representation of $\text{SO}_2(\mathbb{R})$. Then we may choose an inner product on V such that each $\rho(g)$, for $g \in \text{SO}_2(\mathbb{R})$, is an orthogonal map.*

Proof: $\text{SO}_2(\mathbb{R})$ is compact, so we may simple take the $\text{SO}_2(\mathbb{R})$-invariant inner product $(\cdot, \cdot)$ on V from Proposition 3.1.5. Then for each $g \in \text{SO}_2(\mathbb{R})$, $\rho(g)$ will be orthongonal with respect to $(\cdot, \cdot)$. ∎

Remark 3.3.6. Let (ρ, V) be an n-dimensional representation of $\text{SO}_2(\mathbb{R})$. By Lemma 3.3.5, the image of ρ is a subgroup of $\text{O}_n(\mathbb{R})$. Moreover, since ρ is continuous, and

$SO_2(\mathbb{R})$ is connected, the image of ρ is a connected subgroup of $O_n(\mathbb{R})$. But the only connected subgroup of $O_n(\mathbb{R})$ is $SO_n(\mathbb{R})$, so $\rho(SO_2(\mathbb{R})) \subseteq SO_n(\mathbb{R})$. Thus, we may think of ρ as a map from $SO_2(\mathbb{R})$ to $SO_n(\mathbb{R})$.

Lemma 3.3.7. *Let (ρ, V) be a representation of $SO_2(\mathbb{R})$.*

(1) If V is one-dimensional, then $\rho \cong \rho_0$.

(2) If V is two-dimensional, then $\rho \cong \rho_k$ for some $k \in \mathbb{Z} \setminus \{0\}$.

Proof:

(1) By Remark 3.3.6, we may assume ρ is a map from $SO_2(\mathbb{R})$ to $SO_1(\mathbb{R}) = \{1\}$. Thus $\rho \cong \rho_0$.

(2) By Remark 3.3.6, we may assume ρ is a map from $SO_2(\mathbb{R})$ to $SO_2(\mathbb{R})$. From [4, Proposition 7.1.1], we know that the continuous endomorphisms of $SO_2(\mathbb{R})$ all have the form ρ_k for some $k \in \mathbb{Z}$. So $\rho \cong \rho_k$ for some $k \in \mathbb{Z} \setminus \{0\}$.

∎

Theorem 3.3.8. *Let (ρ, V) be a real representation of $SO_2(\mathbb{R})$. Then there exist non-negative integers $a_0, \ldots, a_k$ such that*

$$\rho = \rho_{a_0} \oplus \cdots \oplus \rho_{a_k}.$$

Proof: Assume that $V = \mathbb{R}^n$. Let

$$\mathcal{S} = \text{span}_{\mathbb{R}}\{\rho(g) : g \in SO_2(\mathbb{R})\}.$$

By Lemma 3.3.7, it suffices to show that we may write

$$V = U_1 \oplus \cdots \oplus U_k$$

where each U_i is a one or two dimensional $\mathcal{S}$-invariant irreducible subspace of V. Now, let $\mathcal{F}$ be a basis of $\mathcal{S}$. We may assume that

$$\mathcal{F} = \{X_1, \ldots, X_d\}$$

where each $X_i \in \mathrm{SO}_n(\mathbb{R})$ (by Remark 3.3.6), and is also a commuting family. We apply Proposition 3.3.4, and this concludes our proof. ∎

3.4 Extreme Points of Convex Hulls of Orbits

Now that we have discussed some of the theory of cones in representation spaces, we devote this section to an interesting example. In this section, we study the set of extreme points of the convex hull of $\mathrm{SL}_2(\mathbb{R})$ highest weight orbits. In [20], the authors prove that for a subgroup G of $\mathrm{SO}_n(\mathbb{R})$, every point in an orbit $G \cdot v$ is an extreme point in the convex hull $\mathrm{conv}(G \cdot v)$. See [20, Proposition 2.2] for a precise formulation. We prove this result for a specific representation of $\mathrm{SL}_2(\mathbb{R})$. We will be interested in the representations which contain a regular cone, i.e. spherical representations. As it turns out, these are the odd dimensional representations, which are all isomorphic to $\mathcal{S}^{2d}(\mathbb{R}^2) \cong \mathrm{Sym}^{2d}(\mathbb{R}^2)$. We first show that all of these representations are orthogonal.

Definition 3.4.1. Let V be a real vector space and let $\omega : V \times V \to \mathbb{R}$ be a bilinear form. We say ω is a *symplectic form* if ω is alternating (i.e. $\omega(v, v) = 0$ for each $v \in V$) and non-degenerate (i.e. $\omega(v, w) = 0$ for all $w \in V$ implies $v = 0$). We say the pair (V, ω) is a *symplectic space*.

Definition 3.4.2. Let G be a Lie group and let (π, V) be a representation of G. If there is a symplectic form ω on V such that

$$\omega(\pi(g)v, \pi(g)w) = \omega(v, w)$$

for all $g \in G$ and $v, w \in V$ then we say that (π, V, ω) is a *symplectic representation* of G. When the representation π and symplectic form ω are clear from the context, we sometimes say V is a *symplectic module*.

First we construct a symplectic form on the standard representation $V_1 \cong \mathbb{R}^2$ of $\mathrm{SL}_2(\mathbb{R})$. Define the real bilinear form ω on $\mathbb{R}^2$ by

$$\omega\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) := \det\begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

We let

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Notice that we have the relations

$$\omega(e_1, e_2) = 1, \omega(e_2, e_1) = -1, \omega(e_1, e_1) = \omega(e_2, e_2) = 0.$$

Remark 3.4.3. Note that ω is a symplectic form on $\mathbb{R}^2$.

Lemma 3.4.4. *Let $(\pi, \mathbb{R}^2)$ be the standard representation of $\mathrm{SL}_2(\mathbb{R})$. Then $(\pi, \mathbb{R}^2, \omega)$ is a symplectic representation of $\mathrm{SL}_2(\mathbb{R})$.*

Proof: In view of the above remark, it suffices to show that ω preserves the action of $\mathrm{SL}_2(\mathbb{R})$. Let

$$x = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}), v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2.$$

Then

$$\begin{aligned}\omega(\pi(x)v,\pi(x)w) &= \omega\left(\begin{pmatrix} ax_1+bx_2 \\ cx_1+dx_2\end{pmatrix},\begin{pmatrix} ay_1+by_2 \\ cy_1+dy_2\end{pmatrix}\right) \\ &= \det\begin{pmatrix} ax_1+bx_2 & ay_1+by_2 \\ cx_1+dx_2 & cy_1+dy_2\end{pmatrix} \\ &= (ax_1+bx_2)(cy_1+dy_2)-(cx_1+dx_2)(ay_1+by_2) \\ &= (ad-bc)x_1y_2+(bc-da)x_2y_1 \\ &= x_1y_2-x_2y_1 \\ &= \omega(v,w)\end{aligned}$$

where we use the fact that $ad-bc=1$ and $bc-da=-1$ since $\det x=1$ (as $x\in \mathrm{SL}_2(\mathbb{R})$). ∎

We extend ω to $\mathcal{T}^{2d}(\mathbb{R}^2)$ in the following way, and continue to denote the form on $\mathcal{T}^{2d}(\mathbb{R}^2)$ by ω. For simple tensors $v_1\otimes\cdots\otimes v_{2d}, w_1\otimes\cdots\otimes w_{2d}\in\mathcal{T}^{2d}(\mathbb{R}^2)$, define

$$\omega(v_1\otimes\cdots\otimes v_{2d}, w_1\otimes\cdots\otimes w_{2d}) := \omega(v_1,w_1)\cdots\omega(v_{2d},w_{2d}). \tag{3.4.1}$$

The form ω induces a form on the subspace $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ of $\mathcal{T}^{2d}(\mathbb{R}^2)$.

Lemma 3.4.5. $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ *is an orthogonal* $\mathrm{SL}_(\mathbb{R})$*-module with respect to* ω*.*

Proof: We need to show that on $\mathrm{Sym}^{2d}(\mathbb{R}^2)$, ω is symmetric, non-degnerate and preserves the action of $\mathrm{SL}_2(\mathbb{R})$. The fact that ω is symmetric on $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ follows from the fact that ω is alternating on $\mathbb{R}^2$ and that $2d$ is even. The non-degeneracy of ω on $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ follows from Equation 3.4.1 and the fact that ω is non-degenerate on $\mathbb{R}^2$. Finally, ω preserves the action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ by Equation 3.4.1 and the fact that ω preserves the action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{R}^2$. ∎

Recall that for a real vector space V and a simple tensor $x = v_1 \otimes \cdots \otimes v_k \in \mathcal{T}^k(V)$, we define

$$\mathrm{Sym}(x) = \frac{1}{k!} \sum_{\sigma \in S_k} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}.$$

For vectors $v_1, \ldots, v_k \in V$, we also use the shorthand

$$v_1 \ldots v_k = \mathrm{Sym}(v_1 \otimes \cdots \otimes v_k)$$

and for a vector $v \in V$, we write v^k to denote $vv \cdots v$ (k-times).

Proposition 3.4.6. *There is a basis $w_{-d}, \ldots, w_d$ of $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ such that $\omega(w_i, w_i) = (-1)^{d+i}$ and $\omega(w_i, w_j) = 0$ for all $i, j \in \{-d, \ldots, d\}$ when $i \neq j$.*

Proof: For each $-d \leq i \leq d$, define

$$v_i := e_1^{d-i} e_2^{d+i} \pm e_1^{d+i} e_2^{d-i}$$

where the sign is $+$ if $i \leq 0$ and $-$ if $i > 0$. We claim that $\{v_{-d}, \ldots, v_d\}$ is an orthogonal basis of $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ with respect to ω. Notice that for $i \neq j$ we have

$$\omega(v_i, v_j) = 0.$$

We claim that $\omega(v_i, v_i) \neq 0$ for all $-d \leq i \leq d$. Indeed, fix $-d \leq i \leq d$. Then

$$\begin{aligned}
\omega(v_i, v_i) &= \omega(e_1^{d-i} e_2^{d+i} + e_1^{d+i} e_2^{d-i}, e_1^{d-i} e_2^{d+i} + e_1^{d+i} e_2^{d-i}) \\
&= \omega(e_1^{d-i} e_2^{d+i}, e_1^{d+i} e_2^{d-i}) + \omega(e_1^{d+i} e_2^{d-i}, e_1^{d-i} e_2^{d+i}) \\
&= 2\omega(e_1^{d-i} e_2^{d+i}, e_1^{d+i} e_2^{d-i})
\end{aligned}$$

For each $\sigma \in S_{2d}$, let v_σ denote the tensor in $\mathcal{T}^{2d}(\mathbb{R}^2)$ given by

$$v_\sigma = \sigma \cdot (e_1 \otimes \cdots \otimes e_1 \otimes e_2 \otimes \cdots \otimes e_2)$$

where there are $d - i$ e_1's, and $d + i$ e_2's. Notice that for each $\sigma \in S_{2d}$, there are exactly $(d - i)! \cdot (d + i)!$ elements $\mu \in S_{2d}$ such that $\omega(v_\sigma, v_\mu) \neq 0$. In this case, $\omega(v_\sigma, v_\mu) = (-1)^{d+i}$.

We compute $\omega(e_1^{d-i}e_2^{d+i}, e_1^{d+i}e_2^{d-i})$ as follows. Notice that

$$\begin{aligned}
\omega(e_1^{d-i}e_2^{d+i}, e_1^{d+i}e_2^{d-i}) &= \left(\frac{1}{(2d)!}\right)^2 \sum_{\sigma \in S_{2d}} \sum_{\mu \in S_{2d}} \omega(v_\sigma, v_\mu) \\
&= \left(\frac{1}{(2d)!}\right)^2 (2d)!(d-i)!(d+i)!(-1)^{d+i} \\
&= \frac{1}{(2d)!}(d-i)!(d+i)!(-1)^{d+i}
\end{aligned}$$

so

$$\omega(v_i, v_i) = 2\frac{1}{(2d)!}(d-i)!(d+i)!(-1)^{d+i} = \binom{2d}{d-i}^{-1}(-1)^{d+i}.$$

So we normalize the basis and define

$$w_i := \frac{1}{\sqrt{|\omega(v_i, v_i)|}} v_i.$$

Then $\omega(w_i, w_i) = (-1)^{d+i}$. ∎

We define the real quadratic form Q on $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ by $Q(v) = \omega(v, v)$. We consider the basis $\{w_{-d}, \ldots, w_d\}$ of

$$\mathrm{Sym}^{2d}(\mathbb{R}^2)$$

above and define $Q(x_{-d}, \ldots, x_d) := Q(\sum x_i w_i)$. Notice, then, that

$$Q(x_{-d}, \ldots, x_d) = x_{-d}^2 - x_{-d+1}^2 + \cdots \pm x_0^2 \mp x_1^2 + \cdots + x_d^2.$$

We have thus proven the following result.

Proposition 3.4.7. *Let $d \geq 1$. Then the quadratic form Q on $\mathrm{Sym}^{2d}(\mathbb{R}^2)$ given by*

$$Q(x_{-d}, \ldots, x_d) := x_{-d}^2 - x_{-d+1}^2 + \cdots \pm x_0^2 \mp x_1^2 + \cdots + x_d^2$$

is $\mathrm{SL}_2(\mathbb{R})$-invariant.

Example 3.4.8. Let $d = 1$ and let $\mathcal{O}$ be the $\mathrm{SL}_2(\mathbb{R})$ orbit of the highest weight vector. Then Proposition 3.4.7 implies that every point $v = (x_{-1}, x_0, x_1)$ in $\mathcal{O}$ satisfies

$$Q(v) = 0 \text{ i.e. } x_{-1}^2 - x_0^2 + x_1^2 = 0.$$

In other words, every point $v \in \mathcal{O}$ lies in the *light cone*, i.e. the points (x_{-1}, x_0, x_1) satisfying

$$x_{-1}^2 - x_0^2 + x_1^2 = 0.$$

Moreover, since $\mathcal{O}$ does not contain 0 and $\mathcal{O}$ is connected, all points in $\mathcal{O}$ must lie on one half of the light cone. But we may realize the representation on the space of 2×2 symmetric matricies with (x_{-1}, x_0, x_1) corresponding to $\begin{pmatrix} x_0 & x_{-1} \\ x_{-1} & x_1 \end{pmatrix}$, where we know the highest weight vector is $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. So all points on $\mathcal{O}$ satisfy $x_0 > 0$. Moreover, one can show by an elementary linear algebra argument that the action is transitive, and $\mathcal{O}$ the half of the light cone with $x_0 > 0$, i.e.

$$\mathcal{O} = \{(x_{-1}, x_0, x_1) \in \mathrm{Sym}^{2d}(\mathbb{R}^2) : x_{-1}^2 - x_0^2 + x_1^2 = 0 \text{ and } x_0 > 0\}.$$

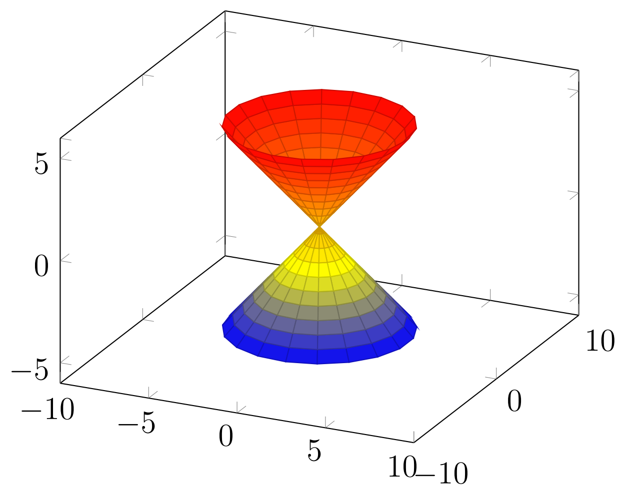


Figure 3.1 The Light Cone

Chapter 4

Defining Equations of the Highest Weight Orbit

The main goal of this chapter is to prove a real version of Kostant's theorem, which gives a set of quadratic equations which define the Zariski closure of the highest weight orbit. This result is proven in Section 4.2.1.

4.1 Orbits in Complex Vector Spaces

We begin with some results on orbits in complex vector spaces. We do this because we need results from algebraic geometry, which are only true over an algebraically closed field.

For simplicity in this section, we assume that $G = \mathrm{SL}_n(\mathbb{R})$ and $G_{\mathbb{C}} = \mathrm{SL}_n(\mathbb{C})$, but the results of this section hold in a much more general context which involves the theory of semisimple algebraic groups. Let (π, V) be a finite dimensional complex representation of $G_{\mathbb{C}}$. Then there is an action of $G_{\mathbb{C}}$ on the projective space $\mathbb{P}(V)$

given by

$$g.[v] := [\pi(g)v] \qquad g \in G_{\mathbb{C}}, v \in V.$$

Proposition 4.1.1. *Let (π, V) be an irreducible complex representation of $G_{\mathbb{C}}$, and let $v \in V$ be a highest weight vector. Then the complex orbit $G_{\mathbb{C}} \cdot [v]$ in $\mathbb{P}(V)$ is Zariski closed.*

Proof: Let $B_{\mathbb{C}}$ be the Borel subgroup of $G_{\mathbb{C}}$ which stabilizes $[v]$. Such a Borel subgroup exists because v is a highest weight vector. Then note that the map $f : G_{\mathbb{C}} \to \mathbb{P}(V)$ given by $f(g) = g.[v]$ is a morphism of algebraic varieties. We know from [2, §11.1] that $G_{\mathbb{C}}/B_{\mathbb{C}}$ is a projective algebraic variety. It follows from the universal property of quotients of algebraic groups that the induced map $\overline{f} : G_{\mathbb{C}}/B_{\mathbb{C}} \to \mathbb{P}(V)$ is a morphism of algebraic varieties. Since $\overline{f}$ is a morphism, we know that its image in $\mathbb{P}(V)$ is Zariski closed (see [22, §5.2]). ∎

In the proof of the following proposition, we use the fact that all the Borel subgroups of $G_{\mathbb{C}}$ are conjugate, i.e. if $B_{\mathbb{C}}$ and $B'_{\mathbb{C}}$ are two Borel subgroups of $G_{\mathbb{C}}$ then there exists $g_0 \in G_{\mathbb{C}}$ such that $g_0 B_{\mathbb{C}} g_0^{-1} = B'_{\mathbb{C}}$. See [11, p. 135] for a proof.

Proposition 4.1.2. *Let (π, V) be an irreducible complex representation of $G_{\mathbb{C}}$, and let $v \in V$ be nonzero. If the complex orbit $G_{\mathbb{C}} \cdot [v]$ is Zariski closed in $\mathbb{P}(V)$, then v is a highest weight vector.*

Proof: Suppose that $G_{\mathbb{C}}.[v]$ is Zariski closed. Then, let $X = G_{\mathbb{C}}.[v]$, and let $B_{\mathbb{C}}$ be the Borel subgroup of $G_{\mathbb{C}}$ consisting of complex upper triangular matricies. Then $B_{\mathbb{C}}$ acts on X and since X is Zariski closed, the Borel Fixed Point Theorem (see [16, §3.4.3]) implies that there exists a fixed point $g_0.[v]$ in X of this action, i.e.

$$B_{\mathbb{C}}.(g_0.[v]) = g_0.[v]$$

which implies

$$g_0^{-1} B_{\mathbb{C}} g_0.[v] = [v].$$

Thus, v is a highest weight vector with respect to the Borel subgroup $g_0^{-1} B_{\mathbb{C}} g_0$ of $G_{\mathbb{C}}$. ∎

4.2 Highest Weight Orbits

We begin by defining the object of interest for this section. Let $G = \mathrm{SL}_n(\mathbb{R})$, as in Section 4.1 and let $\mathfrak{g} = \mathrm{Lie}(G)$. Let $\mathfrak{h}$ denote the Cartan subalgebra of diagonal matricies in $\mathfrak{g}$. Recall from Proposition 3.1.12 that if G is a Lie group and (π, V) is a real irreducible representation of G, then V is a highest weight representation.

Definition 4.2.1. Let G be a real split semisimple Lie group. Let (π, V_λ) be a real, finite dimensional irreducible representation of G of highest weight λ. Let $v_\lambda \in V_\lambda$ be a chosen highest weight vector. Then we set

$$X_\lambda := G.v_\lambda \cup \{0\} = \{\pi(x)v_\lambda : x \in G\} \cup \{0\} \tag{4.2.1}$$

Recall that we say $\alpha \prec \beta$ for two weights $\alpha, \beta \in \mathfrak{h}^*$ if $\beta - \alpha$ is a linear combination of simple roots with non-negative coefficients.

Definition 4.2.2. Let G be a real split semisimple Lie group. Let (π, V_λ) be a real finite dimensional irreducible representation of G of highest weight λ, and let $v_\lambda \in V_\lambda$ be a highest weight vector. Let W denote the Weyl group of G. We choose a basis $\mathcal{B}$ of V_λ consisting of weight vectors, such that $w.v_\lambda \in \mathcal{B}$ for every $w \in W$. For every $v \in V_\lambda$, we write v as a linear combination of vectors in $\mathcal{B}$, and we denote the coefficient of $w.v_\lambda$ by $c(v, w)$. Next, we set

$$E_{\lambda,w} := \{v \in V_\lambda : c(v, w) > 0 \text{ and } c(v, w') = 0 \text{ if } w\lambda \prec w'\lambda\}.$$

Finally, set

$$E_\lambda := \{0\} \cup \bigcup_{w \in W} E_{\lambda,w}.$$

It is clear that E_λ is a semialgebraic set.

We now state Kostant's Theorem. Recall that $\mathbf{C}$ is the Casimir element of $\mathfrak{g}$ and for a weight μ, $C(\mu)$ is the scalar with which $\mathbf{C}$ acts on the irreducible highest weight representation V_μ of weight μ.

Theorem 4.2.3. *(Kostant) Let G be a real split semisimple Lie group and let $\mathfrak{g} =$ Lie(G). Let (π, V_λ) be a real irreducible representation of G of highest weight λ and with a chosen highest weight vector of v_λ. If $-v_\lambda \in G \cdot v_\lambda$, then X_λ is given by*

$$X_\lambda = \{v \in V_\lambda : \mathbf{C}(v \otimes v) - C(2\lambda) v \otimes v = 0\}.$$

If $-v_\lambda \notin G \cdot v_\lambda$, then X_λ is given by the intersection of the set given by the above equations and the set E_λ.

For the sake of simplicity, we only prove this result for $G = \mathrm{SL}_n(\mathbb{R})$. Our proof is based on the proof given in [17, Chapter 10, §6.6], but as in Chapter 3, we need to address the issues that arise from the difference between $\mathbb{R}$ and $\mathbb{C}$.

Remark 4.2.4. From Example 3.4.8 it follows that in the case of $d = 1$, the intersection of the complex orbit $\mathrm{SL}_2(\mathbb{C}) \cdot v_\lambda$ with the real subspace V_λ is the entire light cone, which is strictly larger than the real orbit $\mathrm{SL}_2(\mathbb{R}) \cdot v_\lambda$. This example clarifies the necessity of the constraints from E_λ.

4.2.1 Kostant's Theorem for $\mathrm{SL}_n(\mathbb{R})$

The proof of Theorem 4.2.3 is an immediate consequence of Proposition 4.2.5, Proposition 4.2.12 and Proposition 4.2.13, which will be proved below.

Let E be a finite dimensional real vector space, and as usual, let $\mathbb{P}(E)$ denote the projective space of E. We give E and $\mathbb{P}(E)$ the following topologies, both of which we call the *real Zariski topology*. Recall that the topology of E is the one in which the closed sets are precisely the loci of ideals of $\mathcal{P}(E)$. The topology of $\mathbb{P}(E)$ is the one in which the closed sets are the loci of homogeneous ideals of $\mathcal{P}(E)$. Recall that $G = \mathrm{SL}_n(\mathbb{R})$ and that X_λ is defined in equation 4.2.1. In this section, H and B will denote the standard Cartan and Borel subgroups of G, respectively. We let $H_{\mathbb{C}}$ and $B_{\mathbb{C}}$ denote the standard Cartan and Borel subgroups of $G_{\mathbb{C}} = \mathrm{SL}_n(\mathbb{C})$.

Proposition 4.2.5. *Assume the setting of Theorem 4.2.3. If $-v_\lambda \in X_\lambda$, then we have $G_{\mathbb{C}} \cdot v_\lambda \cap V_\lambda = G \cdot v_\lambda$. If $-v_\lambda \notin X_\lambda$, then $G \cdot v_\lambda$ is the subset of $G_{\mathbb{C}} \cdot v_\lambda$ consisting of vectors which satisfy the constraints given by E_λ.*

Proof: Let $V := V_\lambda$ and let $V_{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C}$ be the complexification of V. Then $V_{\mathbb{C}}$ is a complex highest weight module of G with highest weight vector v_λ. Let N^- and $N^-_{\mathbb{C}}$ denote the subgroups of lower unipotent elements in G and $G_{\mathbb{C}}$, respectively. First assume that $-v_\lambda \in X_\lambda$. We want to show that $G \cdot v_\lambda = G_{\mathbb{C}} \cdot v_\lambda \cap V$.

Note that G and $G_{\mathbb{C}}$ have the same Weyl group W. Moreover, we have the Bruhat decompositions

$$G = \coprod_{w \in W} N^- w B$$

and

$$G_{\mathbb{C}} = \coprod_{w \in W} N^-_{\mathbb{C}} w B_{\mathbb{C}}.$$

Then it is enough to show that for each $w \in W$, we have

$$N^- w B v_\lambda = (N^-_{\mathbb{C}} w B_{\mathbb{C}} v_\lambda) \cap V.$$

The inclusion $N^- w B v_\lambda \subseteq (N^-_{\mathbb{C}} w B_{\mathbb{C}} v_\lambda) \cap V$ is clear. For the other inclusion, note that our assumption implies that $-v_\lambda \in G \cdot v_\lambda$. This implies that $B \cdot v_\lambda = (\mathbb{R} \setminus \{0\}) v_\lambda =$

$(\mathbb{C} \setminus \{0\} v_\lambda) \cap V = B_{\mathbb{C}} \cdot v_\lambda \cap V$. Now note that

$$N_{\mathbb{C}}^- w B_{\mathbb{C}} v_\lambda = \bigcup_{\alpha \in \mathbb{C}^\times} \alpha N_{\mathbb{C}}^- w.v_\lambda.$$

First assume $\alpha \in \mathbb{R}$. Then set $v' = \alpha w v_\lambda$. By [14, Lemma 7.1], we have $N_{\mathbb{C}}^-.v' \cap V = N^- v' = \alpha N^- w v_\lambda$. On the other hand, suppose $\alpha \in \mathbb{C} \setminus \mathbb{R}$. We claim that $(\alpha N_{\mathbb{C}}^- w v_\lambda) \cap V = \emptyset$. Suppose $x \in \alpha N_{\mathbb{C}}^-.v_\lambda \cap V$. Then $x = \alpha g.(w v_\lambda)$ for some $g \in N_{\mathbb{C}}^-$. But $g.(w v_\lambda) = w v_\lambda + \sum_{\eta \neq w\lambda} v_\eta$ where $w v_\lambda \in V(w\lambda)$ and $v_\eta \in V_{\mathbb{C}}(\eta)$. But $\alpha w v_\lambda \notin V$ as well, a contradiction.

The reasoning in the case where $-v_\lambda \notin G \cdot v_\lambda$ is similar. The only difference is that this time,

$$N_{\mathbb{C}}^- w v_\lambda \cap V = N^- w v_\lambda \cup (-N^- w v_\lambda).$$

The constraints from E_λ discard the extra piece $-N^- w v_\lambda$. ∎

Let E be a finite dimensional real vector space. Let $\mathcal{S}(E)$, $\mathcal{P}(E)$ and $\mathcal{D}(E)$ denote the symmetric algebra, polynomial algebra and algebra of constant coefficient differential operators on E, respectively. Let $\mathcal{S}^m(E)$, $\mathcal{P}^m(E)$ and $\mathcal{D}^m(E)$ denote their m-th graded components. There is a natural isomorphism $\mathcal{S}^m(E) \cong \mathcal{D}^m(E)$ given by

$$w_1 \cdots w_m \mapsto \partial_{w_1} \cdots \partial_{w_m}$$

where

$$\partial_v f(x) = \lim_{h \to 0} \frac{f(x + hv) - f(x)}{h}$$

for $v \in E, f \in \mathcal{P}(E)$ and $x \in E$. Furthermore, there is a non-degenerate bilinear form

$$\langle \cdot, \cdot \rangle : \mathcal{D}(E) \times \mathcal{P}(E) \to \mathbb{R}$$

given by

$$\langle D, p \rangle = Dp(0).$$

The form $\langle \cdot, \cdot \rangle$ induces isomorphisms $\mathcal{D}^m(E) \cong \mathcal{P}^m(E)^*$ and $\mathcal{S}^m(E) \cong \mathcal{D}^m(E)$. Thus, we have

$$\mathcal{S}^m(E) \cong \mathcal{D}^m(E) \cong \mathcal{P}^m(E)^*.$$

Proposition 4.2.6. *For any $m \geq 0$, the isomorphism $\mathcal{S}^m(E) \cong \mathcal{P}^m(E)^*$ is given by* $\mathrm{sym}^m(v) \mapsto \eta_v$ *where $\eta_v(p) = p(v)$.*

Proof: It is enough to verify the statement for polynomials $x \mapsto \phi(x)^m$ where $\phi \in E^*$. Then we get that the image of $\mathrm{sym}^m(v)$ in $\mathcal{D}^m(E)$ is $\frac{1}{m!}\partial_v^m$ and

$$\partial_v \phi(x) = \lim_{t \to 0} \frac{1}{t}(\phi(x + tv) - \phi(x)) = \lim_{t \to 0} \frac{1}{t}(\phi(x) + t\phi(v) - \phi(x)) = \phi(v).$$

And it therefore follows from the Leibniz rule that $\frac{1}{m!}\partial_v^m(\phi^m) = \phi(v)^m$. ∎

Let

$$I = \{\phi \in \mathcal{P}(V_\lambda) : \phi|_{X_\lambda} = 0\}.$$

Recall that $\mathcal{P}(V_\lambda)$ is a G-module via the action

$$g.\phi(v) = \phi(g^{-1}.v).$$

Lemma 4.2.7. *I is a homogeneous and G-invariant ideal of $\mathcal{P}(V_\lambda)$.*

Proof: We know that I is G-invariant since X_λ is G-invariant. We now show that I is a homogeneous ideal. It suffices to show that each element $\phi \in I$ is a sum of homogeneous polynomials in I. Write $\phi = \phi_0 + \cdots + \phi_k$ where each ϕ_i is a homogeneous polynomial. We now show that each ϕ_i is in I.

Let $a_1, \ldots, a_n$ be integers such that $\sum_{i=1}^{n} = 0$ and $\sum_{i=1}^{n} a_i\lambda_i = N \in \mathbb{Z}\setminus\{0\}$. For every $g \in G$ and every $\mathbf{t}_a = \operatorname{diag}(t^{a_1}, \ldots, t^{a_n}) \in H$, we have

$$(g\mathbf{t}_a g^{-1}).g(v_\lambda) = g\mathbf{t}_a.v_\lambda = t^{\sum \lambda_i a_i} g.v_\lambda = t^N g.v_\lambda.$$

Set $v := g.v_\lambda$. Then

$$0 = \phi((g\mathbf{t}_a g^{-1}).v) = \phi(t^N v) = \phi_0(v) + \sum_{i=1}^{k} t^{iN}\phi_i(v).$$

Then define $f : \mathbb{R} \to \mathbb{R}$ by $f(t) = \phi((g\mathbf{t}_a g^{-1}).v)$. Note that f is a polynomial in t with coefficients $\phi_0(v), \ldots, \phi_k(v)$. But $f(t) = 0$ for all $t \in \mathbb{R}$, so $f = 0$, i.e. $\phi_i(v) = 0$ for each i. So each $\phi_i \in I$. ∎

For each $k \geq 0$, let

$$A_k := \{\phi|_{X_\lambda} : \phi \in \mathcal{P}^k(V_\lambda)\} = \mathcal{P}^k(V_\lambda)/(I \cap \mathcal{P}^k(V_\lambda)).$$

Since I is G-invariant, A_k is a G-module, and hence an $\mathfrak{h}$-module, where $\mathfrak{h} = \operatorname{Lie}(H)$. Moreover, A_k is finite dimensional and thus has a weight space decomposition

$$A_k = \bigoplus_{\mu \in \mathfrak{h}^*} A_k(\mu).$$

Lemma 4.2.8. *Fix $\mu \in \mathfrak{h}^*$. Suppose $\phi \in A_k(\mu)$ is such that $\phi(v_\lambda) \neq 0$. Then $\mu = -k\lambda$.*

Proof: Write $\mu = \mu_1\varepsilon_1 + \cdots + \mu_n\varepsilon_n$. Let $x = \operatorname{diag}(t_1, \ldots, t_n) \in H$. Then

$$\phi(x^{-1}.v_\lambda) = (x.\phi)(v_\lambda) = {t_1}^{\mu_1} \cdots {t_n}^{\mu_n}\phi(v_\lambda).$$

Note that we can choose $a_1, \ldots, a_n \in \mathbb{Z}$ such that $\sum a_i = 0$ but $N = \sum a_i\lambda_i \neq 0$. Let $t \in \mathbb{R}^*$ and set $\mathbf{t}_a := \mathrm{diag}(t^{a_1}, \ldots, t^{a_n})$. Then

$$t^{-kN}\phi(v_\lambda) = \phi(t^{-N}v_\lambda) = \phi({\mathbf{t}_a}^{-1}.v_\lambda) = t^{\sum a_i\mu_i}\phi(v_\lambda).$$

Since $t \in \mathbb{R}$ is arbitrary, $\sum a_i\mu_i = -kN$. Now let $\mathbf{t} = (t_1, \ldots, t_n) \in H$ be such that ${t_1}^{\lambda_1} \cdots {t_n}^{\lambda_n} = 1$. Then

$$t_1 \cdots t_n = 1 \text{ and } {t_1}^{\lambda_1} \cdots {t_n}^{\lambda_n} = 1.$$

Hence we have proven that for any element $\mathrm{diag}(t_1, \ldots, t_n) \in H$, if ${t_1}^{\lambda_1} \cdots {t_n}^{\lambda_n} = 1$ then ${t_1}^{\mu_1} \cdots {t_n}^{\mu_n} = 1$. Notice that $t_n = \frac{1}{t_1 \cdots t_{n-1}}$. Therefore, if

$$t_1^{\lambda_1-\lambda_n} \cdots t_{n-1}^{\lambda_{n-1}-\lambda_n} = 1$$

then

$$t_1^{\mu_1-\mu_n} \cdots t_{n-1}^{\mu_{n-1}-\mu_n} = 1$$

for every $\mathrm{diag}(t_1, \ldots, t_n) \in H$. Next we write $t_j = e^{2\pi i\theta_j}$, $a_j = \lambda_j - \lambda_n$, and $b_j = \mu_j - \mu_n$. Then if $e^{2\pi i \sum a_j\theta_j} = 1$ then $e^{2\pi i \sum b_j\theta_j} = 1$ for any $\theta \in [0, 2\pi)$, which means that if $\sum a_j\theta_j$ is an integer, then so is $\sum b_j\theta_j$. Thus if $\sum a_j\theta_j = 0$ then $\sum b_j\theta_j = 0$ and hence $(b_1, \ldots, b_n) = k(a_1, \ldots, k_n)$ for some $k \in \mathbb{Q}$ (given that all $a_j, b_j \in \mathbb{Z}$). Thus for each j we have $\mu_j - \mu_n = k(\lambda_j - \lambda_n)$ so $\mu_j = k\lambda_j + (\mu_n - k\lambda_n)$. ∎

Lemma 4.2.9. *As a G-module,*

$$\mathcal{S}^k(V_\lambda) = V_{k\lambda} \oplus \bigoplus_{\mu \prec k\lambda} V_\mu.$$

Proof: We prove the result for $k = 2$. The idea is the same for larger k.

We first show that each weight in $\mathcal{S}^2(V_\lambda)$ is at most 2λ in the ordering of the weights. Let $\{v_1, \ldots, v_n\}$ be a basis of weight vectors of V_λ, where each v_i has weight λ_i. Let $v = \sum a_{ij} v_i v_j$ be a weight vector of weight μ in $\mathcal{S}^2(V_\lambda)$. Let $H \in \mathfrak{h}^*$. Then

$$H \cdot v = \sum a_{ij} \mu(H) v_i v_j$$

and

$$H \cdot v = \sum a_{ij} (\lambda_i + \lambda_j)(H) v_i v_j.$$

Since H is arbitrary and there must exist i, j such that $a_{ij} \neq 0$, we have $\mu = \lambda_i + \lambda_j$ for some i, j. Since $\lambda_i, \lambda_j \preceq \lambda$, we get that $\mu \preceq 2\lambda$.

We now show that the 2λ weight space of $\mathcal{S}^2(V_\lambda)$ is one-dimensional. Indeed, let $v = \sum a_{ij} v_i v_j$ be a weight vector in $\mathcal{S}^2(V_\lambda)$ of weight 2λ. Then for any $H \in \mathfrak{h}^*$,

$$H \cdot v = \sum a_{ij} (\lambda_i + \lambda_j)(H) v_i v_j = \sum a_{ij} 2\lambda(H) v_i v_j.$$

Hence, each $a_{11} \neq 0$ and $a_{ij} = 0$ for $(i, j) \neq (1, 1)$. So $v = a_{11} v_1 v_1 \in \mathbb{R} v_\lambda v_\lambda$. So the 2λ weight space has dimension 1 in $\mathcal{S}^2(V_\lambda)$.

Finally, by Weyl's theorem of complete reducibility ([21, p. 46]), we see that we may decompose $\mathcal{S}^2(V_\lambda)$ as

$$\mathcal{S}^2(V_\lambda) = V_{2\lambda} \oplus V_{\mu_1} \oplus \cdots \oplus V_{\mu_l}$$

where $V_{\mu_1}, \ldots, V_{\mu_l}$ are highest weight modules of weights $\mu_1, \ldots, \mu_l$, respectively. By the above work, for each $1 \leq i \leq l$, we have $\mu_i \preceq 2\lambda$. But the 2λ weight space is one-dimensional, and hence each $\mu_i \prec 2\lambda$. This concludes our proof.

∎

From the isomorphism $\mathcal{P}^k(V_\lambda) \cong \mathcal{S}^k(V_\lambda)^*$ we make the identification

$$\mathcal{P}^k(V_\lambda) = V_{k\lambda}^* \oplus \bigoplus_{\mu \prec k\lambda} V_\mu^*. \tag{4.2.2}$$

Lemma 4.2.10. *For every V_μ^* in (4.2.2) with $\mu \neq k\lambda$, the image of V_μ^* under the map*

$$\mathcal{P}^k(V_\lambda) \to A_k$$

is zero.

Proof: We have a weight space decomposition

$$V_\mu{}^* = \bigoplus V_\mu^*(\eta).$$

From Lemma 4.2.9, we see that each η in this decomposition satisfies $\eta \neq -k\lambda$. Thus if $\phi \in V_\mu{}^*(\eta)$ we have $\phi(v_\lambda) = 0$. Since V_μ^* is G-invariant, the set

$$T := \{x \in V_\lambda : \phi(x) = 0 \text{ for all } \phi \in V_\mu^*\}$$

is also G-invariant. Since $v_\lambda \in T$, we then obtain $X_\lambda \subseteq T$. ∎

Proposition 4.2.11. *As a G-module, $A_k \cong V_{k\lambda}^*$.*

Proof: From Lemma 4.2.10, we know that the map $P^k(V_\lambda) \to A_k$ induces a map $V_{k\lambda}^* \to A_k$. Since $V_{k\lambda}^*$ is irreducible and this map is surjective, either $A_k = 0$ or the map is an isomorphism. But $A_k \neq 0$, since one can choose an element $f \in A_k$ such that $f(v_\lambda) \neq 0$. ∎

Proposition 4.2.12. *Let $v \in V_\lambda$ and let X_λ be as in Theorem 4.2.5. Then $v \in X_\lambda$ if and only if $v \otimes v \in V_{2\lambda}$.*

Proof: Suppose $v \in X_\lambda$. Then $v = g.v_\lambda$ for some $g \in G$. Then $v \otimes v = (g.v_\lambda) \otimes (g.v_\lambda) = g.(v_\lambda \otimes v_\lambda)$. But $v_\lambda \otimes v_\lambda \in V_{2\lambda}$, and $V_{2\lambda}$ is G-invariant. So $v \otimes v \in V_{2\lambda}$.

Conversely, suppose $v \in V_\lambda$ and $v \otimes v \in V_{2\lambda}$. Then we need to show that for all $\phi \in I$ we have $\phi(v) = 0$. But we know from the last proposition that

$$I = \bigoplus_{k=0}^{\infty} (I \cap \mathcal{P}^k(V_\lambda)) = \bigoplus_{k=0}^{\infty} \bigoplus_{\mu \prec k\lambda} V_\mu^*.$$

Thus, to show that $v \in X$, we let $\phi \in V_\mu^*$ where $V_\mu^* \subseteq \mathcal{P}^k(V_\lambda)$ is one of the components above, and we need to show $\phi(v) = 0$. To see this, recall we have the G-invariant non-degenerate pairing

$$\mathcal{S}^k(V_\lambda) \times \mathcal{P}^k(V_\lambda) \to \mathbb{R}.$$

Then $\phi(v) = \langle \mathrm{sym}^k(v), \phi \rangle$. But we know that this form, when restricted to $V_{k\lambda} \times V_\mu^*$, must vanish. Therefore $\phi(v) = 0$. ∎

The following theorem shows that X_λ is a locus of quadratic equations.

Proposition 4.2.13. *Let $\mathbf{C}$ be the Casimir element of $\mathrm{sl}_n(\mathbb{R})$ and let (π, V_λ) be an irreducible highest weight G-module with highest weight λ. Let $V_{2\lambda}$ be the highest weight representation of $\mathfrak{sl}_n(\mathbb{R})$ with weight 2λ. Let $C(2\lambda)$ be value that $\mathbf{C}$ acts on $V_{2\lambda}$ by (see Theorem 3.2.12). Then we have*

$$V_{2\lambda} = \{a \in V_\lambda \otimes V_\lambda : \mathbf{C}a = C(2\lambda)a\}.$$

Proof: We first note that if $\mu \prec \lambda$ are dominant weights, then $C(\mu) < C(\lambda)$.

Indeed, $C(\lambda) - C(\mu) = (\lambda+\rho, \lambda+\rho) - (\mu+\rho, \mu+\rho)$. One can write $\mu = \lambda - \gamma$, where γ is a positive combination of positive roots. Then $(\lambda+\rho, \lambda+\rho) - (\mu+\rho, \mu+\rho) = (\lambda+\mu+2\rho, \gamma)$. Since $\lambda+\mu+2\rho$ is a regular dominant weight, and γ is a nonzero sum of positive roots, we get $(\lambda+\mu+2\rho, \gamma) > 0$. Thus, $C(\mu) < C(\lambda)$.

Now, we can decompose $V_\lambda \otimes V_\lambda = V_{2\lambda} \oplus \bigoplus_{\mu \prec \lambda} V_\mu$. This shows that the elements of $V_\lambda \otimes V_\lambda$ which are eigenvectors of $\mathbf{C}$ (recall Definition 1.3.9) with eigenvalue $C(2\lambda)$ are precisely the vectors in $V_{2\lambda}$. ∎

As mentioned in the begining of this section, Kostant's theorem (Theorem 4.2.3) is an immediate consequence of Proposition 4.2.5, Proposition 4.2.12 and Proposition 4.2.13.

4.2.2 Equations for the Highest Weight Orbit for $\mathrm{SL}_2(\mathbb{R})$

In this section we use Proposition 4.2.13 to write down the quadratic equations which determine X_λ for the 5-dimensional irreducible highest weight module of $\mathrm{SL}_2(\mathbb{R})$.

Let V_d be the highest weight representation of $\mathrm{SL}_2(\mathbb{R})$ as defined in Section 3.2. Recall that our Casimir operator for $\mathrm{SL}_2(\mathbb{R})$ is given by

$$\mathbf{C} = \frac{1}{2}FE + \frac{1}{4}H + \frac{1}{8}H^2.$$

Let $\{v_0, \ldots, v_d\}$ be the basis of V_d as given in Proposition 3.2.2. Let $0 \le i, j \le d$.

Then

$$\begin{aligned}
FE(v_i \otimes v_j) &= F(Ev_i \otimes v_j + v_i \otimes Ev_j) \\
&= F((i-d)v_{i+1} \otimes v_j + (j-d)v_i \otimes v_{j+1}) \\
&= (i-d)(Fv_{i+1} \otimes v_j + v_{i+1} \otimes Fv_j) + (j-d)(Fv_i \otimes v_{j+1} + v_i \otimes Fv_{j+1}) \\
&= ((d-i)(i+1) + (d-j)(j+1))v_i \otimes v_j \\
&+ (d-i)jv_{i+1} \otimes v_{j-1} + (d-j)iv_{i-1} \otimes v_{j+1}
\end{aligned}$$

and $\dfrac{1}{4}H(v_i \otimes v_j) = \frac{1}{2}(i+j-d)v_i \otimes v_j$ and $\dfrac{1}{8}H^2(v_i \otimes v_j) = \frac{1}{2}(i+j-m)^2 v_i \otimes v_j$. so we get

$$\begin{aligned}
\mathbf{C}(v \otimes v) &= \sum\sum x_i x_j C(v_i \otimes v_j) \\
&= \frac{1}{2}\sum\sum x_i x_j (2ij - id - jd + d^2 + d)v_i \otimes v_j \\
&+ \sum\sum x_i x_j (d-i)jv_{i+1} \otimes v_{j-1} \\
&+ \sum\sum x_i x_j (d-j)iv_{i-1} \otimes v_{j+1} \\
&= \sum\sum \frac{1}{2}(x_i x_j (2ij - id - jd + d^2 + d) + x_{i-1}x_{j+1}(d-i+1)(j+1) \\
&\quad + x_{i+1}x_{j-1}(d-j+1)(i+1))v_i \otimes v_j
\end{aligned}$$

so our equation $\mathbf{C}(v \otimes v) - C(2\lambda)(v \otimes v) = 0$ reduces to

$$\begin{aligned}
&\sum\sum \frac{1}{2}(x_i x_j (2ij - id - jd + d^2 + d - C(2\lambda)) + \\
&\quad x_{i-1}x_{j+1}(d-i+1)(j+1) + x_{i+1}x_{j-1}(d-j+1)(i+1))v_i \otimes v_j = 0
\end{aligned}$$

For each $0 \leq i, j \leq d$ let

$$\begin{aligned}\alpha_{ij} &= \frac{1}{2}(2ij - id - jd + d^2 + d - C(2\lambda))\\ \beta_{ij} &= \frac{1}{2}((d - i + 1)(j + 1))\\ \gamma_{ij} &= \frac{1}{2}((d - j + 1)(i + 1))\end{aligned}$$

For each $0 \leq i, j \leq d$, let

$$P_{ij}(x_0, \ldots, x_d) := \alpha_{ij} x_i x_j + \beta_{ij} x_{i-1} x_{j+1} + \gamma_{ij} x_{i+1} x_{j-1}.$$

When working in $\mathcal{S}^2(V_d)$, we see the following. If $i \neq j$, then the coefficient of $v_i v_j$ in $\mathbf{C}(v \otimes v) - C(2\lambda)$ is

$$P_{ij}(x_0, \ldots, x_d) + P_{ji}(x_0, \ldots, x_d) = (\alpha_{ij} + \alpha_{ji}) x_i x_j + (\beta_{ij} + \gamma_{ji}) x_{i-1} x_{j+1} + (\gamma_{ij} + \beta_{ji}) x_{i+1} x_{j-1}.$$

If $i = j$, then the coefficient of $v_i v_j = v_i v_i$ is $P_{ij}(x_0, \ldots, x_d)$.

Example 4.2.14. Let $d = 4$. Then the equations for X_λ are given by

1. $v_0 v_0 : 0 = 0$
2. $v_0 v_1 : 0 = 0$
3. $v_0 v_2 : 3x_1^2 - 8x_0 x_2 = 0$
4. $v_0 v_3 : 2x_1 x_2 - 12x_0 x_3 = 0$
5. $v_0 v_4 : x_1 x_3 - 16x_0 x_4 = 0$
6. $v_1 v_1 : 8x_0 x_2 - 3x_1^2 = 0$
7. $v_1 v_2 : 12x_0 x_3 - 2x_1 x_2 = 0$
8. $v_1 v_3 : 4x_2^2 + 16x_0 x_4 - 10x_1 x_3 = 0$
9. $v_1 v_4 : 2x_2 x_3 - 12x_1 x_4 = 0$

10. $v_2v_2 : 9x_1x_3 - 4x_2^2 = 0$
11. $v_2v_3 : 12x_1x_4 - 2x_2x_3 = 0$
12. $v_2v_4 : 3x_3^2 - 8x_2x_4 = 0$
13. $v_3v_3 : 8x_2x_4 - 3x_3^2 = 0$
14. $v_3v_4 : 0 = 0$
15. $v_4v_4 : 0 = 0$

and the constraints from E_λ are given

$$x_4 > 0 \text{ or } (x_4 = 0 \text{ and } x_0 > 0).$$

Chapter 5

The Convex Hull of the Highest Weight Orbit

In the previous chapter, we have seen that X_λ is a semialgebraic over $\mathbb{R}$, and thus, by Proposition 1.2.6, the set $\mathrm{conv}(X_\lambda)$ is semi-algebraic. The goal of this section is to study this set for $G = \mathrm{SL}_2(\mathbb{R})$.

5.1 The Relationship Between Convex Hulls of the G and K-orbits

Throughout this section, G is a split real semisimple Lie group with finite center, and K is a maximal compact subgroup of G (see Section 2.2). In this section, we state and prove some preliminary results about the convex hull of X_λ. In doing this, we explain how $G.v_\lambda$ is related to the convex hull of the orbit $K.v_\lambda$, where K is the maximal compact subgroup of G. This will be useful, since orbits of compact groups can be easier to understand than those of non-compact groups. Our main reference is [7]. All representations will be real. We let $\mathfrak{g} = \mathrm{Lie}(G)$ and we let $\mathfrak{h}$ be a chosen Cartan subalgebra of $\mathfrak{g}$. Also, we denote the Cartan involution of G by Θ.

Lemma 5.1.1. *Let (π, V) be an irreducible representation of G. Then there exists an inner product $(\cdot,\cdot)$ on V such that*

$$\pi(g)^* = \pi(\Theta(g)^{-1})$$

for all $g \in G$.

Proof: Let $\mathfrak{g}_c = \mathfrak{k} \oplus i\mathfrak{p}$, and let $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$. Then we know that $\mathfrak{g}_c$ corresponds to a compact group G_c, since the Killing form on $\mathfrak{g}_c$ is negative definite. The real representation π of $\mathfrak{g}$ on V induces a complex representation $\pi' : \mathfrak{g}_c \to \mathfrak{gl}(V_{\mathbb{C}})$ where $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$ and π' is defined by $\pi'(X + iY) = \pi(X) + i\pi(Y)$ for all $X \in \mathfrak{k}$ and $Y \in \mathfrak{p}$. We can consider π' as a representation of G_c as well, and hence, by the compactness of G_c, there exists a K-invariant inner product $\langle\cdot,\cdot\rangle$ on $V_{\mathbb{C}}$. This implies that for all $X \in \mathfrak{k}, Y \in \mathfrak{p}, v, w \in V$ we have

$$\langle \pi'(X + iY)v, w\rangle = -\langle v, \pi'(X + iY)w\rangle$$

i.e.

$$\langle \pi'(X)v, w\rangle = -\langle v, \pi'(X)w\rangle \text{ and } \langle \pi'(Y)v, w\rangle = \langle v, \pi'(Y)w\rangle.$$

Then define the inner product $(\cdot,\cdot)$ on V by $(v, w) = \Re\langle v, w\rangle$. This inner product satisfies the desired relation. ∎

When given a representation (π, V) of G or $\mathfrak{g}$, we always equip V with the inner product $(\cdot,\cdot)$.

Proposition 5.1.2. *Let (π, V) be a finite dimensional irreducible representation of G. Then* $\dim V^K \leq 1$.

Proof: This argument is similar to [19, Proposition 4.2]. Suppose $\dim V^K \geq 2$. Since V is irreducible, it is a highest weight module of some weight $\lambda \in \mathfrak{h}^*$. Let

$$V' = \bigoplus_{\mu \neq \lambda} V_\mu.$$

Then since $\dim V_\lambda = 1$ and $\dim V^K \geq 2$, we must have $V^K \cap V' \neq 0$. Now let $v \in V^K \cap V'$ be a nonzero weight vector. Then the PBW theorem (Theorem 1.3.8) and the Iwasawa decomposition (Theorem 2.1.8) imply that

$$V = \mathcal{U}(\mathfrak{g})v = \mathcal{U}(\mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{k})v = \mathcal{U}(\mathfrak{n}^-)v \subseteq V'$$

which is a contradiction. ∎

Definition 5.1.3. We say a finite dimensional irreducible representation (π, V) of G is *spherical* if $\dim V^K = 1$.

By Proposition 5.1.2, we see that (π, V) being spherical is equivalent to the existence of a nonzero K-fixed vector $u \in V$.

Definition 5.1.4. Let V be a real vector space. We say a set $C \subseteq V$ is a *cone* if, for any $x \in C$ and any $r > 0$, $rx \in C$.

Definition 5.1.5. Let V be a vector space, and let $C \subseteq V$ be a cone. We say C is *pointed* if $C \cap (-C) = \{0\}$. We say C is *generating* if $\operatorname{span}(C) = \mathbb{R}^n$. If C is pointed, generating, and closed, we say C is *regular*.

Definition 5.1.6. For a cone $C \subseteq \mathbb{R}^n$, we define the *dual cone* of C to be the set

$$C^* := \{v \in V : (v, w) \geq 0 \text{ for all } w \in C \setminus 0\}.$$

Remark 5.1.7. If G is a Lie group and (π, V) is a representation of G, we say C is *G-invariant* if $\pi(g)v \in C$ for all $g \in G$ and $v \in C$. If the group G and representation (π, V) are understood from the context, we simply say C is *invariant*.

Proposition 5.1.8. *Let (π, V) be an irreducible representation of G. If V contains an invariant regular cone, then (π, V) is spherical.*

Proof: Equip V with the inner product of Proposition 5.1.1. Let C be an invariant regular cone in V. There exists $v \in C^*$ such that $(u, v) > 0$ for all $u \in C \setminus \{0\}$. Fix $u \in C \setminus 0$. Then $(\pi(k)u, v) > 0$ for all $k \in K$. Thus, the vector

$$u_K := \int_K \pi(k)udk$$

is a member of C and is K-fixed (we are integrating with the Haar measure on K). Note that $u_K \neq 0$ since

$$(u_K, v) = (\int_K \pi(k)u, v) = \int_K (\pi(k)u, v)dk > 0$$

so $u_K \neq 0$. Thus (π, V) is spherical. ∎

Corollary 5.1.9. *Let (π, V) be an irreducible representation of G with the inner product of Proposition 5.1.1. If C is an invariant regular cone in V, then C contains a K-fixed unit vector.*

Proof: This is the vector u_K from Propositon 5.1.8. ∎

If (π, V) is spherical, then there are exactly two unit vectors in V^K. Fix one, and call it u_0. The other unit vector is $-u_0$. We say a cone $C \subseteq V$ is *maximal* if it is maximal with respect to inclusion in the collection of all invariant regular cones containing u_0. Similarly, we say a cone is *minimal* if it is minimal with respect to inclusion in the collection of all invariant regular cones containing u_0. Since V is irreducible, it is a highest weight module with some highest weight $\lambda \in \mathfrak{h}^*$. We fix a highest weight vector $v_\lambda \in V$ which satisfies $(u_0, v_\lambda) > 0$.

Theorem 5.1.10. *Let (π, V) be a spherical irreducible representation of G. Then there exists a unique invariant minimal cone C_{min} given by*

$$C_{min} = \overline{\mathbb{R}^+ \mathrm{conv}(G \cdot u_0)}.$$

Proof: Let $C_0 = \overline{\mathbb{R}_+ \mathrm{conv}(G \cdot u_0)}$. We first show that C_0 is a regular invariant cone. Firstly, it is clear that C_0 is closed and invariant. Note that since V is irreducible, and $\mathrm{span}(C_0)$ is an invariant subspace of V, we have that $\mathrm{span}(C_0) = V$, i.e. C_0 is generating. We now show that C_0 is pointed. Now let $g \in G$. By the polar decomposition $G = KP$, we can write $g = kp$ where $k \in K$ and $p = e^X$ where $X \in \mathfrak{p}$. Thus,

$$\begin{aligned}(\pi(g)u_0, u_0) &= (\pi(k)\pi(e^X)u_0, u_0) \\ &= (\pi(e^X)u_0, \pi(k^{-1})u_0) \\ &= (\pi(e^X)u_0, u_0) \\ &= (\pi(e^{1/2X}u_0, \pi(e^{1/2X})u_0) > 0\end{aligned}$$

Let $v_1 = g_1 \cdot u_0$ and $v_2 = g_2 \cdot u_0$ be members of $G \cdot u_0$, where $g_1, g_2 \in G$. Let

$x = \Theta^{-1}(g_2)$. Then

$$\begin{aligned}(v_1, v_2) &= (\pi(g_1)u_0, \pi(g_2)u_0) \\ &= (u_0, \pi(\Theta^{-1}(g_1)g_2)u_0) \\ &= (u_0, \pi(\Theta^{-1}(g_1x)u_0) \\ &= (\pi(g_1x)u_0, u_0) > 0\end{aligned}$$

where the inequality follows from the above fact that $(\pi(g)u_0, u_0) > 0$ for all $g \in G$. So if $v \in C \cap (-C)$, then $(v, -v) \geq 0$, which implies $v = 0$. So C_0 is pointed. Thus, C_0 is a regular invariant cone. Therefore, $C_{min} \subseteq C_0$. To see the opposite inclusion, note that C_{min} contains $G \cdot u_0$, is closed and convex, and therefore $C_0 \subseteq C_{min}$. ∎

Lemma 5.1.11. *Let (π, V_λ) be an irreducible representation of G with highest weight λ, and let $v_\lambda \in V_\lambda$ be a highest weight vector. Then*

$$\text{conv}(G \cdot v_\lambda \cup \{0\}) = \mathbb{R}_{\geq 0}\text{conv}(K.v_\lambda).$$

Proof: Let $C_0 = \text{conv}(G \cdot v_\lambda \cup \{0\})$ and $C_1 = \mathbb{R}_{\geq 0}\text{conv}(K.v_\lambda)$. We prove the inclusions $C_1 \subseteq C_0$ and $C_0 \subseteq C_1$. To begin, notice that we have the Iwasawa decomposition $G = KAN$. Recall that $AN.v_\lambda = \mathbb{R}_+ v_\lambda$. Therefore, the nonzero points of $\text{conv}(G \cdot v_\lambda \cup \{0\})$ have the form

$$x = \sum_i \lambda_i c_i x_i$$

where $\sum_i \lambda_i \leq 1, \lambda_i \geq 0, c_i \in \mathbb{R}^+$ and $x_i \in K.v_\lambda$. One can write such a point as

$$\left(\sum_i \lambda_i c_i\right)\left(\sum_i \mu_i x_i\right) \qquad \text{where } \mu_i = \frac{\lambda_i c_i}{\sum \lambda_i c_i}$$

which shows that $x \in \mathbb{R}_{\geq 0}\mathrm{conv}(K.v_\lambda)$. The other inclusion is clear. ∎

Lemma 5.1.12. *Let (π, V_λ) be an irreducible representation of G with highest weight λ, and let $v_\lambda \in V_\lambda$ be a highest weight vector. The cone* $\mathrm{conv}(G \cdot v_\lambda \cup \{0\})$ *is closed, generating, and invariant.*

Proof: Let $C_0 = \mathrm{conv}(G \cdot v_\lambda \cup \{0\})$. We have the Iwasawa decomposition KAN, and we know that $AN \cdot v_\lambda = \mathbb{R}_+ v_\lambda$. It thus follows that C_0 is indeed a cone. It is also clear that C_0 is invariant, and thus $\mathrm{span}(C_0)$ is an invariant subspace of V. Since V is irreducible and $C_0 \neq \{0\}$, it follows that $\mathrm{span}(C_0) = V$, so C_0 is generating.

We now show that C_0 is closed. Indeed, let (x_n) be a sequence in C_0 converging to a point $x \in V$. By Lemma 5.1.11, for each n, we can write

$$x_n = \alpha_n y_n$$

where $\alpha_n \geq 0$ and y_n is a sequence in $\mathrm{conv}(K \cdot v_\lambda)$. For each n, we can write $y_n = \sum t_{i,n} \pi(k_{i,n}) v_\lambda$ where $t_{i,n} > 0$, $\sum t_{i,n} = 1$ and $k_{i,n} \in K$. Since K is compact, so is $K \cdot v_\lambda$ (as our action is continuous). Moreover, we know that the convex hull of a compact set is compact, so $\mathrm{conv}(K \cdot v_\lambda)$ is compact. Thus, by passing to a subsequence of (y_n), we may assume that (y_n) converges to a point $y \in \mathrm{conv}(K \cdot v_\lambda)$.

Now, since (x_n) converges, we know it is bounded. Moreover,

$$\begin{aligned}(x_n, u_0) &= \alpha_n \sum t_{i,n}(\pi(k_{i,n})v_\lambda, u_0)\\ &= \alpha_n \sum t_{i,n}(\pi(k_{i,n})v_\lambda, \pi(k_{i,n})u_0)\\ &= \alpha_n \sum t_{i,n}(v_\lambda, u_0)\\ &= \alpha_n(v_\lambda, u_0).\end{aligned}$$

Thus, by the Cauchy-Schwarz theorem,

$$|\alpha_n| = \frac{|(a_n, u_0)|}{|(v_\lambda, u_0)|} \leq \frac{||x_n||^{1/2}||u_0||^{1/2}}{|(v_\lambda, u_0)|}$$

and so $|\alpha_n|$ is bounded. Thus, by the Bolzano-Weierstrass theorem, we may pass to a subsequence of (α_n) and assume that (α_n) converges to a point $\alpha \geq 0$. Thus, since both (α_n) and (y_n) converge, the sequence (x_n) defined by $x_n = \alpha_n y_n$ converges to αy. So C_0 is indeed closed. ∎

Theorem 5.1.13. *Let (π, V_λ) be an irreducible spherical representation of G with highest weight λ, and let $v_\lambda \in V_\lambda$ be a highest weight vector. Then the minimal cone is given by $C_{min} = \mathrm{conv}(G \cdot v_\lambda \cup \{0\})$.*

Proof: Let $C_0 = \mathbb{R}_{\geq 0}\mathrm{conv}(K.v_\lambda)$. By Corollary 5.1.9, we know that C_0 contains the spherical vector u_0. Moreover, we know that C_0 is closed (by Lemma 5.1.12) and is invariant. Thus,

$$\overline{\mathbb{R}_+\mathrm{conv}(G \cdot u_0)} \subseteq C_0$$

i.e. $C_{min} \subseteq C_0$. It remains to prove the opposite inclusion.

We fix an orthonormal basis of $\mathfrak{h}$-weight vectors $e_1, \ldots, e_d$ for V, such that $e_1 = v_\lambda$, where each e_i has weight λ_i, and $\lambda_1 = \lambda$. Then for $H \in \mathfrak{h}^+$ (the fundamental

Weyl chamber) we have

$$0 < e^{\lambda_i(H)} < e^{\lambda_1(H)}$$

for $i \geq 2$. Now we write

$$u_0 = \sum_{i=1}^{d} (u_0, e_i) e_i.$$

Thus

$$\pi(e^H) u_0 = \sum_{i=1}^{d} (u_0, e_i) e^{\lambda_i(H)} e_i.$$

Therefore,

$$\lim_{t \to \infty} e^{-t\lambda(H)} \pi(e^{tH}) u_0 = (u_0, v_\lambda) v_\lambda$$

and consequently $v_\lambda \in C_{min}$, as C_{min} is closed. This implies the desired inclusion by the invariance of C_{min}. ∎

Remark 5.1.14. In Example 5.1.15 and Proposition 5.1.16, we use the fact that for a spherical representation (π, V), we have $C_{max} = C_{min}^*$. See [7, Theorem II.2.2] for a proof.

Example 5.1.15. Let $G = \mathrm{SL}_n(\mathbb{R})$ and let $V = \mathrm{Sym}(n, \mathbb{R})$ be the space of $n \times n$ symmetric matrices, equipped with the inner product $(\cdot, \cdot)$ given by

$$(u, v) = \mathrm{tr}(uv).$$

Let $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{R})$ and let $\mathfrak{h}$ be the Cartan subalgebra of $\mathfrak{g}$ consisting of diagonal matricies in $\mathfrak{g}$. Let (π, V) be the representation given by

$$\pi(g)v = gvg^t$$

where g^t is the transpose of g. Recall that we have the maps $\varepsilon_i \in \mathfrak{h}^*$, $1 \leq i \leq n$, given by

$$\varepsilon_i(\mathrm{diag}(a_1, \ldots, a_n)) := a_i.$$

We can see that this representation is irreducible with highest weight $2\varepsilon_1$ (see [7, p. 25]). The subgroup $K = SO(n)$ is a maximal compact subgroup of G. The representation π is spherical with a K-fixed vector of $u_0 = I_n$ (identity matrix). It is not too difficult to see that $C_{min} = \overline{\mathbb{R}_+ \mathrm{conv}(G \cdot u_0)}$ consists of all positive semidefinite matricies in $\mathrm{Sym}(n, \mathbb{R})$. Moreover, this cone is self-dual, i.e.

$$C_{min} = C_{max}.$$

The next proposition shows that the situation of Example 5.1.15 is special.

Proposition 5.1.16. *Let $d \geq 4$ be even. Then the minimal and maximal cones of the irreducible $(d+1)$-dimensional representation V of $\mathrm{SL}_2(\mathbb{R})$ satisfy*

$$C_{min} \neq C_{max}.$$

Proof: Recall that $C_{max} = C^*_{min}$, i.e.

$$C_{max} = \{v \in V : (v, w) \geq 0 \text{ for all } w \in C_{min}\}.$$

Our strategy is to find an element $v \in C_{max} \setminus C_{min}$.

Recall that we have the basis $\{v_0, \ldots, v_d\}$ of V given by $v_i = x^{d-i}y^i$. We first show that if $v = \sum_{i=0}^d c_i v_i \in C_{min}$, then if $c_0 \neq 0$ and $c_d \neq 0$, then $c_{d/2} \neq 0$. It is sufficient to prove this for each vector $v \in X = (\mathrm{SL}_2(\mathbb{R}).y^d) \cup \{0\}$. Indeed, let $g \in \mathrm{SL}_2(R)$ and suppose

$$g^{-1} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}).$$

Then

$$g.y^d = (\gamma x + \delta y)^d = \sum_{k=0}^{d} \binom{d}{k} \gamma^k \delta^{d-k} x^k y^{d-k}.$$

Setting $v := g.y^d$, we get that for $0 \leq k \leq d$, $c_k = \binom{d}{k}\gamma^k \delta^{d-k}$. If $c_0 \neq 0$ and $c_d \neq 0$, then $\gamma^d \neq 0$ and $\delta^d \neq 0$. Thus, $\gamma \neq 0$ and $\delta \neq 0$. Consequently, $c_{d/2} = \binom{d}{d/2}\gamma^{d/2}\delta^{d/2} \neq 0$.

From the above reasoning, we note that if $v \in X$, then either $c_0 > 0$ or $c_d > 0$. This also holds if $v \in C_{min}$.

We now find a vector $x = \sum_{i=0}^{d} x_i v_i \in C_{max}$ which satisfies $x_0 \neq 0$ and $x_d \neq 0$, but not $x_{d/2} \neq 0$. Indeed, set $x_0 = x_d = 1$ and $x_i = 0$ for $0 < i < d$. Then, for $v = \sum_{i=0}^{d} c_i v_i \in C_{min}$, we get

$$\sum_{i=0}^{d} x_i c_i = x_0 c_0 + x_d c_d = c_0 + c_d > 0.$$

So $x \in C_{max}$ but $x \notin C_{min}$. ∎

5.2 Spherical $\mathrm{SL}_2(\mathbb{R})$ Modules

Recall from Section 3.2 that for each $d \geq 0$, there is an $\mathrm{SL}_2(\mathbb{R})$-module (π_d, V_d), where V_d is the space of degree-d homogeneous polynomials in x and y with coefficients in $\mathbb{R}$, and π_d is given by $\pi_d(g)f(x,y) = f(g^{-1}(x,y))$ where $g \in \mathrm{SL}_2(\mathbb{R})$.

We may restrict this representation to $\mathrm{SO}_2(\mathbb{R})$. For $\theta \in \mathbb{R}$, define

$$g_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Note that

$$\mathrm{SO}_2(\mathbb{R}) = \{g_\theta : 0 \leq \theta < 2\pi\}.$$

We may restrict the action of $\mathrm{SL}_2(\mathbb{R})$ on V_d to $\mathrm{SO}_2(\mathbb{R})$ on V_d, and we obtain

$$\pi_d(g_\theta)f(x,y) = f((\cos\theta)x + (\sin\theta)y, -(\sin\theta)x + (\cos\theta)y)$$

for all $g_\theta \in \mathrm{SO}_2(\mathbb{R})$.

Note that $\mathrm{SO}_2(\mathbb{R})$ also acts on the complexification $V_d \otimes_{\mathbb{R}} \mathbb{C}$. For each $0 \leq k \leq d$, define

$$q_k(x,y) := (x+iy)^k(x-iy)^{d-k}$$

and note that for $0 \leq k \leq d$, $q_{d-k} = \overline{q_k}$. For each $0 \leq k \leq d$, $q_k \in V_d \otimes_{\mathbb{R}} \mathbb{C}$. Then

$$\pi_d(g_\theta)q_k(x,y) = e^{i\theta(d-2k)}q_k(x,y).$$

There exist $A_k, B_k \in V_d$ such that $q_k = A_k + iB_k$, where $a, b \in \mathbb{R}$. Let $e^{i\theta(d-2k)} = a+ib$. Then

$$\pi_d(g_\theta)q_k = (aA_k - bB_k) + i(aB_k + bA_k).$$

Hence,

$$\begin{aligned}\pi_d(g_\theta)A_k &= \pi_d(g_\theta)(\frac{1}{2}(q_k + q_{d-k})) \\ &= aA_k - bB_k\end{aligned}$$

and

$$\begin{aligned}\pi_d(g_\theta)B_k &= \pi_d(g_\theta)(\frac{1}{2i}(q_k - q_{d-k})) \\ &= aB_k + bA_k\end{aligned}$$

We define

$$W_k = \mathrm{span}_{\mathbb{R}}\{A_k, B_k\}$$

and see that the above calculations imply that W_k is $\mathrm{SO}_2(\mathbb{R})$-invariant. Note that if d is even, then $q_{\frac{d}{2}}(x, y) = A_{\frac{d}{2}}(x, y)$, so $B_{\frac{d}{2}}(x, y) = 0$.

Lemma 5.2.1. *We may decompose V_d as*

$$V_d = W_0 \oplus \cdots \oplus W_{\lfloor \frac{d}{2} \rfloor}.$$

Proof: Note that $\{q_0, \ldots, q_d\}$ is a basis of $V_d \otimes_{\mathbb{R}} \mathbb{C}$. Thus, $\{q_0, \ldots, q_d\}$ is linearly independent over $\mathbb{C}$. If d is even, define $S := \{A_0, \ldots, A_{\frac{d}{2}}, B_0, \ldots, B_{\frac{d}{2}-1}\}$ and if d is odd, define $S := \{A_0, \ldots, A_{\frac{d-1}{2}}, B_0, \ldots, B_{\frac{d-1}{2}}\}$. In both cases, $|S| = d + 1$ and

$$\mathrm{span}_{\mathbb{C}} S = \mathrm{span}_{\mathbb{C}}\{q_0, \ldots, q_d\}.$$

Thus, S is a basis of $V_d \otimes_{\mathbb{R}} \mathbb{C}$ and hence is linearly independent over $\mathbb{C}$. Thus, S is linearly independent over $\mathbb{R}$. Since S is linearly independent over $\mathbb{R}$ and $|S| = d + 1 = \dim_{\mathbb{R}} V_d$, we get that S is a basis of V_d. Since $W_k = \mathrm{span}_{\mathbb{R}}\{A_k, B_k\}$ for each $0 \leq k \leq \lfloor \frac{d}{2} \rfloor$, our claim is proven. ∎

Proposition 5.2.2. *Let $d \geq 2$ be even. Then the highest weight vector $y^d \in V_d$ has a non-zero component in each W_k, $0 \leq k \leq \frac{d}{2}$.*

Proof: We may write y^d as

$$\begin{aligned} y^d &= \frac{1}{i^d 2^d}((x + iy) - (x - iy))^d \\ &= \frac{1}{i^d 2^d} \sum_{k=0}^{d} \binom{d}{k} q_k(x, y). \end{aligned}$$

But $\binom{d}{k} = \binom{d}{d-k}$ and $q_k = \overline{q_{d-k}}$ for all $0 \le k \le d$. Moreover, $q_{\frac{d}{2}}(x,y) = A_{\frac{d}{2}}(x,y)$. So

$$y^d = \frac{1}{i^d 2^{d-1}} \sum_{k=0}^{\frac{d}{2}-1} \binom{d}{k} A_k + \frac{1}{i^d 2^d} \binom{d}{\frac{d}{2}} A_{\frac{d}{2}} \tag{5.2.1}$$

which has a non-zero component in each W_k, for $0 \le k \le \frac{d}{2}$. ∎

Remark 5.2.3. Recall from Section 3.3 that the irreducible real representations of $SO_2(\mathbb{R})$ have the form (ρ_k, U_k) where $k \in \mathbb{Z}$, and $U_k = \mathbb{R}$ if $k = 0$ and $U_k = \mathbb{R}^2$ if $k \neq 0$. If $k \neq 0$, then ρ_k is defined by $\rho_k(g_\theta) := g_{k\theta}$, for $\theta \in [0, 2\pi)$. We define ρ_0 to be the trivial representation.

Definition 5.2.4. For every $(d+1)$-tuple $A = (a_0, \ldots, a_d)$ of integers satisfying $0 \le a_0 \le \cdots \le a_d$, we define

$$\rho_A := \rho_{a_0} \oplus \cdots \oplus \rho_{a_d}$$

on $U_A := U_{a_0} \oplus \cdots \oplus U_{a_d}$.

Definition 5.2.5. Let $A = (a_1, \ldots, a_d)$ be a d-tuple of integers where $0 < a_1 \le \cdots \le a_d$, and consider the representation (ρ_A, U_A) of $SO_2(\mathbb{R})$. Let $(1,0)^d$ be the vector $((1,0),(1,0),\ldots,(1,0)) \in U_A$. We define C_A by

$$C_A := \operatorname{conv}(\rho_A(SO_2(\mathbb{R})) \cdot (1,0)^d) \subseteq U_A = \mathbb{R}^2 \oplus \cdots \oplus \mathbb{R}^2 \cong (\mathbb{R}^2)^d.$$

If $A = (1, 2, \ldots, d)$ then we write C_d instead of C_A, and we call C_d a *universal Carathéodory orbitope*.

Remark 5.2.6. One can easily see that C_A is the convex hull of the set

$$\{(\cos(a_1\theta), \sin(a_1\theta), \ldots, \cos(a_d\theta), \sin(a_d\theta)) : \theta \in [0, 2\pi)\}.$$

Proposition 5.2.7. *Let $d \geq 2$ be even and consider the realization of the representation (π_d, V_d) of $\mathrm{SL}_2(\mathbb{R})$ on the space of homogeneous polynomials in the variables x and y with coefficients in $\mathbb{R}$. Then the convex set*

$$C = \mathrm{conv}(\mathrm{SO}_2(\mathbb{R}).y^d)$$

is affinely isomorphic to $C_{\frac{d}{2}}$.

Proof: Let $\mathcal{O}_1 = \mathrm{SO}_2(\mathbb{R}).y^d$, let $\mathcal{O}_2$ be the orbit of $(1,0)^{\frac{d}{2}}$ under $\rho_{(2,4,\ldots,d)}$ and let $\mathcal{O}_3$ be the orbit of $(1,0)^{\frac{d}{2}}$ under $\rho_{(1,2,\ldots,\frac{d}{2})}$. We note that if two subsets of finite dimensional vector spaces are affinely isomorphic, then their convex hulls must be affinely isomorphic as well. Thus, it suffices to show that $\mathcal{O}_1$ is affinely isomorphic to $\mathcal{O}_2$, and $\mathcal{O}_2$ is affinely isomorphic to $\mathcal{O}_3$.

We first show that $\mathcal{O}_1$ is affinely isomorphic to $\mathcal{O}_2$. Write $V_d = W_{\frac{d}{2}} \oplus \cdots \oplus W_0$. By the proof of Proposition 5.2.2, we may write

$$y^d = \frac{1}{i^d 2^{d-1}} \sum_{k=0}^{\frac{d}{2}-1} \binom{d}{k} A_k + \frac{1}{i^d 2^d} \binom{d}{\frac{d}{2}} A_{\frac{d}{2}}.$$

For each $1 \leq k \leq \frac{d}{2}$, we have an isomorphism of $\mathrm{SO}_2(\mathbb{R})$-modules

$$\varphi_k : W_{\frac{d}{2}-k} \to U_{2k}$$

given by by

$$\varphi(A_k) := \frac{i^d 2^{d-1}}{\binom{d}{\frac{d}{2}-k}}(1,0) \qquad \varphi(B_k) := \frac{i^d 2^{d-1}}{\binom{d}{\frac{d}{2}-k}}(0,1).$$

For any point $(x_0, x_1, \ldots, x_{\frac{d}{2}}) \in \mathcal{O}_1$, we have $x_0 = \frac{1}{i^d 2^d}\binom{d}{\frac{d}{2}} A_{\frac{d}{2}}$. Define the map $\varphi : \mathcal{O}_1 \to \mathcal{O}_2$ given by $\varphi(x_0, x_1, \ldots, x_{\frac{d}{2}}) = (\varphi_1(x_1), \ldots, \varphi_{\frac{d}{2}}(x_{\frac{d}{2}}))$. Note that since each φ_k is $\mathrm{SO}_2(\mathbb{R})$-equivariant, so is φ. We thus have $\varphi(y^d) = (1,0)^d$ and $\varphi(\mathcal{O}_1) = \mathcal{O}_2$. Since each φ_k is injective, so is φ. Thus, φ is an isomorphism.

It now suffices to show that $\mathcal{O}_3$ is affinely isomorphic to $\mathcal{O}_2$. Indeed, for each $1 \leq k \leq d$, we have a linear isomorphism $\mu_k : U_k \to U_{2k}$ given by

$$\mu_k(\rho_k(g_\theta)(1,0)) := \rho_{2k}(g_\theta)(1,0).$$

Then define the map $\mu : \mathcal{O}_3 \to \mathcal{O}_2$ by

$$\mu(x_1, \dots, x_{\frac{d}{2}}) = (\mu_1(x_1), \dots, \mu_{\frac{d}{2}}(x_{\frac{d}{2}})).$$

Finally, $\Psi_{\frac{d}{2}} := \mu^{-1} \circ \varphi$ is an affine isomorphism from $\mathcal{O}_1$ to $\mathcal{O}_3$. ∎

Definition 5.2.8. Let $\Psi_{\frac{d}{2}}$ denote the affine isomorphism we constructed in the proof of Proposition 5.2.7 from $\mathrm{conv}(\mathrm{SO}_2(\mathbb{R}).y^d)$ to $C_{\frac{d}{2}}$.

Definition 5.2.9. Let $C \subseteq \mathbb{R}^d$ be a convex set. We define $\widehat{C}$ by

$$\widehat{C} := \{(\delta, a_1, \dots, a_d) : \delta + \sum_{i=1}^{d} a_i b_i \geq 0 \text{ for all } (b_1, \dots, b_d) \in C\}.$$

Remark 5.2.10. Let $C \subseteq \mathbb{R}^d$ be a convex set of dimension d. Then we observe that a point $(a_1, \dots, a_d) \in \mathbb{R}^d$ belongs to C if and only if

$$\delta + \sum_{k=1}^{d} a_k c_k \geq 0$$

for all $(\delta, c_1, \dots, c_d) \in \widehat{C}$. This follows from the separating hyperplane theorem. See [3, §2.5.1].

We know from Proposition 1.2.6 that C_d is a semialgebraic set. It turns out that there is an easy way of determining the inequalities which define this set. Before proving this claim, we prove a preliminary proposition.

Proposition 5.2.11. *Let $d \geq 1$, and let $\delta, c_1, \ldots, c_d \in \mathbb{C}$. Define the Laurent polynomial*

$$R(z) = \sum_{k=-d}^{d} u_k z^k$$

with the coefficients defined as follows. Let $u_0 = \delta$, and if $k > 0$, define $u_k = c_k$ and $u_{-k} = \overline{u_k}$. Further, we assume that $R(z) \geq 0$ when $|z| = 1$. Then there exists some $H \in \mathbb{C}[z]$ of degree d such that

$$R(z) = \overline{H}(z^{-1}) \cdot H(z).$$

Proof: We begin by noticing that $R(z) = \overline{R}(\overline{z}^{-1})$. Note that the roots of R come in pairs $\alpha, \overline{\alpha}^{-1}$. Thus, once we have shown that the roots lying on the unit circle $\mathbb{T}$ have even multiplicity, we will be done.

Indeed, suppose $z_0 = e^{it_0}$ is a root of R having odd multiplicity m. Let $\phi : \mathbb{R} \to \mathbb{R}$ be defined by $\phi(t) = R(e^{it})$. Then by Taylor's theorem, in some neighborhood of t_0, ϕ is represented by

$$\phi(t) = \phi(t_0) + \phi'(t_0)(t - t_0) + \cdots + \frac{\phi^{m-1}(t_0)}{(m-1)!}(t - t_0)^{m-1} + \frac{\phi^m(t_0)}{m!}(t - t_0)^m$$

for some $\xi \in [t_0, t]$. Since the multiplicity of z_0 is m, we have $\phi^{(k)}(t_0) = 0$ for $0 \leq k \leq m - 1$. Thus $\phi(t) = \dfrac{\phi^m(\xi)}{m!}(t - t_0)^m$. Since m is odd, this function changes sign around t_0, a contradiction. ∎

Theorem 5.2.12. *The Carathéodory orbitope $C = C_d$ is equal to the set of all vectors $(s_1, t_1, \ldots, s_d, t_d) \in \mathbb{R}^{2d}$ such that the matrix*

$$M_d = \begin{pmatrix} 1 & s_1+\sqrt{-1}t_1 & \cdots & s_{d-1}+\sqrt{-1}t_{d-1} & s_d+\sqrt{-1}t_d \\ s_1-\sqrt{-1}t_1 & 1 & \cdots & s_{d-2}+\sqrt{-1}t_{d-2} & s_{d-1}+\sqrt{-1}t_{d-1} \\ \vdots & & & & \\ s_{d-1}-\sqrt{-1}t_{d-1} & s_{d-2}-\sqrt{-1}t_{d-2} & \cdots & 1 & s_1+\sqrt{-1}t_1 \\ s_d+-\sqrt{-1}t_d & s_{d-1}-\sqrt{-1}t_{d-1} & \cdots & s_1-\sqrt{-1}t_1 & 1 \end{pmatrix}$$

is positive semidefinite.

Before beginning the proof, let us say a few words about the inner product structure on our spaces. The first space we are working with is $\mathbb{C}^{2d+1}$. We equip this space with the basis $e_{-d},\dots,e_0,\dots,e_d$, where e_i is the column vector in $\mathbb{C}^{2d+1}$ with 0's everywhere except the i^{th} position. It will be made clear in the proof why we use this notation. We equip $\mathbb{C}^{2d+1}$ with the standard inner product given by

$$\langle x, y\rangle = \sum_{k=-d}^{d} x_k\overline{y_k}.$$

Let

$$e^{-d},\dots,e^0,\dots,e^d \tag{5.2.2}$$

be the dual basis with respect to this inner product. Then, of course, $e^i = (e_i)^t$ for each i.

Secondly, we have the space $M_{d+1}(\mathbb{C})$ of $(d+1)\times(d+1)$ matrices with entries in $\mathbb{C}$. We give $M_{d+1}(\mathbb{C})$ the basis E_{ij}, $0 \le i,j \le d$, where E_{ij} is the matrix with 1 in the $(i,j)-$th position and 0 everywhere else. This space is equipped with the inner product $\langle A, B\rangle = \operatorname{tr}(AB^*)$ where B^* is the conjugate transpose of B. For a matrix $X \in M_{d+1}(\mathbb{C})$, let $\widetilde{X} \in M_{d+1}(\mathbb{C})^*$ be the linear functional defined by $\widetilde{X}(A) = \langle A, X\rangle$. From linear algebra, we know that $\widetilde{E_{ij}}$ behaves in the following way. For

$X \in M_{d+1}(\mathbb{C})$,

$$\widetilde{E_{ij}}(X) = x_{ij}$$

where x_{ij} is the (i,j)-th entry of X.

We are now ready to give the proof of Theorem 5.2.12.

Proof: Recall from Definition 5.2.9 that

$$\widehat{C}_d = \left\{ (\delta, a_1, b_1, \dots, a_d, b_d) : \delta + \sum_{k=1}^{d} (a_k \cos(k\theta) + b_k \sin(k\theta)) \geq 0 \right\}.$$

We identify each point $(\delta, a_1, b_1, \dots, a_d, b_d) \in \mathbb{R}^{2d+1}$ with the Laurent polynomial

$$R(z) = \sum_{k=-d}^{d} u_k z^k \in \mathbb{C}[z, z^{-1}]$$

where $u_0 = \delta$, $u_k = \frac{1}{2}(a_k - \sqrt{-1}b_k)$ and $u_{-k} = \overline{u_k}$ for $1 \leq k \leq d$. Note that $R \in \widehat{C}_d$ if and only if R is nonnegative on the unit circle $\mathbb{T} \subseteq \mathbb{C}$. By Proposition 5.2.11, we have a factorization

$$R(z) = \overline{H}(z^{-1}) \cdot H(z).$$

Now let $\gamma_d : \mathbb{C} \to \mathbb{C}^{d+1}$ be defined by $\gamma_d(z) = (1, z, \dots, z^d)^T$. Thus, there is a vector $h \in \mathbb{C}^{d+1}$ such that

$$R(z) = \gamma_d(z^{-1})^T \cdot \overline{h}h^T \cdot \gamma_d(z).$$

Now recall that a point $(c_1, s_1, \dots, c_d, s_d) \in \mathbb{R}^{2d}$ belongs to C_d if and only if

$$\delta + \sum_{k=1}^{d} a_k c_k + b_k s_k \geq 0 \qquad \text{for all } (\delta, a_1, b_1, \dots, a_d, b_d) \in \widehat{C}_d.$$

Let $\zeta = (x, 1, y)$, with $x_k = s_k + \sqrt{-1}t_k$ and $y_k = s_k - \sqrt{-1}t_k$. Now there is a linear map $\pi : M_{d+1}(\mathbb{C}) \to \mathbb{C}^{2d+1}$ such that $u = \pi(\overline{h}h^T)$, where $u = (u_{-d}, \dots, u_{-1}, \delta, u_1, \dots, u_d)$.

Indeed, we see that $R(z) = \sum\sum \overline{h_k} h_\ell z^{\ell-k}$ and

$$\overline{h}h^T = \begin{pmatrix} |h_0|^2 & \overline{h_0}h_1 & \cdots & \overline{h_0}h_d \\ \overline{h_1}h_0 & |h_1|^2 & \cdots & \overline{h_1}h_d \\ \vdots & \vdots & \vdots & \\ \overline{h_d}h_0 & \overline{h_d}h_1 & \cdots & |h_d|^2 \end{pmatrix}.$$

Thus, π is given by

$$\pi(A)_j = \sum_{\ell-k=j} a_{k,\ell}$$

where $-d \leq k \leq d$. Next, in the standard basis of $(\mathbb{C}^{2d+1})^*$ we write

$$\zeta = \sum_{i=-d}^{-1} x_{-i}e^i + e^0 + \sum_{i=1}^{d} y_i e^i.$$

By the above definition of π, one then sees that for $-d \leq i \leq -1$, $\pi^*(e^i)$ is a matrix with 1's on the i-th super-diagonal, and for $0 \leq i \leq d$, $\pi^*(e^i)$ is a matrix with 1's on the i-th sub-diagonal. It then follows that $M_d = \pi^*(\zeta)$. Thus we have

$$\begin{aligned} \delta + \sum_{k=1}^{d} a_k c_k + b_k s_k &= \langle \zeta, \pi(\overline{h}h^T)\rangle \\ &= \langle \pi^*(\zeta), \overline{h}h^T\rangle \\ &= \operatorname{tr}(M_d \cdot \overline{h}h^T) \\ &= h^T \cdot M_d \cdot \overline{h} \end{aligned}$$

So $(s_1, t_1, \ldots, s_d, t_d) \in C_d$ if and only if $h^T \cdot M_d \cdot \overline{h} \geq 0$ for all $h \in \mathbb{C}^{d+1}$, i.e. if and only if M_d is positive semidefinite. ∎

The following proposition is from [13, p. 566].

Proposition 5.2.13. *(Sylvester's Criterion) A Hermitian matrix M is positive semidefinite if and only if all of its principal minors are non-negative.*

Remark 5.2.14. Note that Sylvester's criterion gives us polynomial inequalities in the variables $s_1, \dots, s_d, t_1, \dots, t_d$.

Theorem 5.2.15. *Let $d \geq 2$ be even and consider the representation (π_d, V_d) of $\mathrm{SL}_2(\mathbb{R})$, realized on the space of homogeneous polynomials of degree d in x and y with coefficients in $\mathbb{R}$. Let $v = y^d$ be the highest weight vector corresponding to the standard Borel subgroup of $\mathrm{SL}_2(\mathbb{R})$. In addition, let $\Psi_{\frac{d}{2}}$ be the affine isomorphism of Definition 5.2.8. Then we have*

$$\mathrm{conv}(X_\lambda) = \{x \in \mathbb{R}^d : x = \alpha y \text{ for some } \alpha \in \mathbb{R}_+, y \in \Psi_{\frac{d}{2}}^{-1}(C_{\frac{d}{2}})\}.$$

Proof: We have

$$\mathrm{conv}(X_\lambda) = \mathrm{conv}(\mathrm{SL}_2(\mathbb{R}) \cdot y^d) = \mathbb{R}_+ \, \mathrm{conv}(\mathrm{SO}_2(\mathbb{R}) \cdot y^d)$$

where the last equality follows from Lemma 5.1.11. The result then follows from the definition of $\Psi_{\frac{d}{2}}$. ∎

Bibliography

[1] J. Bochnak, M. Coste, and M-F. Roy. *Real Algebraic Geometry*. Springer, 1998.

[2] A. Borel. *Linear Algebraic Groups*. Graduate Texts in Mathematics. Springer New York, 2012.

[3] S.P. Boyd and L. Vandenberghe. *Convex Optimization*. Berichte über verteilte messysteme. Cambridge University Press, 2004.

[4] A. Deitmar. *A First Course in Harmonic Analysis*. Universitext. Springer New York, 2013.

[5] D.S. Dummit and R.M. Foote. *Abstract Algebra*. Wiley, 2004.

[6] K. Erdmann and M.J. Wildon. *Introduction to Lie Algebras*. Springer Undergraduate Mathematics Series. Springer London, 2006.

[7] J. Faraut, S. Kaneyuki, A. Koranyi, Q. Lu, and G. Roos. *Analysis and Geometry on Complex Homogeneous Domains*. Progress in Mathematics. Birkhäuser Boston, 1999.

[8] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. Crm Proceedings & Lecture Notes. American Mathematical Society, 1978.

[9] Joachim Hilgert and Karl-Hermann Neeb. *Structure and geometry of Lie groups*. Springer Monographs in Mathematics. Springer, New York, 2012.

[10] J. E. Humphreys. *Introduction to Lie Algebras and Representation Theory*. Graduate Texts in Mathematics. Springer New York, 1994.

[11] J.E. Humphreys. *Linear Algebraic Groups*. Graduate Texts in Mathematics. Springer, 1975.

[12] Anthony W. Knapp. *Lie groups beyond an introduction*, volume 140 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, second edition, 2002.

[13] C.D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2000.

[14] Calvin C. Moore. Decomposition of unitary representations defined by discrete subgroups of nilpotent groups. *Annals of Mathematics*, 82(1):146–182, 1965.

[15] Y.A. Neretin. *Categories of Symmetries and Infinite-dimensional Groups*. L. M. S. monographs. Clarendon Press, 1996.

[16] V.L. Popov, A.N. Parshin, T.A. Springer, I.R. Shafarevich, and E.B. Vinberg. *Algebraic Geometry IV: Linear Algebraic Groups Invariant Theory*. Encyclopaedia of Mathematical Sciences. Springer Berlin Heidelberg, 2012.

[17] C. Procesi. *Lie Groups: An Approach through Invariants and Representations*. Universitext. Springer New York, 2007.

[18] R.T. Rockafellar. *Convex Analysis*. Princeton landmarks in mathematics and physics. Princeton University Press, 1997.

[19] Siddhartha Sahi and Hadi Salmasian. The capelli problem for gl(m—n) and the spectrum of invariant differential operators. *Advances in Mathematics*, 303:1 – 38, 2016.

[20] Raman Sanyal, Frank Sottile, and Bernd Sturmfels. Orbitopes. *Mathematika*, 57(02):275314, 2011.

[21] J.P. Serre. *Lie Algebras and Lie Groups: 1964 Lectures Given at Harvard University*. Lecture Notes in Mathematics. Springer, 1965.

[22] I.R. Shafarevich. *Basic Algebraic Geometry*. Number bk. 1 in Basic Algebraic Geometry. Springer-Verlag, 1994.

[23] Peter Woit. Topics in representation theory: Su(2) representations and their applications.