STATE UNIVERSITY OF MARINGÁ

EXACT SCIENCES CENTER

MATHEMATICS DEPARTMENT

POSTGRADUATE PROGRAM IN MATHEMATICS

(Doctorate)

BRUNO ALEXANDRE RODRIGUES

Lie Theory on Some Control Problems

Teoria de Lie em Alguns Problemas de Controle

Maringá-PR

2023

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

STATE UNIVERSITY OF MARINGÁ

EXACT SCIENCES CENTER

MATHEMATICS DEPARTMENT

POSTGRADUATE PROGRAM IN MATHEMATICS

Lie Theory on Some Control Problems

Teoria de Lie em Alguns Problemas de Controle

BRUNO ALEXANDRE RODRIGUES

Doctoral thesis submitted to the Graduate Program in Mathematics of Mathematics Department, Exact Sciences Center of the State University of Maringá, as a requirement to obtain the title of Doctor of Mathematics.
Concentration area: Geometry and Topology.

Doctoral supervisor: Prof. Dr. Alexandre José Santana.

Maringá-PR

2023

Dados Internacionais de Catalogação na Publicação (CIP)
(Biblioteca Setorial BSE-DMA-UEM, Maringá, PR, Brasil)

Rodrigues, Bruno Alexandre
R696L Lie theory on some control problems = Teoria de Lie em alguns problemas de controle / Bruno Alexandre Rodrigues. -- Maringá, 2023.
152 f. : il.

Orientador: Prof°. Dr°. Alexandre José Santana.
Tese (Doutorado)- Universidade Estadual de Maringá, Centro de Ciências Exatas,Programa de Pós-Graduação em Matemática - Área de Concentração: Geometria e Topologia, 2023.

1. Controlabilidade. 2. Álgebras e Lie. 3. Ações de grupos. 4. Controlabilidade de campos e vetores. 5. Controllability. 6. Matrix Lie algebras. 7. Group actions. 8. Controllability vector fields. I. Santana, Alexandre José, orient. II. Universidade Estadual de Maringá. Centro de Ciências Exatas. Programa de Pós-Graduação em Matemática - Área de Concentração: Geometria e Topologia. III. Título.

CDD 22.ed. 512.55
Edilson Damasio CRB9-1.123

BRUNO ALEXANDRE RODRIGUES

LIE THEORY ON SOME CONTROL PROBLEMS

Tese apresentada ao Programa de Pós-Graduação em Matemática do Departamento de Matemática, Centro de Ciências Exatas da Universidade Estadual de Maringá, como parte dos requisitos necessários para a obtenção do título de Doutor em Matemática tendo a Comissão Julgadora composta pelos membros:

COMISSÃO JULGADORA:

Prof. Dr. Alexandre José Santana - PMA/UEM (Presidente)

Profa. Dra. Rachida El Assoudi - Normandie Université/França

Profa. Dra. Viviana Jorgelina del Barco - IMECC/UNICAMP

Prof. Dr. Luiz Antonio Barrera San Martin - IMECC/UNICAMP

Profa. Dra. Maria Elenice Rodrigues Hernandes - PMA/UEM

Prof. Dr. Ryuichi Fukuoka - PMA/UEM

Aprovada em: 05 de outubro de 2023.

Local de defesa: Videoconferência - Google Meet (https://meet.google.com/xmp-aefs-qtu)

Agradecimentos

Embora este trabalho seja fruto de muita perseverança, paciência e empenho, nada seria possível sem o apoio e a colaboração de inúmeras pessoas que contribuíram direta ou indiretamente para que este objetivo fosse alcançado.

Devo, portanto, expressar minha gratidão ao Professor Alexandre José Santana, que durante quase uma década me orientou pelos caminhos da Matemática de maneira excepcional, sempre com notável tranquilidade e sabedoria.

Palavras são praticamente insuficientes para manifestar o quanto sou grato ao Professor Luiz Antonio Barrera San Martin por todo o conhecimento gentilmente compartilhado conosco nos últimos anos, por todas as dúvidas sanadas, por todas as brilhantes sugestões e por tornar as tardes de quinta-feira especiais, sempre com boas risadas, muita Matemática e muita inspiração.

Aos professores membros da Comissão Julgadora agradeço pelo tempo dedicado à avaliação desta tese e por todas as contribuições e sugestões realizadas.

À minha família agradeço por todo o amor e suporte concedido, em especial às mulheres que constituem a base de minha vida, minha esposa e companheira de uma década, Thaís Fernanda Cabral dos Santos, e minhas mães e pai, Nadir Andretto, Laíde Aparecida Andretto, Elaine Aparecida Andretto e Paulo Roberto Rodrigues.

A todas as professoras e professores que me instruíram e inspiraram desde a infância, destaco imensa gratidão. De modo muito especial agradeço à Professora Maria Imamura Koyashiki, que me inspirou na escolha pela Graduação em Matemática e me propiciou uma base consistente, possibilitando assim que tal escolha fosse exitosa. Aos professores Luci Harue Fatori, Neuza Teramon, Regina Célia Guapo Pasquini, Ricardo Cezar Ferreira, Olívio Augusto Weber e Albo Carlos Cavalheiro agradeço por me proporcionarem uma introdução suave e completa ao universo da Matemática.

Agradeço sinceramente a todos os amigos, amigas e colegas que ao longo desta jornada me deram apoio, sugestões, discussões enriquecedoras, desde as amizades com laços mais antigos até as amizades construídas recentemente, elemento fundamental para tornar qualquer caminhada mais leve e agradável. Em especial menciono os amigos Ademir, Anderson, João Paulo, Jéssica, Juliana e Richard que diretamente colaboraram com a realização deste trabalho, tanto com discussões matemáticas valiosíssimas quanto compartilhando os momentos felizes e as certezas e incertezas da vida.

Agradeço a Deus pela vida a mim concedida.

Agradeço à CAPES pelo suporte financeiro que estabeleceu condições para que este trabalho fosse realizado.

Por fim, dedico este trabalho a todas as pessoas que acreditam que o Conhecimento é uma dádiva concedida à humanidade e que somente por meio do acesso de todas as pessoas ao Conhecimento poderemos transformar o mundo em um lugar justo e digno para todos.

"Água mole em pedra dura tanto bate até que fura."

Provérbio popular.

ABSTRACT

The Lie group $\mathrm{Sl}(n, \mathbb{H})$ is the Lie group of order n quaternionic matrices g such that $|\det g| = 1$. This thesis establishes conditions for a subsemigroup with nonempty interior, $S \subset \mathrm{Sl}(n, \mathbb{H})$, to coincide with the entire group $\mathrm{Sl}(n, \mathbb{H})$. From this, we set conditions on matrices $A, B \in \mathfrak{sl}(n, \mathbb{H})$ ensuring controllability for the invariant control system $\dot{g} = Ag + uBg$ on $\mathrm{Sl}(n, \mathbb{H})$. We also prove that these conditions are generic in the sense that we obtain an open and dense set of controllable pairs $(A, B) \in \mathfrak{sl}(n, \mathbb{H})^2$.

Subsequently, the Lie saturate technique is used to establish controllability criteria for bilinear control systems on $\mathrm{Sl}(n, \mathbb{C})$ and $\mathrm{Sl}(n, \mathbb{H})$, as well as on certain semidirect products. Our study also employs quaternions to explore invariant control sets for vector fields induced by $\mathrm{SO}(1, 4)$ and $\mathrm{SU}(1, 2)$ on the unit quaternion sphere S^3, their flag manifold. Throughout this thesis, these investigations deepen the understanding of controllability conditions for control systems on classical real Lie groups and their geometric characteristics.

Keywords: controllability, matrix Lie algebras, group actions, controllability vector fields.

MSC 2020. 93B05, 15B30, 57M60, 57R27.

Resumo

O grupo Lie $\mathrm{Sl}\,(n,\mathbb{H})$ é o grupo de Lie das matrizes $n\times n$ quaterniônicas tais que $|\det g| = 1$. Nesta tese apresentamos condições para um subsemigrupo com interior não-vazio, $S \subset \mathrm{Sl}\,(n,\mathbb{H})$, coincidir com todo o grupo $\mathrm{Sl}\,(n,\mathbb{H})$. A partir disso, são determinadas condições sobre as matrizes $A, B \in \mathfrak{sl}\,(n,\mathbb{H})$ para garantir a controlabilidade do sistema de controle invariante $\dot{g} = Ag + uBg$ em $\mathrm{Sl}\,(n,\mathbb{H})$. Provamos também que essas condições são genéricas no sentido de que obtemos um conjunto aberto e denso de pares controláveis $(A, B) \in \mathfrak{sl}\,(n,\mathbb{H})^2$.

Posteriormente, a técnica do saturado de Lie é utilizada para estabelecer critérios de controlabilidade para sistemas de controle bilineares em $\mathrm{Sl}(n,\mathbb{C})$ e $\mathrm{Sl}(n,\mathbb{H})$, bem como em certos produtos semidiretos. Nosso estudo também emprega quatérnios para explorar conjuntos de controle invariantes para campos de vetores induzidos por $\mathrm{SO}(1,4)$ e $\mathrm{SU}(1,2)$ na esfera S^3, variedade flag destes últimos. Ao longo desta tese, essas investigações aprofundam a compreensão das condições de controlabilidade para sistemas de controle em grupos Lie reais clássicos e suas características geométricas.

Palavras-chave: controlabilidade, álgebras de Lie de matrizes, ações de grupos, controlabilidade de campos de vetores.

MSC 2020. 93B05, 15B30, 57M60, 57R27.

CONTENTS

INTRODUCTION

The control theory on Lie groups has its roots in the late 1960s and early 1970s with a series of papers that include the works of V. Jurdjevic and H. J. Sussmann [14], who first achieved controllability results on compact Lie groups. Further developments were made by V. Jurdjevic and I. Kupka [12, 13], in the first half of the 1980s, where the authors established conditions for a control system evolving on a Lie group to be transitive, introducing the technique based on the concept of Lie saturate.

Some improvements to the results in [12, 13] were obtained by J. P. Gauthier and G. Bornard [8] for bilinear control systems evolving on the Lie group $\mathrm{Sl}(n, \mathbb{R})$. Subsequently, in the 1980s, numerous results concerning controllability of invariant control systems on semisimple Lie groups were obtained. For example, J. P. Gauthier, I. Kupka and G. Sallet [9] considered bilinear systems evolving on Lie groups whose Lie algebras are real normal forms (also called split real forms) of complex simple Lie algebras of type A_l, D_l, E_6, E_7 or E_8. A few years later R. El Assoudi and J. P. Gauthier [2] addressed right-invariant control systems on simple Lie groups of B_l, C_l, F_4 and G_2 types, also considering real normal forms. Controllability of bilinear systems on simple Lie groups whose Lie algebras are real normal forms of complex simple Lie algebras was also studied by F. Silva Leite and P. E. Crouch [32] and later by R. El Assoudi and J. P. Gauthier and I. Kupka [3, 4].

Since the 1990s, L. A. B. San Martin and collaborators worked on a series of papers about controllability (transitivity) and a certain type of local controllability (control sets) for semigroup actions of semisimple Lie groups on homogeneous manifolds. See, for example, L. A. B. San Martin [24], L. A. B. San Martin and P. A. Tonelli [29], O. G. do Rocio, L. A. B. San Martin and A. J. Santana [19], V. Ayala and L. A. B. San Martin [5]. This approach has proven to be very useful for studying the controllability of control systems. Specifically, in A. L. Dos Santos and L. A. B. San Martin [30], the authors employed the previously mentioned approach (Lie Theory) and the topology of flag manifolds to study the controllability of bilinear control systems and invariant control systems on semisimple Lie groups.

In this thesis, we also address the topic of global controllability, in the sense of controlling the entire state space, for invariant control systems in $\mathrm{Sl}(n, \mathbb{H})$, systems in $\mathrm{Sl}(n, \mathbb{C})$, in the semi-direct products $\mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$ and $\mathfrak{sl}(2, \mathbb{C}) \rtimes \mathbb{C}^2$, and we also study local controllability, in the sense of determining the invariant control sets, for certain control systems induced on S^3 by the action of $\mathrm{SO}(1, 4)$ and $\mathrm{SU}(1, 2)$.

Part of our work aims to investigate controllability conditions for control systems on certain classical Lie groups whose Lie algebras are non-compact and non-split real forms of complex semisimple Lie algebras. For the complex Lie algebras of type A_l, the non-compact and non-split real forms are $\mathfrak{sl}(n, \mathbb{H})$ and $\mathfrak{su}(p, q)$ (see L. A. B. San Martin [26] or S. Helgason [10] for a complete classification).

Semigroups and controllability on the special linear quaternionic Lie group are the subjects of Chapter 2 (see B. A. Rodrigues, L. A. B. San Martin and A. J. Santana [22]). The central result of this chapter shows that a subsemigroup $S \subset \mathrm{Sl}\,(n, \mathbb{H})$ with nonempty interior must be the whole group in case it contains a certain subgroup of $\mathrm{Sl}\,(n, \mathbb{H})$ isomorphic to $\mathrm{Sl}\,(2, \mathbb{H})$ (see Theorem 2.2.1).

This approach was already applied in A. L. Dos Santos and L. A. B. San Martin [30, 31]. In these papers the same problem was dealt for connected complex semisimple Lie groups and some real Lie groups whose Lie algebras are split real forms of complex Lie algebras. As the Lie algebra $\mathfrak{sl}(n, \mathbb{H})$ is not in any of these classes, the results in [30] and [31] do not include the group worked out here. But our strategy is also based in the following result proved in [29]. If $S \subset G$ is a proper semigroup with nonempty interior in a semisimple Lie group G, then there exists some G-flag manifold $\mathbb{F}_\Theta$ such that the unique S-invariant control set $C_\Theta \subset \mathbb{F}_\Theta$ is contained in a contractible set. In fact, to prove that $S = \mathrm{Sl}(n, \mathbb{H})$ we first show that several $\mathcal{G}$-orbits are non contractible 4-spheres where $\mathcal{G}$ is isomorphic to $\mathrm{Sl}(2, \mathbb{H})$ (see Section 2.2 for a clear description of $\mathcal{G}$), and after we prove that some of these orbits are contained in the S-invariant control set.

Subsequently we apply this theory to the controllability of

$$\dot{g} = Ag + u(t)Bg.$$

More precisely, we find conditions on the matrices A and B to have $\mathcal{G}$ contained in the system semigroup (generated by $\exp t(A+uB)$, $t \geq 0$, $u \in \mathbb{R}$). Therefore if we assume also the Lie algebra rank condition then our results can be applied to obtain that the system semigroup is the system group (generated by $\exp t(A+uB)$, $t, u \in \mathbb{R}$), implying the controllability of the above system).

Further, we prove that the controllability for invariant systems on $\mathrm{Sl}\,(n, \mathbb{H})$ is a generic property in the sense that there is an open and dense set $C \subset \mathfrak{sl}\,(n, \mathbb{H})^2$ such that the control system $\dot{g} = Ag + uBg$ with unrestricted controls is controllable for all pairs $(A, B) \in C$.

In Chapter 3 we study controllability from another perspective. The results therein are inspired by [8], which in turn improves the controllability results present in the classical papers by V. Jurdjevic and I. Kupka, [12] and [13], for the case $\mathrm{Sl}(n, \mathbb{R})$. We obtain similar results for the cases $\mathrm{Sl}(n, \mathbb{C})$ and $\mathrm{Sl}(n, \mathbb{H})$ and using the Lie saturate technique we also get controllability results for the semidirect products $\mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$ and $\mathfrak{sl}(2, \mathbb{C}) \rtimes \mathbb{C}^2$.

In many situations we consider as a control system on a Lie group G a subset Γ of the Lie algebra $\mathfrak{g}$ of G. A system Γ is said to satisfy the Lie algebra rank condition (LARC) when it generates $\mathfrak{g}$ as a Lie algebra, that is, $\mathrm{Lie}(\Gamma) = \mathfrak{g}$. The Lie algebra rank condition is a necessary condition for controllability, but in general it is not sufficient. The Lie saturate technique for controllability consists of a extension process that constructs from Γ a subset $\mathrm{LS}(\Gamma) \subset \mathrm{Lie}(\Gamma)$, such that the controllability of Γ on G turns out to be equivalent to the condition $\mathrm{LS}(\Gamma) = \mathfrak{g}$. The set $\mathrm{LS}(\Gamma)$ is called Lie saturate or Lie wedge of Γ.

Using this approach we prove that if a bilinear control system $\Gamma = A + \mathbb{R}B$ on $\mathrm{Sl}(n, \mathbb{C})$ is such that $A = (a_{ij})$ with $a_{1n}, a_{n1} \neq 0$ and $B = \mathrm{diag}(b_1, b_2, \cdots, b_n)$ is a strongly regular element (see Definition 3.2.1) satisfying

$$\mathrm{Re}(b_1) > \mathrm{Re}(b_2) > \cdots > \mathrm{Re}(b_n) \text{ and } \mathrm{Im}(b_i) - \mathrm{Im}(b_j) \neq 0 \text{ whenever } i \neq j,$$

then the controllability of Γ is equivalent to the irreducibility of A.

The idea behind the proof is to use the characterization of the irreducibility of A by means of strongly connected graphs in way to avoid the zero entries of A when showing that a basis $\mathcal{B}$ and its opposite $-\mathcal{B}$ are contained in the Lie saturate $\mathrm{LS}(\Gamma)$. The set $\mathrm{LS}(\Gamma)$ has a nice geometric characterization, that is, it is a topologically closed, convex and positive cone, which implies that the condition $\pm\mathcal{B} \subset \mathrm{LS}(\Gamma)$ ensures that $\mathrm{LS}(\Gamma) = \mathfrak{sl}(n, \mathbb{C})$.

Continuing to employ the Lie wedge method, we present a second theorem that establishes conditions for the controllability of bilinear controls systems on $\mathrm{Sl}(n, \mathbb{H})$ (see Theorem 3.3.4). It is actually a weaker version of Theorem 2.3.1, as even if the Lie algebra rank condition is not required, it immediately follows from the equality $\mathrm{LS}(\Gamma) = \mathfrak{sl}(n, \mathbb{H})$. However, if we keep in mind that the Lie algebra rank condition is usually difficult to verify, even with the assistance of computational devices, the theorem turns out to be interesting on its own, since all of its hypotheses can be quite easily verified.

As a final application of this method to ensure controllability, we state and prove three theorems concerning controllability on the semidirect products $\mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$ and $\mathfrak{sl}(2, \mathbb{C}) \rtimes \mathbb{C}^2$. Just like in the previous case, the conditions obtained depends only on the entries of matrices that define the control systems, rendering them practical tools for verifying controllability in such situations.

Quaternions prove to be highly useful in Chapters 4 and 5, where we explore invariant control sets for vector fields induced by $\mathrm{SO}(1, 4)$ and $\mathrm{SU}(1, 2)$ on the sphere S^3, the sphere of unit quaternions (see B. A. Rodrigues, L. A. B. San Martin and A. J. Santana [21]). A Cartan decomposition of the Lie algebra $\mathfrak{so}(1, 4)$ can be identified as the direct sum $\mathfrak{so}(4) \oplus \mathbb{H}$, in which $\mathfrak{so}(4)$ corresponds to the compact component, and $\mathbb{H}$ stands for the symmetric one. The maximal abelian subalgebra contained in the symmetric component is one-dimensional, implying that there is only one flag manifold for $\mathfrak{so}(1, 4)$, which is precisely the sphere S^3. For symmetric elements, the vector fields given by the infinitesimal action of $\mathfrak{so}(1, 4)$ on S^3 are gradient vector fields of height functions, and elements in the compact component $\mathfrak{so}(4)$ give rise to vector fields defined by right and left multiplication by imaginary quaternions.

We provide a characterization for the invariant control sets on S^3 for control systems with $\mathbf{1} \in \mathbb{H}$ as a drift and control vector fields corresponding to pure quaternions. Such control sets appear as spherical domes in some cases, while in others, they are described as geodesically convex closures of the set of attractor points for the vector fields corresponding to the control system. This observation seems to be true in general for appropriate families of vector fields whose trajectories follow geodesics. In this context, we establish that if the set E of attractors for a family Γ of geodesic vector fields on a differentiable manifold M is both closed and geodesically convex, then E is the invariant control set for Γ.

Following the same general lines of Chapter 4, on Chapter 5 we will embark on exploring the geometrical aspects inherent to the Lie group $\mathrm{SU}(1,2)$. The compact and symmetric components in a Cartan decomposition for $\mathfrak{su}(1,2)$ are identified respectively with $\mathfrak{u}(2)$ and $\mathbb{H}$, and for an Iwasawa decomposition for $\mathfrak{su}(1,2)$ the identification is

$$\mathfrak{su}(1,2) = \mathfrak{u}(2) \oplus \mathfrak{a} \oplus Heis,$$

where $\mathfrak{a}$ denotes the maximal abelian subalgebra in the symmetric component of the Cartan decomposition and $Heis$ is the Heisenberg Lie algebra. Just like $\mathfrak{so}(1,4)$, the Lie algebra $\mathfrak{su}(1,2)$ stands as a real rank 1 Lie algebra, featuring a four-dimensional symmetric part in its Cartan decompositions, ultimately leading to the identification of the sphere S^3 as its sole flag manifold.

An element $q \in \mathfrak{k} = \mathfrak{u}(2) = \mathfrak{su}(2) \oplus \mathfrak{z}$ decomposes as $q = z + a$, where $z \in \mathfrak{su}(2)$ and $a \in \mathfrak{z}$, being $\mathfrak{z}$ the one-dimensional center of the unitary Lie algebra $\mathfrak{u}(2)$. The vector field induced on S^3 by q is expressed as the combination of left multiplication by z (considered as an imaginary quaternion) and right multiplication by $a\mathbf{i}$.

Nonetheless, the infinitesimal action on S^3 takes a different shape for symmetric elements, since the occurrence of the positive root 2α implies that the Borel metric does not coincides with the canonical metric given by the immersion of S^3 in $\mathbb{H}$.

As at the beginning of every journey, before taking the first step it is important to prepare a backpack with the essential items for the journey. In Chapter 1, we briefly gather the most important prerequisites trying to make this work as self-contained as possible. Within this chapter, a concise overview of the theory regarding the classification of real semisimple Lie algebras is presented in Section 1.1, some results about the realification and complexification of vector spaces are left for the Appendix A. In Section 1.2 we provide the definitions of the Cartan and Iwasawa decompositions. Some basic notions about flag manifolds are discussed in Section 1.3, and the concepts from the Control Theory that will be important for this work are outlined in 1.4.

The remainder of this work is organized as follows. Chapter 2 focuses on the study transitivity for subsemigroups of $\mathrm{Sl}(n,\mathbb{H})$, with the results applied to the controllability of invariant control systems evolving on this group. In Chapter 3, the Lie saturate technique is employed to establish controllability conditions for bilinear control systems on the Lie groups $\mathrm{Sl}(n,\mathbb{C})$, $\mathrm{Sl}(2,\mathbb{H})$, $\mathrm{Sl}(n,\mathbb{H})$ as well as on the semidirect products $\mathrm{Sl}(2,\mathbb{R}) \rtimes \mathbb{R}^2$ and $\mathrm{Sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$. Chapter 4 provides geometrical insights into the structure of $\mathrm{SO}(1,4)$ and determines invariant control sets for certain control systems induced by $\mathrm{SO}(1,4)$ on the sphere S^3. Chapter 5 follows the same general lines as Chapter 4 and begins exploring the geometry inherent to $\mathrm{SU}(1,2)$. Section 5.2.4 describes the vector fields induced by $\mathrm{SU}(1,2)$ on S^3 and offers a preliminary understanding of the behavior of such vector fields.

CHAPTER 1

PRELIMINARIES

In this chapter we collect and summarize the main concepts and results that will appear throughout our work. Most of the definitions and proofs can be found in [26] and [28], which we highly recommend for a precise and detailed construction of the theory. Some of the real semisimple Lie algebras constitute the environment where we are going to study invariant control sets and controllability for control systems. A few formulas and notations concerning the quaternions should be discussed in the examples. The Cartan and Iwasawa decompositions are essential in the study of the real semisimple Lie algebras, specially when defining the flag manifolds, that will stand as state space of control systems given by vector fields induced by infinitesimal actions.

1.1 Classical Lie algebras and their real forms

A Lie algebra is said to be solvable when its derived series eventually vanishes and it is said to be nilpotent when its central descending series eventually vanishes. Of major importance to this work are the semisimple Lie algebras. A Lie algebra $\mathfrak{g}$ is said to be:

(i) simple if $\dim \mathfrak{g} \neq 1$ and its only ideals are 0 and $\mathfrak{g}$ itself.

(ii) semisimple if it has no solvable ideals other than 0.

Definition 1.1.1. *A* ***Cartan subalgebra*** *of a Lie algebra $\mathfrak{g}$ is a nilpotent subalgebra $\mathfrak{h} \subset \mathfrak{g}$ that coincides with its own normalizer $N_{\mathfrak{g}}(\mathfrak{h})$ in $\mathfrak{g}$.*

Example 1.1.2. *Consider the quaternion algebra $\mathbb{H} = \langle \mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle$. Remember that the product of two quaternions is defined on the basis $\{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ by the following rules:*

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \quad \mathbf{ij} = \mathbf{k}, \quad \mathbf{ji} = -\mathbf{k}.$$

Given $p \in \mathbb{H}$, we write $p = p_0 + p_1\mathbf{i} + p_2\mathbf{j} + p_3\mathbf{k} = (p_0, p_1, p_2, p_3)$ and we define the vector part of p as $\mathbf{p} = p_1\mathbf{i} + p_2\mathbf{j} + p_3\mathbf{k}$. With these assumptions, $p \in \mathbb{H}$ can be seen as the sum of a scalar with a vector, that is, $p = p_0 + \mathbf{p}$. The sum in $\mathbb{H}$ is defined termwise while the product of two quaternions $p = p_0 + p_1\mathbf{i} + p_2\mathbf{j} + p_3\mathbf{k}$ and $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$ can be written as

$$p \cdot q = p_0 q_0 - \mathbf{p} \cdot \mathbf{q} + p_0 \mathbf{q} + q_0 \mathbf{p} + \mathbf{p} \times \mathbf{q}. \tag{1.1}$$

Taking into account this formula for the product, the canonical inner product on $\mathbb{H}$ assumes the form $\langle p, q \rangle = {}^1\!/\!{}_2(p\overline{q} + q\overline{p})$, and thus the norm is given simply by $|p|^2 = p\overline{p}$. These formulas are quite useful and they will show their own importance in the next chapters.

The subspace formed by the immaginary quaternions $\mathrm{Im}\mathbb{H}$ endowed with the Lie brackets $[z, w] = zw - wz$ is a simple Lie algebra isomorphic to $\mathfrak{su}(2)$. In fact, in $\mathrm{Im}\mathbb{H}$ we have

$$[\mathbf{i}, \mathbf{j}] = 2\mathbf{k}, \quad [\mathbf{j}, \mathbf{k}] = 2\mathbf{i} \quad \textit{and} \quad [\mathbf{i}, \mathbf{k}] = -2\mathbf{j},$$

while in $\mathfrak{su}(2)$ for the basis

$$A = \begin{bmatrix} 0 & \mathbf{i} \\ \mathbf{i} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{bmatrix},$$

we have

$$[A, B] = 2C, \quad [B, C] = 2A \quad \textit{and} \quad [A, C] = -2B.$$

This implies that $\mathrm{Im}\mathbb{H} = \mathfrak{su}(2)$. Also, $\mathfrak{h} = \langle \mathbf{k} \rangle$ is a Cartan subalgebra of $\mathrm{Im}(\mathbb{H})$. In fact, $\mathfrak{h}$ is nilpotent (it is abelian) and for any $z = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ we have

$$[z, \mathbf{k}] = [a\mathbf{i} + b\mathbf{j} + c\mathbf{k}, \mathbf{k}] = a[\mathbf{i}, \mathbf{k}] + b[\mathbf{j}, \mathbf{k}] = -2a\mathbf{j} + 2b\mathbf{i},$$

and for $[z, \mathbf{k}]$ to be in $\mathfrak{h}$ both the coefficients a and b must be zero. In other words, it holds $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$.

The characteristic polynomial of $\mathrm{ad}(X)$, $X \in \mathfrak{g}$, has the form

$$p_X(t) = t^n + p_{n-1}(X)t^{n-1} + \cdots + p_1(X)t + p_0(X)$$

where n is the dimension of $\mathfrak{g}$. If X has coordinates $X_1, X_2, \ldots, X_n$ with respect to some fixed basis for $\mathfrak{g}$, we can see each $p_k(\cdot)$ as a polynomial function of the n variables $X_1, X_2, \ldots, X_n$, which is of degree $n-k$ in $X_1, X_2, \ldots, X_n$.

Definition 1.1.3. *The* ***rank*** *of a Lie algebra $\mathfrak{g}$ is the less integer r such that the polynomial function p_r is not identically zero.*

Obviously the rank of $\mathfrak{g}$ can be at most n. Further, $\mathfrak{g}$ has rank exactly n if and only if $\mathfrak{g}$ is nilpotent. In fact, if the rank of $\mathfrak{g}$ is n, then $p_X(t) = t^n$, for every $X \in \mathfrak{g}$. So, by the Cayley-Hamilton theorem, $p_X(\mathrm{ad}(X)) = \mathrm{ad}(X)^n = 0$, which means that $\mathrm{ad}(X)$ is nilpotent, for every $X \in \mathfrak{g}$. It follows from the Engel's theorem that $\mathfrak{g}$ is a nilpotent Lie algebra. On the other hand, if $\mathfrak{g}$ is nilpotent, then $\mathrm{ad}(X)$ is nilpotent for every $X \in \mathfrak{g}$. But nilpotent operators does not have zero eigenvalues, showing that for an arbitrary $X \in \mathfrak{g}$, $\mathrm{ad}(X)$ has characteristic polynomial of the form $p_X(t) = t^n$, ensuring that the rank of $\mathfrak{g}$ is n.

If X is a nonzero element of $\mathfrak{g}$, then $\mathrm{ad}(X) \cdot X = [X, X] = 0$, that is, 0 is an eigenvalue of $\mathrm{ad}(X)$. In this way, if $\mathfrak{g} \neq 0$, then $p_0 = 0$, that is, the rank of $\mathfrak{g}$ must be at least 1.

Definition 1.1.4. *An element $X \in \mathfrak{g}$ is called* ***regular*** *if the algebraic multiplicity of 0 as an eigenvalue of $\mathrm{ad}(X)$ is minimal (when compared with all the other elements of $\mathfrak{g}$). Otherwise, X is said to be a* ***singular*** *element.*

Evaluating the characteristic polynomial of $\mathrm{ad}(X)$, we see that X is regular exactly when $p_r(X) \neq 0$, where r is the rank of $\mathfrak{g}$.

It is possible to show that finite dimensional Lie algebras always admit Cartan subalgebras, and every Cartan subalgebra of a Lie algebra $\mathfrak{g}$ has the dimension equals the rank of $\mathfrak{g}$ (see [26], Theorem 4.3 and its corollary).

Example 1.1.5. *Consider for the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ the canonical basis*

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

The subalgebra $\mathfrak{h} = \langle H \rangle$ is a Cartan subalgebra of $\mathfrak{sl}(2, \mathbb{C})$. In fact, $\mathfrak{h}$ is abelian (since it is one-dimensional), and hence it is nilpotent. Further, given $Z = aX + bH + cY \in \mathfrak{g}$ we have

$$[Z, H] = [aX + bH + cY, H] = a[X, H] + c[Y, H] = -2aX + 2cY,$$

and so these brackets will vanish if and only if $a = c = 0$. This means that $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$, proving that $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{sl}(2, \mathbb{C})$ and so $\mathfrak{sl}(2, \mathbb{C})$ has rank 1.

Given $Z = aX + bH + cY$ in $\mathfrak{sl}(2, \mathbb{C})$, we have

$$\mathrm{ad}(Z) = \begin{bmatrix} 2b & -2a & 0 \\ -c & 0 & a \\ 0 & 2c & -2b \end{bmatrix},$$

whose characteristic polynomial is $p_Z(t) = t^3 - 4(b^2 + ac)t$. This tells us that Z is regular if and only if $b^2 + ac \neq 0$. In other words, $Z \in \mathfrak{sl}(2, \mathbb{C})$ is regular if and only if $\det(Z) \neq 0$.

Example 1.1.6. *Let us consider in the Lie algebra*

$$\mathfrak{so}(3, \mathbb{C}) = \{A \in \mathfrak{sl}(3, \mathbb{C}) \mid A^T = -A\}$$

the canonical basis

$$E = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

It is easy to see that $[E, F] = -G$, $[E, G] = F$ and $[F, G] = -E$. And from this follows that the subalgebras of dimension 2 or 3 cannot be abelian. Furthermore, the subalgebra $\mathfrak{h} = \langle G \rangle$ is a Cartan subalgebra of $\mathfrak{so}(3, \mathbb{C})$, since $X = aE + bF + cG$, for $a, b, c \in \mathbb{C}$, then

$$[X, G] = [aE + bF + cG, G] = a[E, G] + b[F, G] = aF - bE,$$

and these brackets belong to $\mathfrak{h}$ if and only if $a = b = 0$, that is, $X = cG$, and hence $\mathfrak{h}$ coincides with its normalizer in $\mathfrak{g}$. It is easy to see that $\mathfrak{h}_1 = \langle E \rangle$ and $\mathfrak{h}_2 = \langle F \rangle$ are also Cartan subalgebras of $\mathfrak{so}(3, \mathbb{C})$.

Now, if we consider the quadratic form given by the matrix

$$J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

we have that $g^T \mathrm{Id} g = J$*, where*

$$g = \frac{1}{\sqrt{2\mathbf{i}}}\begin{bmatrix} \sqrt{2\mathbf{i}} & 0 & 0 \\ 0 & \mathbf{i} & 1 \\ 0 & 1 & \mathbf{i} \end{bmatrix}, \quad \text{with inverse} \quad g^{-1} = \frac{\sqrt{2\mathbf{i}}}{2}\begin{bmatrix} \sqrt{2\mathbf{i}}/\mathbf{i} & 0 & 0 \\ 0 & -\mathbf{i} & 1 \\ 0 & 1 & -\mathbf{i} \end{bmatrix}.$$

And so $\mathfrak{so}(3,\mathbb{C})$ *is isomorphic to* $g\mathfrak{so}(3,\mathbb{C})g^{-1}$*, whose matrices are of the form*

$$\begin{bmatrix} 0 & b & c \\ -c & a & 0 \\ -b & 0 & -a \end{bmatrix},$$

being this last Lie algebra characterized by the condition $AJ + JA^T = 0$*. It is straightforward to check that* $g\mathfrak{h}g^{-1} = \langle D \rangle$*, where*

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

in way that this subalgebra can be seen as a Cartan subalgebra of $\mathfrak{so}(3,\mathbb{C})$*.*

Consider $X = aE + bF + cG$*. Evaluating the matrix of* $\mathrm{ad}(X)$ *gives*

$$\mathrm{ad}(X) \cdot E = 0E - cF + bG$$

$$\mathrm{ad}(X) \cdot F = cE + 0F - aG$$

$$\mathrm{ad}(X) \cdot G = -bE + aF + 0G,$$

and so

$$\mathrm{ad}(X) = \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix},$$

whose characteristic polynomial is $p_X(t) = -t^3 - (a^2 + b^2 + c^2)t$*. This shows that* $\mathfrak{so}(3,\mathbb{C})$ *has rank* 1 *(agreeing with the one-dimensional Cartan subalgebra found) and* X *is a regular element if and only if* $a^2 + b^2 + c^2 \neq 0$*.*

Finally, note that $\mathfrak{sl}(2,\mathbb{C})$ *and* $\mathfrak{so}(3,\mathbb{C})$ *are isomorphic. To see this just define the isomor-*

phism $\varphi : \mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{so}(3,\mathbb{C})$ *by*

$$\varphi(X) = \begin{bmatrix} 0 & \mathbf{i} & 0 \\ -\mathbf{i} & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \quad \varphi(H) = \begin{bmatrix} 0 & 0 & -2\mathbf{i} \\ 0 & 0 & 0 \\ 2\mathbf{i} & 0 & 0 \end{bmatrix},$$

$$\varphi(Y) = \begin{bmatrix} 0 & -\mathbf{i} & 0 \\ \mathbf{i} & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

The matrix of φ *is given by*

$$\varphi = \begin{bmatrix} \mathbf{i} & 0 & -\mathbf{i} \\ 0 & -2\mathbf{i} & 0 \\ 1 & 0 & 1 \end{bmatrix}, \text{ whose inverse is } \varphi^{-1} = \begin{bmatrix} -\mathbf{i}/2 & 0 & 1/2 \\ 0 & \mathbf{i}/2 & 0 \\ \mathbf{i}/2 & 0 & 1/2 \end{bmatrix}.$$

As $\varphi^{-1}(E) = \mathbf{i}/2(Y - X)$, $\varphi^{-1}(F) = \mathbf{i}/2H$ *and* $\varphi^{-1}(G) = 1/2(X + Y)$, *we see that*

$$\begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}$$

also generate Cartan subalgebras of $\mathfrak{sl}(2,\mathbb{C})$.

Example 1.1.7. *Writing* E_{ij} *for the matrix having* 1 *in the* ij*-entry and* 0 *elsewhere, the Lie algebra* $\mathfrak{so}(5,\mathbb{C})$ *has as basis*

$$e_1 = (E_{21} - E_{12}),\ \ e_2 = (E_{31} - E_{13}),\ \ e_3 = (E_{41} - E_{14}),\ \ e_4 = (E_{51} - E_{15}),$$

$$e_5 = (E_{32} - E_{23}),\ \ e_6 = (E_{42} - E_{24}),\ \ e_7 = (E_{52} - E_{25}),$$

$$e_8 = (E_{43} - E_{34}),\ \ e_9 = (E_{53} - E_{35}),\ \ e_{10} = (E_{54} - E_{45}).$$

As in the previous example one can show that $\mathfrak{h} = \langle e_1, e_8 \rangle$ *is a Cartan subalgebra of* $\mathfrak{so}(5,\mathbb{C})$, *and the set generated by the matrices of the form*

$$H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Lambda & 0 \\ 0 & 0 & -\Lambda \end{bmatrix}, \text{ where } \Lambda = \operatorname{diag}(a,b),$$

can also be seen as a Cartan subalgebra of $\mathfrak{so}(5,\mathbb{C})$.

The Cartan-Killing form is a very powerful tool to characterize semisimple Lie algebras. Its non-degeneracy translates into the semisimplicity of the Lie algebra.

Definition 1.1.8 (Invariant forms). *Let* $\mathfrak{g}$ *be a Lie algebra over* $\mathbb{K}$. *A bilinear form* $\beta : \mathfrak{g}\times\mathfrak{g} \to \mathbb{K}$ *is said to be* ***invariant*** *if for every* $X, Y, Z \in \mathfrak{g}$ *holds*

$$\beta([X,Y],Z) = \beta(X,[Y,Z]).$$

For instance, we can equip $\mathfrak{gl}(V)$ with a natural invariant bilinear form defined by

$$\beta(X,Y) = \mathrm{tr}(XY).$$

To see that β is invariant, just remember that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$:

$$\begin{aligned} \beta([X,Y],Z) = \mathrm{tr}([X,Y]Z) &= \mathrm{tr}(XYZ - YXZ) \\ &= \mathrm{tr}(XYZ) - \mathrm{tr}(YXZ) \\ &= \mathrm{tr}(XYZ) - \mathrm{tr}(XZY) \\ &= \mathrm{tr}(XYZ - XZY) \\ &= \mathrm{tr}(X(YZ - ZY)) = \mathrm{tr}\,(X[Y,Z]) = \beta(X,[Y,Z]). \end{aligned}$$

Now let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism and β an invariant form in $\mathfrak{h}$. We can induce an invariant form β_ϕ in $\mathfrak{g}$ in the following way

$$\beta_\phi(X,Y) = \beta(\phi(X),\phi(Y)), \quad X,Y \in \mathfrak{g}.$$

As ϕ is a Lie algebra homomorphism and as β is invariant in $\mathfrak{h}$, it follows that β_ϕ is an invariant form in $\mathfrak{g}$:

$$\begin{aligned} \beta_\phi([X,Y],Z) &= \beta(\phi([X,Y]),\phi(Z)) \\ &= \beta([\phi(X),\phi(Y)],\phi(Z)) \\ &= \beta(\phi(X),[\phi(Y),\phi(Z)]) \\ &= \beta(\phi(X),\phi([Y,Z])) = \beta_\phi(X,[Y,Z]). \end{aligned}$$

If $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation of $\mathfrak{g}$ on V, it follows that the trace in $\mathfrak{gl}(V)$ induces an invariant form in $\mathfrak{g}$ by

$$\beta_\rho(X, Y) = \operatorname{tr}(\rho(X)\rho(Y)).$$

Definition 1.1.9. *The* ***Cartan-Killing form*** κ *in* $\mathfrak{g}$ *is defined by setting* ρ *as being the adjoint representation, that is,*

$$\kappa(X, Y) = \operatorname{tr}(\operatorname{ad}(X)\operatorname{ad}(Y)).$$

Proposition 1.1.10. *If* $\mathfrak{h}$ *is an ideal of the Lie algebra* $\mathfrak{g}$*, then the restriction of* κ *to* $\mathfrak{h} \times \mathfrak{h}$ *is the Cartan-Killing form of* $\mathfrak{h}$*.*

Proof: See [26], Section 3.2. □

The orthogonal complement of $W \subset \mathfrak{g}$ with respect to κ is defined as being the subset $W^\perp$ given by

$$W^\perp = \{X \in \mathfrak{g} \mid \kappa(X, Y) = 0, \forall\, Y \in W\}.$$

Note that if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{h}^\perp$ is also an ideal of $\mathfrak{g}$. In fact, given $X \in \mathfrak{h}^\perp$, $Y \in \mathfrak{g}$ and $Z \in \mathfrak{h}$ we have

$$\kappa([X, Y], Z) = \kappa(X, [Y, Z]) = 0,$$

since $[Y, Z] \in \mathfrak{h}$. Hence, $[X, Y] \in \mathfrak{h}^\perp$, showing that $\mathfrak{h}^\perp$ is an ideal of $\mathfrak{g}$.

Definition 1.1.11. *Let* $\mathfrak{g}$ *be a finite dimensional Lie algebra and* $\beta : \mathfrak{g} \times \mathfrak{g} \to \mathbb{K}$ *a bilinear invariant form. We say that* β *is* ***non-degenerate*** *if its* ***radical*** S *is* 0*, where* $S = \mathfrak{g}^\perp$*, that is,*

$$S = \{X \in \mathfrak{g} \mid \beta(X, Y) = 0 \;\; \forall\, Y \in \mathfrak{g}\}.$$

Theorem 1.1.12. *A Lie algebra* $\mathfrak{g}$ *is semisimple if and only if its Cartan-Killing form is non-degenerate.*

Proof: See [26], Theorem 3.8. □

A Lie algebra $\mathfrak{g}$ is the direct sum of ideals $\mathfrak{h}_1, \mathfrak{h}_2, \ldots \mathfrak{h}_k$ simply if $\mathfrak{g}$ is the direct sum of $\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_k$ as sum of vector subspaces. In special we have that $\mathfrak{h}_i \cap \mathfrak{h}_j = 0$ if $i \neq j$.

Also, $[\mathfrak{h}_i, \mathfrak{h}_j] = 0$ since $[\mathfrak{h}_i, \mathfrak{h}_j] \subset \mathfrak{h}_i \cap \mathfrak{h}_j$. The Lie brackets in the direct sum is naturally defined componentwise.

Theorem 1.1.13. *If $\mathfrak{g}$ is a semisimple Lie algebra, then there are simple ideals $\mathfrak{h}_1$, $\mathfrak{h}_2$,...,$\mathfrak{h}_k$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{h}_1 \oplus \mathfrak{h}_2 \oplus \cdots \oplus \mathfrak{h}_k$. Each simple ideal of $\mathfrak{g}$ coincides with some component $\mathfrak{h}_i$ and the Cartan-Killing form of $\mathfrak{h}_i$ is the restriction of κ to $\mathfrak{h}_i \times \mathfrak{h}_i$, for $i = 1, 2, \cdots, k$.*

Proof: See [26], Theorem 3.10. □

The converse of this statement is also true, that is, the direct sum of simple Lie algebras (or even semisimple) is also semisimple. To see this, suppose $\mathfrak{g} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$ with both components semisimple and let $\mathfrak{a}$ be an abelian ideal of $\mathfrak{g}$. If π_1 is the projection on $\mathfrak{h}_1$, then $\pi_1(\mathfrak{a})$ is an abelian ideal of $\mathfrak{h}_1$. Thus, $\pi_1(\mathfrak{a}) = 0$, that is, $\mathfrak{a} \subset \mathfrak{h}_2$. Similarly for the projection on $\mathfrak{h}_2$, we get $\mathfrak{a} \subset \mathfrak{h}_1$ which follows that $\mathfrak{a} = 0$.

This theorem simplifies the classification of the semisimple Lie algebras, as it implies that it is sufficient to classify all the simple ones. At the end of this section we present the classification of the classical real simple Lie algebras.

A real Lie algebra $\mathfrak{g}$ is said to admit a complex structure J when

(i) J is a complex structure on the vector space $\mathfrak{g}$ and

(ii) $\mathrm{ad}(X) \circ J = J \circ \mathrm{ad}(X)$, for every $X \in \mathfrak{g}$.

Note that the last condition implies that $[JX, JY] = -[X, Y]$. In fact, we have that

$$\begin{aligned} [X, JY] &= \mathrm{ad}(X) \circ J(Y) = J \circ \mathrm{ad}(X)(Y) = J[X, Y] = -J[Y, X] \\ &= -J \circ \mathrm{ad}(Y)(X) = -\mathrm{ad}(Y) \circ J(X) = -[Y, JX] \\ &= [JX, Y], \end{aligned}$$

that is, the condition (ii) tells us that $[X, JY] = [JX, Y] = J[X, Y]$, and hence,

$$[JX, JY] = J^2[X, Y] = -[X, Y].$$

A real vector space V endowed with a complex structure J can be regarded as a complex vector space $V^{\mathbb{C}}$ with the scalar multiplication

$$(a + \mathbf{i}b)X = aX + bJX, \quad X \in V, \ a, b \in \mathbb{R}.$$

The space $V^{\mathbb{C}}$ is called complex vector space associated to V, and it is different from the complexified space $V_{\mathbb{C}}$, obtained from V via complexification (see Appendix A for details).

Definition 1.1.14. *Let $\mathfrak{g}_c$ be a complex Lie algebra. A* ***real form*** *of $\mathfrak{g}_c$ is a subalgebra $\mathfrak{g}$ of the real Lie algebra $\mathfrak{g}_c^{\mathbb{R}}$ such that $\mathfrak{g}_c^{\mathbb{R}} = \mathfrak{g} + \mathbf{i}\mathfrak{g}$ (direct sum).*

Each $Z \in \mathfrak{g}_c$ can be uniquely written in the form $Z = X + \mathbf{i}Y$, $X, Y \in \mathfrak{g}$. The conjugation $\sigma : \mathfrak{g}_c \to \mathfrak{g}_c$ defined by $X + \mathbf{i}Y \mapsto X - \mathbf{i}Y$ is an automorphism of $\mathfrak{g}_c^{\mathbb{R}}$.

An antilinear and invertible transformation σ satisfying

$$[\sigma X, \sigma Y] = \sigma[X, Y]$$

is called **antiautomorphism**. It is clear that an antiautomorphism of $\mathfrak{g}_c$ is an automorphism of the realified space $\mathfrak{g}_c^{\mathbb{R}}$. It is also straightforward that the real subspace of $\mathfrak{g}_c$ formed by the fixed points of σ is a subalgebra of the realified $\mathfrak{g}_c^{\mathbb{R}}$, actually, if $\sigma(X) = X$ and $\sigma(Y) = Y$, then

$$\sigma[X, Y] = [\sigma X, \sigma Y] = [X, Y],$$

showing that the Lie brackets are also fixed under σ.

Hence, a real form of $\mathfrak{g}_c$ can also be defined as being a subalgebra of the realified $\mathfrak{g}_c^{\mathbb{R}}$ which is the subspace of fixed points of a conjugation σ that is also an antiautomorphism. And this means that the complex Lie algebra $\mathfrak{g}_c$ is the complexification of the real form. In fact, as σ is a linear transformation of $\mathfrak{g}_c^{\mathbb{R}}$ and an involution ($\sigma^2 = 1$), we have $\mathfrak{g}_c^{\mathbb{R}} = \mathfrak{g}_1 \oplus \mathfrak{g}_{-1}$, where $\mathfrak{g}_1$ and $\mathfrak{g}_{-1}$ are the eigenspaces associated to the eigenvalues 1 and -1, respectively. Now, as σ is antilinear in $\mathfrak{g}_c$, for a given $X \in \mathfrak{g}_1$ we have $\sigma(\mathbf{i}X) = -\mathbf{i}X$, that is, $\mathbf{i}X \in \mathfrak{g}_{-1}$. But this tells us that $J(\mathfrak{g}_1) \subset \mathfrak{g}_{-1}$, being J the complex structure of $\mathfrak{g}_c^{\mathbb{R}}$. Analogously, if $X \in \mathfrak{g}_{-1}$, then $\sigma(\mathbf{i}X) = -\mathbf{i}\sigma(X) = -\mathbf{i}(-X) = \mathbf{i}X$, that is, $J(\mathfrak{g}_{-1}) \subset \mathfrak{g}_1$. Since J is injective, it follows that $\dim \mathfrak{g}_1 = \dim \mathfrak{g}_{-1}$. Thus $\mathfrak{g}_c$ is the complexification of $\mathfrak{g}_1$.

A real Lie algebra $\mathfrak{g}$ is said to be **compact** when its Cartan-Killing form,

$$\langle X, Y \rangle_{\mathfrak{g}} = \mathrm{tr}\left(\mathrm{ad}(X)\mathrm{ad}(Y)\right),$$

is negative definite. This definition is justified by the fact that a real semisimple Lie algebra $\mathfrak{g}$ is the Lie algebra of a compact Lie group if and only if it is compact in the above sense (note that a compact Lie algebra is semisimple since its Cartan-Killing form is nondegenerate). A more detailed discussion about compact Lie algebras can be found in [28], Chapter 11.

Theorem 1.1.15. *Every complex semisimple Lie algebra $\mathfrak{g}_c$ admits a compact real form $\mathfrak{u}$.*

Proof: See [26], Theorem 12.13. □

Proposition 1.1.16. *If $\mathfrak{g}_0$ and $\mathfrak{g}_1$ are real forms of $\mathfrak{g}_c$ with σ_0 and σ_1 corresponding conjugations such that $\sigma_0\sigma_1 = \sigma_1\sigma_0$, then $\mathfrak{g}_1 = (\mathfrak{g}_1 \cap \mathfrak{g}_0) \oplus (\mathfrak{g}_1 \cap \mathbf{i}\mathfrak{g}_0)$.*

Proof: See [26], Proposition 12.15. □

This proposition implies that two compact real forms $\mathfrak{u}_0$ and $\mathfrak{u}_1$ of $\mathfrak{g}_c$ are equal if and only if the corresponding conjugations commute. In fact, if the conjugations commute, we have $\mathfrak{u}_1 = (\mathfrak{u}_1 \cap \mathfrak{u}_0) \oplus (\mathfrak{u}_1 \cap \mathbf{i}\mathfrak{u}_0)$. Given $X \in \mathfrak{u}_1 \cap \mathbf{i}\mathfrak{u}_0$ of the form $X = \mathbf{i}Y$, with $Y \in \mathfrak{u}_1$, the negative definiteness of the Cartan-Killing forms of $\mathfrak{u}_0$ and $\mathfrak{u}_1$ shows that

$$0 \geq \langle X, X \rangle = -\langle Y, Y \rangle \geq 0.$$

This means that $X = 0$, hence $\mathfrak{u}_1 \subset \mathfrak{u}_0$. A repetition of this argument interchanging the roles of $\mathfrak{u}_0$ and $\mathfrak{u}_1$ concludes the proof.

Theorem 1.1.17. *Let $\mathfrak{u}$ be a compact real form of the complex semisimple Lie algebra $\mathfrak{g}_c$ and let $\mathfrak{g}_0$ be any real form of $\mathfrak{g}_c$, with corresponding conjugation σ. Then there exists an inner automorphism ϕ of $\mathfrak{g}_c$ such that σ commutes with the conjugation corresponding to the compact real form $\phi(\mathfrak{u})$.*

Proof: See [26], Theorem 12.18. □

From this theorem the uniqueness of the compact real forms can be stated in the following sense: if $\mathfrak{u}_1$ and $\mathfrak{u}_2$ are two compact real forms of $\mathfrak{g}_c$, then there exists an automorphism ϕ of $\mathfrak{g}$ such that $\phi(\mathfrak{u}_2) = \mathfrak{u}_2$.

The following theorem summarizes the classification of the real forms of the classical simple complex Lie algebras:

Theorem 1.1.18 (Onishchik and Vinberg [17], Theorem 6, p. 233)**.** *Any real form of a classical simple complex Lie algebra $\mathfrak{g}$ is isomorphic to exactly one of the following real forms:*

1. $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$, $n \geq 2$

   (i) $\mathfrak{sl}(n, \mathbb{R})$

   (ii) $\mathfrak{sl}(n, \mathbb{H})$, $n = 2m$

   (iii) $\mathfrak{su}(p, q)$, $p + q = n$, $p = 0, 1, \ldots, [n/2]$

2. $\mathfrak{g} = \mathfrak{so}(n, \mathbb{C})$, $n = 3$ *or* $n \geq 5$

   (i) $\mathfrak{so}(p, q)$, $p + q = n$, $p = 0, 1, \ldots, [n/2]$

   (ii) $\mathfrak{u}^*_m(\mathbb{H})$, $n = 2m$.

3. $\mathfrak{g} = \mathfrak{sp}(n, \mathbb{C})$, $n = 2m \geq 2$

   (i) $\mathfrak{sp}(n, \mathbb{R})$, $n = 2m$

   (ii) $\mathfrak{sp}(p, q)$, $q + p = m$, $p = 0, 1, \ldots, [m/2]$.

The next theorem exhaust, up to isomorphisms, all the non-abelian real simple Lie algebras. For a detailed classification of the real forms of the exceptional simple complex Lie algebras we refer San Martin [26] and Helgason [10].

Theorem 1.1.19 ([26], Theorem 12.11)**.** *Let $\mathfrak{g}$ be a real simple Lie algebra. The possibilities for $\mathfrak{g}$ are:*

(i) $\mathfrak{g}$ *is the real form of a simple complex Lie algebra, or*

(ii) $\mathfrak{g}$ *is the realified of a simple complex Lie algebra.*

Note that this theorem completely classifies the real semisimple Lie algebras, since any semisimple Lie algebra decomposes as the direct sum of simple Lie algebras (see Theorem 1.1.13). In this work we will be specially interested in the Lie algebras $\mathfrak{sl}(n, \mathbb{H})$, $\mathfrak{so}(1, 4)$ and $\mathfrak{su}(1, 2)$, for their very rich intrinsic geometry.

1.2 Cartan and Iwasawa decompositions

Cartan decompositions

Let $\mathfrak{g}$ be a noncompact semisimple Lie algebra over $\mathbb{R}$, $\mathfrak{g}_{\mathbb{C}}$ its complexification and σ the conjugation in $\mathfrak{g}_{\mathbb{C}}$ with respect to $\mathfrak{g}$. A direct sum decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}$ is called a **Cartan decomposition** if there is a compact real form $\mathfrak{u}$ such that

$$\sigma(\mathfrak{u}) \subset \mathfrak{u}, \quad \mathfrak{g} \cap \mathfrak{u} = \mathfrak{k}, \quad \mathfrak{g} \cap \mathbf{i}\mathfrak{u} = \mathfrak{s}.$$

As $\mathfrak{u}$ is a subalgebra, we have $[\mathfrak{u}, \mathbf{i}\mathfrak{u}] \subset \mathbf{i}\mathfrak{u}$ and $[\mathbf{i}\mathfrak{u}, \mathbf{i}\mathfrak{u}] \subset \mathfrak{u}$. Note that the Lie brackets between the components of the Cartan decomposition satisfy

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{s}] \subset \mathfrak{s}, \quad [\mathfrak{s}, \mathfrak{s}] \subset \mathfrak{k}.$$

This means that $\mathfrak{k}$ is a subalgebra such that $\mathfrak{s}$ is left invariant under its adjoint representation. Note also that $\mathfrak{s}$ is not a subalgebra, otherwise we would have $[\mathfrak{s}, \mathfrak{s}] = 0$, and $\mathfrak{s}$ would be an abelian ideal, contradicting the semisimplicity of $\mathfrak{g}$.

If $\mathfrak{g}_0$ is a noncompact real form of the complex semisimple Lie algebra $\mathfrak{g}_c$ with corresponding conjugation σ, then for a given real compact form $\mathfrak{u}$ with conjugation τ, we can assume, without loss of generality, that $\tau\sigma = \sigma\tau$ (Theorem 1.1.17), and this means that $\mathfrak{g}_0$ is invariant under τ while $\mathfrak{u}$ is invariant under σ. Thus, Lemma 1.1.16 implies that

$$\mathfrak{g}_0 = \mathfrak{k} \oplus \mathfrak{s}, \quad \text{where} \quad \mathfrak{k} = \mathfrak{g}_0 \cap \mathfrak{u} \text{ and } \mathfrak{s} = \mathfrak{g}_0 \cap \mathbf{i}\mathfrak{u},$$

which is the Cartan decomposition for $\mathfrak{g}_0$. Of course the Cartan decomposition for $\mathfrak{g}_0$ depends on the choice of the real compact form $\mathfrak{u}$.

Proposition 1.2.1. *Given a Cartan decomposition $\mathfrak{g}_0 = \mathfrak{k} \oplus \mathfrak{s}$, the involutive automorphism θ defined by $\theta(X) = X$ if $X \in \mathfrak{k}$, and $\theta(Y) = -Y$ if $Y \in \mathfrak{s}$, is such that the bilinear form*

$$B_\theta(X, Y) = -\langle X, \theta Y \rangle$$

is an inner product in $\mathfrak{g}_0$. On the other hand, given an automorphism θ such that the bilinear form defined like above is an inner product, then its eigenspaces determine a Cartan decompo-

sition. The automorphism θ is called Cartan involution.

Proof: See [26], Proposition 12.21. $\square$

Proposition 1.2.2. *Let $\mathfrak{g}_0 = \mathfrak{k} \oplus \mathfrak{s}$ be a Cartan decomposition with involution θ. Then $\mathrm{ad}(X)$, $X \in \mathfrak{k}$ is skew-symmetric with respect to B_θ while $\mathrm{ad}(Y)$, $Y \in \mathfrak{s}$, is symmetric. Also, $\mathfrak{k}$ and $\mathfrak{s}$ are orthogonal with respect to both the Cartan-Killing form $\langle \cdot, \cdot \rangle$ and B_θ.*

Proof: See [26], Proposition 12.22. $\square$

In the case that the real semisimple Lie algebra is the realified of a complex Lie algebra, then its Cartan decompositions are the decompositions of the complex Lie algebra into real and immaginary parts with respect to the compact real forms:

Proposition 1.2.3. *Let $\mathfrak{g}_c$ be a complex semisimple Lie algebra and $\mathfrak{u}$ a compact real form. Then $\mathfrak{g}_c^{\mathbb{R}} = \mathfrak{u} \oplus \mathbf{i}\mathfrak{u}$ is a Cartan decomposition of the realified of $\mathfrak{g}_c$.*

Proof: See [26], Proposition 12.23. $\square$

Theorem 1.2.4. *Let $\mathfrak{g}_0$ be a real semisimple Lie algebra and consider $\mathfrak{g}_0 = \mathfrak{k}_1 \oplus \mathfrak{s}_1$ and $\mathfrak{g}_0 = \mathfrak{k}_2 \oplus \mathfrak{s}_2$ two Cartan decompositions of $\mathfrak{g}_0$. Then, there is an automorphism ϕ of $\mathfrak{g}_0$ such that $\phi(\mathfrak{k}_1) = \mathfrak{k}_2$ and $\phi(\mathfrak{s}_1) = \mathfrak{s}_2$.*

Proof: See [26], Theorem 12.24. $\square$

Example 1.2.5. *Consider $\mathfrak{g}_c = \mathfrak{sl}(n, \mathbb{C})$ and the conjugation σ defined by $\sigma(A) = -\overline{A}^T$. It is clear that σ is an antiautomorphism. The Lie algebra of fixed points of σ is*

$$\mathfrak{su}(n, \mathbb{R}) = \{A \in \mathfrak{sl}(n, \mathbb{C}) \mid A = -\overline{A}^T\}.$$

For the real form $\mathfrak{g}_0 = \mathfrak{sl}(n, \mathbb{R})$, we have that $\mathfrak{g}_0 \cap \mathfrak{u}$ is the subalgebra formed by the real skew-hermitian matrices, that is, the algebra $\mathfrak{so}(n)$, of the skew-symmetric matrices. On the other hand, $\mathfrak{g}_0 \cap \mathbf{i}\mathfrak{u}$ is the subspace formed by the real matrices X such that $\mathbf{i}X$ is skew-hermitian, that is, the subspace $\mathfrak{s}$ of the symmetric matrices. Hence, $\mathfrak{sl}(n, \mathbb{R}) = \mathfrak{so}(n, \mathbb{R}) \oplus \mathfrak{s}$ is a Cartan decomposition of $\mathfrak{sl}(n, \mathbb{R})$. The corresponding Cartan involution is given by $\theta(X) = -X^t$, since $\theta = 1$ in $\mathfrak{so}(n, \mathbb{R})$ and $\theta = -1$ in $\mathfrak{s}$.

Example 1.2.6. *For the Lie algebra of real symplectic matrices*

$$\mathfrak{sp}(n,\mathbb{R}) = \{A \in \mathfrak{gl}(2n,\mathbb{R}) \mid AJ + JA^T = 0\},$$

where J has the $n \times n$ block form

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

a Cartan decomposition is

$$\mathfrak{k} = \left\{ \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \;\middle|\; A + A^T = B - B^T = 0 \right\}$$

and

$$\mathfrak{s} = \left\{ \begin{bmatrix} A & B \\ B & -A \end{bmatrix} \;\middle|\; A - A^T = B - B^T = 0 \right\},$$

that is, skew-symmetric and symmetric matrices in $\mathfrak{sp}(n,\mathbb{R})$.

The following theorem states the Cartan decomposition at the Lie group level. It shows that a noncompact semisimple Lie group G can be written as $G = KS = SK$, with $K = \exp\mathfrak{k}$ and $S = \exp\mathfrak{s}$, where $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}$ is a Cartan decomposition of the Lie algebra $\mathfrak{g}$ of G.

Theorem 1.2.7 (Global Cartan Decomposition)**.** *Let G be a connected semisimple Lie group and $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}$ a Cartan decomposition for its Lie algebra. Write $K = \langle \exp\mathfrak{k} \rangle$ and $S = \exp\mathfrak{s}$. Then,*

1. *$G = SK = KS$ and every $g \in G$ is uniquely written as $g = sk$ or $g = ks$, where $k \in K$ and $s \in S$.*
2. *S is an embedded submanifold of G diffeomorphic to $\mathfrak{s}$ under the embedding $\exp : \mathfrak{s} \to S$.*
3. *The maps $K \times S \to G$ given by $(k, s) \mapsto ks$ and $(k, s) \mapsto sk$ are diffeomorphisms.*
4. *The center $Z(G)$ of G is contained in K.*
5. *$K = \exp\mathfrak{k}$ and K is compact if and only if $Z(G)$ is finite.*

Proof: See [28], Theorem 12.4. □

Iwasawa decompositions

As in the case of the complex semisimple Lie algebras, the real ones also admit root spaces decompositions. In the real case, instead of considering Cartan subalgebras, the root space decompositions are obtained from maximal abelian subalgebras contained in the symmetric part of the Cartan decomposition.

Let $\mathfrak{g}$ be a real semisimple Lie algebra, $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}$ a Cartan decomposition and $\mathfrak{a} \subset \mathfrak{s}$ a **maximal abelian** subalgebra, in the sense that it is not contained in any other abelian subalgebra of $\mathfrak{s}$. It is easy to ensure the existence of such subalgebras, since the unidimensional subspaces in $\mathfrak{s}$ are abelian subalgebras of $\mathfrak{g}$ contained in $\mathfrak{s}$.

For instance, considering the Cartan decomposition $\mathfrak{sl}(n,\mathbb{R}) = \mathfrak{so}(n) \oplus \mathfrak{s}$, where $\mathfrak{s}$ is the subspace of symmetric matrices, we have that the subalgebra $\mathfrak{a}$ of diagonal matrices is a maximal abelian subalgebra contained in $\mathfrak{s}$. Note that $\mathfrak{a}$ is also a Cartan subalgebra, even though this fact does not occurs in general. A general fact is that a maximal abelian subalgebra $\mathfrak{a}$ is always contained in a Cartan subalgebra (often being different from it).

Proposition 1.2.8. *Let $\mathfrak{a}$ be a maximal abelian subalgebra contained in $\mathfrak{s}$. There is a maximal abelian subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ containing $\mathfrak{a}$. The subalgebra $\mathfrak{h}$ is a Cartan subalgebra and it decomposes as a direct sum $\mathfrak{h} = (\mathfrak{h} \cap \mathfrak{k}) \oplus \mathfrak{a}$.*

Proof: See [26], Proposition 12.25. □

Let α be a real linear functional on $\mathfrak{a}$ and consider the subspace

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

Such a functional $\alpha \neq 0$ is called a **restricted root** of $\mathfrak{g}$ with respect to $\mathfrak{a}$ if $\mathfrak{g}_\alpha \neq 0$.

The zero functional appears as a weight of the adjoint representation of $\mathfrak{a}$ in $\mathfrak{g}$, since $\mathfrak{a}$ is abelian, and the subspace associated to the zero weight is the centralizer of $\mathfrak{a}$, $\mathfrak{z}(\mathfrak{a})$, because $\mathrm{ad}(H)$, $H \in \mathfrak{a}$, is diagonalizable. So, $\mathfrak{g}$ decomposes as

$$\mathfrak{g} = \mathfrak{z}(\mathfrak{a}) \oplus \sum_\alpha \mathfrak{g}_\alpha.$$

Considering the Cartan decomposition of the elements of $\mathfrak{z}(\mathfrak{a})$, writing $\mathfrak{m} = \mathfrak{z}(\mathfrak{a}) \cap \mathfrak{k}$, then $\mathfrak{z}(\mathfrak{a}) = \mathfrak{m} \oplus \mathfrak{a}$ (see [28], Proposition 12.5) and the previous decomposition can be rewritten as

$$\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{a} \oplus \sum_{\alpha} \mathfrak{g}_{\alpha}.$$

The set of restricted roots is a finite set of linear functionals in $\mathfrak{a}$, and thus the subset

$$\overline{\mathfrak{a}} = \{H \in \mathfrak{a} \,|\, \alpha(H) \neq 0 \text{ for every root } \alpha\}$$

is an open and dense subset of $\mathfrak{a}$. An element $H \in \overline{\mathfrak{a}}$ is called **real regular**. The centralizer of a real regular element is $\mathfrak{z}(\mathfrak{a})$ and from this follows that if $X \in \mathfrak{s}$ is such that $[H, X] = 0$ for some real regular element H, then $X \in \mathfrak{a}$. In other words, the set of elements in $\mathfrak{s}$ that commute with a real regular element H is exactly $\mathfrak{a}$.

The above decompositions depend upon the choice of the maximal abelian subalgebra $\mathfrak{a}$, but this choice dos not affect the generality, since two maximal abelian subalgebras $\mathfrak{a}_1, \mathfrak{a}_2 \subset \mathfrak{s}$ are obtained from one another by an element of K_{ad} (see [28], Proposition 12.6 and its corollary).

The inclusion $[\mathfrak{k}, \mathfrak{s}] \subset \mathfrak{s}$ shows that the subspace $\mathfrak{s}$ is left invariant under the subgroup K_{ad}. Thus, if $\mathfrak{a} \subset \mathfrak{s}$ is maximal abelian and $k \in K_{\text{ad}}$, then $k(\mathfrak{a}) \subset \mathfrak{s}$ is also a maximal abelian subalgebra. Every single maximal abelian subalgebra is obtained in this way by conjugations under the elements of K_{ad}:

The common dimension of the maximal abelian subalgebras $\mathfrak{a} \subset \mathfrak{s}$ is called **real rank** of $\mathfrak{g}$, and in general the real rank differs from the Lie algebra's rank (the dimension of its Cartan subalgebras).

To define an Iwasawa decomposition for $\mathfrak{g}$, we start choosing:

- a Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}$,
- a maximal abelian subalgebra $\mathfrak{a} \subset \mathfrak{s}$ and
- a real regular element $H \in \mathfrak{a}$.

From the decomposition of $\mathfrak{g}$ into the root spaces of $\mathfrak{a}$, we define

$$\mathfrak{n} = \mathfrak{n}_H^+ = \sum_{\alpha(H)>0} \mathfrak{g}_{\alpha},$$

which is the sum of the eigenspaces of $\mathrm{ad}(H)$ associated with positive eigenvalues. With these choices, the decomposition

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$$

is called an **Iwasawa decomposition** for $\mathfrak{g}$.

Theorem 1.2.9. *The Iwasawa decomposition is a direct sum decomposition.*

Proof: See [28], Theorem 12.8. □

The component $\mathfrak{k}$ in an Iwasawa decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ is a compact subalgebra while $\mathfrak{a}$ is abelian. With respect to $\mathfrak{n}$ and $\mathfrak{a} \oplus \mathfrak{n}$ we have the following result:

Proposition 1.2.10. *The component $\mathfrak{n}$ of the Iwasawa decomposition is a nilpotent subalgebra and $\mathfrak{a} \oplus \mathfrak{n}$ is a solvable Lie algebra.*

Proof: See [28], Proposition 12.9. □

Example 1.2.11. *For $\mathfrak{sl}(n, \mathbb{R})$, choosing the real regular element $H = \mathrm{diag}\{a_1, \ldots, a_n\}$, with $a_1 > a_2 > \cdots > a_n$, we get*

- $\mathfrak{k} = \mathfrak{so}(n)$*;*
- $\mathfrak{a}$ *is the algebra of traceless diagonal matrices;*
- $\mathfrak{n}$ *is the algebra of upper triangular matrices with zero diagonal;*

Finally, at the Lie group level the Iwasawa decompositions are described by the following theorem.

Theorem 1.2.12 ([28], Theorem 12.12)**.** *Let G be a connected semisimple Lie group and $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ an Iwasawa decomposition for its Lie algebra $\mathfrak{g}$. Then, $G = KAN$, where $K = \exp \mathfrak{k}$, $A = \exp \mathfrak{a}$ and $N = \exp \mathfrak{n}$. The map*

$$\phi = K \times A \times N \to KAN,$$

given by $\phi(k, a, n) = kan$, is a diffeomorphism. The groups A, N and AN are simply connected and diffeomorphic to euclidean spaces.

The subgroup K is the same of the Cartan decomposition. The abelian subgroup $A = \exp \mathfrak{a}$ is closed, since $S = \exp \mathfrak{s}$ is closed in G, $\mathfrak{a}$ is closed in $\mathfrak{s}$ and $\exp \mathfrak{s} \to S$ is a diffeomorphism.

Example 1.2.13. *Let $G = \mathrm{Sl}(n, \mathbb{R})$ and the Iwasawa decomposition of its Lie algebra $\mathfrak{sl}(n, \mathbb{R}) = \mathfrak{so}(n) \oplus \mathfrak{a} \oplus \mathfrak{n}$, where $\mathfrak{a}$ is the algebra of diagonal matrices and $\mathfrak{n}$ the algebra of upper triangular matrices. We get then the Iwasawa decomposition $\mathrm{Sl}(n, \mathbb{R}) = \mathrm{SO}(n)AN$, where A is the group formed by the diagonal matrices with positive entries and determinant 1 and N is the group formed by the upper triangular matrices having 1 at each diagonal entry.*

1.3 Flag manifolds

Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}^+$ an Iwasawa decomposition and let $\mathfrak{m} = \mathfrak{g}_0 \cap \mathfrak{k}$ be the centralizer of $\mathfrak{a}$ in $\mathfrak{k}$. As $\mathfrak{a}$ is maximal abelian in $\mathfrak{s}$, we have $\mathfrak{g}_0 = \mathfrak{m} \oplus \mathfrak{a}$. The subalgebra $\mathfrak{p} = \mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n}^+$ is called standard minimal parabolic subalgebra. It is in fact a subalgebra since $[\mathfrak{m}, \mathfrak{n}^+] \subset \mathfrak{n}^+$ and $[\mathfrak{a}, \mathfrak{n}^+] \subset \mathfrak{n}^+$. If g is an automorphism of $\mathfrak{g}$, then $g \cdot \mathfrak{g}$ is a subalgebra.

We define a parabolic minimal subalgebra of $\mathfrak{g}$ as being a subalgebra $\mathfrak{q}$ that is conjugate to the standard minimal parabolic subalgebra via an automorphism g of $\mathfrak{g}$. In other words, $\mathfrak{q}$ is parabolic minimal if for some automorphism g of $\mathfrak{g}$ it holds $\mathfrak{q} = g \cdot \mathfrak{p}$. A parabolic subalgebra is a subalgebra of $\mathfrak{g}$ containing a minimal parabolic subalgebra.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. A parabolic subgroup of G is the normalizer in G of a parabolic subalgebra, that is, if $\mathfrak{p} \subset \mathfrak{g}$ is a parabolic subalgebra then

$$P = \{g \in G \,|\, \mathrm{Ad(g)}\mathfrak{p} = \mathfrak{p}\}$$

is the parabolic subgroup of G associated to $\mathfrak{p}$. Note that P is a Lie subgroup of G, since it is a closed subgroup.

A flag manifold of a Lie group G is a coset space G/Q with Q a parabolic subgroup of G. For the group $SU(1, 2)$ the only flag manifold is given by the standard minimal parabolic subgroup P as above. Since a parabolic subgroup Q is always conjugate to a standard parabolic subgroup P_Θ, it is enough consider standard parabolic subgroups to obtain the flag manifolds for a given Lie group G. Given a simple root system Σ, the subsets of Σ are ordered by inclusion and this induces an ordering of the flag manifolds as well. We say that a flag manifold $\mathbb{F}_{\Theta_1} = G/P_{\Theta_1}$ is bigger than $\mathbb{F}_{\Theta_2} = G/P_{\Theta_2}$ if and only if $\Theta_1 \subset \Theta_2$. Thus, under this ordering we get only one maximal flag manifold, corresponding to $\Theta = \emptyset$. There are $\mathrm{rank}_{\mathbb{R}}\mathfrak{g} - 1$ minimal flag manifolds, which are exactly those corresponding to the subsets Θ that are complementary to the sets having only one element.

Flag manifolds are connected (because G is connected) and the following proposition ensures that flag manifolds are compact as well.

Proposition 1.3.1. *K acts transitively on any flag manifold $\mathbb{F}_{\Theta}$.*

Proof: First of all, we know that $\mathbb{F}_{\Theta} = G/P_{\Theta}$ has Lie algebra $\mathfrak{g}/\mathfrak{p}_{\Theta}$, and $\dim \mathbb{F}_{\Theta} = \dim \mathfrak{g} - \dim \mathfrak{p}_{\Theta}$. Let $\pi : G \to G/P_{\Theta}$ be the canonical projection. The orbit $K \cdot b_0$ of the origin $b_0 \in G/P_{\Theta}$ is a submanifold of G/P_{Θ} with tangent space $d\pi_1(\mathfrak{k})$, and $\dim d\pi_1(\mathfrak{k}) = \dim \mathfrak{k} - \dim(\mathfrak{k} \cap \mathfrak{p})$. Now, the Iwasawa decomposition gives us $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}_{\Theta}$, hence $\dim \mathfrak{g} = \dim \mathfrak{k} + \dim \mathfrak{p}_{\Theta} - \dim(\mathfrak{k} \cap \mathfrak{p}_{\Theta})$. This implies that the dimension of the orbit $K \cdot b_0$ coincides with $\dim(G/P_{\Theta})$, and hence the orbit is an open submanifold in G/P_{Θ}. Since K is compact, we get that $K \cdot b_0$ is closed as well. Finally, the connectedness of $\mathbb{F}_{\Theta} = G/P_{\Theta}$ implies that $K \cdot b_0 = \mathbb{F}_{\Theta}$, that is, K acts transitively on the flag manifold $\mathbb{F}_{\Theta}$. $\square$

1.4 Control systems and controllability

Here we denominate control system a family of a complete vector fields Γ on an n-dimensional manifold M. A trajectory of Γ is a continuous curve γ from an interval $[0, T]$, $T \geq 0$ of the real line into M such that for some partition $0 < t_1 < t_2 < \cdots < t_n = T$ there exist vector fields $X_1, \ldots, X_n$ in Γ such that the restriction of γ to each interval $[t_{i-1}, t_i)$ is an integral curve of X_i. A special case studied here are the control systems given by the equations

$$\dot{x}(t) = X_1(x) + \Sigma_{i=2}^{l} u_i X_i(x),$$

where $X_j \in \Gamma$ and

$$\mathcal{U} = \{u \in L^\infty(\mathbb{R}, \mathbb{R}^m) \mid u(t) \in U \text{ for almost all } t\}.$$

We assume that the control range $U \subset \mathbb{R}^m$ is nonempty and that for every initial state $x \in M$ and every $u \in \mathcal{U}$ there exists a unique solution denoted by $\phi(t, x, u), t \in \mathbb{R}$, satisfying $\phi(0, x, u) = x$. We can also admit piecewise constant controls, that is,

$$\mathcal{U}_{pc} = \{u : \mathbb{R} \to U \subset \mathbb{R}^m \mid u \text{ picewise constant}\}.$$

The positive orbit of $x \in M$ at time exactly $t > 0$ is the set

$$\mathcal{O}_t^+(x) = \{y \in M \mid \exists u \in \mathcal{U} \text{ with } y = \phi(t, x, u)\}.$$

Similarly, the negative orbit of $x \in M$ at time $t > 0$ is

$$\mathcal{O}_t^-(x) = \{y \in M \mid \exists u \in \mathcal{U} \text{ with } x = \phi(t, y, u)\}.$$

The positive and negative orbits of $x \in M$ up to time T are defined as

$$\begin{aligned} \mathcal{O}_{\leq T}^+(x) &= \bigcup\nolimits_{0 \leq t \leq T} \mathcal{O}_t^+(x) \\ &= \{y \in M, \text{ such that there are } 0 \leq t \leq T \text{ and } u \in \mathcal{U} \text{ with } y = \phi(t, x, u)\} \end{aligned}$$

and

$$\begin{aligned} \mathcal{O}_{\leq T}^-(x) &= \bigcup\nolimits_{0 \leq t \leq T} \mathcal{O}_t^-(x) \\ &= \{y \in M, \text{ such that there are } 0 \leq t \leq T \text{ and } u \in \mathcal{U} \text{ with } x = \phi(t, y, u)\} \end{aligned}$$

The positive and negative orbits of $x \in M$ are

$$\mathcal{O}^+(x) = \bigcup\nolimits_{T>0} \mathcal{O}_{\leq T}^+(x) = \bigcup\nolimits_{t>0} \mathcal{O}_t^+(x)$$

and

$$\mathcal{O}^-(x) = \bigcup\nolimits_{T>0} \mathcal{O}_{\leq T}^-(x) = \bigcup\nolimits_{t>0} \mathcal{O}_t^-(x).$$

A key concept for this work is that one of controllability of a control system Γ, which roughly speaking means that the orbit of every single point of M under Γ covers the whole manifold M.

Definition 1.4.1. *A control system Γ is controllable from $x \in M$ when $\mathcal{O}^+(x) = M$, and it is said to be controllable when it is controllable from every $x \in M$.*

When a control system Γ fails to be controllable, one can ask for the maximal subsets of M where controllability holds. These are the control sets.

Definition 1.4.2. *A nonvoid set $C \subset M$ is an invariant control set of a control system Γ on M if it has the following properties:*

(i) for all $x \in C$ there is a control $u \in \mathcal{U}$ such that $\phi(t, x, u) \in C$ for all $t \geq 0$,

(ii) for all $x \in C$ one has $\mathrm{cl}C = \mathrm{cl}(\mathcal{O}^+(x))$, and

(iii) C is maximal with these properties, that is, if $C' \supset C$ satisfies conditions (i) and (ii), then $C' = C$.

As special cases of control systems that occur very often in the literature we have:

1. The bilinear control systems with unrestricted controls,

$$\dot{x} = Ax + \sum_{i=1}^{m} u_i(t) B_i x,$$

where $A, B_1, B_2, \ldots, B_m \in M_n(\mathbb{R})$ and $u(t) = (u_1(t), \ldots, u_m(t)) \in \mathbb{R}^m$. This control system is given by the family

$$\Gamma = \{X_u = A + u_1 B_1 + \cdots + A_m B_m \,|\, u = (u_1, \ldots, u_m) \in \mathbb{R}^m\}$$

of vector fields, where $X_u(x) = Ax + \sum_{i=1}^{m} u_i(t) B_i x$, $x \in \mathbb{R}^n$. These are also called control-affine systems.

2. The affine systems,

$$\dot{x} = Ax + a + \sum_{i=1}^{m} u_i(t)(B_i x + b_i), \qquad x \in \mathbb{R}^n,$$

being $A, B_1, \ldots, B_m \in M_n(\mathbb{R})$, $a, b_1, \ldots, b_m \in \mathbb{R}^n$, $u(t) = (u_1(t), \ldots, u_m(t)) \in \mathbb{R}^m$. As before, this system can be described by the family

$$\Gamma = \{X_u = (A, a) + u_1(B_1, b_1) + \cdots + u_m(B_m, b_m) \,|\, u = (u_1, \ldots, u_m) \in \mathbb{R}^m\},$$

where $X_u(x) = Ax + a + \sum_{i=1}^{m} u_i(t)(B_i x + b_i)$, $x \in \mathbb{R}^n$.

Another important class of control systems are the invariant control systems, given by a family of (right or left) invariant vector fields on a Lie group G, that is, a subset Γ of the Lie algebra $\mathfrak{g}$ of G.

Definition 1.4.3. *An invariant control system* $\Gamma \subset \mathfrak{g}$ *is said to satisfy the Lie algebra rank condition (LARC) if the Lie algebra generated by* Γ, $\mathrm{Lie}(\Gamma)$, *is the whole* $\mathfrak{g}$, *that is,*

$$\mathrm{Lie}(\Gamma) = \mathfrak{g}.$$

Complete controllability can occur only on connected Lie groups and the Lie algebra rank condition is a necessary condition for controllability, although in general it is not sufficient.

In this work we are going to give sufficient conditions for the controllability of some invariant control systems on certain Lie groups and study the invariant control sets for several non controllable systems.

CHAPTER 2

CONTROLLABILITY AND SEMIGROUPS OF INVARIANT CONTROL SYSTEMS ON $\mathrm{Sl}(n, \mathbb{H})$

Let $\mathrm{Sl}(n, \mathbb{H})$ be the Lie group of $n \times n$ quaternionic matrices g with $|\det g| = 1$. In this chapter we prove that a subsemigroup $S \subset \mathrm{Sl}(n, \mathbb{H})$ with nonempty interior is equal to $\mathrm{Sl}(n, \mathbb{H})$ if S contains a special subgroup isomorphic to $\mathrm{Sl}(2, \mathbb{H})$. From this, we give sufficient conditions on $A, B \in \mathfrak{sl}(n, \mathbb{H})$ to ensure that the invariant control system $\dot{g} = Ag + uBg$ is controllable on $\mathrm{Sl}(n, \mathbb{H})$. We prove also that these conditions are generic in the sense that we obtain an open and dense set of controllable pairs $(A, B) \in \mathfrak{sl}(n, \mathbb{H})^2$.

2.1 Lie theoretical setting for $\mathrm{Sl}(n, \mathbb{H})$

In this section we establish some necessary notations, concepts and results. For details see Rabelo and San Martin [18], San Martin [27, 28] and San Martin and Tonelli [29].

Let G be a semi-simple connected and noncompact Lie group with finite center and denote by $\mathfrak{g}$ its Lie algebra. We describe the flag manifolds of G from the simple roots of $\mathfrak{g}$. Choose an Iwasawa decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}^+$ and take Π the set of roots of the pair $(\mathfrak{g}, \mathfrak{a})$. Denote by Π^+ and Σ the set of positive and simple roots, respectively, which correspond to the nilpotent component

$$\mathfrak{n}^+ = \sum_{\alpha \in \Pi^+} \mathfrak{g}_\alpha,$$

with $\mathfrak{g}_\alpha$ standing for the α-root space. Consider $\mathfrak{m}$ the centralizer of $\mathfrak{a}$ in $\mathfrak{k}$. Then the standard minimal parabolic subalgebra of $\mathfrak{g}$ is given by $\mathfrak{p} = \mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n}^+$. We can define the maximal flag manifold $\mathbb{F}$ of G as G/P where P is the minimal parabolic subgroup defined as the normalizer of $\mathfrak{p}$ in G. Recall that $P = MAN$ where $A = \exp \mathfrak{a}$, $N = \exp \mathfrak{n}^+$ and M is the centralizer of A in $K = \exp \mathfrak{k}$.

Given a subset $\Theta \subset \Sigma$, let $\mathfrak{n}^-(\Theta)$ be the subalgebra spanned by the root spaces $\mathfrak{g}_{-\alpha}$, $\alpha \in \langle \Theta \rangle$ where $\langle \Theta \rangle$ is the set of positive roots generated by Θ. We define the parabolic subalgebra associated to Θ by

$$\mathfrak{p}_\Theta = \mathfrak{n}^-(\Theta) \oplus \mathfrak{p}.$$

Then we have the homogeneous space $\mathbb{F}_\Theta = G/P_\Theta$, also called as (partial) flag manifold where P_Θ is the normalizer of $\mathfrak{p}_\Theta$ in G, that is, $P_\Theta = \{g \in G : \mathrm{Ad}(g)\, \mathfrak{p}_\Theta = \mathfrak{p}_\Theta\}$ and called parabolic subgroup (associated to Θ).

Let

$$\mathfrak{a}^+ = \{H \in \mathfrak{a} : \alpha(H) > 0 \text{ for all } \alpha \in \Sigma\}$$

be the Weyl chamber associated to Σ. We say that $X \in \mathfrak{g}$ is regular in case $X = \mathrm{Ad}(g)(H)$ for some $g \in G$ and $H \in \mathfrak{a}^+$. Analogously, $x \in G$ is said to be regular in case $x = ghg^{-1}$ with $h \in A^+ = \exp \mathfrak{a}^+$, that is, $x = \exp X$ with X regular in $\mathfrak{g}$.

The Weyl group $\mathcal{W}$ associated to $\mathfrak{a}$ is the finite group generated by the reflections over the root hyperplanes $\alpha = 0$ contained in $\mathfrak{a}$, $\alpha \in \Sigma$. For $\Theta \subset \Sigma$ there is a subgroup $\mathcal{W}_\Theta \subset \mathcal{W}$ generated by the reflections w.r.t. $\alpha \in \Theta$. Let $b_\Theta \in \mathbb{F}_\Theta$ be the origin in the sense of [18]. Then the Bruhat decomposition of $\mathbb{F}_\Theta$ is given by

$$\mathbb{F}_\Theta = \bigcup_{w \in \mathcal{W}/\mathcal{W}_\Theta} N \cdot w b_\Theta.$$

Hence, a Schubert cell in $\mathbb{F}_\Theta$ can be defined as $\mathrm{cl}(N \cdot w b_\Theta)$.

Given two subsets $\Theta_1 \subset \Theta_2 \subset \Sigma$, the corresponding parabolic subgroups satisfy $P_{\Theta_1} \subset P_{\Theta_2}$, then there is a canonical fibration $G/P_{\Theta_1} \to G/P_{\Theta_2}$. In particular, $\mathbb{F} = \mathbb{F}_\emptyset$ projects onto every flag manifold $\mathbb{F}_\Theta$.

Take a semigroup S with $\mathrm{int} S \neq \emptyset$. Consider the action of S in the flag manifolds of G. It was proved in [29] that S is not transitive in $\mathbb{F}_\Theta$ unless $S = G$. Moreover, there exists just one closed invariant subset $C_\Theta \subset \mathbb{F}_\Theta$ such that Sx is dense in C_Θ for all $x \in C_\Theta$. This subset is called the invariant control set of S in C_Θ. Since S is not transitive, $C_\Theta \neq \mathbb{F}_\Theta$.

There exists $\Theta \subset \Sigma$ such that $\pi_\Theta^{-1}(C_\Theta) \subset \mathbb{F}$ is the invariant control set in the maximal flag manifold. Among the subsets Θ satisfying this property, there is one which is maximal, in the sense that it contains all the others. We denote this subset by $\Theta(S)$ and say that it is the flag type of S. Alternatively, we denote this type of S by the corresponding flag manifold $\mathbb{F}(S) = \mathbb{F}_{\Theta(S)}$. Furthermore, $C_{\Theta(S)}$ is contractible in $\mathbb{F}_{\Theta(S)}$ (see e.g. [27] and references therein).

Now, we collect and specialize some of the above concepts in case of $\mathrm{Sl}(n,\mathbb{H})$. In a matrix Lie algebra, a Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}$ is given by skew symmetric and symmetric (or hermitian) matrices. Hence, the natural Cartan decomposition of $\mathfrak{sl}(n,\mathbb{H})$ is given by

$$\mathfrak{k} = \{X \in M_{n\times n}(\mathbb{H}) : X = -\overline{X}^T\} \text{ and } \mathfrak{s} = \{X \in M_{n\times n}(\mathbb{H}) : X = \overline{X}^T\}$$

where $\overline{X}$ is a quaternionic conjugation. The algebra $\mathfrak{k}$ of the quaternionic skew hermitian matrices is denoted by $\mathfrak{k} = \mathfrak{sp}(n)$ and is the compact real form of $\mathfrak{sp}(n,\mathbb{C})$. Note that a real form of a complex Lie algebra is compact if its Cartan-Killing form is negative definite (see e.g. [28]).

The maximal abelian subalgebra $\mathfrak{a} \subset \mathfrak{s}$ is given by the diagonal matrices $\Lambda = \mathrm{diag}\{a_1,\ldots,a_n\}$ with $a_i \in \mathbb{R}$ and $\mathrm{tr}\Lambda = 0$. The roots of $\mathfrak{a}$ are the following linear functionals

$$\alpha_{rs}(\Lambda) = (\lambda_r - \lambda_s)(\Lambda) = a_r - a_s \qquad r \neq s.$$

The vector space $\mathfrak{g}_{\alpha_{rs}}$, corresponding to the root α_{rs}, is given by the quaternionic matrices with non zero entries only in the position rs. Then all roots have multiplicity 4. The set of simple roots is given by

$$\begin{aligned}\Sigma &= \{\lambda_1 - \lambda_2, \ldots, \lambda_{n-1} - \lambda_n\} \\ &= \{\alpha_{12}, \ldots, \alpha_{(n-1)n}\}.\end{aligned}$$

With this choice, the set of positive roots is formed by α_{rs} with $r < s$. Hence, an Iwasawa decomposition is $\mathfrak{sl}(n,\mathbb{H}) = \mathfrak{sp}(n) \oplus \mathfrak{a} \oplus \mathfrak{n}^+$, where $\mathfrak{n}^+$ is the Lie algebra of upper triangular quaternionic $n \times n$ matrices with zero entries in the diagonal.

For $d = 1, \ldots, n-1$, we denote by $\mathrm{Gr}_d(\mathbb{H})$ the grassmannian of d-dimensional quaternionic subspaces of $\mathbb{H}^n$. The group $\mathrm{Sl}(n, \mathbb{H})$ acts transitively on each $\mathrm{Gr}_d(\mathbb{H})$ and the compact subgroup $\mathrm{Sp}(n) \subset \mathrm{Sl}(n, \mathbb{H})$ also acts transitively on $\mathrm{Gr}_d(\mathbb{H})$.

Theorem 2.1.1. *Let $S \subset \mathrm{Sl}(n, \mathbb{H})$ be a proper subsemigroup with $\mathrm{int} S \neq \emptyset$. Then there are $d \in \{1, \ldots, n-1\}$ and a subset $C_d \subset \mathrm{Gr}_d(\mathbb{H})$ satisfying*

1. *C_d is closed, has nonempty interior and it is invariant by the action of S. (C_d is the unique invariant control set of S in $\mathrm{Gr}_d(\mathbb{H})$).*
2. *C_d is contractible in $\mathrm{Gr}_d(\mathbb{H})$ in the sense that there exists $H \in \mathfrak{sl}(n, \mathbb{H})$ such that $e^{tH} C_d$ shrinks to a point as $t \to +\infty$.*

In the context of the above theorem, the grassmannian $\mathrm{Gr}_d(\mathbb{H})$ is the flag type of the semigroup S.

2.2 Transitivity of a subsemigroup of $\mathrm{Sl}(n, \mathbb{H})$

In this section, following the same constructions and notations developed in the previous sections, we prove our central result that gives sufficient conditions for a semigroup S to be equal to $\mathrm{Sl}(n, \mathbb{H})$. The proof of this result is based on the existence of a flag type of a proper semigroup S with $\mathrm{int} S \neq \emptyset$. By Theorem 2.1.1, we get that $S = \mathrm{Sl}(n, \mathbb{H})$ if we can prove that S does not leave invariant contractible subsets in the grassmannians $\mathrm{Gr}_d(\mathbb{H})$, $d = 1, \ldots, n-1$.

In the Lie group $\mathrm{Sl}(n, \mathbb{H})$, the connected Lie subgroup $\mathcal{G} = \langle \exp \mathfrak{g}_{\pm\alpha} \rangle$ is described in the following way. For a ordered pair (r, s), $1 \leq r < s \leq n$, let $\mathrm{Sl}(2, \mathbb{H})_{r,s}$ be the subgroup of $\mathrm{Sl}(n, \mathbb{H})$ (isomorphic to $\mathrm{Sl}(2, \mathbb{H})$) of the matrices in the space $\mathrm{span}\{e_r, e_s\}$ where $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{H}^n$, plugged into the $n \times n$ matrices. That is, $\mathrm{Sl}(2, \mathbb{H})_{r,s} = \langle \exp \mathfrak{sl}(2, \mathbb{H})_{r,s} \rangle$ where the Lie algebra $\mathfrak{sl}(2, \mathbb{H})_{r,s}$ is given by the matrices in $\mathfrak{sl}(n, \mathbb{H})$ having nonzero entries only at the positions (a, b) with $a, b \in \{r, s\}$.

Now, we are able to state the main result of this section.

Theorem 2.2.1. *Let $S \subset \mathrm{Sl}(n, \mathbb{H})$ be a subsemigroup with $\mathrm{int} S \neq \emptyset$ and suppose that $\mathrm{Sl}(2, \mathbb{H})_{r,s} \subset S$ for some pair of indices (r, s), $1 \leq r < s \leq n$. Then $S = \mathrm{Sl}(n, \mathbb{H})$.*

Before proving this theorem, we need some remarks and lemmas. In the hypothesis of the above theorem we consider separately the case where $(r, s) = (1, n)$, that is, $\mathrm{Sl}(2, \mathbb{H})_{1,n} \subset S$. Then the strategy to prove this case will be based in Theorem 2.1.1.

Since we have $\mathrm{int}(S) \neq \emptyset$, the idea is to show that for all $d \in \{1, \ldots, n-1\}$ the S-invariant closed C_d sets with $\mathrm{int}(C_d) \neq \emptyset$ are not contractible in the sense of the item 2 of Theorem 2.1.1.

Let $C \subset \mathrm{Gr}_d(\mathbb{H})$. Note that if there exists $H \in \mathfrak{sl}(n, \mathbb{H})$ such that $e^{tH}C$ shrinks to a point as $t \to \infty$, then C is contractible in the usual sense. To see this, denote by $id : C \to C$ the identity map of C. We need to construct an homotopy between id and a constant map. Let $H \in \mathfrak{sl}(n, \mathbb{H})$ be such that

$$\lim_{t \to \infty} e^{tH} C = \{c_0\},$$

where $c_0 \in \mathrm{Gr}_d(\mathbb{H})$. Define $F : C \times I \to C$ by setting $F(x, t) = e^{\left(\frac{t}{1-t}\right)H} x$ if $t \in [0, 1)$ and $F(x, 1) = c_0$ for all $x \in C$. Then F is a homotopy between id and the constant map c_0. In fact, $F(x, 0) = e^{0 \cdot H} x = x$ and $F(x, 1) = c_0$ for all $x \in C$. Also F is a continuous map, since

$$\lim_{t \to 1} e^{\left(\frac{t}{1-t}\right)H} x = c_0, \quad \forall x \in C.$$

This means that if C is not contractible in the usual sense, then C is not contractible in the sense of item 2 of Theorem 2.1.1.

So to prove our theorem, we will show that for all $d \in \{1, \ldots, n-1\}$ the S-invariant closed C_d sets with $\mathrm{int}(C_d) \neq \emptyset$ contains an orbit that is not contractible in the usual sense. This implies that the orbit is not contractible in the sense of item 2 of Theorem 2.1.1, and so C_d cannot be contractible in this sense.

In this way we will proceed in some steps.

The next lemma describes the non-contractible $\mathrm{Sl}(2, \mathbb{H})_{1,n}$-orbits in the grassmannians that are proved to be contained in the invariant control sets C_d. They are 4-dimensional spheres.

Lemma 2.2.2. *For $d = 1, \ldots, n-1$ let V_d be the subspace of $\mathbb{H}^n$ spanned by the first d basic vectors,*

$$V_d = \{(q_1, \ldots, q_d, 0, \ldots, 0) : q_r \in \mathbb{H}\}.$$

Then the $\mathrm{Sl}(2, \mathbb{H})_{1,n}$*-orbit in* $\mathrm{Gr}_d(\mathbb{H})$ *through* V_d *is diffeomorphic to* S^4.

Proof: We have that the orbit $\mathrm{Sl}(2, \mathbb{H})_{1,n} V_d$ is diffeomorphic to the homogeneous space $\mathrm{Sl}(2, \mathbb{H})_{1,n} / P$, where

$$P = \{g \in \mathrm{Sl}(2, \mathbb{H})_{1,n} \mid gV_d = V_d\}$$

is the isotropy subgroup of V_d under the action of $\mathrm{Sl}(2, \mathbb{H})_{1,n}$ in $\mathrm{Gr}_d(\mathbb{H})$. By a direct computation one sees that P is the subgroup of upper triangular matrices in $\mathrm{Sl}(2, \mathbb{H})_{1,n}$, which is a parabolic subgroup of $\mathrm{Sl}(2, \mathbb{H})_{1,n}$. This means that $\mathrm{Sl}(2, \mathbb{H})_{1,n} / P$ is a flag of $\mathrm{Sl}(2, \mathbb{H})_{1,n}$. As we know, $\mathrm{Sl}(2, \mathbb{H})$ is a real rank 1 group so that it has just one flag manifold which is diffeomorphic to a sphere. That is, the orbit $\mathrm{Sl}(2, \mathbb{H})_{1,n} V_d$ is diffeomorphic to a sphere. Finally,

$$\dim\left(\mathrm{Sl}(2, \mathbb{H})_{1,n} / P\right) = \dim\left(\mathrm{Sl}(2, \mathbb{H})_{1,n}\right) - \dim(P) = 15 - 11 = 4.$$

Therefore $\mathrm{Sl}(2, \mathbb{H})_{1,n} / P$ as well as the $\mathrm{Sl}(2, \mathbb{H})_{1,n}$-orbit through V_d is a sphere S^4. □

The next step is to check that for any $d = 1, \ldots, n-1$ the orbit $\mathrm{Sl}(2, \mathbb{H})_{1,n} V_d \approx S^4$ is not contractible in $\mathrm{Gr}_d(\mathbb{H})$ in the usual sense, that is, the identity map

$$id : \mathrm{Sl}(2, \mathbb{H})_{1,n} V_d \to \mathrm{Sl}(2, \mathbb{H})_{1,n} V_d$$

is not homotopic to a point.

In other words, we are required to prove that the 4-sphere $\mathrm{Sl}(2, \mathbb{H})_{1,n} V_d$ is not a representative of the identity of the homotopy group $\pi_4(\mathrm{Gr}_d(\mathbb{H}))$. To this purpose we recall the cellular decomposition of $\mathrm{Gr}_d(\mathbb{H})$ given in [18]. From that decomposition the homology $H_*(\mathrm{Gr}_d(\mathbb{H}))$ of a grassmannian $\mathrm{Gr}_d(\mathbb{H})$ is freely generated by the Schubert cells and $H_r(\mathrm{Gr}_d(\mathbb{H})) = \{0\}$ if r is not a multiple of 4. In $\mathrm{Gr}_d(\mathbb{H})$ there is just one 4-dimensional cell which is the orbit $\mathrm{Sl}(2, \mathbb{H})_{d,d+1} V_d$. Here,

$$\mathrm{Sl}(2, \mathbb{H})_{d,d+1} = \langle \exp \mathfrak{sl}(2, \mathbb{H})_{d,d+1} \rangle \approx \mathrm{Sl}(2, \mathbb{H})$$

and $\mathfrak{sl}(2, \mathbb{H})_{d,d+1}$ is the algebra of matrices with nonzero entries only in the entries (d, d), $(d, d+1)$, $(d+1, d)$ and $(d+1, d+1)$. Analogous to Lemma 2.2.2 we have that $\mathrm{Sl}(2, \mathbb{H})_{d,d+1} V_d$ is diffeomorphic to S^4.

Now, by the Hurewicz homomorphism we have $\pi_4(\mathrm{Gr}_d(\mathbb{H})) \approx H_4(\mathrm{Gr}_d(\mathbb{H}))$ because the homology is trivial in degrees less than 4. It follows that $\pi_4(\mathrm{Gr}_d(\mathbb{H})) \approx \mathbb{Z}$ and the equivalence class of the orbit $\mathrm{Sl}(2,\mathbb{H})_{d,d+1} V_d \approx S^4$ is a generator of $\pi_4(\mathrm{Gr}_d(\mathbb{H}))$. The next lemma shows that $\mathrm{Sl}(2,\mathbb{H})_{1,n} V_d \approx S^4$ is a generator as well.

Lemma 2.2.3. *The orbits* $\mathrm{Sl}(2,\mathbb{H})_{d,d+1} V_d \approx S^4$ *and* $\mathrm{Sl}(2,\mathbb{H})_{1,n} V_d \approx S^4$ *are homotopic to each other.*

Proof: The homotopy is performed by the product of two one-parameter subgroups. Let $A, B \in \mathfrak{sl}(n,\mathbb{H})$ be the matrices such that $Ae_1 = e_d$, $Ae_d = -e_1$, $Be_{d+1} = e_n$, $Be_n = -e_{d+1}$ and $Ae_r = Be_r = 0$ elsewhere. Put $P(t) = e^{tA}e^{tB}$. Then for all t, $P(t) V_d = V_d$ and $P(\pi/2)$ permutes the subspaces spanned by $\{e_d, e_{d+1}\}$ and $\{e_1, e_n\}$ so that $P(\pi/2)\,\mathrm{Sl}(2,\mathbb{H})_{d,d+1} P(\pi/2)^{-1} = \mathrm{Sl}(2,\mathbb{H})_{1,n}$. Hence,

$$\begin{aligned} P(\pi/2)\,\mathrm{Sl}(2,\mathbb{H})_{d,d+1} V_d &= P(\pi/2)\,\mathrm{Sl}(2,\mathbb{H})_{d,d+1} P(\pi/2)^{-1} P(\pi/2) V_d \\ &= \mathrm{Sl}(2,\mathbb{H})_{1,n} V_d, \end{aligned}$$

showing that the map $t \mapsto P(t)\,\mathrm{Sl}(2,\mathbb{H})_{d,d+1} V_d$ is a homotopy between the orbits $\mathrm{Sl}(2,\mathbb{H})_{d,d+1} V_d$ and $\mathrm{Sl}(2,\mathbb{H})_{1,n} V_d$. □

Now, we can start the proof of Theorem 2.2.1 in the case when $\mathrm{Sl}(2,\mathbb{H})_{1,n} \subset S$. Denote by N the nilpotent group of lower triangular matrices in $\mathrm{Sl}(n,\mathbb{H})$ having 1's at the diagonal. It is well known (and easy to prove) that NV_d is an open and dense set in $\mathrm{Gr}_d(\mathbb{H})$. Hence, as $\mathrm{int}C_d \neq \emptyset$ we have $NV_d \cap C_d \neq \emptyset$ where C_d is the invariant control set of S in $\mathrm{Gr}_d(\mathbb{H})$.

The assumption $\mathrm{Sl}(2,\mathbb{H})_{1,n} \subset S$ of Theorem 2.2.1 implies that $gC_d \subset C_d$ for any $g \in \mathrm{Sl}(2,\mathbb{H})_{1,n}$. Since C_d is closed it follows that any limit $\lim g_l x$, with $x \in C_d$ and $g_l \in \mathrm{Sl}(2,\mathbb{H})_{1,n}$, also belongs to C_d.

Now, take $x = gV_d \in NV_d \cap C_d$ with $g \in N$ and

$$h = \mathrm{diag}\{\lambda, 1, \ldots, 1, \lambda^{-1}\} \in \mathrm{Sl}(2,\mathbb{H})_{1,n}$$

with $\lambda > 1$. As $l \to +\infty$ the sequence of conjugations $h^l g h^{-l}$ converges to the matrix $g_1 \in N$ that has zeros at the first column and the last row outside the diagonal. We have $h^{-l}V_d = V_d$ so that $h^l x = h^l g V_d = h^l g h^{-l} V_d$ implying that $W = \lim h^l V_d = g_1 V_d \in C_d$.

Therefore the orbit $\mathrm{Sl}(2,\mathbb{H})_{1,n} W$ is entirely contained in C_d.

The next step is to prove that the orbit $\mathrm{Sl}(2,\mathbb{H})_{1,n} W \subset C_d$ is a sphere S^4 homotopic to $\mathrm{Sl}(2,\mathbb{H})_{1,n} V_d$. By the zeros in the first column and the last row of g_1, the subspace $W = g_1 V_d$ is the direct sum $\langle e_1 \rangle \oplus W_1$ where W_1 is a $(d-1)$-dimensional subspace of $\mathrm{span}_{\mathbb{H}}\{e_2, \ldots, e_{n-1}\}$. If $d = 1$ then $W = V_d$ and we are done. Otherwise, let $G = \mathrm{Sl}(n-2,\mathbb{H})_{2,\ldots,n-1} \approx \mathrm{Sl}(n-2,\mathbb{H})$ be the subgroup of matrices in $\mathrm{Sl}(n,\mathbb{H})$ where the restriction to $\mathrm{span}_{\mathbb{H}}\{e_1, e_n\}$ is the identity. Then G commutes with $\mathrm{Sl}(2,\mathbb{H})_{1,n}$ so that if $g \in G$ then $g\mathrm{Sl}(2,\mathbb{H})_{1,n} W = \mathrm{Sl}(2,\mathbb{H})_{1,n} gW$, that is, the image under $g \in G$ of the orbit $\mathrm{Sl}(2,\mathbb{H})_{1,n} W$ is again an orbit of $\mathrm{Sl}(2,\mathbb{H})_{1,n}$. Moreover, G acts transitively in the grassmannian of $(d-1)$-dimensional subspaces of $\mathrm{span}_{\mathbb{H}}\{e_2, \ldots, e_{n-1}\}$. Hence, there exists $g \in G$ such that $gW_1 = \mathrm{span}_{\mathbb{H}}\{e_2, \ldots, e_d\}$ so that $gW = V_d$. It follows that the orbit $\mathrm{Sl}(2,\mathbb{H})_{1,n} W$ is diffeomorphic to $\mathrm{Sl}(2,\mathbb{H})_{1,n} V_d$ and then is a sphere S^4. Furthermore, G is connected so that there is a continuous curve $g_t \in G$ with $g_0 = 1$ and $g_1 = g$. Therefore $t \mapsto g_t\mathrm{Sl}(2,\mathbb{H})_{1,n} W$ is a homotopy between $\mathrm{Sl}(2,\mathbb{H})_{1,n} W$ and $\mathrm{Sl}(2,\mathbb{H})_{1,n} V_d$, therefore $\mathrm{Sl}(2,\mathbb{H})_{1,n} W \approx S^4$ is not contractible.

We proved that the invariant control set C_d of the semigroup S in $\mathrm{Gr}_d(\mathbb{H})$ contains a non-contractible sphere S^4. Hence, C_d is not contractible. Since $d = 1, \ldots, n-1$ is arbitrary, S cannot be a proper semigroup, concluding the proof of Theorem 2.2.1 in the case when $\mathrm{Sl}(2,\mathbb{H})_{1,n} \subset S$.

Corollary 2.2.4. *Let $T \subset \mathrm{Sl}(n,\mathbb{H})$ be a semigroup with nonempty interior. Suppose that $\mathrm{Sl}(2,\mathbb{H})_{r,s} \subset T$ for some pair (r,s), $r \neq s$. Then $T = \mathrm{Sl}(n,\mathbb{H})$.*

Proof: Let P be a matrix in $\mathrm{Sl}(n,\mathbb{H})$ that permutes the subspaces $\langle e_1 \rangle$ and $\langle e_r \rangle$ and the subspaces $\langle e_n \rangle$ and $\langle e_s \rangle$. Then $P\mathrm{Sl}(2,\mathbb{H})_{r,s} P^{-1} = \mathrm{Sl}(2,\mathbb{H})_{1,n}$ so that the semigroup PTP^{-1} contains $\mathrm{Sl}(2,\mathbb{H})_{1,n}$. Since $\mathrm{int}PTP^{-1} \neq \emptyset$ we conclude, by Theorem 2.2.1, that $PTP^{-1} = \mathrm{Sl}(n,\mathbb{H})$, hence $T = \mathrm{Sl}(n,\mathbb{H})$. □

2.3 Application to Controllability

In this section we apply Theorem 2.2.1 to show the following result that gives sufficient conditions for controllability of the control system

$$\dot{g} = Ag + u(t)Bg, \quad A, B \in \mathfrak{sl}(n,\mathbb{H}). \tag{2.1}$$

Theorem 2.3.1. *The system (2.1) is controllable if the following conditions are satisfied.*

H1. The pair (A, B) *generates* $\mathfrak{sl}(n, \mathbb{H})$ *as a Lie algebra (Lie algebra rank condition).*

H2. $B = \operatorname{diag}\{a_1 + ib_1, \ldots, a_n + ib_n\}$ *with* $a_1 > a_2 \geq \cdots \geq a_{n-1} > a_n$*,* $b_n \neq 0 \neq b_1$ *and* b_1/b_n *is irrational.*

H3. Denote the $1, n$ *and* $n, 1$ *entries of the matrix* A *by* $p \in \mathbb{H}$ *and* $q \in \mathbb{H}$*, respectively. Let* $\mathbb{H}_{1,i}$ *and* $\mathbb{H}_{j,k}$ *be the (real) subspaces of* $\mathbb{H}$ *spanned by* $\{1, i\}$ *and* $\{j, k\}$ *respectively. Then* p *and* q *do not belong to* $\mathbb{H}_{1,i} \cup \mathbb{H}_{j,k}$*.*

The proof of Theorem 2.3.1 will be made throughout this section and it follows from Theorem 2.2.1 combined with the following proposition. Although the Lie algebra rank condition will not be needed for this proposition, it allows us to conclude the proof of the Theorem 2.3.1 by ensuring that the system semigroup S has nonempty interior, leading us to the conditions required for Theorem 2.2.1.

Proposition 2.3.2. *Under the conditions H2 and H3 of Theorem 2.3.1, the semigroup* S *of the system contains the group* $\mathrm{Sl}(2, \mathbb{H})_{1,n}$*.*

To prove this proposition, let S be the system semigroup of the invariant system (2.1) and write

$$\mathfrak{c}(S) = \{X \in \mathfrak{sl}(n, \mathbb{H}) : \forall t \geq 0,\ e^{tX} \in \mathrm{cl}S\}$$

for the Lie wedge of S (see [12], [13] and Hilgert, Hofmann and Lawson [11]). The main properties of $\mathfrak{c}(S)$ are:

1) $\mathfrak{c}(S)$ is a closed convex cone in the Lie algebra $\mathfrak{sl}(n, \mathbb{H})$;

2) $\mathfrak{c}(S) \cap (-\mathfrak{c}(S))$ is a Lie subalgebra and

3) If $X \in \mathfrak{c}(S) \cap (-\mathfrak{c}(S))$ then $e^{\mathrm{ad}(X)}\mathfrak{c}(S) = \mathfrak{c}(S)$.

By definition of S we have that $A + uB \in \mathfrak{c}(S)$ for all $u \in \mathbb{R}$ (since we consider unrestricted controls). Hence, $A \in \mathfrak{c}(S)$ and if $u \neq 0$ then

$$\frac{1}{|u|}A + \frac{u}{|u|}B = \frac{1}{|u|}(A + uB) \in \mathfrak{c}(S).$$

Taking limits as $u \to \pm\infty$ we see that $\pm B \in \mathfrak{c}(S)$, that is, $B \in \mathfrak{c}(S) \cap (-\mathfrak{c}(S))$. It follows that $e^{t\mathrm{ad}(B)}A \in \mathfrak{c}(S)$ and hence $e^{-t(a_1 - a_n)}e^{t\mathrm{ad}(B)}A \in \mathfrak{c}(S)$ for all $t \in \mathbb{R}$ where $a_1, \ldots, a_n$ are the real parts of the entries of B.

Now, by assumption we have $a_1 > a_2 > \cdots > a_n$ so that as $t \to +\infty$ the entries $e^{-t(a_1-a_n)}e^{t\mathrm{ad}(B)}A$ converge to 0 except the $(1,n)$-entry. The $(1,n)$-entry of the matrix $e^{-t(a_1-a_n)}e^{t\mathrm{ad}(B)}A$ is $e^{it(b_1-b_n)}p$ where p is as in the statement of the above theorem and $b_1,\ldots,b_n$ are the imaginary parts of the entries of B. Choosing a sequence $t_k \to +\infty$ such that $e^{it(b_1-b_n)} \to 1$ we conclude that

$$X = \begin{bmatrix} 0 & \cdots & p \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in \mathfrak{c}\,(S)\,.$$

Using again the properties of the Lie wedge $\mathfrak{c}\,(S)$ we have that for all $t \in \mathbb{R}$,

$$e^{-t(a_1-a_n)}e^{t\mathrm{ad}(B)}X = \begin{bmatrix} 0 & \cdots & e^{itb_1}pe^{-itb_n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in \mathfrak{c}\,(S)\,. \tag{2.2}$$

The following lemma about conjugation of quaternions shows that $\mathfrak{g}_{\alpha_{1n}} = \mathrm{span}_{\mathbb{H}}\{X\}$ is contained in $\mathfrak{c}\,(S)$.

Lemma 2.3.3. *Consider the action of the circle group $S^1 = \{e^{\mathbf{i}t} \,|\, t \in [0,2\pi]\}$ in the complex plane $\mathbb{C}$ given by rotation $(t,x) = e^{\mathbf{i}t}x$. If $x \neq 0$, then the orbit S^1x is a circle and the closed convex cone generated by S^1x is $\mathbb{C}$.*

Proof: It is immediate that S^1x is a circle of radius $a = |x|$. Denote by $co(S^1x)$ the convex cone generated by S^1x and let $z = re^{\mathbf{i}s_1} \in \mathbb{C}$ be any complex number. Writing $x = ae^{\mathbf{i}s_2}$ and choosing $t = s_1 - s_2$, we have $({}^r\!/\!_a)\, e^{\mathbf{i}t}x \in co(S^1x)$, but

$$\frac{r}{a}e^{\mathbf{i}t}x = \left(\frac{r}{a}e^{\mathbf{i}(s_1-s_2)}\right)\left(ae^{\mathbf{i}s_2}\right) = re^{\mathbf{i}s_1} = z,$$

proving that $co(S^1x) = \mathbb{C}$. □

Lemma 2.3.4. *Consider the torus $\mathbb{T}^2$ acting on the quaternions $\mathbb{H}$ by*

$$\phi\left((t,s),q\right) = e^{\mathbf{i}t}qe^{-\mathbf{i}s}.$$

Write $q = a + b$, with $a = x_1 + x_2\mathbf{i} \in \mathbb{H}_{\{\mathbf{1},\mathbf{i}\}}$ and $b = x_3\mathbf{j} + x_4\mathbf{k} \in \mathbb{H}_{\{\mathbf{j},\mathbf{k}\}}$. Suppose that

$q \notin \mathbb{H}_{\{\mathbf{1},\mathbf{i}\}} \cup \mathbb{H}_{\{\mathbf{j},\mathbf{k}\}}$, that is, $a \neq 0 \neq b$. Then, the orbit $\mathbb{T}^2 q$ is a 2-dimensional torus and $\mathbb{H}$ is the convex cone generated by $\mathbb{T}^2 q$.

Proof: Note initially that ϕ is in fact an action because $\phi((0,0),q) = q$ and

$$\phi((t,s)\cdot(u,v),q) = e^{\mathbf{i}(t+u)}qe^{-\mathbf{i}(s+v)} = e^{\mathbf{i}t}(e^{\mathbf{i}u}qe^{-\mathbf{i}v})e^{-\mathbf{i}s} = \phi\left((t,s),\phi((u,v),q)\right).$$

Now, the restriction of this action to the orbit $\mathbb{T}^2 q$ is a transitive action. Let H_q be the isotropy subgroup of q under ϕ. As $\mathbb{T}^2$ is abelian, then H_q is a normal closed subgroup of $\mathbb{T}^2$. This means that $\mathbb{T}^2/H_q$ is a Lie group diffeomorphic to $\mathbb{T}^2 q$ such that $\dim \mathbb{T}^2/H_q = \dim \mathbb{T}^2 - \dim H_q$. Further, as $\mathbb{T}^2$ is abelian, compact and connected we have that $\mathbb{T}^2/H_q$ is also abelian, compact and connected. This means that the orbit $\mathbb{T}^2 q$ is a torus. In this way, to conclude the first part of the proof all we need to do is show that this orbit is 2-dimensional.

The tangent space of $\mathbb{T}^2 q$ at q is spanned by

$$\frac{\partial}{\partial t}\left(e^{\mathbf{i}t}qe^{-\mathbf{i}s}\right)\Big|_{(0,0)} = \mathbf{i}q \quad \text{and} \quad \frac{\partial}{\partial s}\left(e^{\mathbf{i}t}qe^{-\mathbf{i}s}\right)\Big|_{(0,0)} = -q\mathbf{i},$$

and we have

$$\begin{aligned} \mathbf{i}q &= \mathbf{i}x_1 - x_2 + \mathbf{k}x_3 - \mathbf{j}x_4 = v + w \\ -q\mathbf{i} &= -\mathbf{i}x_1 + x_2 + \mathbf{k}x_3 - \mathbf{j}x_4 = -v + w \end{aligned}$$

with $v = \mathbf{i}x_1 - x_2$ and $w = \mathbf{k}x_3 - \mathbf{j}x_4$. The assumption about q says that $v \neq 0 \neq w$ so that $\{v,w\}$ is linearly independent. Hence $\{\mathbf{i}q, q\mathbf{i}\}$ is linearly independent as well because the linear combination $a(v+w) + b(v-w) = 0$ implies $a+b=0$ and $a-b=0$, that is, $a=b=0$. This shows that the orbit $\mathbb{T}^2 q$ is a 2-dimensional torus.

To the convex cone C generated by $\mathbb{T}^2 q$ take $r = e^{\mathbf{i}t}qe^{-\mathbf{i}s} \in \mathbb{T}^2 q$. Then $-r = e^{\mathbf{i}\pi}r = e^{\mathbf{i}(t+\pi)}qe^{-\mathbf{i}s}$ also belongs to $\mathbb{T}^2 q$. Hence C is a subspace. Also, it is easy to see that C is invariant under left multiplication by $\mathbf{i}$, in fact,

$$\mathbf{i}\lim_{n\to\infty} e^{\mathbf{i}t_n}qe^{-\mathbf{i}s_n} = \lim_{n\to\infty} e^{\mathbf{i}\left(t_n + \frac{\pi}{2}\right)}qe^{-\mathbf{i}s_n},$$

and the same holds for convex combinations of elements in $\mathbb{T}^2 q$. The orbit contains

$e^{\mathbf{i}\pi/2}q = \mathbf{i}q = \mathbf{i}a + \mathbf{i}b$ and $qe^{-\mathbf{i}3\pi/2} = q\mathbf{i} = a\mathbf{i} + b\mathbf{i} = \mathbf{i}a - \mathbf{i}b$. So that C contains $\mathbf{i}a$ and $\mathbf{i}b$ and hence contains a and b. Finally,

$$\frac{x_2}{x_1}a + \mathbf{i}a = x_2 + \frac{x_2^2}{x_1}\mathbf{i} - x_2 + x_1\mathbf{i} = \left(\frac{x_2^2}{x_1} + x_1\right)\mathbf{i} \in C$$

and

$$\mathbf{i}a - \frac{x_1}{x_2}a = -x_2 + x_1\mathbf{i} - \frac{x_1^2}{x_2} - x_1\mathbf{i} = -x_2 - \frac{x_1^2}{x_2} \in C,$$

which implies that $\mathbb{H}_{\{1,\mathbf{i}\}} \subset C$. Analogously, we have $\mathbb{H}_{\{\mathbf{j},\mathbf{k}\}} \subset C$. Thus $\mathbb{H}$ is the cone generated by $\mathbb{T}^2 q$. □

Lemma 2.3.5. *Let C_X and C_Y be the closed convex cones generated by the sets X and Y in $\mathbb{H}$, respectively. If $Y \subset X$ is dense in X, then $C_Y = C_X$.*

Proof: It is clear that $C_Y \subset C_X$. On the other hand, let $x \in C_X$ and let $(x_n)_{n\in\mathbb{N}} \subset co(X)$ be a sequence such that $x_n \to x$ as $n \to \infty$, where $co(X)$ stands for the convex cone generated by X. If $k = \dim(co(X))$, let $\mathcal{B} = \{x^1, \ldots, x^k\}$ be a basis for $co(X)$. Then for each $n \in \mathbb{N}$, x_n can be written as $x_n = \alpha_n^1 x^1 + \cdots + \alpha_n^k x^k$, where $\alpha_n^i \geq 0$, $1 \leq i \leq k$. As Y is dense in X, for each $n \in \mathbb{N}$ and each $1 \leq i \leq k$, choose $y_n^i \in Y$ such that $|x^i - y_n^i| < \frac{1}{n\beta}$, where $\beta = \alpha_n^1 + \cdots + \alpha_n^k$. For each $n \in \mathbb{N}$, set $y_n = \alpha_n^1 y_n^1 + \cdots + \alpha_n^k y_n^k$. Then

$$\begin{aligned} |x_n - y_n| &= |(\alpha_n^1 x^1 + \cdots + \alpha_n^k x^k) - (\alpha_n^1 y_n^1 + \cdots + \alpha_n^k y_n^k)| \\ &= |\alpha_n^1(x^1 - y_n^1) + \cdots + \alpha_n^k(x^k - y_n^k)| \\ &\leq \alpha_n^1 \frac{1}{n\beta} + \cdots + \alpha_n^k \frac{1}{n\beta} = \left(\alpha_n^1 + \cdots + \alpha_n^k\right)\frac{1}{n\beta} = \frac{\beta}{n\beta} = \frac{1}{n}, \end{aligned}$$

and this means that $y_n \to x$ as $n \to \infty$. Since $y_n \in co(Y)$, we conclude that $x \in C_Y$, that is, $C_X \subset C_Y$. □

Corollary 2.3.6. *Let $c_1, c_2 \in \mathbb{R}$ with $c_1c_2 \neq 0$ and c_1/c_2 irrational. Take $q \in \mathbb{H}$ with $q \notin \mathbb{H}_{\{1,\mathbf{i}\}} \cup \mathbb{H}_{\{\mathbf{j},\mathbf{k}\}}$. Then $\mathbb{H}$ is the closed convex cone generated by the curve $e^{\mathbf{i}tc_1}qe^{-\mathbf{i}tc_2}$.*

Proof: Since c_1/c_2 is irrational, the curve $\left(e^{\mathbf{i}tc_1}, e^{\mathbf{i}tc_2}\right)$ is dense in the torus $\mathbb{T}^2$. Hence $t \mapsto e^{\mathbf{i}tc_1}qe^{-\mathbf{i}tc_2}$ is dense in the orbit $\mathbb{T}^2 q$ which implies the corollary. □

Applying this corollary to the curve (2.2) it follows, by the assumption on B in

Theorem 2.3.1, that the subspace $\mathfrak{g}_{\alpha_{1n}} = \mathrm{span}_{\mathbb{H}}\{X\}$ is contained in $\mathfrak{c}(S)$ and hence in $\mathfrak{c}(S) \cap (-\mathfrak{c}(S))$.

By similar arguments we get lower triangular matrices in $\mathfrak{c}(S)$: taking limits as $t \to -\infty$ of $e^{-t(a_1-a_n)}e^{t\mathrm{ad}(B)}A$ it follows that

$$Y = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ q & \cdots & 0 \end{bmatrix} \in \mathfrak{c}(S),$$

hence applying the same idea we conclude that $\mathfrak{g}_{\alpha_{n1}} = \mathrm{span}_{\mathbb{H}}\{Y\}$ is contained in $\mathfrak{c}(S)$ and therefore in $\mathfrak{c}(S) \cap (-\mathfrak{c}(S))$.

Now, the Lie algebra generated by $\mathfrak{g}_{\alpha_{1n}}$ and $\mathfrak{g}_{\alpha_{n1}}$ is $\mathfrak{sl}(2,\mathbb{H})_{1,n}$ then this Lie algebra is contained in $\mathfrak{c}(S)$. It follows that $\mathrm{Sl}(2,\mathbb{H})_{1,n}$ is contained in S, concluding the proof of Proposition 2.3.2.

2.4 Cartan subalgebras and genericity

In this section we prove that controllability for invariant control systems in $\mathrm{Sl}(n,\mathbb{H})$ is a generic property.

We begin discussing the Cartan subalgebra of $\mathfrak{sl}(n,\mathbb{H})$ (see e.g. Warner [36]). Considering the same notations of Section 2, note that the algebra of diagonal matrices

$$\mathfrak{h} = \{\mathrm{diag}\{a_1 + ib_1, \ldots, a_n + ib_n\} : a_r, b_r \in \mathbb{R},\ a_1 + \cdots + a_n = 0\} \tag{2.3}$$

is a Cartan subalgebra of $\mathfrak{sl}(n,\mathbb{H})$ since it is maximal abelian and $\mathrm{ad}(H)$ is semi-simple for any $H \in \mathfrak{h}$.

Next we prove that up to conjugation, $\mathfrak{h}$ is the only Cartan subalgebra of $\mathfrak{sl}(n,\mathbb{H})$. First take the Cartan decomposition $\mathfrak{sl}(n,\mathbb{H}) = \mathfrak{sp}(n) \oplus \mathfrak{s}$ given in Section 2. The subspace $\mathfrak{a} \subset \mathfrak{s}$ is a maximal abelian subalgebra contained in $\mathfrak{s}$.

Observe that the Cartan subalgebra $\mathfrak{h}$ decomposes as $\mathfrak{h} = (\mathfrak{h} \cap \mathfrak{sp}(n)) \oplus \mathfrak{a}$. More generally $\mathfrak{j}$ is said to be a standard Cartan subalgebra if it decomposes as $\mathfrak{j} = \mathfrak{j}_{\mathfrak{k}} \oplus \mathfrak{j}_{\mathfrak{a}}$ with $\mathfrak{j}_{\mathfrak{k}} = \mathfrak{j} \cap \mathfrak{k}$ and $\mathfrak{j}_{\mathfrak{a}} = \mathfrak{j} \cap \mathfrak{a}$. The following statement is a basic fact for the classification of Cartan subalgebras in real semi-simple Lie algebras (Theorem of Kostant-Sugiura).

Proposition 2.4.1. *Any Cartan subalgebra of* $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{H})$ *is conjugate (by an inner automorphism) to a standard Cartan subalgebra* $\mathfrak{j}$.

Proof: See [36], Section 1.3.1. □

In the next proposition, we prove that in $\mathfrak{sl}(n, \mathbb{H})$ there is a unique conjugacy class of Cartan subalgebras. We give a direct proof without relying in the general classification theorem (Theorem of Kostant-Sugiura).

Proposition 2.4.2. *Every Cartan subalgebra of* $\mathfrak{sl}(n, \mathbb{H})$ *is conjugate (by an inner automorphism) to the subalgebra* $\mathfrak{h}$ *defined in (2.3).*

Proof: Let $\mathfrak{j} = \mathfrak{j}_{\mathfrak{k}} \oplus \mathfrak{j}_{\mathfrak{a}}$ be a standard Cartan subalgebra. The following simple arguments show that $\mathfrak{j}_{\mathfrak{k}}$ is a Cartan subalgebra of $\mathfrak{sp}(n)$ and $\mathfrak{j}_{\mathfrak{a}} = \mathfrak{a}$. We have $\dim \mathfrak{j}_{\mathfrak{a}} \leq \dim \mathfrak{a} = n-1$. Also, $\dim \mathfrak{j}_{\mathfrak{k}} \leq \mathrm{rank}\mathfrak{sp}(n) = n$ because $\mathfrak{j}_{\mathfrak{k}}$ is an abelian subalgebra of $\mathfrak{sp}(n)$ and hence is contained in a Cartan subalgebra of $\mathfrak{sp}(n)$, whose dimension is $\mathrm{rank}\mathfrak{sp}(n)$. On the other hand, $\dim \mathfrak{j} = \mathrm{rank}\mathfrak{sl}(n, \mathbb{H}) = \dim \mathfrak{h} = 2n-1$. Hence, we must have $\dim \mathfrak{j}_{\mathfrak{k}} = n$ and $\dim \mathfrak{j}_{\mathfrak{a}} = n-1$. By the first equality, $\mathfrak{j}_{\mathfrak{k}}$ is a Cartan subalgebra of $\mathfrak{sp}(n)$ while the second equality shows that $\mathfrak{j}_{\mathfrak{a}} = \mathfrak{a}$.

Now, $\mathfrak{j}_{\mathfrak{k}}$ commutes with $\mathfrak{a}$ and then is contained in the algebra $\mathfrak{m} \approx \mathfrak{sp}(1)^n$ of diagonal matrices with entries in the imaginary quaternions $\mathrm{Im}\mathbb{H}$. Since $\dim \mathfrak{j}_{\mathfrak{k}} = n = \mathrm{rank}\mathfrak{sp}(1)^n$, there is an inner automorphism $g = e^{\mathrm{ad}(X)}$, $X \in \mathfrak{m}$, such that $g(\mathfrak{j}_{\mathfrak{k}}) = \mathfrak{h}_k$ and g fixes $\mathfrak{a}$. Therefore, $g(\mathfrak{j}) = \mathfrak{h}$ showing that any standard Cartan subalgebra is conjugate to $\mathfrak{h}$. By the above proposition, $\mathfrak{h}$ is a representative of the unique conjugacy class of Cartan subalgebras of $\mathfrak{sl}(n, \mathbb{H})$. □

Denote by $\mathfrak{a}^+$ the Weyl chamber of real diagonal matrices

$$\mathrm{diag}\{a_1, \ldots, a_n\} \qquad \text{with } a_1 > \cdots > a_n.$$

A matrix B satisfying the second condition of Theorem 2.3.1 belongs to $\mathfrak{a}^+ + \mathfrak{h}_{\mathfrak{k}}$ where $\mathfrak{h}_{\mathfrak{k}}$ is as above the space of diagonal matrices with entries in $i\mathbb{R}$. Denote by $D_0 \subset \mathfrak{a}^+ + \mathfrak{h}_{\mathfrak{k}}$ the set of the matrices B satisfying that condition. By definition, if $H \in \mathfrak{a}^+$ and $X = \mathrm{diag}\{ib_1, \ldots, ib_n\} \in \mathfrak{h}_{\mathfrak{k}}$ then $H + X \in D_0$ if and only if $b_1 b_n \neq 0$ and b_1/b_n is irrational. Hence, D_0 is a dense subset of $\mathfrak{a}^+ + \mathfrak{h}_{\mathfrak{k}}$.

Let $\mathcal{W}$ be the permutation group in n letters (Weyl group) acting on the diagonal matrices by permutation of indices. The set $\mathcal{W}\mathfrak{a}^+ = \{w\mathfrak{a}^+ : w \in \mathcal{W}\}$ is open and dense in $\mathfrak{a}$. Since D_0 is dense in $\mathfrak{a}^+ + \mathfrak{h}_\mathfrak{k}$, it follows that

$$\mathcal{W}D_0 = \{wD_0 : w \in \mathcal{W}\} \subset \mathfrak{a} + \mathfrak{h}_\mathfrak{k} = \mathfrak{h}$$

is dense in $\mathfrak{h}$.

We apply now Proposition 2.4.2 ensuring that every Cartan subalgebra is conjugate to $\mathfrak{h}$. This implies that the set $\{\mathrm{Ad}\,(g)\,\mathfrak{h} : g \in \mathrm{Sl}\,(n,\mathbb{H})\}$ is dense in $\mathfrak{sl}\,(n,\mathbb{H})$ because the set of regular elements is dense and each regular element is contained in a Cartan subalgebra. With these facts we get the following density result.

Proposition 2.4.3. *Let D be the set of conjugates of the matrices B satisfying the second condition of Theorem 2.3.1. Then D is dense in $\mathfrak{sl}\,(n,\mathbb{H})$.*

Proof: Take an open set $U \subset \mathfrak{sl}\,(n,\mathbb{H})$. Then there exists a regular element X of $\mathfrak{sl}\,(n,\mathbb{H})$ with $X \in U$. Let $\mathfrak{h}_X$ be the unique Cartan subalgebra containing X. By Proposition 2.4.2 there exists $g \in \mathrm{Sl}\,(n,\mathbb{H})$ such that $\mathrm{Ad}\,(g)\,\mathfrak{h}_X = \mathfrak{h}$. Then $\mathrm{Ad}\,(g)\,U \cap \mathfrak{h}$ is a nonempty open set of $\mathfrak{h}$ and hence $\mathrm{Ad}\,(g)\,U \cap \mathcal{W}D_0 \neq \emptyset$. This means that U meets $\mathrm{Ad}\,(g^{-1})\,\mathcal{W}D_0 \subset D$. Therefore as U is arbitrary it follows that D is dense. □

Now, we can show the main result of this section.

Theorem 2.4.4. *There is an open and dense set $C \subset \mathfrak{sl}\,(n,\mathbb{H})^2$ such that the control system $\dot{g} = A\,(g) + uB\,(g)$ with unrestricted controls ($u \in \mathbb{R}$) is controllable for all pairs $(A,B) \in C$.*

To prove this theorem, first note that the union $\mathbb{H}_{1,i} \cup \mathbb{H}_{j,k}$ is a nowhere dense subset of $\mathbb{H}$, which implies that its complement is an open and dense subset of $\mathbb{H}$. Consequently, the set of matrices A satisfying the third condition of Theorem 2.3.1 is open and dense in $\mathfrak{sl}\,(n,\mathbb{H})$.

Remark 2.4.5. *For $\Omega \subset M \times N$ open, the set $\pi_1\left(\Omega \cap \pi_2^{-1}(b)\right) \subset M$ is open in M for any $b \in N$. Here, M and N are arbitrary metric spaces and $\pi_1 : M \times N \to N$ and $\pi_2 : M \times N \to N$ are the canonical projections in the first and second coordinates, respectively. To see this take the continuous map $i_b : M \to M \times N$, $i_b(x) = (x,b)$, and observe that*

$$\pi_1\left(\Omega \cap \pi_2^{-1}(b)\right) = \{\pi_1(x,b) \mid (x,b) \in \Omega\} = \{x \in M \mid i_b(x) \in \Omega\} = (i_b)^{-1}(\Omega).$$

Now we can prove that the set $C \subset \mathfrak{sl}(n, \mathbb{H})^2$ of the conjugates of pairs satisfying the three conditions of Theorem 2.3.1 is dense in $\mathfrak{sl}(n, \mathbb{H})^2$.

Let O be an open subset of $\mathfrak{sl}(n, \mathbb{H})^2$. Since the set of pairs (A, B) satisfying H1 is open and dense in $\mathfrak{sl}(n, \mathbb{H})^2$, there is $(A, B) \in O$ satisfying H1. Moreover, there exists $O' \ni (A, B)$ for which every pair belonging to O' satisfies H1. Without loss of generality we can assume $O' \subset O$. Now, $\pi_2(O')$ is open in $\mathfrak{sl}(n, \mathbb{H})$ and by Proposition 2.4.3 we can choose $\tilde{B} \in \pi_2(O') \cap D$. As the set $\pi_1 \left(O' \cap \pi_2^{-1}(\tilde{B}) \right)$ is open in $\mathfrak{sl}(n, \mathbb{H})^2$, by the above considerations we can take $\tilde{A} \in \pi_1 \left(\pi_2^{-1}(\tilde{B}) \cap O' \right)$ satisfying H3. Thus the pair $(\tilde{A}, \tilde{B})$ has the following properties:

i) $(\tilde{A}, \tilde{B}) \in O' \subset O$.

ii) $(\tilde{A}, \tilde{B})$ is conjugate to a pair satisfying H1, H2 and H3.

That is, $(\tilde{A}, \tilde{B}) \in O \cap C$ proving that C is dense in $\mathfrak{sl}(n, \mathbb{H})^2$. Finally, as invariant systems remain controllable, under small perturbations, we can slightly enlarge the dense set C to get the open and dense set, as claimed in Theorem 2.4.4.

CHAPTER 3

THE LIE SATURATE TECHNIQUE FOR CONTROLLABILITY

Now we shall study controllability from a slightly different perspective. The results in this chapter are inspired by Gauthier and Bornard [8], which in turn improves the controllability results present in the classical papers by Jurdjevic and Kupka, [12] and [13]. We first recall the Lie saturate and its main properties, and subsequently we state our controllability result on $\mathrm{Sl}(n, \mathbb{C})$, adapting the results in [8], that were made considering the case $\mathrm{Sl}(n, \mathbb{R})$. The same method for controllability via the Lie saturate is applied to $\mathrm{Sl}(n, \mathbb{H})$ and some semidirect products.

3.1 The Lie saturate and irreducible matrices

If G is a connected Lie group with Lie algebra $\mathfrak{g}$, we will denote by $\mathrm{Lie}(\Gamma)$ the Lie algebra generated by $\Gamma \subset \mathfrak{g}$, that is, the smallest Lie subalgebra of $\mathfrak{g}$ containing Γ.

Given a right invariant control system Γ, let $\mathcal{A}_\Gamma$ denote its attainable set from the identity.

Definition 3.1.1. *The **Lie saturate** of* Γ*, written as* $\mathrm{LS}(\Gamma)$*, is the set*

$$\mathrm{LS}(\Gamma) = \{A \in \mathrm{Lie}(\Gamma) \mid \exp(tA) \in \mathrm{cl}(\mathcal{A}_\Gamma) \ \ \forall t \geq 0\}.$$

Some useful properties of the Lie saturate are listed in the next proposition.

Proposition 3.1.2 (cf. [12])**.** *Let G be a connected Lie group and $\Gamma \subset \mathfrak{g}$ a right invariant control system. Then*

1. $\mathrm{LS}(\Gamma)$ *is topologically closed;*

2. $\mathrm{LS}(\Gamma)$ *is convex:*

$$X, Y \in \mathrm{LS}(\Gamma) \;\; \Rightarrow \;\; \alpha X + (1-\alpha)Y \in \mathrm{LS}(\Gamma) \;\; \forall \alpha \in [0,1].$$

3. $\mathrm{LS}(\Gamma)$ *is a positive cone:*

$$X \in \mathrm{LS}(\Gamma) \;\; \Rightarrow \;\; \alpha X \in \mathrm{LS}(\Gamma) \;\; \forall \alpha \geq 0.$$

Thus,

$$X, Y \in \mathrm{LS}(\Gamma) \;\; \Rightarrow \;\; \alpha X + \beta Y \in \mathrm{LS}(\Gamma) \;\; \forall \alpha, \beta \geq 0.$$

4. *For any* $\pm X, Y \in \mathrm{LS}(\Gamma)$ *and any* $t \in \mathbb{R}$*,*

$$\exp(t\,\mathrm{ad}X)Y = \exp(tX)\,Y\exp(-tX) \in \mathrm{LS}(\Gamma).$$

5. $\pm X, \pm Y \in \mathrm{LS}(\Gamma)$ *implies that* $\pm[X,Y] \in \mathrm{LS}(\Gamma)$*.*

The following theorem will be the main tool in way to achieve controllability under certain circumstances.

Theorem 3.1.3. *A right invariant control system* $\Gamma \subset \mathfrak{g}$ *is controllable on a connected Lie group* G *if and only if* $\mathrm{LS}(\Gamma) = \mathfrak{g}$*.*

Proof: See [13], Proposition 2. □

A characterization of the concept of irreducibility by means of the graph theory will be quite useful in the present context, so let us explain it briefly. The matrices to be considered can have real, complex or even quaternionic entries, since the main idea relies on the entries being 0 or not.

Definition 3.1.4. *An* $n \times n$ *matrix* A *(*$n \geq 2$*), whose entries are real, complex or quaternionic, is called* ***reducible*** *if there exists an* $n \times n$ *permutation matrix* P *such that*

$$PAP^t = \begin{bmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{bmatrix},$$

where $A_{1,1}$ *is an* $r \times r$ *submatrix, with* $1 \leq r < n$*. Otherwise,* A *is said* ***irreducible****. If* $n = 1$*, then* A *is reducible if its single entry is* 0*, and irreducible otherwise.*

Definition 3.1.5. *A **finite directed graph** associated to the $n \times n$ matrix $A = (a_{ij})$ is a set of n points in the plane, say $P_1, \cdots, P_n$ (called **nodes**), together with directed arcs connecting these points in such a way that the point P_i is joined to the point P_j by means of a line directed from P_i to P_j if and only if $a_{ij} \neq 0$.*

A directed path from the node P_i to the node P_j in the graph is simply a collection of concatenated directed arcs starting at P_i and ending at P_j, that is,

$$\overrightarrow{P_i P_{l_1}}, \quad \overrightarrow{P_{l_1} P_{l_2}}, \cdots, \overrightarrow{P_{l_{r-1}} P_{l_r}}, \quad \text{with } l_r = j.$$

Definition 3.1.6. *A directed graph is said to be **strongly connected** if for any ordered pair of nodes P_i and P_j there exists a directed path connecting P_i to P_j.*

The following simple result characterizes completely the irreducibility of a given square matrix A via the strongly connectedness of its directed graph. For completeness we prove it here, although a more detailed approach and further applications can be found in [35].

Theorem 3.1.7 (cf. [35], Theorem 1.6)**.** *A square matrix A of order n is irreducible if and only if its directed graph $g(A)$ is strongly connected.*

Proof: Suppose that A is reducible, then there exist $\tilde{A}$ and a permutation matrix P such that

$$A = P\tilde{A}P^t = P \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} P^t,$$

where A_{11} is of order r and A_{22} is of order $(n-r)$, for some $1 \leq r < n$. Set $v_1, \ldots, v_n$ for the nodes of $g(A)$ and $\tilde{v}_1, \ldots, \tilde{v}_n$ for the nodes of $g(\tilde{A})$. In $g(\tilde{A})$ observe that there does not exist a directed path from $\tilde{v}_i$ to $\tilde{v}_j$ if $r < i \leq n$ and $1 \leq j \leq r$. Hence $g(\tilde{A})$ is not strongly connected. As the directed graph of $\tilde{A}$ is obtained from that of A just by renumbering the nodes and this operation does not affect the connectedness of the directed graph, it follows that $g(A)$ is not strongly connected.

Conversely, suppose $g(A)$ is not strongly connected. Then, there exists nonempty sets of vertices S_1 and S_2 of $g(A)$ such that no directed path from v_i to v_j exists if $v_i \in S_2$ and $v_j \in S_1$. Let $|S_1| = r$ and $|S_2| = n-r$. Relabel the vertices of $g(A)$ as $\tilde{v}_1, \ldots, \tilde{v}_n$ where $\tilde{v}_1, \ldots, \tilde{v}_r \in S_1$ and $\tilde{v}_{r+1}, \ldots, \tilde{v}_n \in S_2$. Permute the matrix A in the same way to create

$\tilde{A}$. Thus, the directed graph created from relabeling the vertices of $g(A)$ is precisely those of $g(\tilde{A})$. Since there are no directed paths in $g(\tilde{A})$ from the vertices in S_2 to the vertices in S_1, it means that $\tilde{A}$ must have a block of zeros with size $(n-r) \times r$ in the low left corner. In other words, $\tilde{A}$ is reducible. Being A and $\tilde{A}$ similar by permutations, it follows that A is reducible as well. □

3.2 Controllability on $\mathrm{Sl}(n, \mathbb{C})$

In [30] the authors improve the main theorem of [13], giving sufficient conditions for controllability on connected, simple and complex Lie groups. In this section we deal with control systems evolving on $\mathrm{Sl}(n, \mathbb{C})$ and under certain conditions we set necessary and sufficient conditions for the controllability of bilinear control systems on this group. The source of inspiration and the general lines for the proof follows the main theorem in [8].

For the canonical basis of $\mathfrak{sl}(n, \mathbb{C})$ as a $\mathbb{C}$-vector space we take the set $\mathcal{B}$ formed by the matrices E_{ij}, $i \neq j$, $1 \leq i, j \leq n$, and $E_{ii} - E_{11}$, $1 < i \leq n$. Here, E_{ij} stands for the matrix which has 1 in the ij-entry and 0 elsewhere. Given $A \in \mathfrak{sl}(n, \mathbb{C})$ we denote by ${}^{j}A^{i}$ the matrix which the only nonzero column is the j-th, and this column is exactly the i-th column of A. In a similar way, ${}_{j}A_{i}$ stands for the matrix having all lines zero, except for the j-th line, with this line being precisely the i-th line of A.

It is easy to see that $E_{ij}A = {}_{i}A_{j}$, $AE_{ij} = {}^{j}A^{i}$ and $[E_{ij}, A] = {}_{i}A_{j} - {}^{j}A^{i}$.

Definition 3.2.1 (cf. [13]). *A matrix $B \in \mathfrak{sl}(n, \mathbb{C})$ is said **strongly regular** if its eigenvalues $\lambda_1, \ldots, \lambda_n$ are all distinct and the real parts of the eigenvalues satisfy*

$$\mathrm{Re}(\lambda_p) - \mathrm{Re}(\lambda_q) \neq \mathrm{Re}(\lambda_s) - \mathrm{Re}(\lambda_t), \;\; \forall\, p, q, s, t \textit{ such that } \{p, q\} \neq \{s, t\}.$$

Through section $B = \mathrm{diag}(b_1, b_2, \cdots, b_n)$ will indicate a fixed strongly regular element where $b_i \in \mathbb{C}$ for $i = 1, 2, \ldots, n$, and $\mathrm{Re}(b_p) - \mathrm{Re}(b_q) \neq \mathrm{Re}(b_s) - \mathrm{Re}(b_t)$, for all p, q, s, t such that $\{p, q\} \neq \{s, t\}$. We should ask the following additional hypotheses on B:

$$\mathrm{Im}(b_i) - \mathrm{Im}(b_j) \neq 0 \;\; \text{whenever } i \neq j$$

and

$$\mathrm{Re}(b_1) > \mathrm{Re}(b_2) > \cdots > \mathrm{Re}(b_n).$$

Lemma 3.2.2. *If $B \in \mathrm{LS}(\Gamma)$, $\pm E_{ij} \in \mathrm{LS}(\Gamma)$ and $E_{jm} \in \mathrm{LS}(\Gamma)$, $i \neq j$, $j \neq m$, $i \neq m$, then*

$$\mathrm{span}_{\mathbb{C}}(E_{im}) \subset \mathrm{LS}(\Gamma).$$

Proof: We have that

$$\exp(-tE_{ij})E_{jm}\exp(tE_{ij}) = E_{jm} - tE_{im} \in \mathrm{LS}(\Gamma), \qquad \forall t \in \mathbb{R},$$

and this implies that

$$\lim_{t\to\pm\infty} \frac{1}{|t|}(E_{jm} - tE_{im}) = \pm E_{im} \in \mathrm{LS}(\Gamma).$$

Now,

$$\exp(tB)(\pm E_{im})\exp(-tB) = e^{t(b_i-b_m)}(\pm E_{im}) \in \mathrm{LS}(\Gamma), \quad \forall t \in \mathbb{R}.$$

Choosing $t \in \mathbb{R}$ in way that $e^{t(b_i-b_m)} \in \mathbb{R}$ we see that $\mathrm{span}_{\mathbb{R}}(E_{im}) \subset \mathrm{LS}(\Gamma)$. Similarly, taking $t \in \mathbb{R}$ such that $e^{t(b_i-b_m)}$ is purely imaginary we get $\mathrm{span}_{\mathbb{R}}(\mathbf{i}E_{im})$. Putting altogether we conclude that $\mathrm{span}_{\mathbb{C}}(E_{im}) \subset \mathrm{LS}(\Gamma)$. □

As in the previous lemma, if A and $\pm E_{ij}$ belong to $\mathrm{LS}(\Gamma)$, $i \neq j$, then $\pm({}^jA^i - {}_iA_j) \in \mathrm{LS}(\Gamma)$. A simple calculation with the curve $\exp(-tE_{ij})A\exp(tE_{ij})$, $t \in \mathbb{R}$, shows this fact.

Lemma 3.2.3. *If A, $\pm B$ and $\pm E_{ij}$ (with $i \neq j$) are elements in $\mathrm{LS}(\Gamma)$ and if c_{km} (with $k \neq m$) is a nonzero and non-diagonal entry in ${}^jA^i$ or in ${}_iA_j$, then $\mathrm{span}_{\mathbb{C}}(E_{km}) \subset \mathrm{LS}(\Gamma)$.*

Proof: According to the precedent paragraph we have $\pm({}^jA^i - {}_iA_j) \in \mathrm{LS}(\Gamma)$. Then, the curves

$$\pm\exp(tB)({}^jA^i - {}_iA_j)\exp(-tB)$$

are entirely contained ($t \in \mathbb{R}$) in the Lie saturate $\mathrm{LS}(\Gamma)$. The km-entries of these curves are given by

$$\pm({}^jA^i - {}_iA_j)_{km}e^{(b_k-b_m)t},$$

where $({}^jA^i - {}_iA_j)_{km}$ indicates the km-entry of the matrix ${}^jA^i - {}_iA_j$. Let k_0m_0 be the index corresponding to the maximum of the differences $\mathrm{Re}(b_k)-\mathrm{Re}(b_m)$ such that $({}^jA^i - {}_iA_j)_{km} \neq 0$. We can choose a sequence $t_N \to \infty$ in way that $e^{t_N(b_{k_0}-b_{m_0})} \in \mathbb{R}$. Moreover, taking $t_N \in \mathbb{R}$ such that

$$\cos(t_N(\mathrm{Im}(b_{k_0}) - \mathrm{Im}(b_{m_0}))) = 1 \quad \text{and} \quad \sin(t_N(\mathrm{Im}(b_{k_0}) - \mathrm{Im}(b_{m_0}))) = 0,$$

we get

$$e^{t_N(b_{k_0}-b_{m_0})} = e^{t_N(\mathrm{Re}(b_{k_0})-\mathrm{Re}(b_{m_0}))} \in \mathbb{R}.$$

This choice implies that

$$\pm\frac{1}{e^{t_N(b_{k_0}-b_{m_0})}} \exp(tB)({}^jA^i - {}_iA_j)\exp(-tB) \in \mathrm{LS}(\Gamma), \quad \forall N \in \mathbb{N}.$$

The nonzero entries of these two curves have the form

$$\pm c_{km}\frac{e^{t_N(b_k-b_m)}}{e^{t_N(b_{k_0}-b_{m_0})}} = \pm c_{km}\frac{e^{t_N(b_k-b_m)}}{e^{t_N(\mathrm{Re}(b_{k_0})-\mathrm{Re}(b_{m_0}))}}.$$

Taking limits as $t_N \to \infty$ all these entries converges to 0, except that one corresponding to the k_0m_0-entry, which is exactly $c_{k_0m_0}$. In other words,

$$\lim_{t_N\to\infty} \frac{1}{e^{t_N(b_{k_0}-b_{m_0})}} \exp(tB)({}^jA^i - {}_iA_j)\exp(-tB) = c_{k_0m_0}E_{k_0m_0} \in \mathrm{LS}(\Gamma).$$

This implies that $\exp(tB)(c_{k_0m_0}E_{k_0m_0})\exp(-tB) \in \mathrm{LS}(\Gamma)$, for all $t \in \mathbb{R}$. But

$$\exp(tB)(c_{k_0m_0}E_{k_0m_0})\exp(-tB) = e^{t(b_{k_0}-b_{m_0})}c_{k_0m_0}E_{k_0m_0}.$$

We can make four different choices of $t \in \mathbb{R}$ to get

$$\cos(t_N(\mathrm{Im}(b_{k_0}) - \mathrm{Im}(b_{m_0}))) = \pm 1 \;\text{ and }\; \mathbf{i}\sin(t_N(\mathrm{Im}(b_{k_0}) - \mathrm{Im}(b_{m_0}))) = \pm\mathbf{i},$$

and this shows that $\mathrm{span}_{\mathbb{R}}(c_{k_0m_0}E_{k_0m_0}) \subset \mathrm{LS}(\Gamma)$ and $\mathrm{span}_{\mathbb{R}}(\mathbf{i}c_{k_0m_0}E_{k_0m_0}) \subset \mathrm{LS}(\Gamma)$, that is,

$$\mathrm{span}_{\mathbb{C}}(c_{k_0m_0}E_{k_0m_0}) = \mathrm{span}_{\mathbb{C}}(E_{k_0m_0}) \subset \mathrm{LS}(\Gamma).$$

We can then consider the element $({}^jA^i - {}_iA_j) - c_{k_0m_0}E_{k_0m_0}$ and iterate the preceding procedure. At this point the maximum over the numbers $\mathrm{Re}(b_k) - \mathrm{Re}(b_m)$ will have decreased and this iteration will successively show that all the needed inclusions are true. $\square$

We can now state a controllability theorem on $\mathfrak{sl}(n, \mathbb{C})$.

Theorem 3.2.4. *Let $B = \mathrm{diag}(b_1, b_2, \cdots, b_n) \in \mathfrak{sl}(n, \mathbb{C})$ a strongly regular element satisfying*

$$\mathrm{Re}(b_1) > \mathrm{Re}(b_2) > \cdots > \mathrm{Re}(b_n) \;\; \textit{and} \;\; \mathrm{Im}(b_i) - \mathrm{Im}(b_j) \neq 0 \;\; \textit{whenever } i \neq j.$$

If $A \in \mathfrak{sl}(n, \mathbb{C})$ is such that $a_{1n} \neq 0 \neq a_{n1}$, then the invariant control system $\Gamma = A + uB$ is controllable on $\mathrm{Sl}(n, \mathbb{C})$ if and only if A is irreducible.

Proof: At first we prove that the irreducibility of A is a necessary condition for controllability. To see this we suppose A reducible and show that the system Γ is not controllable. If the matrix A is reducible, then there exists a permutation matrix P such that

$$PAP^t = \begin{bmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{bmatrix}. \tag{3.1}$$

Being B a diagonal matrix, it is easy to see that any permutation matrix leaves B in the diagonal form (by conjugation as above), and this means that both A and B are conjugate to matrices of the form (3.1). The set formed by the matrices having the form (3.1) is a Lie subalgebra of $\mathfrak{sl}(n, \mathbb{C})$. Consequently, $\mathrm{Lie}(\Gamma) \subsetneq \mathfrak{sl}(n, \mathbb{C})$ and this implies that $\mathrm{LS}(\Gamma) \neq \mathfrak{sl}(n, \mathbb{C})$.

Now we prove the sufficiency, in other words, we show that the system Γ under the conditions imposed on A and B is controllable.

In this way, our first aim is to show that $\pm E_{1n} \in \mathrm{LS}(\Gamma)$ and $\pm E_{n1} \in \mathrm{LS}(\Gamma)$.

As $\mathrm{LS}(\Gamma)$ is a closed convex positive cone, we have $\exp(t\,B) A \exp(t\,B) \in \mathrm{LS}(\Gamma)$ for all $t \in \mathbb{R}$. Taking $(t_N)_{N \in \mathbb{N}} \subset \mathbb{R}$ a sequence such that $e^{\pm t_N(b_n - b_1)} \in \mathbb{R}$, simple calculations together with limits show us that

$$e^{t_N(b_n - b_1)} \in \mathbb{R} \exp(t\,B) A \exp(t\,B) \in \mathrm{LS}(\Gamma) \longrightarrow a_{1n} E_{n1} = A_{n1} \in \mathrm{LS}(\Gamma)$$

and

$$e^{-t_N(b_n - b_1)} \in \mathbb{R} \exp(t\,B) A \exp(t\,B) \in \mathrm{LS}(\Gamma) \longrightarrow a_{1n} E_{1n} = A_{1n} \in \mathrm{LS}(\Gamma).$$

Knowing that $A_{1n}, A_{n1} \in \mathrm{LS}(\Gamma)$, we have

$$\exp(t\,\mathrm{ad}(B))A_{1n} \in \mathrm{LS}(\Gamma) \quad \text{and} \quad \exp(t\,\mathrm{ad}(B))A_{n1} \in \mathrm{LS}(\Gamma),$$

for all $t \in \mathbb{R}$. But it is easy to see that

$$\exp(t\,\mathrm{ad}(B))A_{1n} = \exp(t\,B)A_{1n}\exp(-t\,B) = e^{t(b_1-b_n)}A_{1n}$$

and

$$\exp(t\,\mathrm{ad}(B))A_{n1} = \exp(t\,B)A_{n1}\exp(-t\,B) = e^{t(b_n-b_1)}A_{n1}.$$

Calling $M = \mathrm{Im}(b_1) - \mathrm{Im}(b_n)$ and letting $t = \frac{\pi}{M}$ we have

$$e^{t(b_1-b_n)} = e^{\frac{\pi(b_1-b_n)}{M}} = e^{\pi\frac{\mathrm{Re}(b_1-b_n)}{\mathrm{Im}(b_1-b_n)}}(\cos(\pi) + \mathbf{i}\sin(\pi)) = -e^{\pi\frac{\mathrm{Re}(b_1-b_n)}{\mathrm{Im}(b_1-b_n)}}$$

and

$$e^{t(b_n-b_1)} = e^{\frac{\pi(b_n-b_1)}{M}} = e^{\pi\frac{\mathrm{Re}(b_n-b_1)}{\mathrm{Im}(b_n-b_1)}}(\cos(\pi) + \mathbf{i}\sin(\pi)) = -e^{\pi\frac{\mathrm{Re}(b_n-b_1)}{\mathrm{Im}(b_n-b_1)}}.$$

This choice together with $t = 0$ shows us that $\mathrm{span}_{\mathbb{R}}\{A_{1n}, A_{n1}\} \subset \mathrm{LS}(\Gamma)$, but this is equivalent to

$$\mathrm{span}_{\mathbb{R}}\{E_{1n}, E_{n1}\} \subset \mathrm{LS}(\Gamma).$$

In the same fashion, choosing $t \in \mathbb{R}$ such that $\cos(t) = 0$ and $\sin(t) = \pm 1$ we get $\mathrm{span}_{\mathbb{R}}\{\mathbf{i}A_{1n}, \mathbf{i}A_{n1}\} \subset \mathrm{LS}(\Gamma)$. And these considerations allow us to conclude that

$$\mathrm{span}_{\mathbb{C}}\{E_{1n}\} \subset \mathrm{LS}(\Gamma) \quad \text{and} \quad \mathrm{span}_{\mathbb{C}}\{E_{n1}\} \subset \mathrm{LS}(\Gamma).$$

We proceed to prove that $\mathrm{LS}(\Gamma) = \mathfrak{sl}(n, \mathbb{C})$. Being A irreducible, Theorem 3.1.7 says us that for any $m \neq n$ exists a directed path on the graph of A from the node P_n to the node P_m. We can choose this path free of loops, that is, a path

$$\overrightarrow{P_n P_{i_1}},\ \ \overrightarrow{P_{i_1} P_{i_2}}, \cdots, \overrightarrow{P_{i_p} P_m}\ .$$

without repetition of index. In particular,

$$P_{i_k} \neq P_n, \quad P_{i_k} \neq P_m \quad \forall k = 1, 2, \ldots, p.$$

Further, we are interested in the case which this path does not pass through P_1. The only two obvious possibilities are:

- This path in fact does not pass through P_1 or
- $P_{i_k} = P_1$ for some k.

If the second case occurs, taking the greatest k such that $P_{i_k} = P_1$, we get the path

$$P_1\overrightarrow{P}_{i_{k+1}}, \quad P_{i_{k+1}}\overrightarrow{P}_{i_{k+2}}, \cdots, P_{i_p}\overrightarrow{P}_m,$$

such that $P_{i_{k+h}} \neq P_1$, $P_{i_{k+h}} \neq P_m$ and $P_{i_{k+h}} \neq P_n$ for every h. In every case we get either

1. a path from the node P_n to the node P_m not passing through P_1

   or

2. a path from the node P_1 to the node P_m not passing through P_n.

Note that the former situation can be obtained from the latter just by interchanging 1 and n, and this would not affect the generality of our reasoning. Therefore we can assume that the path is from P_n to P_m not passing through P_1, and the arcs of this path correspond respectively to $a_{ni_1}, a_{i_1i_2}, \ldots, a_{i_pm}$ (nonzero entries of A).

As $\pm E_{1n} \in \mathrm{LS}(\Gamma)$ and $\pm(a_{ni_1}E_{1i_1})_{1i_1}$ is a nonzero and non diagonal (because $i_k \neq 1$) entry in $\pm_1 A_n$, an application of Lemma 3.2.3 shows that $\mathrm{span}_{\mathbb{C}}\{E_{1i_1}\} \subset \mathrm{LS}(\Gamma)$. A simple iteration of this reasoning shows also that

$$\mathrm{span}_{\mathbb{C}}\{E_{1i_2}, \ldots, E_{1m}\} \subset \mathrm{LS}(\Gamma).$$

Being $\pm E_{n1}$ elements of $\mathrm{LS}(\Gamma)$, as $E_{1m} \in \mathrm{LS}(\Gamma)$ the Lemma 3.2.2 shows us that $\mathrm{span}_{\mathbb{C}}\{E_{nm}\} \subset \mathrm{LS}(\Gamma)$.

We now take a path from the node P_m to the node P_n with corresponding nonzero and non diagonal entries $a_{mj_1}, a_{j_1j_2}, \ldots, a_{j_pn}$. We make a procedure like the previous one but in the opposite sense, that is, starting by the last arc on the path, that corresponds to the entry $a_{j_pn}(E_{j_p1})_{j_p1}$ of ${}^1A^n$ we conclude that $\mathrm{span}_{\mathbb{C}}\{E_{j_p1}\} \subset \mathrm{LS}(\Gamma)$. Then we follow with the iterations until the first arc on the path. At this point we have that

$a_{mj_1}(E_{m1})_{m1}$ is a nonzero and nondiagonal entry on ${}^1A^{j_1}$, thus Lemma 3.2.3 implies that $\mathrm{span}_{\mathbb{C}}\{E_{m1}\} \subset \mathrm{LS}(\Gamma)$.

Since $\mathrm{span}_{\mathbb{C}}\{E_{m1}\} \subset \mathrm{LS}(\Gamma)$ and $\mathrm{span}_{\mathbb{C}}\{E_{1n}\} \subset \mathrm{LS}(\Gamma)$ Lemma 3.2.2 ensures that $\mathrm{span}_{\mathbb{C}}\{E_{mn}\} \subset \mathrm{LS}(\Gamma)$.

Let $k \neq 1, n, m$. Applying the same reasoning as before we get

$$\mathrm{span}_{\mathbb{C}}\{E_{km}\} \subset \mathrm{LS}(\Gamma) \quad \text{and} \quad \mathrm{span}_{\mathbb{C}}\{E_{mk}\} \subset \mathrm{LS}(\Gamma).$$

This means that $\mathrm{span}_{\mathbb{C}}\{E_{km}\} \subset \mathrm{LS}(\Gamma)$ for all pairs of indices (k, m) with $k \neq m$, but these subspaces generate $\mathfrak{sl}(n, \mathbb{C})$ as a Lie algebra. Consequently $\mathrm{LS}(\Gamma) = \mathfrak{sl}(n, \mathbb{C})$ and this implies that Γ is controllable on $\mathrm{Sl}(n, \mathbb{C})$. □

3.3 Controllability on $\mathrm{Sl}(n, \mathbb{H})$

The Lie algebra $\mathfrak{sl}(2, \mathbb{H})$ is the real simple Lie algebra formed by matrices having trace with zero real part and it complexifies to a complex Lie algebra isomorphic to $\mathfrak{sl}(4, \mathbb{C})$. It is a fifteen dimensional Lie algebra whose basis vectors will be denoted as follows.

Set E_{ij} to be the matrix whose the (k, m) entry is given by

$$(E_{ij})_{km} = \delta_{ik} \cdot \delta_{jm}, \quad \text{where} \quad \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j, \end{cases}$$

and let E_{ij}^q denote the 2×2 matrix whose the (k, m) entry is given by

$$(E_{ij}^q)_{km} = q \cdot \delta_{ik} \cdot \delta_{jm}, \quad \text{where} \quad \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

Then, the set $\mathcal{B}$ formed by the matrices

$$E_{22} - E_{11}, E_{12}, E_{21}, E_{12}^{\mathbf{i}}, E_{12}^{\mathbf{j}}, E_{12}^{\mathbf{k}}, E_{21}^{\mathbf{i}}, E_{21}^{\mathbf{j}}, E_{21}^{\mathbf{k}}, E_{11}^{\mathbf{i}}, E_{11}^{\mathbf{j}}, E_{11}^{\mathbf{k}}, E_{22}^{\mathbf{i}}, E_{22}^{\mathbf{j}}, E_{22}^{\mathbf{k}}$$

will be considered as the canonical basis of $\mathfrak{sl}(2, \mathbb{H})$ as a vector space over $\mathbb{R}$.

For sake of convenience we reproduce here the Lemma 2.3.4 and its Corollary 2.3.6 that were proved in Chapter 1.

Lemma 3.3.1. *Consider the torus* $\mathbb{T}^2$ *acting on the quaternions* $\mathbb{H}$ *by*

$$\phi\left((t, s), q\right) = e^{\mathbf{i}t} q e^{-\mathbf{i}s}.$$

Write $q = a + b$, *with* $a = x_1 + x_2\mathbf{i} \in \mathbb{H}_{\{\mathbf{1},\mathbf{i}\}}$ *and* $b = x_3\mathbf{j} + x_4\mathbf{k} \in \mathbb{H}_{\{\mathbf{j},\mathbf{k}\}}$. *Suppose that* $q \notin \mathbb{H}_{\{\mathbf{1},\mathbf{i}\}} \cup \mathbb{H}_{\{\mathbf{j},\mathbf{k}\}}$, *that is,* $a \neq 0 \neq b$. *Then, the orbit* $\mathbb{T}^2 q$ *is a* 2*-dimensional torus and* $\mathbb{H}$ *is the convex cone generated by* $\mathbb{T}^2 q$.

Corollary 3.3.2. *Consider* $c_1, c_2 \in \mathbb{R}$ *with* $c_1 c_2 \neq 0$ *and* c_1/c_2 *irrational. Take* $q \in \mathbb{H}$ *with* $q \notin \mathbb{H}_{\{1,\mathbf{i}\}} \cup \mathbb{H}_{\{\mathbf{j},\mathbf{k}\}}$. *Then* $\mathbb{H}$ *is the closed convex cone generated by the curve* $e^{\mathbf{i}tc_1} q e^{-\mathbf{i}tc_2}$.

With this results and notations in mind we are able to state a controllability theorem that slightly improves Theorem 2.3.1 for the Lie group $\mathrm{Sl}(2, \mathbb{H})$, in the sense that under certain conditions we get a necessary and sufficient condition for controllability.

Theorem 3.3.3. *Let* $A, B \in \mathfrak{sl}(2, \mathbb{H})$ *be matrices such that* $B = \mathrm{diag}\{a + \mathbf{i}b_1, -a + \mathbf{i}b_2\}$ *with* $b_1 b_2 \neq 0$ *and* b_1/b_2 *irrational. Denote the* $(1, 2)$ *and* $(2, 1)$ *entries of the matrix* A *by* $p \in \mathbb{H}$ *and* $q \in \mathbb{H}$, *respectively. Let* $\mathbb{H}_{1,\mathbf{i}}$ *and* $\mathbb{H}_{\mathbf{j},\mathbf{k}}$ *be the (real) subspaces of* $\mathbb{H}$ *spanned by* $\{1, \mathbf{i}\}$ *and* $\{\mathbf{j}, \mathbf{k}\}$, *respectively, and suppose that* p *and* q *do not belong to* $\mathbb{H}_{1,\mathbf{i}} \cup \mathbb{H}_{\mathbf{j},\mathbf{k}}$. *Then the right invariant control system with unrestricted controls* $\Gamma = A + \mathbb{R}B$ *is controllable on* $\mathrm{Sl}(2, \mathbb{H})$ *if and only if* p *and* q *are both nonzero quaternions.*

Proof: We start by proving the necessary condition. Since we are dealing with matrices of order 2, the condition $p \neq 0 \neq q$ is equivalent to A being an irreducible matrix. So, if we suppose A reducible it would exist a permutation matrix P such that

$$PAP^T = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}. \tag{3.2}$$

As B is diagonal, it is easy to see that any permutation matrix leaves B in the diagonal form (by conjugation as above), and this means that both A and B are conjugate to matrices of the form (3.2). The set formed by the matrices having the form (3.2) is a Lie subalgebra of $\mathfrak{sl}(2, \mathbb{H})$. Consequently, $\mathrm{Lie}(\Gamma) \subsetneq \mathfrak{sl}(2, \mathbb{H})$ and this implies that $\mathrm{LS}(\Gamma) \neq \mathfrak{sl}(2, \mathbb{H})$.

By the other hand, to prove that Γ is controllable it suffices to show that $\mathrm{LS}(\Gamma) = \mathfrak{sl}(2, \mathbb{H})$. First of all, since $\mathrm{LS}(\Gamma)$ is a closed convex positive cone, then $A + tB \in \mathrm{LS}(\Gamma)$

implies that $\pm B \in \mathrm{LS}(\Gamma)$ as a consequence of the following convergences

$$\frac{A + tB}{|t|} \longrightarrow \pm B \text{ as } t \to \pm\infty.$$

Thus, $A, \pm B \in \mathrm{LS}(\Gamma)$ and we can apply Proposition 3.1.2 item (4) with $X = B$ and $Y = A$ to conclude that $\exp(t\,\mathrm{ad}B)A \in \mathrm{LS}(\Gamma)$ for all $t \in \mathbb{R}$.

Direct calculations shows us that

$$e^{-2at}\exp(t\,\mathrm{ad}B)A = \begin{bmatrix} e^{-2at}s & e^{\mathbf{i}t(b_1-b_2)}p \\ e^{-4at}e^{\mathbf{i}t(b_1-b_2)}q & e^{-2at}r \end{bmatrix}, \text{ where } A = \begin{bmatrix} s & p \\ q & r \end{bmatrix}$$

and the quaternions p and q are as in the statement of the theorem. It is clear that when $t \to \infty$ all the entries of this matrix converges to 0 except for the $(1,2)-$entry.

Choosing a sequence $t_k \to \infty$ such that $e^{\mathbf{i}t_k(b_1-b_2)} \to 1$ we conclude that

$$X = \begin{bmatrix} 0 & p \\ 0 & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Similarly, we have

$$e^{2at}\exp(t\,\mathrm{ad}B)A = \begin{bmatrix} e^{2at}s & e^{4at}e^{\mathbf{i}t(b_1-b_2)}p \\ e^{\mathbf{i}t(b_1-b_2)}q & e^{2at}r \end{bmatrix} \in \mathrm{LS}(\Gamma) \quad \forall t \in \mathbb{R}$$

and taking limits as $t \to -\infty$ we get

$$Y = \begin{bmatrix} 0 & 0 \\ q & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Again by the item (4) of Proposition 3.1.2 we have that

$$e^{-2at}\exp(tB)X\exp(tB) = \begin{bmatrix} 0 & e^{\mathbf{i}tb_1}pe^{-\mathbf{i}tb_2} \\ 0 & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma)$$

and

$$e^{2at}\exp(tB)Y\exp(tB) = \begin{bmatrix} 0 & 0 \\ e^{\mathbf{i}tb_2}qe^{-\mathbf{i}tb_1} & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Applying Corollary 3.3.2 to these curves we conclude that $\mathrm{span}_{\mathbb{H}}\{X\}$ and $\mathrm{span}_{\mathbb{H}}\{Y\}$ are contained in $\mathrm{LS}(\Gamma)$. In other words, this means that

$$\mathrm{span}_{\mathbb{R}}\{E_{12}, E_{21}, E_{12}^{\mathbf{i}}, E_{12}^{\mathbf{j}}, E_{12}^{\mathbf{k}}, E_{21}^{\mathbf{i}}, E_{21}^{\mathbf{j}}, E_{21}^{\mathbf{k}}\} \subset \mathrm{LS}(\Gamma).$$

Now, by the item (5) of Proposition 3.1.2 we have that

$$\pm[E_{21}, E_{12}] = \pm E_{22} - E_{11} \in \mathrm{LS}(\Gamma).$$

Also,

$$\pm[E_{12}^{\mathbf{j}}, E_{21}^{\mathbf{k}}] = \pm \begin{bmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{i} \end{bmatrix} \in \mathrm{LS}(\Gamma) \text{ and } \pm [E_{21}, E_{12}^{\mathbf{k}}] = \pm \begin{bmatrix} -\mathbf{i} & 0 \\ 0 & \mathbf{i} \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Adding these terms together will give us

$$\pm E_{11}^{\mathbf{i}} \in \mathrm{LS}(\Gamma) \text{ and } \pm E_{22}^{\mathbf{i}} \in \mathrm{LS}(\Gamma).$$

And we go on with similar computations:

$$\pm[E_{12}^{\mathbf{i}}, E_{21}^{\mathbf{k}}] = \pm \begin{bmatrix} -\mathbf{j} & 0 \\ 0 & -\mathbf{j} \end{bmatrix} \in \mathrm{LS}(\Gamma) \text{ and } \pm [E_{21}, E_{12}^{\mathbf{j}}] = \pm \begin{bmatrix} -\mathbf{j} & 0 \\ 0 & \mathbf{j} \end{bmatrix} \in \mathrm{LS}(\Gamma)$$

implies that

$$\pm E_{11}^{\mathbf{j}} \in \mathrm{LS}(\Gamma) \text{ and } \pm E_{22}^{\mathbf{j}} \in \mathrm{LS}(\Gamma).$$

Finally,

$$\pm[E_{12}^{\mathbf{i}}, E_{21}^{\mathbf{j}}] = \pm \begin{bmatrix} \mathbf{k} & 0 \\ 0 & \mathbf{k} \end{bmatrix} \in \mathrm{LS}(\Gamma) \text{ and } \pm [E_{21}, E_{12}^{\mathbf{k}}] = \pm \begin{bmatrix} -\mathbf{k} & 0 \\ 0 & \mathbf{k} \end{bmatrix} \in \mathrm{LS}(\Gamma),$$

what leads us to conclude that

$$\pm E_{11}^{\mathbf{k}} \in \mathrm{LS}(\Gamma) \text{ and } \pm E_{22}^{\mathbf{k}} \in \mathrm{LS}(\Gamma).$$

Consequently, $\pm\mathcal{B} \subset \mathrm{LS}(\Gamma)$ proving $\mathrm{LS}(\Gamma) = \mathfrak{sl}(2, \mathbb{H})$. □

Using Theorem 3.1.3 we also can state a similar theorem about the controllability of $\Gamma = A + \mathbb{R}B$, for specific $A, B \in \mathfrak{sl}(n, \mathbb{H})$. The reader should note that it is a weaker version of Theorem 2.3.1, because even if the Lie algebra rank condition is not asked, it follows immediately from the equality $\mathrm{LS}(\Gamma) = \mathfrak{sl}(n, \mathbb{H})$. However, if we keep in mind that the Lie algebra rank condition is usually hard to check even with the help of computational devices, the following theorem turns out to be interesting on its own, since all of its hypotheses can be quite readily verified.

Theorem 3.3.4. *Let $A, B \in \mathfrak{sl}(n, \mathbb{H})$ be matrices satisfying the following conditions*

1. $B = \mathrm{diag}\{\alpha_1 + \mathrm{i}\beta_1, \ldots, \alpha_n + \mathrm{i}\beta_n\}$ *with*

   i) $\alpha_1 > \alpha_2 > \cdots > \alpha_{n-1} > \alpha_n$;

   ii) $\alpha_i - \alpha_j \neq \alpha_k - \alpha_m$ *if* $(i, j) \neq (k, m)$.

   iii) $\beta_i \neq 0$ *for* $i = 1, \ldots, n$;

   iv) $\beta_i/\beta_j \notin \mathbb{Q}$ *if* $i \neq j$.

2. *Denote by a_{ij} the entries of the matrix A. Let $\mathbb{H}_{1,\mathbf{i}}$ and $\mathbb{H}_{\mathbf{j},\mathbf{k}}$ be the (real) subspaces of $\mathbb{H}$ spanned by $\{1, \mathbf{i}\}$ and $\{\mathbf{j}, \mathbf{k}\}$ respectively. Then*

   i) $a_{i1}, a_{1i} \notin \mathbb{H}_{1,\mathbf{i}} \cup \mathbb{H}_{\mathbf{j},\mathbf{k}}$ *for* $i = 2, \ldots, n$;

   ii) $a_{in}, a_{ni} \notin \mathbb{H}_{1,\mathbf{i}} \cup \mathbb{H}_{\mathbf{j},\mathbf{k}}$ *for* $i = 1, \ldots, n-1$.

Under these conditions, the bilinear control system $\Gamma = A + \mathbb{R}B$ is controllable.

Proof: As $A + uB \in \mathfrak{c}(S)$ for all $u \in \mathbb{R}$, it is immediate that $A \in \mathrm{LS}(\Gamma)$ and if $u \neq 0$ then

$$\frac{1}{|u|}A + \frac{u}{|u|}B = \frac{1}{|u|}(A + uB) \in \mathrm{LS}(\Gamma).$$

Taking limits as $u \to \pm\infty$ we see that $\pm B \in \mathrm{LS}(\Gamma)$. It follows that $e^{t\mathrm{ad}(B)}A \in \mathrm{LS}(\Gamma)$ and hence

$$e^{-t(\alpha_1-\alpha_n)}e^{t\mathrm{ad}(B)}A \in \mathrm{LS}(\Gamma)$$

for all $t \in \mathbb{R}$ where $\alpha_1, \ldots, \alpha_n$ are the real parts of the entries of B. Now by assumption $\alpha_1 > \alpha_2 > \cdots > \alpha_n$ so that as $t \to +\infty$ the entries $e^{-t(\alpha_1-\alpha_n)}e^{t\mathrm{ad}(B)}A$ converge to 0 except for the $(1, n)$-entry. The $(1, n)$-entry of $e^{-t(\alpha_1-\alpha_n)}e^{t\mathrm{ad}(B)}A$ is $e^{\mathbf{i}t(\beta_1-\beta_n)}p$ where p is as in the statement of the theorem and $\beta_1, \ldots, \beta_n$ are the imaginary parts of the entries of B. Choosing a sequence $t_k \to +\infty$ such that $e^{\mathbf{i}t(\beta_1 - \beta_n)} \to 1$ we conclude that

$$X = \begin{bmatrix} 0 & \cdots & p \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Using again the properties of the Lie saturate we have that for all $t, s \in \mathbb{R}$,

$$e^{-t(\alpha_1 - \alpha_n)} e^{t\mathrm{ad}(B)} X = \begin{bmatrix} 0 & \cdots & e^{\mathbf{i}t\beta_1} p e^{-\mathbf{i}t\beta_n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Applying Corollary 3.3.2 to this curve it follows that the subspace $\mathrm{span}_{\mathbb{H}}\{X\}$ is contained in $\mathrm{LS}(\Gamma)$. By taking limits as $t \to -\infty$ of $e^{-t(\alpha_1 - \alpha_n)} e^{t\mathrm{ad}(B)} A$ it follows that

$$Y = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ q & \cdots & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma),$$

and applying again the Corollary 3.3.2 we conclude that $\mathrm{span}_{\mathbb{H}}\{Y\} \subset \mathrm{LS}(\Gamma)$.

Claim 1. $\pm({}^nA^1 - {}_1A_n), \pm({}^1A^n - {}_nA_1) \in \mathrm{LS}(\Gamma)$.

In fact, since $\mathrm{span}_{\mathbb{H}}\{X\} \subset \mathrm{LS}(\Gamma)$, we have that $\pm E_{1n} \in \mathrm{LS}(\Gamma)$, this means that

$$\alpha(t) = \exp(-tE_{1n}) A \exp(tE_{1n}) \in \mathrm{LS}(\Gamma), \quad \forall t \in \mathbb{R}.$$

But

$$\begin{aligned} \alpha(t) &= (Id - tE_{1n}) A (Id + tE_{1n}) \\ &= A + tAE_{1n} - tE_{1n}A - t^2 E_{1n} A E_{1n} \\ &= A + t({}^nA^1) - t({}_1A_n) - t^2 a_{nn} E_{1n}. \end{aligned}$$

Again, as $\mathrm{span}_{\mathbb{H}}\{X\} \subset \mathrm{LS}(\Gamma)$, we have that $\beta(t) := \alpha(t) + t^2 a_{nn} E_{1n} \in \mathrm{LS}(\Gamma)$ for all $t \in \mathbb{R}$.

Thus

$$\frac{1}{|t|}\beta(t) = \frac{1}{|t|}A + \frac{t}{|t|}({}^{n}A^{1} - {}_{1}A_{n}) \in \mathrm{LS}(\Gamma), \quad \forall t \in \mathbb{R}.$$

By taking limits as $t \to \pm\infty$ we get that $\pm({}^{n}A^{1} - {}_{1}A_{n}) \in \mathrm{LS}(\Gamma)$. The same proof as above with E_{n1} rather than E_{1n} shows us that $\pm({}^{1}A^{n} - {}_{n}A_{1}) \in \mathrm{LS}(\Gamma)$. □

Claim 2. $\mathrm{span}_{\mathbb{H}}\{E_{1i}, E_{ni}, E_{i1}, E_{in}\} \subset \mathrm{LS}(\Gamma)$ *for* $i = 2, \ldots, n-1$.

In fact, we already know that $\mathrm{span}_{\mathbb{H}}\{E_{1n}\} \subset \mathrm{LS}(\Gamma)$. So, let us consider the matrix $C := ({}^{n}A^{1} - {}_{1}A_{n}) - (a_{11} - a_{nn})E_{1n}$ which has null $(1, n)$-entry. Then the curves

$$\gamma_1(t) = \pm\exp(t\mathrm{ad}B)C \quad (t \in \mathbb{R}),$$

are contained in $\mathrm{LS}(\Gamma)$. Further, the (k, m)-entries of these curves are given by

$$\gamma_1(t)_{km} = \pm e^{t(\alpha_k - \alpha_m)} e^{\mathbf{i}t(\beta_k - \beta_m)} c_{km}.$$

Let (k_0, m_0) be a non diagonal entry such that $\alpha_{k_0} - \alpha_{m_0}$ is the maximum (or minimum) of the differences $\alpha_k - \alpha_m$ such that $c_{km} \neq 0$. Thus, as $t \to \infty$ $(-\infty)$, all the entries of $e^{-t(\alpha_{k_0} - \alpha_{m_0})}\gamma(t)$ goes to 0 but the (k_0, m_0)-entry, which is given by $e^{\mathbf{i}t(\beta_{k_0} - \beta_{m_0})}c_{k_0 m_0}$. Taking a sequence $t_l \to \infty$ such that $e^{\mathbf{i}t_l(\beta_{k_0} - \beta_{m_0})} \to 1$ we obtain that $c_{k_0 m_0}E_{k_0 m_0} \in \mathrm{LS}(\Gamma)$. Recalling that $c_{k_0 m_0}$ is one of that terms listed in the condition 2 of the theorem, we can apply Corollary 3.3.2 and conclude that $\mathrm{span}_{\mathbb{H}}\{c_{k_0 m_0}E_{k_0 m_0}\} \subset \mathrm{LS}(\Gamma)$. That is,

$$\mathrm{span}_{\mathbb{H}}\{E_{k_0 m_0}\} \subset \mathrm{LS}(\Gamma).$$

Now, consider the matrix $C_1 := C - c_{k_0 m_0}E_{k_0 m_0}$. We can iterate the reasoning just made until getting

$$\mathrm{span}_{\mathbb{H}}\{E_{12}, E_{13}, \ldots, E_{1(n-1)}, E_{2n}, E_{3n}, \ldots, E_{(n-1)n}\} \subset \mathrm{LS}(\Gamma).$$

In the same fashion, we define the matrix $D = ({}^{1}A^{n} - {}_{n}A_{1}) - (a_{nn} - a_{11})$ and the curves

$$\gamma_2(t) = \pm\exp(t\mathrm{ad}B)D \quad (t \in \mathbb{R}),$$

and proceed in a similar way as before to conclude that

$$\mathrm{span}_{\mathbb{H}}\{E_{21}, E_{31}, \ldots, E_{(n-1)1}, E_{n2}, E_{n3}, \ldots, E_{n(n-1)}\} \subset \mathrm{LS}(\Gamma).$$

□

We turn now our attention to the diagonal matrices. Direct computations shows us that

$$\pm[E_{1i}, E_{i1}] = \pm(E_{11} - E_{ii}) \in \mathrm{LS}(\Gamma), \quad i = 2, 3, \ldots, n.$$

We also have that

$$\pm[E^{\mathbf{j}}_{1i}, E^{\mathbf{k}}_{i1}] = \pm(E^{\mathbf{i}}_{11} + E^{\mathbf{i}}_{ii}) \in \mathrm{LS}(\Gamma) \quad i = 1, 2, \ldots, n,$$

and

$$\pm[E^{\mathbf{i}}_{1i}, E_{i1}] = \pm(E^{\mathbf{i}}_{11} - E^{\mathbf{i}}_{ii}) \in \mathrm{LS}(\Gamma) \quad i = 1, 2, \ldots, n,$$

which implies that $\pm E^{\mathbf{i}}_{ii} \in \mathrm{LS}(\Gamma)$ for $i = 1, 2, \ldots, n$. Interchanging the roles of $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ in the above calculations we get

$$\pm E^{\mathbf{i}}_{ii},\ \pm E^{\mathbf{j}}_{ii},\ \pm E^{\mathbf{k}}_{ii} \in \mathrm{LS}(\Gamma), \qquad i = 1, 2, \ldots, n.$$

Finally, we can prove that the remaining matrices are in $\mathrm{LS}(\Gamma)$ by the following relations. For $i, j \neq 1$, $i, j \neq n$ and $i \neq j$,

$$\pm[E_{i1}, E_{1j}] = \pm E_{ij},$$

$$\pm[E_{i1}, E^{\mathbf{i}}_{1j}] = \pm E^{\mathbf{i}}_{ij}, \ \ \pm[E_{i1}, E^{\mathbf{j}}_{1j}] = \pm E^{\mathbf{j}}_{ij}, \ \ \pm[E_{i1}, E^{\mathbf{k}}_{1j}] = \pm E^{\mathbf{k}}_{ij}.$$

Consequently, $\mathrm{LS}(\Gamma) = \mathfrak{sl}(n, \mathbb{H})$ proving that the bilinear control system Γ is controllable on $\mathrm{Sl}(n, \mathbb{H})$. This concludes the proof of Theorem 3.3.4

□

3.4 Application to semidirect products

Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras and ρ a representation of $\mathfrak{g}$ on $\mathfrak{h}$ such that $\rho(X) \in \mathrm{Der}\mathfrak{h}$ for all $X \in \mathfrak{g}$. The semidirect product of $\mathfrak{g}$ and $\mathfrak{h}$ is defined to be $\mathfrak{g} \times \mathfrak{h}$ endowed with the Lie brackets

$$[(X_1, Y_1), (X_2, Y_2)] = ([X_1, X_2], \rho(X_1)Y_2 - \rho(X_2)Y_1 + [Y_1, Y_2]) .$$

The Lie algebra $\mathfrak{g} \times \mathfrak{h}$ decomposes as the direct sum of the ideals $\mathfrak{g} \times 0$ and $0 \times \mathfrak{h}$ that are isomorphic to $\mathfrak{g}$ and $\mathfrak{h}$, respectively. A detailed construction of the semidirect product can be found in [28].

Theorem 3.1.3 tells us that we can decide about the controllability of a right invariant control system Γ just by checking a condition at the Lie algebra level. In such a way, following the technique introduced by Gauthier and Bornard in [8], based on the seminal papers of Jurdjevic and Kupka [13] and [12], we get conditions for controllability of right invariant control systems on the semidirect products $\mathrm{Sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$ and $\mathrm{Sl}(2, \mathbb{C}) \rtimes \mathbb{C}^2$. We also refer to [20] for a different approach of controllability on $\mathrm{Sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$.

3.4.1 The case $\mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$

The Lie algebra $\mathfrak{sl}(2, \mathbb{R})$ represents canonically on $\mathbb{R}^2$ via matrix multiplication, that is,

$$\rho(X) : \mathbb{R}^2 \to \mathbb{R}^2, \;\; \rho(X)v = X \cdot v, \quad X \in \mathfrak{sl}(2, \mathbb{R}), \;\; v \in \mathbb{R}^2,$$

and clearly $\rho(X)$ is a derivation for every $X \in \mathfrak{sl}(2, \mathbb{R})$. Thus, we have the semidirect product $\mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$ with the Lie brackets

$$[(A_1, v_1), (A_2, v_2)] = ([A_1, A_2], A_1 v_2 - A_2 v_1). \tag{3.3}$$

Given

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathfrak{sl}(2, \mathbb{R}) \text{ and } v = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2,$$

the pair $(A, v) \in \mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$ can be written as the order 3 matrix in the block form

$$(A, v) = \begin{bmatrix} a_{11} & a_{12} & a_1 \\ a_{21} & a_{22} & a_2 \\ 0 & 0 & 0 \end{bmatrix}.$$

In this way, the Lie brackets (3.3) can be computed directly via the matrix commutator, that is,

$$[(A_1, v_1), (A_2, v_2)] = \begin{bmatrix} [A_1, A_2] & A_1 v_2 - A_2 v_1 \\ 0 & 0 \end{bmatrix}.$$

Since we are dealing with a low dimensional case, the notation for the basis elements can be shortened if we write $E_1 := (E_{22} - E_{11}, 0)$, $E_2 := (E_{12}, 0)$, $E_3 := (E_{21}, 0)$, $E_4 := (0, e_1)$, and $E_5 := (0, e_2)$. Thus we consider $\mathcal{B} = \{E_1, E_2, E_3, E_4, E_5\}$ as the canonical basis for $\mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$. The Lie brackets between these elements are easily computed and are given by

$$[E_1, E_2] = -2E_2, \;\; [E_1, E_3] = 2E_3, \;\; [E_1, E_4] = -E_4, \;\; [E_1, E_5] = E_5,$$

$$[E_2, E_3] = -E_1, \;\; [E_2, E_4] = 0, \;\; [E_2, E_5] = E_4,$$

$$[E_3, E_4] = E_5, \;\; [E_3, E_5] = 0, \;\; [E_4, E_5] = 0.$$

During this section and the next, and if there is no risk of confusion, intending to clarify the notation we can write simply B to indicate the element $(B, 0)$, or even write A for (A, α). It will be specified at the beginning of each proof.

Keeping these considerations in mind, we have

Theorem 3.4.1. *Consider in the Lie algebra $\mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$ the right invariant control system $\Gamma = (A, \alpha)x + u(B, 0)x$, $u \in \mathbb{R}$, defined by (A, α) and $(B, 0)$, such that*

$$(A, \alpha) = \begin{bmatrix} -a & a_{12} & a_1 \\ a_{21} & a & a_2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad (B, 0) = \begin{bmatrix} -b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

with $a, b > 0$ and $a_{12} \cdot a_{21} < 0$. Then, the system Γ is controllable on $\mathrm{Sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$.

Proof: Our aim is to show that $\mathrm{LS}(\Gamma) = \mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$. Let's do that.

First of all, it is easy to see that $A := (A, \alpha) \in \mathrm{LS}(\Gamma)$ and $\pm B := \pm(B, 0) \in \mathrm{LS}(\Gamma)$, and so does

$$\exp(t\,\mathrm{ad}(B))A = \exp(t\,B)A\exp(-t\,B), \quad \forall t \in \mathbb{R}.$$

But

$$\exp(t\,B)A\exp(-t\,B) = \begin{bmatrix} a \\ a_{12}e^{-2tb} \\ a_{21}e^{2tb} \\ a_1 e^{-tb} \\ a_2 e^{tb} \end{bmatrix},$$

and since $\pm B = \pm bE_1$, $b > 0$, we have $\pm E_1 \in \mathrm{LS}(\Gamma)$. Thus,

$$A_t = \begin{bmatrix} 0 \\ a_{12}e^{-2tb} \\ a_{21}e^{2tb} \\ a_1 e^{-tb} \\ a_2 e^{tb} \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Multiplying A_t by e^{2tb} and passing to the limit, we have that

$$\begin{bmatrix} 0 \\ a_{12} \\ 0 \\ 0 \\ 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Analogously, if we multiply A_t by e^{-2tb} and take into limits, we get

$$\begin{bmatrix} 0 \\ 0 \\ a_{21} \\ 0 \\ 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Without loss of generality, suppose that $a_{12} > 0$ and $a_{21} < 0$. Hence the previous inclusions shows us that

$$E_2 \in \mathrm{LS}(\Gamma) \quad \text{and} \quad -E_3 \in \mathrm{LS}(\Gamma).$$

Summing up, we have $E_2 - E_3 \in \mathrm{LS}(\Gamma)$. Now, this element generates a periodic 1-parameter subgroup:

$$\exp(t(E_2 - E_3)) = \begin{bmatrix} \cos(t) & \sin(t) & 0 \\ -\sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and this means that $-(E_2 - E_3) \in \mathrm{LS}(\Gamma)$. By combining these terms we get

$$-(E_2 - E_3) + E_2 = E_3 \in \mathrm{LS}(\Gamma)$$

and

$$-(E_2 - E_3) - E_3 = -E_2 \in \mathrm{LS}(\Gamma).$$

So $\pm E_1, \pm E_2, \pm E_3 \in \mathrm{LS}(\Gamma)$. Now, as $A \in \mathrm{LS}(\Gamma)$, we obtain $a_1E_4 + a_2E_5 \in \mathrm{LS}(\Gamma)$. This implies the following inclusions hold for all $t \in \mathbb{R}$:

$$\exp(tE_2)(a_1E_4 + a_2E_5)\exp(-tE_2) \in \mathrm{LS}(\Gamma)$$

and

$$\exp(tE_3)(a_1E_4 + a_2E_5)\exp(-tE_3) \in \mathrm{LS}(\Gamma).$$

A direct computation shows us that

$$\exp(tE_2)(a_1E_4 + a_2E_5)\exp(-tE_2) = \begin{bmatrix} 0 & 0 & a_1 + a_2t \\ 0 & 0 & a_2 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$\exp(tE_3)(a_1E_4 + a_2E_5)\exp(-tE_3) = \begin{bmatrix} 0 & 0 & a_1 \\ 0 & 0 & a_1t + a_2 \\ 0 & 0 & 0 \end{bmatrix}$$

Choosing in the first equation $t = \dfrac{-2a_1}{a_2}$, we get

$$X = \begin{bmatrix} 0 & 0 & -a_1 \\ 0 & 0 & a_2 \\ 0 & 0 & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma),$$

and if we choose in the second equation $t = \dfrac{-2a_2}{a_1}$, we get that

$$-X = \begin{bmatrix} 0 & 0 & a_1 \\ 0 & 0 & -a_2 \\ 0 & 0 & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

This allows us to compute Lie brackets inside the Lie saturate using these elements. So we have

$$[E_2, X] = [E_2, -a_1E_4 + a_2E_5] = -a_1[E_2, E_4] + a_2[E_2, E_5] = a_2E_4,$$

$$[E_2, -X] = [E_2, a_1E_4 - a_2E_5] = a_1[E_2, E_4] - a_2[E_2, E_5] = -a_2E_4$$

and this implies that $\pm E_4 \in \mathrm{LS}(\Gamma)$. Finally,

$$[E_3, X] = [E_3, -a_1E_4 + a_2E_5] = -a_1[E_3, E_4] + a_2[E_3, E_5] = -a_1E_5,$$

$$[E_3, -X] = [E_3, a_1E_4 - a_2E_5] = a_1[E_3, E_4] - a_2[E_3, E_5] = a_1E_5,$$

which means that $\pm E_5 \subset \mathrm{LS}(\Gamma)$. Consequently, $\pm\mathcal{B} \subset \mathrm{LS}(\Gamma)$ and we conclude that the invariant control system Γ is controllable on $\mathrm{Sl}(2, \mathbb{R})$. □

Theorem 3.4.2. *Consider in the Lie algebra $\mathfrak{sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$ the elements (A, α) and (B, β) such that*

$$(A, \alpha) = \begin{bmatrix} -a & a_{12} & a_1 \\ a_{21} & a & a_2 \\ 0 & 0 & 0 \end{bmatrix} \quad e \quad (B, \beta) = \begin{bmatrix} -b & 0 & b_1 \\ 0 & b & b_2 \\ 0 & 0 & 0 \end{bmatrix},$$

with $a, b > 0$ and $a_{12} \cdot a_{21} < 0$. Assume also that $ba_1 - ab_1 - a_{12}b_2 \neq 0$ and $ba_2 - ab_2 + a_{21}b_1 \neq 0$. Then, the control system $\Gamma = \{(A, \alpha) + u(B, \beta) \mid u \in \mathbb{R}\}$ is controllable on $\mathrm{Sl}(2, \mathbb{R}) \rtimes \mathbb{R}^2$.

Proof: As before, the proof will be accomplished as soon as we show that $\mathrm{LS}(\Gamma) = \mathfrak{sl}(2,\mathbb{R}) \rtimes \mathbb{R}^2$. Without loss of generality, suppose that $a_{12} > 0$ and $a_{21} < 0$.

First of all, it is easy to see that $A := (A, \alpha) \in \mathrm{LS}(\Gamma)$ and $\pm B := \pm(B, \beta) \in \mathrm{LS}(\Gamma)$. Thus

$$\exp(t\,\mathrm{ad}(B))A = \exp(t\,B)A\exp(-t\,B), \quad \forall t \in \mathbb{R}.$$

But

$$\exp(t\,B)A\exp(-t\,B) = \begin{bmatrix} a \\ a_{12}e^{-2tb} \\ a_{21}e^{2tb} \\ \frac{ab_1}{b}e^{-bt}(e^{bt}-1) - \frac{a_{12}b_2}{b}e^{-2bt}(e^{bt}-1) + a_1e^{-bt} \\ -\frac{ab_2}{b}(e^{bt}-1) - \frac{a_{21}b_1}{b}e^{bt}(e^{bt}-1) + a_2e^{bt} \end{bmatrix},$$

multiplying by $\dfrac{e^{2tb}}{a_{12}}$ and passing to the limit, we have that

$$M = \begin{bmatrix} 0 & 1 & 0 & \frac{b_2}{b} & 0 \end{bmatrix}^t = E_2 + \frac{b_2}{b}E_4 \in \mathrm{LS}(\Gamma)$$

Analogously, if we multiply by $\dfrac{-e^{-2tb}}{a_{21}}$ and take into limits, we get

$$N = \begin{bmatrix} 0 & 0 & -1 & 0 & \frac{b_1}{b} \end{bmatrix}^t = -E_3 + \frac{b_1}{b}E_5 \in \mathrm{LS}(\Gamma).$$

Summing up, we have $X = E_2 - E_3 + \frac{b_2}{b}E_4 + \frac{b_1}{b}E_5 \in \mathrm{LS}(\Gamma)$. Now, this element generates a periodic 1-parameter subgroup:

$$\exp(t\,X) = \begin{bmatrix} \cos(t) & \sin(t) & -b_1\cos(t) + b_2\sin(t) + b_1 \\ -\sin(t) & \cos(t) & b_2\cos(t) + b_1\sin(t) - b_2 \\ 0 & 0 & 1 \end{bmatrix},$$

and this means that $-X \in \mathrm{LS}(\Gamma)$. By combining these terms we get $-X + M = -N \in \mathrm{LS}(\Gamma)$ and $-X + N = -M \in \mathrm{LS}(\Gamma)$. Hence

$$\pm B, \pm M, \pm N \in \mathrm{LS}(\Gamma).$$

Now, as $(A, \alpha) \in \mathrm{LS}(\Gamma)$, we obtain

$$b(A,\alpha) - a(B,\beta) + ba_{21}N - ba_{12}M = \begin{bmatrix} 0 & 0 & ba_1 - ab_1 - a_{12}b_2 \\ 0 & 0 & ba_2 - ab_2 + a_{21}b_1 \\ 0 & 0 & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

Let $x = ba_1 - ab_1 - a_{12}b_2$ and $y = ba_2 - ab_2 + a_{21}b_1$. With this notation, we have that the following inclusions hold for all $t \in \mathbb{R}$:

$$\exp(t\,M)(xE_4 + yE_5)\exp(-t\,M) \in \mathrm{LS}(\Gamma)$$

and

$$\exp(t\,N)(xE_4 + yE_5)\exp(-t\,N) \in \mathrm{LS}(\Gamma).$$

A direct computation shows us that

$$\exp(t\,M)(xE_4 + yE_5)\exp(-t\,M) = \begin{bmatrix} 0 & 0 & x + ty \\ 0 & 0 & y \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$\exp(t\,N)(xE_4 + yE_5)\exp(-t\,N) = \begin{bmatrix} 0 & 0 & x \\ 0 & 0 & y - xt \\ 0 & 0 & 0 \end{bmatrix}.$$

Choosing in the first equation $t = \dfrac{-2x}{y}$, we get

$$Y = \begin{bmatrix} 0 & 0 & -x \\ 0 & 0 & y \\ 0 & 0 & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma),$$

and if we choose in the second equation $t = \dfrac{2y}{x}$, we get that

$$-Y = \begin{bmatrix} 0 & 0 & x \\ 0 & 0 & -y \\ 0 & 0 & 0 \end{bmatrix} \in \mathrm{LS}(\Gamma).$$

This allows us to compute Lie brackets inside the Lie saturate using these elements. So we have

$$\pm[M,Y] = \pm\left[E_2 + \frac{b_2}{b}E_4, -xE_4 + yE_5\right] = \pm y[E_2, E_5] = \pm yE_4$$

and this implies that $\pm E_4 \in \mathrm{LS}(\Gamma)$. And we have

$$\pm[M,N] = \pm\left[-E_3 + \frac{b_1}{b}E_5, -xE_4 + yE_5\right] = \pm x[E_3, E_4] = \pm xE_5$$

which means that $\pm E_5 \subset \mathrm{LS}(\Gamma)$. Also,

$$\pm\frac{1}{b}(B,\beta) \mp \frac{b_1}{b}E_4 \mp \frac{b_2}{b}E_5 = \pm E_1 \in \mathrm{LS}(\Gamma),$$

$$\pm M \mp \frac{b_2}{b}E_4 = \pm E_2 \in \mathrm{LS}(\Gamma)$$

and

$$\pm N \mp \frac{b_1}{b}E_5 = \pm E_3 \in \mathrm{LS}(\Gamma).$$

That is

$$\mathcal{B} = \{\pm E_1, \pm E_2, \pm E_3, \pm E_4, \pm E_5\} \subset \mathrm{LS}(\Gamma)$$

which implies that $\mathrm{LS}(\Gamma) = \mathfrak{sl}(2,\mathbb{R}) \rtimes \mathbb{R}^2$, proving that the invariant control system Γ is controllable. □

3.4.2 The case $\mathfrak{sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$

Now we consider the Lie group $\mathrm{Sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$, whose Lie algebra can be described as

$$\mathfrak{sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2 = \left\{ (A,z) := \begin{bmatrix} A & z \\ 0 & 0 \end{bmatrix} \;\middle|\; A \in \mathfrak{sl}(2,\mathbb{C}),\ z \in \mathbb{C}^2 \right\} \subset \mathfrak{sl}(3,\mathbb{C})$$

endowed with the Lie brackets

$$[(A_1,z_1),(A_2,z_2)] = \begin{bmatrix} [A_1,A_2] & A_1z_2 - A_2z_1 \\ 0 & 0 \end{bmatrix}.$$

As before, the set $\mathcal{B} = \{E_1, E_2, E_3, E_4, E_5\}$ will denote the canonical basis for the Lie algebra $\mathfrak{sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$.

Theorem 3.4.3. *Consider on the Lie group* $\mathrm{Sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$ *the right invariant control system* $\Gamma = \{(A,\alpha) + u(B,0) \,|\, u \in \mathbb{R}\}$, *where*

$$(A,\alpha) = \begin{bmatrix} a & a_{12} & a_1 \\ a_{21} & -a & a_2 \\ 0 & 0 & 0 \end{bmatrix} \quad \textit{and} \quad (B,0) = \begin{bmatrix} b & 0 & 0 \\ 0 & -b & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

with both a_{12} *and* a_{21} *nonzero complex numbers and* $\mathrm{Im}(b) \neq 0$. *Then,* Γ *is controllable on* $\mathrm{Sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$.

Proof: Write B to indicate $(B,0)$. As $(A,\alpha) + uB \in \mathrm{LS}(\Gamma)$ for all $u \in \mathbb{R}$, we obtain

$$\lim_{u\to\pm\infty} \frac{1}{|u|}\left((A,\alpha) + uB\right) = \pm B \in \mathrm{LS}(\Gamma),$$

since $\mathrm{LS}(\Gamma)$ is a positive closed convex cone in $\mathfrak{sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$. As a consequence we get $\exp(t\,\mathrm{ad}(B))(A,\alpha) \in \mathrm{LS}(\Gamma)$ for all $t \in \mathbb{R}$. By the formula

$$\exp(t\,\mathrm{ad}(B))(A,\alpha) = \exp(t\,B)(A,\alpha)\exp(-t\,B)$$

we have

$$\exp(t\,B) = \begin{bmatrix} e^{t\,b} & 0 & 0 \\ 0 & e^{-t\,b} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \exp(-t\,B) = \begin{bmatrix} e^{-t\,b} & 0 & 0 \\ 0 & e^{t\,b} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

hence

$$\exp(t\,B)(A,\alpha)\exp(-t\,B) = \begin{bmatrix} a & a_{12}e^{2bt} & a_1e^{bt} \\ a_{21}e^{-2bt} & -a & a_2e^{-bt} \\ 0 & 0 & 0 \end{bmatrix}.$$

Rewriting this element as a column vector in the basis $\mathcal{B}$, we get

$$\exp(t\,B)(A,\alpha)\exp(-t\,B) = \begin{bmatrix} a & a_{12}e^{2bt} & a_{21}e^{-2bt} & a_1e^{bt} & a_2e^{-bt} \end{bmatrix}^t.$$

Being $b \in \mathbb{C}$ with $\mathrm{Im}(b) \neq 0$, we can take a sequence $t_k \to \infty$ such that $e^{2t_k b} \in \mathbb{R}$.

Precisely, choose $t_k = \frac{k\pi}{\mathrm{Im}(b)}$, $k \in \mathbb{Z}$. So,

$$\begin{aligned} e^{2t_k b} &= e^{2t_k \mathrm{Re}(b)} \left(\cos\left(2t_k \mathrm{Im}(b)\right) + \mathbf{i}\sin\left(2t_k \mathrm{Im}(b)\right)\right) \\ &= e^{2k\pi\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}} \left(\cos\left(2k\pi\right) + \mathbf{i}\sin\left(2k\pi\right)\right) \\ &= e^{2k\pi\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}} \in \mathbb{R}, \ \ \forall k \in \mathbb{N}. \end{aligned}$$

As $e^{2k\pi\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}}$, $k \in \mathbb{Z}$, is a positive real number, we get that

$$e^{2k\pi\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}} \exp(t_k\, B)(A,\alpha)\exp(-t_k\, B) \in \mathrm{LS}(\Gamma), \quad \forall k \in \mathbb{Z}.$$

Calling $m = k\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}$ and taking into limits we get

$$\lim_{m\to\infty} e^{-2\pi m} \exp(t_k\, B)(A,\alpha)\exp(-t_k\, B) = \begin{bmatrix} 0 \\ a_{12} \\ 0 \\ 0 \\ 0 \end{bmatrix} =: A_1 \in \mathrm{LS}(\Gamma).$$

Also,

$$\lim_{m\to-\infty} e^{2\pi m} \exp(t_k\, B)(A,\alpha)\exp(-t_k\, B) = \begin{bmatrix} 0 \\ 0 \\ a_{21} \\ 0 \\ 0 \end{bmatrix} =: A_2 \in \mathrm{LS}(\Gamma).$$

Now, as $A_1, A_2 \in \mathrm{LS}(\Gamma)$, we have

$$\exp(t\,\mathrm{ad}(B))A_1 \in \mathrm{LS}(\Gamma) \quad \text{and} \quad \exp(t\,\mathrm{ad}(B))A_2 \in \mathrm{LS}(\Gamma),$$

for all $t \in \mathbb{R}$. But it is easy to see that

$$\exp(t\,\mathrm{ad}(B))A_1 = \exp(t\,B)A_1\exp(-t\,B) = e^{2bt}A_1$$

and

$$\exp(t\,\mathrm{ad}(B))A_2 = \exp(t\,B)A_2\exp(-t\,B) = e^{-2bt}A_2.$$

Choosing $t = \frac{\pi}{2\mathrm{Im}(b)}$ we have

$$e^{2b\,t} = e^{\frac{\pi b}{\mathrm{Im}(b)}} = e^{\pi\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}}\left(\cos(\pi) + \mathbf{i}\sin(\pi)\right) = -e^{\pi\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}}$$

and

$$e^{-2b\,t} = e^{\frac{-\pi b}{\mathrm{Im}(b)}} = e^{-\pi\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}}\left(\cos(-\pi) + \mathbf{i}\sin(-\pi)\right) = -e^{-\pi\frac{\mathrm{Re}(b)}{\mathrm{Im}(b)}}.$$

This choice together with $t = 0$ shows us that $\mathrm{span}_{\mathbb{R}}\{A_1\} \subset \mathrm{LS}(\Gamma)$ and $\mathrm{span}_{\mathbb{R}}\{A_2\} \subset \mathrm{LS}(\Gamma)$, since $\mathrm{LS}(\Gamma)$ is a positive cone.

In the same fashion, choosing $t \in \mathbb{R}$ such that $\cos(t) = 0$ and $\sin(t) = \pm 1$ we get $\mathrm{span}_{\mathbb{R}}\{\mathbf{i}A_1\} \subset \mathrm{LS}(\Gamma)$ and $\mathrm{span}_{\mathbb{R}}\{\mathbf{i}A_2\} \subset \mathrm{LS}(\Gamma)$. And these considerations allow us to conclude that

$$\mathrm{span}_{\mathbb{C}}\{E_2\} \subset \mathrm{LS}(\Gamma) \quad \text{and} \quad \mathrm{span}_{\mathbb{C}}\{E_3\} \subset \mathrm{LS}(\Gamma).$$

Computing the Lie brackets between the basis elements E_2 and E_3 we get

$$\pm[E_2, E_3] = \mp E_1 \in \mathrm{LS}(\Gamma) \quad \text{and} \quad \pm[\mathbf{i}E_2, E_3] = \mp\mathbf{i}E_1 \in \mathrm{LS}(\Gamma),$$

which means that

$$\mathrm{span}_{\mathbb{C}}\{E_1\} \subset \mathrm{LS}(\Gamma).$$

Now, since $A \in \mathrm{LS}(\Gamma)$ and $\mathrm{span}_{\mathbb{C}}\{E_1, E_2, E_3\} \subset \mathrm{LS}(\Gamma)$ we obtain

$$a_1E_4 + a_2E_5 \in \mathrm{LS}(\Gamma),$$

and this yields

$$\exp(t\,\mathrm{ad}(B))(a_1E_4 + a_2E_5) \in \mathrm{LS}(\Gamma), \qquad \forall t \in \mathbb{R}.$$

An easy calculation shows that

$$\exp(t\,B)(a_1E_4 + a_2E_5)\exp(-t\,B) = \begin{bmatrix} 0 & 0 & a_1e^{bt} \\ 0 & 0 & a_2e^{-bt} \\ 0 & 0 & 0 \end{bmatrix}.$$

Choosing a sequence $(t_k)_{k\in\mathbb{Z}} \subset \mathbb{R}$ such that $\cos(t_k\mathrm{Im}(b)) = 1$ and $\sin(t_k\mathrm{Im}(b)) = 0$, we get $e^{\pm b\,t} \in \mathbb{R}$ in such way that

$$\lim_{t_k \mathrm{Re}(b) \to \infty} e^{-bt_k} \exp(t_k\, B)(a_1 E_4 + a_2 E_5) \exp(-t_k\, B) = a_1 E_4 \in \mathrm{LS}(\Gamma)$$

and

$$\lim_{t_k \mathrm{Re}(b) \to -\infty} e^{bt_k} \exp(t_k\, B)(a_1 E_4 + a_2 E_5) \exp(-t_k\, B) = a_2 E_5 \in \mathrm{LS}(\Gamma).$$

Again, we have

$$\exp(t\,\mathrm{ad}(B))(a_1 E_4) = e^{b\,t} a_1 E_4 \in \mathrm{LS}(\Gamma),$$

$$\exp(t\,\mathrm{ad}(B))(a_2 E_5) = e^{-b\,t} a_2 E_5 \in \mathrm{LS}(\Gamma).$$

Suitable choices of $t \in \mathbb{R}$ shows that

$$\pm a_1 E_4 \in \mathrm{LS}(\Gamma), \qquad \pm \mathbf{i} a_1 E_4 \in \mathrm{LS}(\Gamma),$$

$$\pm a_2 E_5 \in \mathrm{LS}(\Gamma) \quad \text{and} \quad \pm\, \mathbf{i} a_2 E_5 \in \mathrm{LS}(\Gamma),$$

which is equivalent to say that

$$\mathrm{span}_{\mathbb{C}}\{E_4\} \subset \mathrm{LS}(\Gamma) \quad \text{and} \quad \mathrm{span}_{\mathbb{C}}\{E_5\} \subset \mathrm{LS}(\Gamma),$$

proving that $\mathrm{LS}(\Gamma) = \mathfrak{sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$, that is, Γ is controllable on $\mathrm{Sl}(2,\mathbb{C}) \rtimes \mathbb{C}^2$. □

Chapter 4

Invariant control sets for some control systems induced by $\mathfrak{so}(1,4)$ on S^3

Now we explore invariant control sets for vector fields induced by $\mathrm{SO}(1,4)$ on the sphere S^3. For symmetric elements, the vector fields given by the infinitesimal action of $\mathfrak{so}(1,4)$ on S^3 are gradient vector fields of height functions, and elements in the compact component $\mathfrak{so}(4)$ give rise to vector fields defined by right and left multiplication by imaginary quaternions.

We provide a characterization for the invariant control sets on S^3 for control systems with $\mathbf{1} \in \mathbb{H}$ as a drift and control vector fields corresponding to pure quaternions. Such control sets appear as spherical domes in some cases, while in others, they are described as geodesically convex closures of the set of attractor points for the vector fields corresponding to the control system.

4.1 General Theorem

A subset C of a differentiable manifold M is said to be geodesically convex if the minimal geodesic segment joining $p, q \in C$ is contained in C. We define also the geodesic convex hull of a subset as the smaller geodesically convex subset $C \subset M$ containing it.

Let $C^\infty(M)$ be the set of all complete $\mathcal{C}^\infty$ vector fields on a differentiable manifold M. Given $X \in C^\infty(M)$, we say that a point $a \in M$ is an attractor fixed point for X in M if the vector field X vanishes in a and $\lim_{t\to\infty} X_t(x) = a$ for every initial state $x \in M$, where X_t denotes the flow corresponding to X. For simplicity we will refer to such points as just the attractors of X.

Theorem 4.1.1. *Let M be a differentiable manifold with an affine connection ∇ and $\Gamma \subset C^\infty(M)$ a family of vector fields on M such that every $X \in \Gamma$ satisfies $\nabla_X X = c_X X$, where $c_X : M \to \mathbb{R}$ is a smooth function on M. Assume that every $X \in \Gamma$ admits only one global attractor $a \in M$ and suppose also that a trajectory γ for X is minimal geodesic between all of its points and the attractor a. If*

$$E = \{a \in M \mid a \text{ is attractor for some } X \in \Gamma\}$$

is closed and geodesically convex, then E is the invariant control set for Γ.

Proof: The condition $\nabla_X X = c_X X$ on the vector fields of the system tells us that the trajectories for the vector fields $X \in \Gamma$ follow the geodesics on M, and in this case we call X a geodesic vector field.

Now, let D be the invariant control set for Γ. We know that

$$D = \bigcap_{x \in M} \mathrm{cl}(S_\Gamma \cdot x),$$

where S_Γ denotes the semigroup of the system Γ. Since every $y \in E$ is an attractor fixed point for some $X \in \Gamma$, we have $y \in \mathrm{cl}(S_\Gamma \cdot x)$, for every $x \in M$. In other words, $E \subseteq D$. Being D a control set, if $x \in D$ then $x \in \mathrm{cl}(S_\Gamma \cdot y)$, for every $y \in D$. In particular, $x \in \mathrm{cl}(S_\Gamma \cdot y)$ for every $y \in E$ (we just proved that $E \subseteq D$). Well, $x \in \mathrm{cl}(S_\Gamma \cdot y)$ implies that we can choose a trajectory γ for the control system Γ with starting point at y and endpoint x_T arbitrarily close from x. Let us say that the very last path γ_T of γ joins $x_{T'} \in E$ to x_T and corresponds to $X_0 \in \Gamma$ with attractor fixed point x_0. By the definition of E we have that $x_0 \in E$ and γ_T is a minimal geodesic segment joining $x_{T'}$ to x_0. Since A is geodesically convex, we must have $\gamma_T \subset E$, that is $x_T \in E$. That is, $x \in \mathrm{cl}E$, and this proves the equality $E = D$. Hence E is the invariant control set for Γ. □

The above theorem assumes the set of attractors E to be geodesically convex. However, there are examples where this convexity does not hold. In this work we present some examples, which suggest that if this occurs then the convex hull of E can be a good candidate for the invariant control set. Our first example will consider a control system evolving on the projective space $\mathbb{P}^{n-1}$ induced by the action of $\mathrm{Sl}(n, \mathbb{R})$.

Example 4.1.2. *Let*

$$v_0 = \begin{bmatrix} \frac{1}{\sqrt{n}} \\ w \end{bmatrix} \in \mathbb{R}^n \text{ with } \frac{1}{n} + |w|^2 = 1$$

and

$$B_i = \begin{bmatrix} 0 & 0 \\ 0 & X_i \end{bmatrix} \in \mathfrak{sl}(n, \mathbb{R}), \quad i = 1, \ldots, m = \dim(\mathfrak{so}(n-1)),$$

where each diagonal block X_i *has order* $n-1$ *and* $\{X_1, \ldots, X_m\}$ *is a basis for* $\mathfrak{so}(n-1)$*. Setting* $A = v_0 v_0^t - \frac{1}{n}\mathrm{Id_n}$*, we will describe the set of attractors* E *and the invariant control set* C *for the control system*

$$\dot{x} = Ax + u_1 B_1 x + \cdots + u_m B_m x, \quad x \in \mathbb{P}^{n-1}. \tag{4.1}$$

Before anything else, note that (4.1) satisfies the Lie algebra rank condition. In fact, to verify this let ad *be the adjoint representation of*

$$\mathfrak{so}(n-1) = \begin{bmatrix} 0 & 0 \\ 0 & \mathfrak{so}(n-1) \end{bmatrix}$$

on $\mathfrak{s} \subset \mathfrak{sl}(n, \mathbb{R})$*, the subspace of symmetric matrices. The restriction of* ad *to the subspace*

$$W = \left\{ \begin{bmatrix} 0 & w^t \\ w & 0 \end{bmatrix}; w \in \mathbb{R}^{n-1} \right\}$$

of $\mathfrak{sl}(n, \mathbb{R})$ *is an irreducible representation of* ad*. In fact, if*

$$B = \begin{bmatrix} 0 & 0 \\ 0 & X \end{bmatrix} \in \mathfrak{so}(n-1) \text{ and } \tilde{w} = \begin{bmatrix} 0 & w^t \\ w & 0 \end{bmatrix}$$

we have

$$[B, \tilde{w}] = \begin{bmatrix} 0 & (Xw)^t \\ Xw & 0 \end{bmatrix}.$$

Similarly, the restriction of ad *to*

$$\mathfrak{s}_0 = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & S \end{bmatrix}; \; S \text{ traceless symmetric} \right\}$$

is irreducible.

Now, we can decompose $A = v_0 v_0^t - \frac{1}{n}\mathrm{Id}_n$ *as*

$$A = \frac{1}{\sqrt{n}} \begin{bmatrix} 0 & w^t \\ w & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & ww^t - \frac{1}{n}\mathrm{Id}_{n-1} \end{bmatrix} =: \frac{1}{\sqrt{n}}\tilde{w} + S_0,$$

and since $w \neq 0$ *we obtain that the subspaces generated by*

$$\mathrm{ad}\left(\mathfrak{so}(n-1)\right)\left(\frac{1}{\sqrt{n}}\tilde{w}\right) \text{ and } \mathrm{ad}\left(\mathfrak{so}(n-1)\right)(S_0)$$

are W *and* $\mathfrak{s}_0$*, respectively. Finally, the skew-symmetric matrices of the form*

$$\begin{bmatrix} 0 & -w^t \\ w & 0 \end{bmatrix}, \quad w \in \mathbb{R}^{n-1},$$

are obtained as Lie brackets between suitable matrices

$$\begin{bmatrix} 0 & u^t \\ u & 0 \end{bmatrix} \in \mathfrak{s}_0 \text{ and } \begin{bmatrix} 0 & 0 \\ 0 & X \end{bmatrix} \in \mathfrak{so}(n-1),$$

proving that (4.1) *satisfies LARC, and this implies that the invariant control set for this system has nonempty interior.*

Now we describe the system (4.1) in a more detailed way. For, let $\mathrm{Gl}(n,\mathbb{R})$ *acting on the projective space* $\mathbb{P}^{n-1}$ *under* $g * [x] := [gx]$*, where* $g \in \mathrm{Gl}(n,\mathbb{R})$ *and* $[x] \in \mathbb{P}^{n-1}$ *indicates the straight line containing* 0 *and* $x \in S^{n-1}$*. We have that*

$$g * x = \left[\frac{gx}{|gx|}\right],$$

and given any $X \in \mathfrak{gl}(n,\mathbb{R})$ *the corresponding infinitesimal action on* $\mathbb{P}^{n-1}$ *is*

$$\tilde{X}[x] = \frac{d}{dt}\left(e^{tX} * x\right)\Big|_{t=0} = \frac{d}{dt}\left(\frac{e^{tX}x}{|e^{tX}x|}\right)\Bigg|_{t=0} = [Xx - \langle Xx, x\rangle x].$$

For instance, if $X = \mathrm{diag}(1,0,\dots,0) \in \mathfrak{gl}(n,\mathbb{R})$ *and* $x = (x_1,\dots,x_n) \in S^{n-1}$*, then* $Xx = (x_1,0,\dots,0)$ *is the projection of* x *along* $e_1 = (1,0,\dots,0)$ *and* $\tilde{X}[x]$ *is given by the orthogonal projection of* Xx *on the one dimensional subspace generated by* x*. We have that* $\tilde{X}[e_1] = 0$ *and if* x *belongs to the orthogonal complement of* e_1 *we also have* $\tilde{X}[x] = 0$*. The trajectories for* $\tilde{X}$ *follow the great circles and* $[e_1]$ *is the only attractor fixed point for* $\tilde{X}$ *while* *the points in the orthogonal complement of* e_1 *are the repeller fixed points for* $\tilde{X}$*.*

In this way, if $\Gamma \subset \mathbb{P}^{n-1}$ *is a closed and geodesically convex set, for every* $[x] \in \Gamma$ *corresponds a vector field* $\tilde{X}_x$ *where* X_x *indicates the projection over the line generated by* x *in* $\mathbb{R}^n$*. If* $\tilde{\Gamma}$ *is the family of such vector fields, then the previous theorem tells us that its invariant control system is exactly* Γ*.*

Now, let $\mathfrak{s} \subset \mathfrak{sl}(n, \mathbb{R})$ *be the subspace of symmetric matrices, and consider the embedding of the projective space* $\mathbb{P}^{n-1}$ *in* $\mathfrak{s}$ *which associates to each* $[v] \in \mathbb{P}^{n-1}$*,* $|v| = 1$*, the symmetric* $n \times n$ *trace free matrix* V *given by*

$$V = vv^t - \frac{1}{n}\mathrm{Id}_n.$$

Given $[v], [w] \in \mathbb{P}^{n-1}$*, with respect to the canonical inner product in* $\mathfrak{s}$*, we have*

$$\begin{aligned}
\langle V, W\rangle &= \mathrm{tr}VW \\
&= \mathrm{tr}\left(vv^tww^t - \frac{1}{n}vv^t - \frac{1}{n}ww^t + \frac{1}{n^2}\mathrm{Id}_n\right) \\
&= \mathrm{tr}\left(\langle v, w\rangle vw^t - \frac{1}{n}vv^t - \frac{1}{n}ww^t + \frac{1}{n^2}\mathrm{Id}_n\right)
\end{aligned}$$

since $\mathrm{tr}(vv^t) = |v|$ *and* $\mathrm{tr}(vw^t) = \langle v, w\rangle$*, we get*

$$\langle V, W\rangle = \langle v, w\rangle^2 - \frac{1}{n} - \frac{1}{n} + \frac{n}{n^2} = \langle v, w\rangle^2 - \frac{1}{n}.$$

For the symmetric matrix $E^0_{11} = e_1e_1^t - \frac{1}{n}\mathrm{Id}_n$*, let* C *be the convex set in* $\mathbb{P}^{n-1}$ *defined by*

$$C = \left\{[v] \in \mathbb{P}^{n-1} \,|\, \langle V, E^0_{11}\rangle \geq 0\right\}.$$

We have

$$C = \left\{[v] \in \mathbb{P}^{n-1} \,|\, \langle v, e_1\rangle^2 - \frac{1}{n} \geq 0\right\} = \left\{[v] \in \mathbb{P}^{n-1} \,|\, \langle v, e_1\rangle \geq \frac{1}{\sqrt{n}}\right\},$$

and if θ *is the angle between* v *and* e_1*, then* C *is the set of* $[v] \in \mathbb{P}^{n-1}$ *such that* $\cos(\theta) \geq 1/\sqrt{n}$*, that is, the spherical dome in* $\mathbb{P}^{n-1}$ *of the elements* $[v]$ *whose* $v_1 \geq 1/\sqrt{n}$*, and* C *can be seen as the intersection of an halfspace in* $\mathfrak{s}$ *with* $\mathbb{P}^{n-1}$ *(using the embedding previously defined).*

Returning to system (4.1), note that $[v_0] \in \mathbb{P}^{n-1}$ *is the attractor fixed point for the vector field on* $\mathbb{P}^{n-1}$ *induced by* $A = v_0 v_0^t - \frac{1}{n}\mathrm{Id}$. *In fact,* $|v_0| = 1$ *and*

$$\begin{aligned}
Av_0 - \langle Av_0, v_0\rangle v_0 &= \left(v_0v_0^t - \frac{1}{n}\mathrm{Id}_n\right) v_0 - \left\langle \left(v_0v_0^t - \frac{1}{n}\mathrm{Id}_n\right) v_0, v_0 \right\rangle v_0 \\
&= v_0v_0^tv_0 - \frac{1}{n}v_0 - \left\langle v_0v_0^tv_0 - \frac{1}{n}v_0, v_0 \right\rangle v_0 \\
&= |v_0|v_0 - \frac{1}{n}v_0 - \left(|v_0|^3 - \frac{1}{n}|v_0|^2\right) v_0 \\
&= 0
\end{aligned}$$

Also, the spherical dome C *is invariant under each* B_i, *since*

$$e^{tB_i} = \begin{bmatrix} 1 & 0 \\ 0 & e^{tX_i} \end{bmatrix}$$

and

$$\left\langle e^{tB_i}v, e^{tB_i}e_1 \right\rangle = \left\langle e^{tB_i}v, e_1 \right\rangle \geq \frac{1}{\sqrt{n}},$$

for every $v \in \mathbb{R}^n$ *such that* $\langle v, e_1\rangle \geq 1/\sqrt{n}$.

Being $\{X_1, \ldots, X_m\}$ *a basis for* $\mathfrak{so}(n-1)$, *choosing adequate controls we see that the attractors for the vector fields induced by* $A + u_1B_1 + \cdots + u_mB_m$ *in* $\mathbb{P}^{n-1}$ *just rotate around the boundary of* C, *meaning that*

$$E = \left\{ [v] \in \mathbb{P}^{n-1} \,|\, \langle v, e_1\rangle = \frac{1}{\sqrt{n}} \right\}$$

is fulfilled with attractors for system (4.1), that is, C *is invariant under* (4.1), *and hence the invariant control set for this system must be contained in* C. *Furthermore, if* $\mathcal{O}^+(x)$ *stands for the positive orbit of* $x \in \mathbb{P}^{n-1}$, *then for every* $y \in E$ *we have*

$$y \in \bigcap_{x\in\mathbb{P}^{n-1}} \mathrm{cl}\mathcal{O}^+(X),$$

ensuring that C *is the invariant control set for* (4.1).

4.2 Lie-theoretic structure of $\mathrm{SO}(1,4)$

The purpose of this section is to present the concepts and structures necessary to describe the control systems for which we will study the invariant control sets. As it is difficult to find bibliographical references that deal with Lie theory in the special case of $\mathrm{SO}(1,4)$, then we will provide here the necessary theory of $\mathrm{SO}(1,4)$ for the purpose of this work, for example, the description of the Cartan decomposition, the flag manifold and the vector fields on this manifold.

4.2.1 Cartan decomposition

The quadratic form with matrix

$$I_{p,q} = \begin{bmatrix} 1_p & 0 \\ 0 & -1_q \end{bmatrix}$$

gives rise to the indefinite special orthogonal Lie algebra of type (p,q). That is to say, the Lie algebra $\mathfrak{so}(p,q)$ consists of the order $(p+q)$ real matrices that are skew-symmetric with respect to $I_{p,q}$, that is,

$$\mathfrak{so}(p,q) = \{X \in \mathfrak{sl}(p+q,\mathbb{R}) \mid I_{p,q}X + X^t I_{p,q} = 0\}.$$

The skew-symmetry with respect to any other quadratic form equivalent to $I_{p,q}$ defines a Lie algebra isomorphic to $\mathfrak{so}(p,q)$. We are interested in the Lie algebra $\mathfrak{so}(1,4)$, and in this case the condition $I_{1,4}X + X^t I_{1,4} = 0$ implies that the matrix X (which is an order 5 matrix) must have the block form

$$X = \begin{bmatrix} 0 & \beta \\ \beta^t & \gamma \end{bmatrix}, \text{ with } \gamma = -\gamma^t, \text{ that is, } \gamma \in \mathfrak{so}(4) = \mathfrak{su}(2) \oplus \mathfrak{su}(2).$$

The corresponding compact real form in this case is the Lie algebra $\mathfrak{so}(5)$ and the associate conjugation in $\mathfrak{so}(5,\mathbb{C})$ is $\theta(X) = -X^t$ (note that it is an involutive automorphism for this Lie algebra). The bilinear form B_θ in $\mathfrak{so}(1,4)$ given by $B_\theta(X,Y) = -\langle X, \theta Y\rangle$ is an inner product in $\mathfrak{so}(1,4)$, as follows from [26] (Lemma 12.21). This implies that

$\mathfrak{so}(1,4)$ decomposes as the direct sum of the eigenspaces

$$V_1 = \{X \in \mathfrak{so}(1,4) \,|\, \theta(X) = X\} \text{ and } V_{-1} = \{X \in \mathfrak{so}(1,4) \,|\, \theta(X) = -X\}.$$

Setting $\mathfrak{k} = V_1$ and $\mathfrak{s} = V_{-1}$, then $\mathfrak{so}(1,4) = \mathfrak{k} \oplus \mathfrak{s}$ is a Cartan decomposition. Explicitly, we have

$$\mathfrak{k} = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & \gamma \end{bmatrix} \in \mathfrak{so}(1,4) \middle| \gamma \in \mathfrak{so}(4) \right\} = \mathfrak{so}(4)$$

and

$$\mathfrak{s} = \left\{ \begin{bmatrix} 0 & \beta \\ \beta^t & 0 \end{bmatrix} \in \mathfrak{so}(1,4) \middle| \beta \in \mathbb{R}^4 \right\}.$$

Also, $\mathfrak{u} = \mathfrak{k} \oplus \mathbf{i}\mathfrak{s}$ is a compact real form for $\mathfrak{so}(5,\mathbb{C})$ (see [26] Proposition 12.27). It is given by

$$\mathfrak{u} = \left\{ \begin{bmatrix} 0 & \mathbf{i}\beta \\ \mathbf{i}\beta^t & \gamma \end{bmatrix} \middle| \beta \in \mathbb{R}^4, \ \gamma \in \mathfrak{so}(4) \right\}.$$

The application $\mathfrak{u} \to \mathfrak{so}(5)$ defined by

$$\begin{bmatrix} 0 & \mathbf{i}\beta \\ \mathbf{i}\beta^t & \gamma \end{bmatrix} \mapsto \begin{bmatrix} 0 & \beta \\ -\beta^t & \gamma \end{bmatrix}$$

is a Lie isomorphism between $\mathfrak{u}$ and $\mathfrak{so}(5)$. A maximal abelian subalgebra $\mathfrak{a}$ contained in $\mathfrak{s}$ is generated by the matrices

$$\begin{bmatrix} 0 & \beta \\ \beta^t & 0 \end{bmatrix}, \quad \text{where } \beta = [1,0,0,0],$$

and thus $\mathfrak{so}(1,4)$ is a real rank 1 non-compact real form of $\mathfrak{so}(5,\mathbb{C})$. When realizing $\mathfrak{so}(1,4)$ by the quadratic form given by

$$J_{1,n} = \begin{bmatrix} -1_3 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

we get a simpler way to compute the restricted roots with respect to a maximal abelian subalgebra in $\mathfrak{s}$. In this realization $X \in \mathfrak{so}(1,4)$ if and only if X has the form

$$\begin{bmatrix} A & B & C \\ C^t & \alpha & 0 \\ B^t & 0 & -\alpha \end{bmatrix}, \quad \text{where } A \in \mathfrak{so}(3).$$

A Cartan decomposition is given by the symmetric and the anti-symmetric matrices of this type.

Precisely,

$$\mathfrak{k} = \left\{ \begin{bmatrix} A & B & -B \\ -B^t & 0 & 0 \\ B^t & 0 & 0 \end{bmatrix} \middle| \, A \in \mathfrak{so}(3),\ B^t \in \mathbb{R}^3 \right\}$$

and

$$\mathfrak{s} = \left\{ \begin{bmatrix} 0 & B & B \\ B^t & \alpha & 0 \\ B^t & 0 & -\alpha \end{bmatrix} \middle| \, B^t \in \mathbb{R}^3,\ \alpha \in \mathbb{R} \right\}.$$

In this realization, a maximal abelian subalgebra contained in $\mathfrak{s}$ is

$$\mathfrak{a} = \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & -\alpha \end{bmatrix} \middle| \, \alpha \in \mathbb{R} \right\}.$$

And the restricted roots are the functionals $\pm\lambda : \mathfrak{a} \to \mathbb{R}$, $\lambda(H) = \alpha$, where $H = \mathrm{diag}(0, \alpha, -\alpha) \in \mathfrak{a}$. The only simple root is λ and it has multiplicity 3. The corresponding root spaces are

$$\begin{aligned} \mathfrak{g}_\lambda &= \{X \in \mathfrak{so}(1,4) \,|\, \mathrm{ad}(H)X = \lambda(X),\ \forall H \in \mathfrak{a}\} \\ &= \left\{ \begin{bmatrix} 0 & 0 & C \\ C^t & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{so}(1,4) \,\middle|\, C^t \in \mathbb{R}^3 \right\}, \end{aligned}$$

similarly we have

$$\mathfrak{g}_{-\lambda} = \left\{ \begin{bmatrix} 0 & B & 0 \\ 0 & 0 & 0 \\ B^t & 0 & 0 \end{bmatrix} \in \mathfrak{so}(1,4) \;\middle|\; B^t \in \mathbb{R}^3 \right\}$$

and

$$\mathfrak{g}_0 = \left\{ \begin{bmatrix} A & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & -\alpha \end{bmatrix} \in \mathfrak{so}(1,4) \;\middle|\; A \in \mathfrak{so}(3),\ \alpha \in \mathbb{R} \right\}.$$

4.2.2 The flag manifold of $\mathrm{SO}(1,4)$

By Proposition 1.3.1 we have that $\mathbb{F}_\Theta = K/K_\Theta$, where the stabilizer is given by $K_\Theta = K \cap P_\Theta$.

Since $\mathfrak{so}(1,4)$ is a rank one real Lie algebra, there is only one flag manifold for $\mathrm{SO}(1,4)$, namely the maximal one. Note that we can identify $\mathbb{F} = {}^{SO(1,4)}/_{P}$ as a K-orbit under the adjoint representation, that is, $\mathrm{Ad}(K)H = {}^{K}/_{K_H}$, where $H \in \mathrm{cl}\mathfrak{a}^+$ and K_H is the centralizer of H in K.

Proposition 4.2.1. *The only flag manifold $\mathbb{F}$ of $SO(1,4)$ embeds in the component $\mathfrak{s}$ of the Cartan decomposition as the $\mathrm{Ad}(SO(4))$-orbit of $H \in \mathfrak{a}^+$.*

Proof: The component $\mathfrak{s}$ of the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s} = \mathfrak{so}(4) \oplus \mathfrak{s}$ is invariant under the adjoint representation of $K = SO(4)$. Since $\mathfrak{a}$ is an one dimensional subalgebra, a Weyl chamber $\mathfrak{a}^+ \subset \mathfrak{a}$ is just a ray starting at the origin. Choosing an element $H \in \mathfrak{a}^+$, the stabilizer of H under the adjoint action of K on $\mathfrak{s}$ is the centralizer K_H of H in K, which is given by $K \cap P$. It follows that the adjoint orbit $\mathrm{Ad}(SO(4))H \subset \mathfrak{s}$ is identified with the coset space K/K_H, and this one is the flag manifold $\mathbb{F} = G/P$. □

Proposition 4.2.2. *The sphere S^3 is the only flag manifold of* $\mathrm{SO}(1,4)$.

Proof: We look at $\mathfrak{so}(1,4)$ realized by $I_{1,4}$, where the maximal abelian subalgebra $\mathfrak{a}$

contained in $\mathfrak{s}$ is of the form

$$\mathfrak{a} = \operatorname{span}\left\{ \begin{bmatrix} 0 & e_1 \\ e_1^t & 0 \end{bmatrix}, \quad e_1 = (1,0,0) \right\}.$$

Let

$$H = \begin{bmatrix} 0 & e_1 \\ e_1^t & 0 \end{bmatrix} \in \mathfrak{a}^+ \quad \text{and} \quad x = \begin{bmatrix} 1 & 0 \\ 0 & \beta \end{bmatrix} \in K, \text{ with } \beta \in \mathrm{SO}(4).$$

Denote by β_{1i} and by β_{i1} the first row and the first column of β respectively. With this notation we have

$$xH = Hx \Leftrightarrow \begin{bmatrix} 0 & e_1 \\ \beta_{i1} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \beta_{1i} \\ e_1^t & 0 \end{bmatrix},$$

from which $\beta_{1i} = e_1$ and $\beta_{i1} = e_1^t$. This means that β is of the form

$$\beta = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}, \quad \gamma \in \mathrm{SO}(3).$$

Hence, $x \in K_H$ if and only if x has the block form

$$x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \gamma \end{bmatrix}, \quad \gamma \in \mathrm{SO}(3).$$

This shows that $K_H = \mathrm{SO}(3)$, and we get $\mathbb{F} = {}^{K}/_{K_H} = {}^{\mathrm{SO}(4)}/_{\mathrm{SO}(3)} = S^3$. □

The previous computations remain basically unchanged for $\mathrm{SO}(1,n)$, and in this case we have $\mathbb{F} = {}^{\mathrm{SO}(n)}/_{\mathrm{SO}(n-1)} = S^{n-1}$.

4.2.3 The infinitesimal action of $\mathfrak{so}(1,4)$ on S^3

Now we describe the vector fields induced on the sphere S^3 by the action of $\mathrm{SO}(1,4)$. Remember that an infinitesimal action of $\mathfrak{so}(1,4)$ on S^3 is a homomorphism $\mathfrak{so}(1,4) \to \Gamma(TS^3)$, where $\Gamma(TS^3)$ stands for the Lie algebra of vector fields on S^3. By means of an infinitesimal action of $\mathfrak{so}(1,4)$ on S^3 one can see the Lie algebra $\mathfrak{so}(1,4)$ as a Lie algebra of vector fields on the sphere S^3.

Theorem 4.2.3. *The infinitesimal action of* $\mathfrak{so}(1,4)$ *induced by the action of* $\mathrm{SO}(1,4)$ *on* S^3 *has as image the vector space formed by the vector fields*

$$X_{(q,z,w)}(x) = \frac{1}{2}(q - x\overline{q}x) + zx + xw, \qquad x \in S^3,$$

where $q \in \mathbb{H} = \mathfrak{s}$ *and* $z, w \in \mathrm{Im}\mathbb{H} = \mathfrak{su}(2)$. *This vector space is a* 10*-dimensional Lie algebra isomorphic to* $\mathfrak{so}(1,4)$.

Proof: We begin by investigating the vector fields corresponding to elements belonging to the $\mathfrak{s}$ component. There exists a K-invariant Riemannian metric such that for every $q \in \mathbb{H} = \mathfrak{s}$ the vector field $\tilde{X}_q$ induced by q on S^3 is the gradient of the height function $f_q(\cdot) = \langle q, \cdot\rangle$ with respect to this K-invariant metric (see [7] and [34] for details). In the present case, since $\alpha(H) = 1$ ($H \in \mathfrak{a}^+$) for every positive root α for which $\alpha(H) \neq 0$ we have that the Borel metric coincides with the metric induced by the immersion of S^3 in $\mathfrak{s}$ (by this reason S^3 is called an immersed flag manifold). The height function f_q is linear on $\mathfrak{s}$, so its gradient vector field evaluated at $p \in S^3$ is obtained from the orthogonal projection of q over p. In fact, $(\mathrm{grad} f_q)_p = d(f_q)_p(v)$ is the cotangent vector ω such that $\omega(v) = \langle q, v\rangle$, that is, the cotangent vector ω such that

$$\langle \omega, v\rangle = \langle q, v\rangle, \ \ \forall v \in T_pS^3.$$

Since

$$\langle q - \langle q, p\rangle p, v\rangle = \langle q, v\rangle \ \ \forall v \in T_pS^3,$$

we get

$$(\mathrm{grad} f_q)_p = q - \langle q, p\rangle p.$$

The vector field $\tilde{X}_q$ is thus given by

$$\begin{aligned} \tilde{X}_q(p) &= q - \langle q, p\rangle p \\ &= q - \frac{1}{2}\left(q\overline{p} + p\overline{q}\right) p \\ &= q - \frac{1}{2}(q\overline{p}p + p\overline{q}p), \qquad p \in S^3. \end{aligned}$$

Since $|p| = p\overline{p} = 1$, we get

$$\tilde{X}_q(p) = \frac{1}{2}(q - p\overline{q}p).$$

Now we turn our attention to the elements in the compact component $\mathfrak{k}$. We know that the adjoint representation

$$\begin{aligned} \mathrm{Ad}: K &\to \mathrm{Gl}(\mathfrak{so}(1,4)) \\ k &\mapsto \mathrm{Ad}(k) = d(C_k)_1. \end{aligned}$$

is differentiable and defines the action

$$K \times \mathfrak{s} \to \mathfrak{s}, \quad (k,x) \mapsto \mathrm{Ad}(k)x.$$

Since the flag S^3 embeds in $\mathfrak{s}$ as an $\mathrm{Ad}(K)$-orbit, we can consider the restriction of the above action to S^3 (viewed as an $\mathrm{Ad}(K)$-orbit) to get an infinitesimal action of $\mathfrak{k}$ on S^3. Thus, for $X \in \mathfrak{k}$ the corresponding infinitesimal action on S^3 is given by

$$\tilde{X}(x) = d(\mathrm{Ad})_1(X)x = \mathrm{ad}(X)x,$$

that is, the induced vector field is given by the adjoint action of $\mathfrak{k}$ on $\mathfrak{s}$.

Now, as the Lie algebra $\mathrm{Im}\mathbb{H}$ is represented in $\mathbb{H}$ through left multiplication and also through right multiplication we have that the Lie algebra $\mathfrak{so}(4) = \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ is isomorphic to the Lie algebra of linear transformations $\{E_z + D_w \mid z, w \in \mathrm{Im}\mathbb{H}\}$.

As the Lie algebra $\mathfrak{so}(4)$ decomposes as the sum of two simple ideals commuting to each other, we describe the adjoint representation of $\mathfrak{k} = \mathfrak{so}(4)$ on $\mathfrak{s}$ looking at each simple ideal. The first component corresponds to left multiplication by immaginary quaternions and the second one corresponds to right multiplication by immaginary quaternions, since these two kinds of quaternionic multiplications commute with each other as we have already seen.

We can check this as follows. Given $\beta = (p,q,r,s) \in \mathbb{R}^4$, we identify an element $S \in \mathfrak{s}$ as a quaternion number in the following way

$$S = \begin{bmatrix} 0 & \beta \\ \beta^t & 0 \end{bmatrix} \in \mathfrak{s} \Leftrightarrow S = p + q\mathbf{i} + r\mathbf{j} + s\mathbf{k} \in \mathbb{H}. \tag{4.2}$$

Set
$$A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } C_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$
With this notation, define the $\mathfrak{so}(4)$ matrices

$$\gamma_{\mathbf{i}} = \begin{bmatrix} A_2 & 0 \\ 0 & -A_2 \end{bmatrix}, \quad {}_{\mathbf{i}}\gamma = \begin{bmatrix} A_2 & 0 \\ 0 & A_2 \end{bmatrix},$$

$$\gamma_{\mathbf{j}} = \begin{bmatrix} 0 & -1_2 \\ 1_2 & 0 \end{bmatrix}, \quad {}_{\mathbf{j}}\gamma = \begin{bmatrix} 0 & C_2 \\ -C_2 & 0 \end{bmatrix},$$

$$\gamma_{\mathbf{k}} = \begin{bmatrix} 0 & A_2 \\ A_2 & 0 \end{bmatrix}, \quad {}_{\mathbf{k}}\gamma = \begin{bmatrix} 0 & -B_2 \\ B_2 & 0 \end{bmatrix}.$$

If we write
$$X_{\mathbf{i}} = \begin{bmatrix} 0 & 0 \\ 0 & \gamma_{\mathbf{i}} \end{bmatrix}, \quad {}_{\mathbf{i}}X = \begin{bmatrix} 0 & 0 \\ 0 & {}_{\mathbf{i}}\gamma \end{bmatrix},$$
and so on, we get the following relations:

$$[X_{\mathbf{i}}, S] = S\mathbf{i} \in \mathbb{H}, \;\; [X_{\mathbf{j}}, S] = S\mathbf{j} \in \mathbb{H}, \;\; [X_{\mathbf{k}}, S] = S\mathbf{k} \in \mathbb{H},$$

$$[{}_{\mathbf{i}}X, S] = \mathbf{i}S \in \mathbb{H}, \;\; [{}_{\mathbf{j}}X, S] = \mathbf{j}S \in \mathbb{H}, \;\; [{}_{\mathbf{k}}X, S] = \mathbf{k}S \in \mathbb{H},$$

where $S\mathbf{i}$, $S\mathbf{j}$, $S\mathbf{k}$, $\mathbf{i}S$, $\mathbf{j}S$ and $\mathbf{k}S$ are well defined from (4.2). Finally, we have that $\langle \gamma_{\mathbf{i}}, \gamma_{\mathbf{j}}, \gamma_{\mathbf{k}} \rangle$ and $\langle {}_{\mathbf{i}}\gamma, {}_{\mathbf{j}}\gamma, {}_{\mathbf{k}}\gamma \rangle$ are ideals of $\mathfrak{so}(4)$, each one isomorphic to $\mathfrak{su}(2)$. Thus, the adjoint representation of these components corresponds to right and left multiplication by immaginary quaternions, as claimed. □

Note that the construction in the above proof gives us a direct way to verify the isomorphism $\mathfrak{so}(4) = \mathfrak{su}(2) \oplus \mathfrak{su}(2)$.

In the remaining of this section we study the gradient vector fields $X_{(q,0,0)}$. To get a description about the behavior of these vector fields, we start by studying the singularities of the vector fields $X_{(q,z,w)}$ on the sphere S^3. Remember that we can consider S^3

as the Lie group $\mathrm{SU}(2)$. So, given a vector field on $\mathrm{SU}(2)$ we define the function

$$F : \mathrm{SU}(2) \longrightarrow \mathfrak{su}(2)$$

by setting

$$F(p) = d(D_{p^{-1}})_p(X(p)) = X(p) \cdot p^{-1}$$

that completely determines the vector field X.

Note that a point $p \in \mathrm{SU}(2)$ is a singular point for X if and only if $F(p) = 0$. Thus the set of singular points of the vector field X is $F^{-1}\{0\}$ and X has no singular points if and only if $F^{-1}\{0\} = \emptyset$, that is, 0 does not belong to the image of the function F.

Now we prove the following lemma.

Lemma 4.2.4.

(i) The vector field X is right-invariant if and only if its corresponding function F is constant. More precisely, $F(p) = X(1)$ for all $p \in \mathrm{SU}(2)$.

(ii) The vector field X is left-invariant if and only if $F(p) = \mathrm{Ad}(p)X(1)$.

Proof: In fact, for the item (i), suppose that X is a right-invariant vector field, that is,

$$d(D_g)_h(X(h)) = X(hg), \quad \forall g, h \in \mathrm{SU}(2).$$

This implies that

$$F(p) = d(D_{p^{-1}})_p(X(p)) = X(pp^{-1}) = X(1) \in \mathrm{SU}(2).$$

By the other hand, if $F(p) = X(1)$ for all $p \in \mathrm{SU}(2)$, we get

$$F(p) = X(p) \cdot p^{-1} = X(1) \Longleftrightarrow X(p) = X(1) \cdot p,$$

that is, X is right-invariant.

For item (ii), suppose first that X is left-invariant, that is,

$$d(E_g)_h(X(h)) = X(gh), \quad \forall g, h \in \mathrm{SU}(2).$$

Note that

$$\mathrm{Ad}(p)X(1) = d(D_{p^{-1}})_p \circ d(E_p)_1(X(1)) = d(D_{p^{-1}})_p(X(p)) = F(p).$$

Reciprocally, if $F(p) = \mathrm{Ad}(p)X(1)$, we have

$$F(p) = X(p)\cdot p^{-1} = \mathrm{Ad}(p)X(1) = p\cdot X(1)\cdot p^{-1} \Longleftrightarrow X(p) = p\cdot X(1),$$

proving that X is left-invariant and concluding the proof. □

Now we take a look at the functions corresponding to the vector fields given in the previous theorem. A vector field

$$X_{(q,z,w)}(p) = \frac{1}{2}(q - p\overline{q}p) + zp + pw, \quad p \in \mathrm{SU}(2),$$

can be written as

$$X_{(q,z,w)} = X_{(q,0,0)} + X_{(0,z,0)} + X_{(0,0,w)}$$

and the corresponding function is given by $F_{(q,z,w)} = F_{(q,0,0)} + F_{(0,z,0)} + F_{(0,0,w)}$, since it is defined by the differential of the right translation by p^{-1}, which is linear. The functions $F_{(q,0,0)}$, $F_{(0,z,0)}$ and $F_{(0,0,w)}$ are given by

1. $X_{(0,z,0)}(p) = zp$ implies that $F_{(0,z,0)}(p) = X_{(0,z,0)}(p)\cdot p^{-1} = zpp^{-1} = z$, in other words, $F_{(0,z,0)}$ is constant. Note that this agrees with the item (i) above, since $X_{(0,z,0)}$ is a right-invariant vector field.

2. Since $X_{(0,0,w)} = pw$ we have $F_{(0,0,w)}(p) = pwp^{-1} = \mathrm{Ad}(p)X(1)$. And again it agrees with item (ii) above, because $X_{(0,0,w)}$ is left-invariant. Furthermore, as $p^{-1} = \overline{p}$, we get $F_{(0,0,w)}(p) = pw\overline{p}$.

3. Finally, for the vector field $X_{(q,0,0)} = \frac{1}{2}(q - p\overline{q}p)$ we have

$$F_{(q,0,0)}(p) = X_{(q,0,0)}(p)\cdot p^{-1} = \frac{1}{2}(q - p\overline{q}p)\overline{p},$$

and so

$$F_{(q,0,0)}(p) = \frac{1}{2}(q\overline{p} - p\overline{q}p\overline{p}) = \frac{1}{2}(q\overline{p} - p\overline{q}) = \mathrm{Im}q\overline{p} = \mathrm{Im}p\overline{q}.$$

It turns out that $F_{(q,0,0)}(p) = \mathrm{Im} p\overline{q} = 0$ if and only if $p\overline{q} = x \in \mathbb{R}$, that is, $p|q|^2 = xq$, which means that the only singularities for $X_{(q,0,0)}$ are the antipodal points $p = \pm\frac{q}{|q|}$.

In the sequel we deal with the image of $F_{(1,0,0)}$. Since $F_{(1,0,0)}(p) = \mathrm{Im}(p)$, we have that $F_{(1,0,0)}(S^3)$ is exactly the unit ball in $\mathrm{Im}\mathbb{H}$. To get a more detailed description of this image we look at the great circles passing by the elements $\pm\mathbf{1} \in S^3$. The great circle C_z is just the intersection of the plane generated by $\{\mathbf{1}, z\}$ with S^3 (here z is an purely immaginary unit quaternion). The computations with $C_{\mathbf{i}}$ give us a good idea of how the general case works. Let $C_{\mathbf{i}}^+$ be the semicircle $C_{\mathbf{i}}^+ = \{p \in C_{\mathbf{i}} \,|\, \langle p, i\rangle \geq 0\}$. As p runs through $C_{\mathbf{i}}^+$ from $\mathbf{1}$ to $-\mathbf{1}$ the values $F_{(1,0,0)}(p)$ cover twice the line segment $[0, \mathbf{i}] = \{t\mathbf{i} \,|\, t \in [0, 1]\}$. In fact, for p going from $\mathbf{1}$ to $\mathbf{i}$ we have $\mathrm{Re}(p) \geq 0$ and $\mathrm{Im}(p)$ goes from 0 to $\mathbf{i}$. When p goes from $\mathbf{i}$ to $-\mathbf{1}$ we have $\mathrm{Re}(p) \leq 0$ and $\mathrm{Im}(p)$ goes from $\mathbf{i}$ to 0. Briefly, as p goes from $\mathbf{1}$ to $-\mathbf{1}$ on $C_{\mathbf{i}}^+$ the image $F_{(1,0,0)}(p)$ goes from 0 to $\mathbf{i}$ and then goes back from $\mathbf{i}$ to 0, always lying on the line segment $[0, \mathbf{i}]$. On the other half $C_{\mathbf{i}}^-$ the situation is quite analogous, in this case however the image is the line segment $[-\mathbf{i}, 0] = \{-t\mathbf{i} \,|\, t \in [0, 1]\}$, also covered twice as p goes from $\mathbf{1}$ to $-\mathbf{1}$. By placing the things together we get

$$F_{(1,0,0)}(C_{\mathbf{i}}) = [-\mathbf{i}, \mathbf{i}] = \{t\mathbf{i} \,|\, t \in [-1, 1]\}.$$

An analogous reasoning works for a general great circle C_z, that is, $F_{(1,0,0)}(C_z) = [-z, z]$, the line segment joining $-z$ and z.

4.3 Invariant control sets

In this section we describe the control sets for a family of control systems on the sphere S^3 having $X_{(1,0,0)}$ as drift and control vector fields corresponding to pure quaternions. Others control systems, that is, considering others drifts and control vector fields the computations clearing become very intricate and the approach should be different to determine the control sets. Specifically the family of system considered here is given by:

(i) $X_{(1,0,0)} + uX_{(z,0,0)}$, where $z \in \mathrm{Im}\mathbb{H}$, $u \in [-1, 1]$.

(i′) $X_{(1,0,0)} + uX_{(z,0,0)}$, $z \in \mathrm{Im}\mathbb{H}$, $u = \pm 1$.

(ii) $X_{(1,0,0)} + uX_{(z_1,0,0)} + vX_{(z_2,0,0)}$, $z_1, z_2 \in \mathrm{Im}\mathbb{H}$ and $(u,v) \in [0,1]^2$.

(ii′) $X_{(1,0,0)} + uX_{(z_1,0,0)} + vX_{(z_2,0,0)}$, $z_1, z_2 \in \mathrm{Im}\mathbb{H}$, $(u,v) = (0,0), (1,0), (0,1)$.

(iii) $X_{(1,0,0)} + uX_{(\mathbf{i},0,0)} + vX_{(\mathbf{j},0,0)} + wX_{(\mathbf{k},0,0)}$, $(u,v,w) \in B[0, \frac{1}{2}] \subset \mathbb{R}^3$.

Note that the Lie algebra $\mathfrak{so}(1,4)$ cannot be generated by less than four symmetric elements, that is, elements in $\mathfrak{s}$. For example, $\mathbf{k} \in \mathfrak{s}$ does not belong to the Lie algebra generated by $\{\mathbf{1}, \mathbf{i}, \mathbf{j}\} \subset \mathfrak{s}$ in $\mathfrak{so}(1,4)$ (here we identify the elements in $\mathfrak{s}$ with quaternions as in Theorem 4.2.3).

Hence we do not have controllability on the identity component of $\mathrm{SO}(1,4)$ for these control systems whose vector fields correspond to elements in the component $\mathfrak{s}$ of the Cartan decomposition of $\mathfrak{so}(1,4)$, that is, vector fields of the form $X_{(q,0,0)}$, $q \in \mathbb{H}$.

Before the main results of this section we introduce the following concepts that will be important in the description of the control sets. A subset $C \subset S^3$ is said to be spherically convex when for any pair of points $p, q \in C$ every minimal geodesic segment joining them are contained in C. Given a subset $S \subset S^3$, denote by K_S the cone spanned by S, that is,

$$K_S = \{tp \,|\, p \in S,\ t \geq 0\} \subset \mathfrak{s},$$

and denote by $\mathrm{co}S$ the conic hull of S,

$$\mathrm{co}S = \left\{ \sum_{i=1}^{n} a_i p_i \,\middle|\, a_i \geq 0,\ p_i \in S,\ n \in \mathbb{N} \right\}.$$

Cones are very useful in characterizing convex sets in the sphere, since it is well known that a non-empty subset $C \subset S^3$ is convex if and only if the cone K_C is convex and pointed (that is, $K_C \cap (-K_C) \subseteq \{0\}$). This means that the proper convex sets on S^3 are the intersections of S^3 with pointed convex cones. In fact, if K is a convex and pointed cone, then $K_C = K \cap S^3$ is convex since K_C is exactly K. This means that the spherical convex hull of a subset $A \in S^3$ is the intersection of its conical hull $\mathrm{co}A$ with S^3.

With respect to the quaternions $\mathbb{H}$, the following notations are quite useful. Given $p \in \mathbb{H}$, we write $p = p_0 + p_1\mathbf{i} + p_2\mathbf{j} + p_3\mathbf{k}$ and we denote the imaginary part of p as a three-dimensional vector $\mathbf{p} = p_1\mathbf{i} + p_2\mathbf{j} + p_3\mathbf{k}$. Under these assumptions, the product of two quaternions $p = p_0 + p_1\mathbf{i} + p_2\mathbf{j} + p_3\mathbf{k}$ and $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$ assumes the form

$$p \cdot q = p_0 q_0 - \mathbf{p} \cdot \mathbf{q} + p_0 \mathbf{q} + q_0 \mathbf{p} + \mathbf{p} \times \mathbf{q}.$$

Our first result is this section determines which spherical domes are invariant by $X_{(1 \pm z, 0, 0)}$.

Proposition 4.3.1. *Given $z \in \mathrm{Im}\mathbb{H}$, let $r_1 = 1/\sqrt{1+|z|^2}$. Then, for every $a \in [0, r_1)$ the spherical dome*

$$S_a = \{p \in S^3 \,|\, \mathrm{Re}(p) \geq a\}$$

is invariant under the family of vector fields $\{X_{(1+z,0,0)}, X_{(1-z,0,0)}\}$.

Proof: For a pure quaternion z

$$X_{(1-z,0,0)}(p) = \frac{1}{2}\left((1-z) - p(1+z)p\right) = \frac{1}{2}\left(1 - z - p^2 - pzp\right).$$

If $p \in S^3 \cap \mathrm{Im}\mathbb{H}$ the above expression becomes

$$X_{(1-z,0,0)}(p) = \frac{1}{2}\left(2 - z - pzp\right).$$

Similarly,

$$X_{(1+z,0,0)}(p) = \frac{1}{2}\left(2 + z + pzp\right).$$

Evaluating the products we get $pz = -\mathbf{p} \cdot \mathbf{z} + \mathbf{p} \times \mathbf{z}$, and hence

$$\begin{aligned} pzp &= -(\mathbf{p} \times \mathbf{z}) \cdot \mathbf{p} - (\mathbf{p} \cdot \mathbf{z})\mathbf{p} + (\mathbf{p} \times \mathbf{z}) \times \mathbf{p} \\ &= -(\mathbf{p} \cdot \mathbf{z})\mathbf{p} + (\mathbf{p} \times \mathbf{z}) \times \mathbf{p} \in \mathrm{Im}\mathbb{H}. \end{aligned}$$

Since both z and pzp are pure quaternions, we can see that $X_{(1-z,0,0)}(p)$ and $X_{(1+z,0,0)}(p)$ have positive real parts (being p a pure quaternion). This fact ensures that the upper half of the sphere S^3 is invariant under the trajectories of the family

$$\left\{X_{(1-z,0,0)}, X_{(1+z,0,0)}\right\}.$$

Now, let $S^3 \ni p = a + w$, with $0 < a < 1$ and $w \in \mathrm{Im}\mathbb{H}$. Note that $|w| = \sqrt{1-a^2}$.

We have then

$$\begin{aligned} X_{(1+z,0,0)}(p) &= X_{(1+z,0,0)}(a+w) \\ &= \frac{1}{2}\left(1+z-(a+w)^2+(a+w)z(a+w)\right) \\ &= \frac{1}{2}\left(1+z-(a+w)^2+a^2z+azw+awz+wzw\right) \\ &= \mathrm{Re}\left(X_{(1+z,0,0)}(a+w)\right)+\mathrm{Im}\left(X_{(1+z,0,0)}(a+w)\right), \end{aligned}$$

where

$$\mathrm{Re}\left(X_{(1+z,0,0)}(a+w)\right)=\frac{1}{2}\left(1-a^2+\mathbf{w}\cdot\mathbf{w}-2a\mathbf{z}\cdot\mathbf{w}\right)$$

and

$$\mathrm{Im}\left(X_{(1+z,0,0)}(a+w)\right)=\frac{1}{2}\left(z-2a\mathbf{w}-\mathbf{w}\times\mathbf{w}+a^2z+wzw\right).$$

Similarly,

$$X_{(1-z,0,0)}(a+w)=\mathrm{Re}\left(X_{(1-z,0,0)}(a+w)\right)+\mathrm{Im}\left(X_{(1-z,0,0)}(a+w)\right),$$

where

$$\mathrm{Re}\left(X_{(1-z,0,0)}(a+w)\right)=\frac{1}{2}\left(1-a^2+\mathbf{w}\cdot\mathbf{w}+2a\mathbf{z}\cdot\mathbf{w}\right)$$

and

$$\mathrm{Im}\left(X_{(1-z,0,0)}(a+w)\right)=\frac{1}{2}\left(-z-2a\mathbf{w}-\mathbf{w}\times\mathbf{w}-a^2z-wzw\right).$$

The real parts of $X_{(1+z,0,0)}(a+w)$ and $X_{(1-z,0,0)}(a+w)$ are thus given by

$$\mathrm{Re}\left(X_{(1-z,0,0)}(a+w)\right)=1-a^2+a\sqrt{(1-a^2)}|z|\cos(t)$$

and

$$\mathrm{Re}\left(X_{(1+z,0,0)}(a+w)\right)=1-a^2-a\sqrt{(1-a^2)}|z|\cos(t),$$

where t denotes the angle between w and z.

Being t the angle formed by w and z, define the number

$$r_t=\frac{1}{\sqrt{1+|z|^2\cos^2(t)}}.$$

Note that $\mathrm{Re}\left(X_{(1-z,0,0)}(a+w)\right)=0$ if and only if $a=r_t$ or $a=1$. For a fixed $t\in[0,2\pi]$,

$t \neq \pi/2$, $t \neq 3\pi/2$, we have

$$\begin{aligned} \mathrm{Re}\left(X_{(1-z,0,0)}(a+w)\right) > 0 \quad \text{if} \quad a \in [0, r_t) \\ \mathrm{Re}\left(X_{(1-z,0,0)}(a+w)\right) < 0 \quad \text{if} \quad a \in (r_t, 1)\,, \end{aligned}$$

If $t = \pi/2$ or $t = 3\pi/2$, then $\mathrm{Re}\left(X_{(1-z,0,0)}(a+w)\right) > 0$ for every $a \in [0,1)$, since in this case $\mathrm{Re}\left(X_{(1-z,0,0)}(a+w)\right) = 1 - a^2$. For a given $z \in \mathrm{Im}\mathbb{H}$ the number r_t is minimum when $\cos(t) = 1$. Observe that

$$r_1 = \frac{1}{\sqrt{1+|z|^2}}$$

is precisely the real part of the singularities of $X_{(1\pm z,0,0)}$. This implies that if $a \in [0, r_1)$ then $\mathrm{Re}\left(X_{(1\pm z,0,0)}(a+w)\right) > 0$ for every possible w. In other words, for every $a \in [0, r_1)$ the spherical dome $S_a = \{p \in S^3 \,|\, \mathrm{Re}(p) \geq a\}$ is invariant under $\{X_{(1\pm z,0,0)}\}$. □

The Figure 4.1 is a three dimensional representation of the trajectories of $X_{(1\pm z,0,0)}$. The trajectories are the great circles containing the antipodal points $\frac{1\pm z}{|1\pm z|}$. The figure shows the attractors (the white and gray dots) while the quaternion **1** is represented by the black dot.

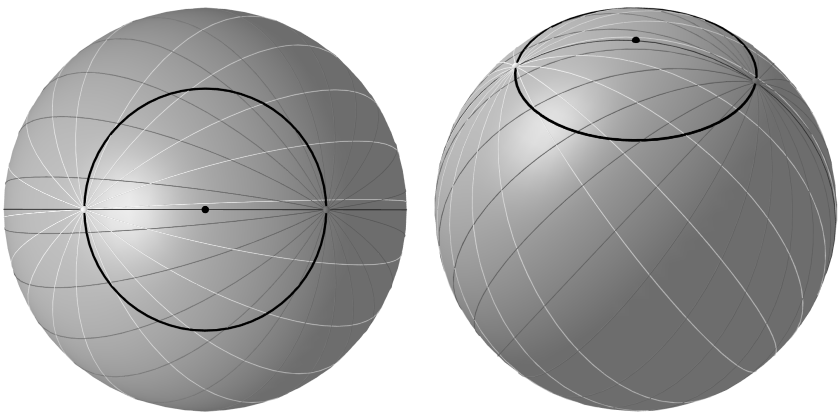


Figure 4.1: A 3D representation of the trajectories of $X_{(1-z,0,0)}$ and $X_{(1+z,0,0)}$

Now we can describe the control sets for the families of control systems given in the beginning of this section.

Case (i):

For the control system with bounded controls

$$\Gamma = \left\{X_{(1,0,0)} + uX_{(z,0,0)} \,|\, z \in \mathrm{Im}\mathbb{H},\ u \in [-1,1]\right\},$$

since all the attractors for such a system lie on (and cover) the minimal geodesic segment C on S^3 joining the attractors of $X_{(1+z,0,0)}$ and $X_{(1-z,0,0)}$, it follows from Teorem 4.1.1 that C is the only invariant control set for Γ.

Case (i'):

Using a geometrical approach we will show now that the invariant control set for the bang-bang control system

$$\{X_{(1,0,0)} \pm X_{(z,0,0)}\} \qquad (z \in \mathrm{Im}\mathbb{H})$$

is still a geodesic segment joining p_1 and p_2, the singularities (attractors) of the vector fields $X_{(1+z,0,0)}$ and $X_{(1-z,0,0)}$, respectively. That is, the spherical closed convex hull C of the set $\{p_1, p_2\}$, which is given by $C = \mathrm{cl}(\mathrm{co}\{p_1, p_2\}) \cap S^3$.

Theorem 4.3.2. *On the sphere S^3 the invariant control set for the family $\{X_{(1\pm z,0,0)}\}$ is the minimal geodesic segment C joining*

$$p_1 = \frac{1+z}{|1+z|} \text{ and } p_2 = \frac{1-z}{|1-z|}.$$

Proof: Denote by D the unique invariant control set for this system (it exists since S^3 is a compact Hausdorff space). Our aim is to show that $C = D$.

Note first that for every pair of points $x, y \in C$, $x, y \neq p_1, p_2$ we have a trajectory of the system joining x and y. In fact, the solutions for this control system follow geodesic curves in S^3. If the solution with initial value x that goes towards p_1 do not contains y, then the solution going towards p_2 does. Furthermore, $p_1 \in \mathrm{cl}\mathcal{O}^+(p_2)$ and $p_2 \in \mathrm{cl}\mathcal{O}^+(p_1)$. In other words, C is contained in the closure of the positive orbit of all its elements. Also, the positive orbit of all its points is contained in C.

Now let $x \in D$. For any $y \in C$ we have x in the closure of the positive orbit of y, that is, we can approximate x by an trajectory of the control system starting at y. This trajectory is the concatenation of solutions, all of them lying on the great circle that contains $\mathbf{1}$, p_1 and p_2, that is, all the trajectory is contained in this great circle. Since C is spherically convex, every path of the trajectory is a minimal geodesic segment between its initial and final points. This implies that $x \in \mathrm{cl}(C) = C$, and the maximality of D ensures that $C = D$. $\square$

Case (ii):

Again, Theorem 4.1.1 ensures that the control system with constrained controls

$$\left\{X_{(1,0,0)} + uX_{(z_1,0,0)} + vX_{(z_2,0,0)} \,|\, u, v \in [0,1]\right\} \qquad (z_1, z_2 \in \mathrm{Im}\mathbb{H})$$

is such that its invariant control set is the spherical closed convex hull of the set $A = \{\mathbf{1}, p_1, p_2, p_3\}$, being p_1 the attractor corresponding to $u = 1$ and $v = 0$, p_2 corresponding to $u = 0$ and $v = 1$ and p_3 to $u = 1$ and $v = 1$. However, if the controls were considered in the interval $[-1, 1]$, we should have add to the set A also the attractors corresponding to the control values $u, v = -1$. Any other attractors would belong to the spherical closed convex hull too, and a similar reasoning works.

Case (ii'):

In this case with only two vector fields for the control, the situation is quite similar to that one in case (i') and a geometrical reasoning still works.

Theorem 4.3.3. *If* $\Gamma = \{X_{(1,0,0)}, X_{(1+z_1,0,0)}, X_{(1+z_2,0,0)}\}$*,* $z_1, z_2 \in \mathrm{Im}\mathbb{H}$*, is a family of vector fields on the sphere* S^3 *such that the attractors* p_1 *and* p_2 *for* $X_{(1+z_1,0,0)}$ *and* $X_{(1+z_2,0,0)}$ *do not belong to the same great circle containing* $\mathbf{1}$*, then the invariant control set for* Γ *in* S^3 *is the spherical closed convex hull of* $A = \{\mathbf{1}, p_1, p_2\}$*.*

Proof: Let $\mathrm{co}A$ be the conic hull of A. The closed convex hull of A is

$$C = S^3 \cap \mathrm{cl}(\mathrm{co}A).$$

Note that C has empty interior in S^3 since it is a two-dimensional submanifold ($\mathrm{co}A$ is a cone generated by three elements). To avoid confusion, write $\mathrm{algint}C$ for the algebraic interior of C. The relative boundary of $\mathrm{co}A$ is the union of the conic hulls of $\{\mathbf{1}, p_1\}$, $\{\mathbf{1}, p_2\}$ and $\{p_1, p_2\}$, that is, the relative boundary of C is union of great circle arcs joining $\mathbf{1}$ to p_1, $\mathbf{1}$ to p_2 and p_1 to p_2.

To see that C is the desired invariant control set, let $x \in \mathrm{algint}C$. Denote by $X_1(x)$, $X_{p_1}(x)$ and $X_{p_2}(x)$ the solutions for the corresponding vector fields with initial value x. Remember that all these solutions follows geodesics on S^3. Let C_1 be the closed convex hull of $\{x, \mathbf{1}, p_1\}$, C_2 the closed convex hull of $\{x, \mathbf{1}, p_2\}$ and C_3 that one of $\{x, p_1, p_2\}$. It is clear that $C = C_1 \cup C_2 \cup C_3$. If $y \in \mathrm{algint}C_1$, then we obtain a trajectory from x to y in the following way. The convexity of C tells us that the solution $X_1(y)$ meets $X_{p_1}(x)$ in $y_1 \in \mathrm{algint}C$ (the intersection $X_1(y) \cap X_{p_1}(x)$ is nonempty since both solutions are contained in the same three dimensional subspace of $\mathfrak{s}$, and the intersection belongs to $\mathrm{algint}C$ by convexity). The solution $X_1(y_1)$ contains y since $X_1(y_1)$ and $X_1(y)$ are contained in the same great circle of S^3. Thus, joining $X_{p_1}(x)$ and $X_1(y_1)$ we get a trajectory from x to y. Analogous reasonings work for $\mathrm{algint}C_2$ and $\mathrm{algint}C_3$, just interchanging the considered solutions. If $y \in \mathrm{algint}C$ does not belong to $\mathrm{algint}C_i$, $i = 1, 2, 3$, then y belongs to $X_1(x)$, $X_{p_1}(x)$ or $X_{p_2}(x)$, and can be trivially reached from x. This proves that C is contained in the closure of the positive orbit of all its elements.

Now, let D be the unique invariant control set for this control system and take $x \in D$. We just proved that $C \subset D$, by the maximality of D. That is, for any $y \in C$ we have $x \in \mathrm{cl}\mathcal{O}^+(y)$. In other words, we can get close of x by trajectories starting at y. Every path of such trajectory is a minimal geodesic segment, and this implies that $x \in \mathrm{cl}C$, by the definition of spherical closed convex hull. Hence $x \in C$, and so $C = D$.

□

Case (iii):

Although very interesting, the control systems considered above does not satisfy the Lie algebra rank condition, as said in the beginning of this section. One can note that all the invariant control sets obtained have empty interior in S^3. In our final theorem we investigate a control system satisfying the Lie algebra rank condition. Note that checking LARC is easier if it is done with matrices instead of vector fields (see Proposition 4.2.3).

Theorem 4.3.4. *Consider the control system with restricted controls defined by the vector fields*

$$X_{(1,0,0)} + uX_{(\mathbf{i},0,0)} + vX_{(\mathbf{j},0,0)} + wX_{(\mathbf{k},0,0)}, \qquad (u, v, w) \in B[0, 1] \subset \mathbb{R}^3,$$

where $B[0, 1]$ stands for the closed ball of radius 1 in $\mathbb{R}^3$ centered at the origin. Then, the invariant control set for this control system is exactly the spherical dome

$$C = \{p \in S^3 \mid \mathrm{Re}(p) \geq 1/\sqrt{2}\}.$$

Proof: To prove this theorem, note first that the invariance follows from computations similar to those we have done for $X_{(1\pm z,0,0)}$ in Proposition 4.3.1. The singularities corresponding to the vector fields $X_{(1\pm\mathbf{i},0,0)}$, $\pm X_{(1\pm\mathbf{j},0,0)}$ and $X_{(1\pm\mathbf{k},0,0)}$ on the invariant hemisphere are

$$\frac{1\pm\mathbf{i}}{\sqrt{2}},\ \frac{1\pm\mathbf{j}}{\sqrt{2}},\ \frac{1\pm\mathbf{k}}{\sqrt{2}}$$

and such singularities lie on the boundary of the spherical dome C, since their real parts are all $1/\sqrt{2}$, and the same statement is true for the singularities corresponding to vector fields

$$X_{(1,0,0)} + uX_{(\mathbf{i},0,0)} + vX_{(\mathbf{j},0,0)} + wX_{(\mathbf{k},0,0)}$$

such that $u^2 + v^2 + w^2 = 1$.

By the other hand, let p in the boundary of C. It must have the form $p = 1/\sqrt{2} + z$, with $z \in \mathrm{Im}\mathbb{H}$. Observe that

$$|p|^2 = \frac{1}{2} + |z|^2 \Longrightarrow |z| = \frac{1}{\sqrt{2}}.$$

Write $z = u_0\mathbf{i} + v_0\mathbf{j} + w_0\mathbf{k}$. Then $u_0^2 + v_0^2 + w_0^2 = 1/2$ and the attractor corresponding to the vector field

$$X_{(1,0,0)} + \sqrt{2}u_0X_{(\mathbf{i},0,0)} + \sqrt{2}v_0X_{(\mathbf{j},0,0)} + \sqrt{2}w_0X_{(\mathbf{k},0,0)}$$

is the point p, as can be easily checked. This proves that every point belonging to the boundary of C is a singularity of a vector field of the control system corresponding to a control (u, v, w) such that $u^2 + v^2 + w^2 = 1$.

The previous reasoning still holds for controls $(u, v, w) \in B[0, a]$, with $a > 0$. In this case, the singularities corresponding to controls (u, v, w) such that $u^2 + v^2 + w^2 = a^2$ are in the boundary of the spherical dome

$$C_a = \left\{ p \in S^3 \;\middle|\; \mathrm{Re}(p) \geq \frac{1}{\sqrt{1+a^2}} \right\},$$

and so we conclude that all of the points belonging to C are singularities corresponding to some vector field of the control system. The result is thus a consequence of Theorem 4.1.1. □

CHAPTER 5

FUTURE DIRECTIONS CONCERNING THE SPECIAL UNITARY LIE GROUP $\mathrm{SU}(1,2)$

Following the same general lines of the previous chapter, the present chapter will be set for start exploring the geometrical aspects of the Lie group $\mathrm{SU}(1,2)$. Just as $\mathfrak{so}(1,4)$, the Lie algebra $\mathfrak{su}(1,2)$ is also a real rank 1 Lie algebra with a four dimensional symmetric part in the Cartan decompositions, leading to the sphere S^3 as its only flag manifold. The infinitesimal action on S^3 is quite different in this case, since the occurrence of the positive root 2α implies that the Borel metric does not coincides with the canonical metric given by the immersion of S^3 in $\mathbb{H}$.

5.1 General structure of $\mathfrak{su}(p,q)$

For a complex matrix M, denote by M^* its conjugate transpose, that is, $M^* = \overline{M}^t$. Let $I_{p,q}$ denote the block diagonal matrix

$$I_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix},$$

where I_p and I_q are $p \times p$ and $q \times q$ identity matrices. The group $\mathrm{SU}(p,q)$ is the group of $(p+q) \times (p+q)$ complex matrices M satisfying $M^* I_{p,q} M = I_{p,q}$ and $\det(M) = 1$:

$$\mathrm{SU}(p,q) = \{M \in M_{p+q}(\mathbb{C}) \mid M^* I_{p,q} M = I_{p,q} \text{ and } \det(M) = 1\}.$$

If we write M as a block matrix of the form

$$M = \begin{bmatrix} A_{p\times p} & B_{p\times q} \\ C_{q\times p} & D_{q\times q} \end{bmatrix},$$

then we have that $M \in \mathrm{SU}(p,q)$ if and only if $\det(M) = 1$ and

$$\begin{aligned} A^*A - C^*C &= I_p \\ B^*B - D^*D &= -I_q \\ A^*B &= C^*D \\ B^*A &= D^*C. \end{aligned}$$

Now, if we take a smooth curve $M(t) \in \mathrm{SU}(p,q)$ with $M(0) = I$, then we can take the derivatives of the expressions $M^*(t)I_{p,q}M(t) = I_{p,q}$ and $\det(M(t)) = 1$ with respect to t. We have the derivative of the determinant given by the Jacobi's formula and

$$\frac{dM^*}{dt} I_{p,q} M(t) + M^*(t) I_{p,q} \frac{dM}{dt} = 0.$$

Evaluating at $t = 0$ and calling $X = \frac{dM}{dt}\big|_{t=0}$, we get

$$X^* I_{p,q} + I_{p,q} X = 0, \qquad \mathrm{trace}(X) = 0,$$

which is the condition for X to be in the Lie algebra $\mathfrak{su}(p,q)$ of $\mathrm{SU}(p,q)$. That is,

$$\mathfrak{su}(p,q) = \{X \in M_{p+q}(\mathbb{C}) \mid X^* I_{p,q} + I_{p,q} X = 0 \text{ and } \mathrm{trace}(X) = 0\}.$$

To see that $\mathfrak{su}(p,q)$ is in fact a real form of $\mathfrak{sl}(p+q,\mathbb{C})$, let

$$\begin{aligned} \sigma : \mathfrak{sl}(p+q,\mathbb{C}) &\longrightarrow \mathfrak{sl}(p+q,\mathbb{C}) \\ X &\longmapsto -I_{p,q} X^* I_{p,q}. \end{aligned}$$

Note that σ is a conjugation on $\mathfrak{sl}(p+q,\mathbb{C})$. In fact, σ is antilinear and

$$\sigma(\sigma(X)) = -I_{p,q}\sigma(X)^* I_{p,q} = -I_{p,q}\left(-I_{p,q} X^* I_{p,q}\right)^* I_{p,q} = I_{p,q}^2 X^{**} I_{p,q}^2 = X,$$

that is, σ is an involution. Moreover, $\mathfrak{su}(p,q)$ is exactly the set of fixed points of σ. If $X \in \mathfrak{su}(p,q)$, then $X^* I_{p,q} = -I_{p,q}X$, and this implies

$$\sigma(X) = -I_{p,q}X^*I_{p,q} = -I_{p,q}\left(-I_{p,q}X\right) = X.$$

By the other hand, if $X \in \mathfrak{sl}(p+q,\mathbb{C})$ is such that $\sigma(X) = X$, then we have

$$-I_{p,q}X^*I_{p,q} = X \iff X^*I_{p,q} = -I_{p,q}X \iff X \in \mathfrak{su}(p,q).$$

This means that $\mathfrak{su}(p,q)$ is a real form of $\mathfrak{sl}(p+q,\mathbb{C})$. Letting A, B, C and D blocks of sizes $p\times p$, $p\times q$, $q\times p$ and $q\times q$, respectively, and writing X as

$$X = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

then the conditions $X^*I_{p,q} + I_{p,q}X = 0$ and $\mathrm{trace}(X) = 0$ lead us to

$$\begin{aligned}\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \cdot \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} + \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} &= 0 \\ \begin{bmatrix} A^* & -C^* \\ B^* & -D^* \end{bmatrix} + \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} &= 0\end{aligned}$$

and $\mathrm{tr}(A) + \mathrm{tr}(D) = 0$. That is,

$$\begin{bmatrix} A^* & -C^* \\ B^* & -D^* \end{bmatrix} = \begin{bmatrix} -A & -B \\ C & D \end{bmatrix}, \quad \mathrm{tr}(A) + \mathrm{tr}(D) = 0.$$

Implying

$$A^* = -A, \quad B^* = C, \quad D^* = -D, \quad \mathrm{tr}(A) + \mathrm{tr}(D) = 0.$$

So, one can see the Lie algebra $\mathfrak{su}(p,q)$ as being the set

$$\mathfrak{su}(p,q) = \left\{ \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \,\middle|\, A^* = -A,\ D^* = -D,\ \mathrm{tr}(A) + \mathrm{tr}(D) = 0 \right\},$$

where A is of size $p\times p$, B is $p\times q$ and D is $q\times q$.

Considering the compact real form $\mathfrak{su}(n)$ of $\mathfrak{sl}(n,\mathbb{C})$, that is,

$$\mathfrak{su}(n) = \{X \in \mathfrak{sl}(n,\mathbb{C}) \mid X^* = -X\},$$

then a Cartan decomposition of the algebra $\mathfrak{su}(p,q)$ is given by its intersection with the algebra $\mathfrak{su}(p+q)$, precisely, $\mathfrak{su}(p,q) = \mathfrak{k} \oplus \mathfrak{s}$ where

$$\mathfrak{k} = \mathfrak{su}(p,q) \cap \mathfrak{su}(p+q) \quad \text{and} \quad \mathfrak{s} = \mathfrak{su}(p,q) \cap \mathbf{i}\mathfrak{su}(p+q)$$

is a Cartan decomposition of $\mathfrak{su}(p,q)$. This means that

$$\mathfrak{k} = \left\{ \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \middle| A^* = -A,\ D^* = -D,\ \operatorname{tr}(A+D) = 0 \right\}$$

and

$$\mathfrak{s} = \left\{ \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} \middle| B \text{ is a } p \times q \text{ matrix} \right\},$$

in fact, a matrix of the form

$$\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix},$$

belongs to $\mathfrak{su}(p,q)$ and also to $\mathbf{i}\mathfrak{su}(p+q)$, since

$$\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} = \mathbf{i}X, \text{ where } X = \begin{bmatrix} 0 & -\mathbf{i}B \\ -\mathbf{i}B^* & 0 \end{bmatrix},$$

so that $X^* = -X$. Note also that if $X \in \mathfrak{k}$, then we can write

$$\begin{aligned}
X &= \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \\
&= \begin{bmatrix} A - \frac{\operatorname{tr}(A)}{p} I_p + \frac{\operatorname{tr}(A)}{p} I_p & 0 \\ 0 & D - \frac{\operatorname{tr}(D)}{q} I_q + \frac{\operatorname{tr}(D)}{q} I_q \end{bmatrix} \\
&= \begin{bmatrix} A - \frac{\operatorname{tr}(A)}{p} I_p & 0 \\ 0 & D - \frac{\operatorname{tr}(D)}{q} I_q \end{bmatrix} + \begin{bmatrix} \frac{\operatorname{tr}(A)}{p} I_p & 0 \\ 0 & \frac{\operatorname{tr}(D)}{q} I_q \end{bmatrix} \\
&= \begin{bmatrix} A - \frac{\operatorname{tr}(A)}{p} I_p & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & D - \frac{\operatorname{tr}(D)}{q} I_q \end{bmatrix} + \begin{bmatrix} \frac{\operatorname{tr}(A)}{p} I_p & 0 \\ 0 & \frac{\operatorname{tr}(D)}{q} I_q \end{bmatrix},
\end{aligned}$$

and this shows that $\mathfrak{k}$ is not semisimple and isomorphic to

$$\mathfrak{su}(p) \oplus \mathfrak{su}(q) \oplus \mathfrak{z},$$

where the center $\mathfrak{z}$ is given by the matrices having the form

$$\begin{bmatrix} \frac{\operatorname{tr}(A)}{p} I_p & 0 \\ 0 & \frac{\operatorname{tr}(D)}{q} I_q \end{bmatrix}.$$

Now we describe a maximal abelian subalgebra $\mathfrak{a}$ of $\mathfrak{s}$. If $B_1, B_2 \in \mathfrak{s}$ are of the form

$$B_1 = [\ \Lambda_1 \quad 0\ ] \quad \text{and} \quad B_2 = [\ \Lambda_2 \quad 0\ ],$$

with λ_i diagonal $p \times p$ real matrices, then we have

$$B_1 B_2^t - B_2 B_1^t = 0,$$

and this means that the matrices B_1 and B_2 commute. So, a maximal abelian subalgebra $\mathfrak{a}$ of $\mathfrak{s}$ is

$$\mathfrak{a} = \left\{ \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} \,\middle|\, B = [\ \Lambda \quad 0\ ], \text{ where } \Lambda = \operatorname{diag}(a_1, \ldots, a_p),\ a_j \in \mathbb{R} \right\}.$$

From this we obtain that the real rank of the Lie algebra $\mathfrak{su}(p,q)$ is exactly $p = \min\{p,q\}$ (since we consider $p \leq q$). For a distinguished Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{su}(p,q)$, let $\mathfrak{h}_{\mathfrak{k}}$ be

$$\mathfrak{h}_{\mathfrak{k}} = \left\{ \begin{bmatrix} \mathbf{i}A & 0 \\ 0 & \mathbf{i}D \end{bmatrix} \,\middle|\, \begin{array}{ll} A = \operatorname{diag}(a_1, \ldots, a_p), & a_j, b_j \in \mathbb{R}, \\ D = \operatorname{diag}(b_1, \ldots, b_q), & \operatorname{tr}(A) + \operatorname{tr}(D) = 0 \end{array} \right\}.$$

And then the distinguished Cartan subalgebra is given by $\mathfrak{h} = \mathfrak{h}_{\mathfrak{k}} \oplus \mathfrak{a}$, that is,

$$\mathfrak{h} = \left\{ \begin{bmatrix} \mathbf{i}A & B \\ B^* & \mathbf{i}D \end{bmatrix} \,\middle|\, \begin{array}{ll} A = \operatorname{diag}(a_1, \ldots, a_p), & B = [\ \Lambda \quad 0\ ], \\ \Lambda = \operatorname{diag}(b_1, \ldots, b_p), & \operatorname{tr}(A) + \operatorname{tr}(D) = 0, \\ D = \operatorname{diag}(c_1, \ldots, c_q), & a_j, b_j, c_j \in \mathbb{R} \end{array} \right\}.$$

Two hermitian forms are unitarily equivalent if and only if they have the same characteristic roots, that is, if and only if they are similar.

Let $\mathfrak{h}_I$ and $\mathfrak{g}_J$ be the Lie algebras given by

$$\mathfrak{g}_I = \{A \in \mathfrak{sl}(n,\mathbb{C}) \mid \overline{A}^t I + IA = 0\}$$

and

$$\mathfrak{g}_J = \{A \in \mathfrak{sl}(n,\mathbb{C}) \mid \overline{A}^t J + JA = 0\}.$$

If I and J are unitarilly equivalent hermitian forms, there exists U unitary such that $I = \overline{U}^t JU$. Taking X such that $\overline{X}^t I + IX$, then

$$\begin{aligned} 0 = \overline{X}^t I + IX &= \overline{X}^t \overline{U}^t JU + \overline{U}^t JUX \\ &= U\overline{X}^t \overline{U}^t JU\overline{U}^t + U\overline{U}^t JUX\overline{U}^t \\ &= U\overline{X}^t \overline{U}^t J + JUX\overline{U}^t \\ &= \overline{(UX\overline{U}^t)}^t J + J(UX\overline{U}^t) \end{aligned}$$

which means that if $X \in \mathfrak{g}_I$, then $(UX\overline{U}^t) \in \mathfrak{g}_J$, and the Lie algebras $\mathfrak{g}_I$ and $\mathfrak{g}_J$ are isomorphic via $\varphi : X \mapsto (UX\overline{U}^t)$, since

$$\varphi[X,Y] = U(XY - YX)\overline{U}^t = UX\overline{U}^t UY\overline{U}^t - UY\overline{U}^t UX\overline{U}^t = [\varphi(X), \varphi(Y)].$$

Now, the hermitian forms

$$I_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} \quad \text{and} \quad J_{p,q} = \begin{bmatrix} 0 & I_p & 0 \\ I_p & 0 & 0 \\ 0 & 0 & -I_{q-p} \end{bmatrix}$$

have the same characteristic roots, and so we can also see the Lie algebra $\mathfrak{su}(p,q)$ as the set of the matrices that are skew-hermitian with respect to the hermitian form on $\mathbb{C}^{p+q}$ given by $J_{p,q}$. In other words,

$$\mathfrak{su}(p,q) = \{M \in M_{p+q}(\mathbb{C}) \mid MJ_{p,q} + J_{p,q}M^* = 0,\ \mathrm{tr}(M) = 0\}.$$

It is a real form of $\mathfrak{sl}(p+q,\mathbb{C})$ with associate conjugation given by

$$\sigma(M) = -J_{p,q}M^*J_{p,q}, \quad M \in \mathfrak{sl}(p+1,\mathbb{C}).$$

If M is a matrix in $M_{p+q}(\mathbb{C})$, writing M in a block form corresponding to the hermitian form $J_{p,q}$,

$$M = \begin{bmatrix} A & B & E \\ C & D & F \\ X & Y & Z \end{bmatrix},$$

then we have

$$M^* = \begin{bmatrix} A^* & C^* & X^* \\ B^* & D^* & Y^* \\ E^* & F^* & Z^* \end{bmatrix},$$

$$MJ_{p,q} = \begin{bmatrix} B & A & -E \\ D & C & -F \\ Y & X & -Z \end{bmatrix} \quad \text{and} \quad J_{p,q}M^* = \begin{bmatrix} B^* & D^* & Y^* \\ A^* & C^* & X^* \\ -E^* & -F^* & -Z^* \end{bmatrix}.$$

Thus the condition $MJ_{p,q} + J_{p,q}M^* = 0$ yields

$$B = -B^*, \quad D = -A^*, \quad Y = E^*, \quad A = -D^*, \quad C = -C^*, \quad X = F^*,$$

$$E = Y^*, \quad F = X^*, \quad Z = -Z^*.$$

This means that M must have the block form

$$M = \begin{bmatrix} A & B & Y^* \\ C & -A^* & X^* \\ X & Y & Z \end{bmatrix}, \tag{5.1}$$

where B, C and Z are skew-hermitian matrices of respective orders being p, p and $q-p$. Consequently, M is in $\mathfrak{su}(p,q)$ if and only if it has the block form (5.1) and

$$B, C \in \mathfrak{u}(p), \quad Z \in \mathfrak{su}(q-p) \quad \text{and} \quad \mathrm{Im}(\mathrm{tr}(A)) = 0.$$

A matrix $M \in \mathfrak{su}(p,q)$ will also belong to $\mathfrak{su}(p+q)$ if and only if it has the form

$$M = \begin{bmatrix} A & B & -X^* \\ B & A & X^* \\ X & -X & Z \end{bmatrix}, \text{ with } A, B \in \mathfrak{u}(p),\ Z \in \mathfrak{su}(q-p) \text{ and } \mathrm{tr}(A) = 0.$$

Similarly, a matrix $M \in \mathfrak{su}(p,q)$ belongs also to $\mathbf{i}\mathfrak{su}(p+q)$ if and only if M has the form

$$M = \begin{bmatrix} A & B & X^* \\ -B & -A & X^* \\ X & X & 0 \end{bmatrix}, \quad \text{with} \quad A = A^* \text{ and } B \in \mathfrak{u}(p).$$

Hence, a Cartan decomposition $\mathfrak{su}(p,q) = \mathfrak{k} \oplus \mathfrak{s}$ is such that

$$\mathfrak{k} = \left\{ \begin{bmatrix} A & B & -X^* \\ B & A & X^* \\ X & -X & Z \end{bmatrix} \middle| \; A \in \mathfrak{su}(p),\; B \in \mathfrak{u}(p),\; Z \in \mathfrak{su}(q-p) \right\}$$

and

$$\mathfrak{s} = \left\{ \begin{bmatrix} A & B & X^* \\ -B & -A & X^* \\ X & X & 0 \end{bmatrix} \middle| \; A = A^* \text{ and } B \in \mathfrak{u}(p) \right\}.$$

In this realization of $\mathfrak{su}(p,q)$ a maximal abelian subalgebra in $\mathfrak{s}$ is given by

$$\mathfrak{a} = \left\{ \begin{bmatrix} \Lambda & 0 & 0 \\ 0 & -\Lambda & 0 \\ 0 & 0 & 0 \end{bmatrix} \middle| \; \Lambda = \operatorname{diag}(a_1, \ldots, a_p),\; a_j \in \mathbb{R} \right\},$$

since in general non-diagonal matrices do not commute with respect to the usual Lie bracket given by the commutator. Once again we see that the real rank of $\mathfrak{su}(p,q)$ is $p = \dim \mathfrak{a}$. A Cartan subalgebra in $\mathfrak{su}(p,q)$ is given by

$$\mathfrak{h} = \left\{ \begin{bmatrix} D & 0 & 0 \\ 0 & -\overline{D} & 0 \\ 0 & 0 & \mathbf{i}T \end{bmatrix} \middle| \; \begin{array}{c} D = \operatorname{diag}(z_1, \ldots, z_p),\; z_j \in \mathbb{C}, \\ T = \operatorname{diag}(a_1, \ldots, a_{q-p}),\; a_j \in \mathbb{R}, \\ \operatorname{tr}(2\operatorname{Im}(D)) + \operatorname{tr}(T) = 0 \end{array} \right\}.$$

So we can take $\mathfrak{h}_{\mathfrak{k}} = \mathfrak{h} \cap \mathfrak{k}$:

$$\mathfrak{h}_{\mathfrak{k}} = \left\{ \begin{bmatrix} \mathbf{i}\Lambda & 0 & 0 \\ 0 & \mathbf{i}\Lambda & 0 \\ 0 & 0 & \mathbf{i}T \end{bmatrix} \middle| \; \begin{array}{cc} \Lambda = \operatorname{diag}(a_1, \ldots, a_p), & a_j, b_j \in \mathbb{R}, \\ T = \operatorname{diag}(b_1, \ldots, b_{q-p}), & \operatorname{tr}(\Lambda) + \operatorname{tr}(T) = 0 \end{array} \right\}.$$

So, the distinguished Cartan subalgebra is given by $\mathfrak{h} = \mathfrak{h}_{\mathfrak{k}} \oplus \mathfrak{a}$, and it complexifies into the Cartan subalgebra $\mathfrak{h}_{\mathbb{C}}$ of $\mathfrak{sl}(p+q,\mathbb{C})$ formed by the diagonal matrices. As we know, the roots of $\mathfrak{h}_{\mathbb{C}}$ are given by the differences of the diagonal coordinate linear functionals in $\mathfrak{h}_{\mathbb{C}}$. Thus, if λ_j is the linear functional on $\mathfrak{h}_{\mathbb{C}}$ given by $\lambda_j(H) = a_j$, where $H = \mathrm{diag}(a_1,\ldots,a_{p+q}) \in \mathfrak{h}_{\mathbb{C}}$, we have the root system

$$\Pi_{\mathbb{C}} = \{\alpha_{j,k} = \lambda_j - \lambda_k,\ j \neq k\}.$$

For $\alpha \in \mathfrak{h}_{\mathbb{C}}^*$ we set $H_\alpha \in \mathfrak{h}_{\mathbb{C}}$ as $\alpha(\cdot) = \langle H_\alpha, \cdot\rangle$ and we denote by $\mathfrak{h}_{\mathbb{R}}$ the real vector subspace of $\mathfrak{h}_{\mathbb{C}}$ generated by H_α, $\alpha \in \Pi_{\mathbb{C}}$. Since $\langle \alpha, \beta\rangle \in \mathbb{R}$, the roots $\alpha \in \Pi_{\mathbb{C}}$ are real valued in $\mathfrak{h}_{\mathbb{R}}$. We have

$$\mathfrak{h}_{\mathbb{R}} = \{H \in \mathfrak{h}_{\mathbb{C}} \,|\, \forall \alpha \in \Pi_{\mathbb{C}},\ \alpha(H) \in \mathbb{R}\},$$

and the direct sum decomposition $\mathfrak{h}_{\mathbb{R}} = \mathfrak{a} \oplus \mathbf{i}\mathfrak{h}_{\mathfrak{k}}$, where the real subspaces $\mathfrak{a}$ and $\mathbf{i}\mathfrak{h}_{\mathfrak{k}}$ are orthogonal with respect to the Cartan-Killing form. By definition, a root $\alpha \in \Pi_{\mathbb{C}}$ is said real if it vanishes in $\mathfrak{h}_{\mathfrak{k}}$ and it is said imaginary if it vanishes in $\mathfrak{a}$. We denote the sets of real and imaginary roots in $\Pi_{\mathbb{C}}$ by Π_{Re} and Π_{Im}, respectively. The real roots are given by $\Pi_{\mathrm{Re}} = \{\pm\alpha_{1,p+1},\ldots,\pm\alpha_{p,2p}\}$. The imaginary roots are

$$\Pi_{\mathrm{Im}} = \{\pm\alpha_{j,k} \,|\, 2p+1 \leq j \neq k \leq p+q,\ \text{with } q \geq p+2\}.$$

If $q = p$ or $q = p+1$ there are no imaginary roots. The restricted roots are the restrictions of $\alpha_{j,k}$, $j \neq k$, to the maximal abelian subalgebra $\mathfrak{a} = \{\mathrm{diag}\{\Lambda, -\Lambda, 0\} \,|\, \Lambda = \mathrm{diag}\{a_1,\ldots,a_p\}\}$. Precisely, considering $\Pi_{\mathbb{C}} = \{\lambda_j - \lambda_k,\ j \neq k\}$,

- if $p < q$, then the restricted roots are

$$\{\pm(\lambda_r - \lambda_s), \pm(\lambda_r + \lambda_s),\ r \neq s,\ 1 \leq r,s \leq p\} \cup \{\pm\lambda_r, \pm 2\lambda_r,\ 1 \leq r \leq p\},$$

- and if $p = q$, the restricted roots are given by

$$\{\pm(\lambda_r - \lambda_s), \pm(\lambda_r + \lambda_s),\ r \neq s,\ 1 \leq r,s \leq p\} \cup \{\pm 2\lambda_r,\ 1 \leq r \leq p\},$$

From these sets we obtain the sets of simple roots

- $\Sigma = \{\lambda_1 - \lambda_2, \dots, \lambda_{p-1} - \lambda_p, \lambda_p\}$ if $p < q$, and
- $\Sigma = \{\lambda_1 - \lambda_2, \dots, \lambda_{p-1} - \lambda_p, 2\lambda_p\}$ if $p = q$.

The set of the real regular elements $\overline{\mathfrak{a}} \subset \mathfrak{a}$ is defined by

$$\overline{\mathfrak{a}} = \{H \in \mathfrak{a} \,|\, \alpha(H) \neq 0, \text{ for every restricted root } \alpha\}.$$

In the case $p < q$, if $H = \operatorname{diag}(a_1, \dots, a_p, -a_1, \dots, -a_p, 0, \dots, 0) \in \mathfrak{a}$, it is easy to see that for $1 \leq r, s \leq 2p$, $\pm(\lambda_r - \lambda_s)(H) \neq 0$ only when $a_1, \dots, a_p$ are all different from each other and all of them are nonzero. In a similar way, $\pm(\lambda_r - \lambda_s)(H) \neq 0$ only if $a_1, \dots, a_p$ are all different from each other's opposite and all of them are nonzero. Lastly, $\pm\lambda_r(H) \neq 0$ for all $r = 1, \dots, p$ only if $a_1, \dots, a_p$ are all different from zero. Writing $\Lambda = (a_1, \dots, a_p)$, the set of real regular elements is thus given by

$$\overline{\mathfrak{a}} = \{H = \operatorname{diag}(\Lambda, -\Lambda, 0 \dots, 0) \,|\, a_i \neq \pm a_j,\ a_i \neq 0, \text{ for } 1 \leq i, j \leq p\}.$$

In the case $p = q$ the calculations are completely analogous and the set of real regular elements is

$$\overline{\mathfrak{a}} = \{H = \operatorname{diag}(\Lambda, -\Lambda) \,|\, a_i \neq \pm a_j,\ a_i \neq 0, \text{ for } 1 \leq i, j \leq p\}.$$

Associated to the simple root system $(p < q)$

$$\Sigma = \{\lambda_1 - \lambda_2, \dots, \lambda_{p-1} - \lambda_p, \lambda_p\}$$

we get the set of positive roots

$$\Pi^+ = \{(\lambda_r - \lambda_s) \,|\, 1 \leq r < s \leq p\} \cup \{(\lambda_r + \lambda_s),\ \lambda_r \,|\, 1 \leq r, s \leq p\}$$

and the set of negative roots

$$\Pi^- = \{(\lambda_s - \lambda_r) \,|\, 1 \leq r < s \leq p\} \cup \{(-\lambda_r - \lambda_s),\ -\lambda_r \,|\, 1 \leq r, s \leq p\}.$$

In a similar way, for the simple root system $\Sigma = \{\lambda_1 - \lambda_2, \dots, \lambda_{p-1} - \lambda_p, 2\lambda_p\}$ we have the positive roots

$$\Pi^+ = \{(\lambda_r - \lambda_s)\,|\,1 \le r < s \le p\} \cup \{(\lambda_r + \lambda_s),\, 2\lambda_r\,|\,1 \le r, s \le p\}$$

and the negative roots

$$\Pi^- = \{(\lambda_s - \lambda_r)\,|\,1 \le r < s \le p\} \cup \{(-\lambda_r - \lambda_s),\, -2\lambda_r\,|\,1 \le r, s \le p\}.$$

A Weyl chamber with positive elements is given by

$$\mathfrak{a}^+ = \{H = \mathrm{diag}(\Lambda, -\Lambda, 0, \ldots, 0)\,|\,a_1 > \cdots > a_p > 0\},$$

for the case $p < q$ and

$$\mathfrak{a}^+ = \{H = \mathrm{diag}(\Lambda, -\Lambda)\,|\,a_1 > \cdots > a_p > 0\},$$

for the case $p = q$. Now we turn our attention to the restricted root spaces. We have a few cases to study:

1. For the roots $\pm(\lambda_r - \lambda_s)$, $r \neq s$, $1 \le r, s \le p$, the corresponding root space is formed by matrices of the form

$$\begin{bmatrix} A & 0 & 0 \\ 0 & -A^* & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

   where A has only one non-zero entry, which is non-diagonal.

2. For the roots $\pm(\lambda_r + \lambda_s)$, $r \neq s$, $1 \le r, s \le p$, the corresponding root space is formed by

$$\begin{bmatrix} 0 & B & 0 \\ C & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

   where there is only one non-zero entry (this one being also non-diagonal) in the block B or in the block C.

3. In the case $\pm 2\lambda_r$, $1 \le r \le p$, we have the root space formed by matrices of the

form

$$\begin{bmatrix} 0 & B & 0 \\ C & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where there is only one non-zero entry in the block B or in the block C, and such entry must be diagonal.

4. By last, for $\pm\lambda_r$, $1 \leq r \leq p$, the root space is given by

$$\begin{bmatrix} 0 & 0 & Y^* \\ 0 & 0 & X^* \\ X & Y & 0 \end{bmatrix},$$

where the block X or the block Y has only one non-zero entry.

This allows us to describe the algebras $\mathfrak{n}^+$ and $\mathfrak{n}^-$:

$$\mathfrak{n}^+ = \left\{ \begin{bmatrix} A & B & Y^* \\ 0 & -A^* & 0 \\ 0 & Y & 0 \end{bmatrix} \middle| \begin{array}{c} A \text{ is strictly upper triangular} \\ B^* = -B \\ Y \text{ arbitrary} \end{array} \right\}.$$

$$\mathfrak{n}^- = \left\{ \begin{bmatrix} A & 0 & 0 \\ C & -A^* & X^* \\ X & 0 & 0 \end{bmatrix} \middle| \begin{array}{c} A \text{ is strictly lower triangular} \\ C^* = -C \\ X \text{ arbitrary} \end{array} \right\}.$$

We also have

$$\mathfrak{g}_0 = \mathfrak{z}(\mathfrak{a}) = \left\{ \begin{bmatrix} A & 0 & 0 \\ 0 & -A^* & 0 \\ 0 & 0 & Z \end{bmatrix} \middle| \begin{array}{c} A \text{ is diagonal} \\ Z^* = -Z \end{array} \right\}.$$

In the previous calculations we are considering the general case $p < q$. Note that in the case $p = q$ the blocks X, Y and Z on the third row and on the third column disappear in the above matrices. An Iwasawa decomposition for $\mathfrak{su}(p,q)$ is thus given by

$$\mathfrak{su}(p,q) = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}^+.$$

Since $\mathfrak{k} = \mathfrak{su}(p) \oplus \mathfrak{su}(q) \oplus \mathbb{R}$, we can consider $K = \mathrm{SU}(p) \times \mathrm{SU}(q) \times S^1$. By taking exponentials of elements in $\mathfrak{a}$ we get

$$\mathrm{A} = \left\{ \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \middle| \ \alpha \text{ diagonal.} \right\}.$$

In the same way, we have

$$\mathrm{N} = \exp\left(\mathfrak{n}^{+}\right) = \left\{ \begin{bmatrix} * & * & * \\ 0 & * & 0 \\ 0 & * & 1 \end{bmatrix} \right\},$$

where the sizes of the blocks in this description of N agree with the size of the blocks in $\mathfrak{n}^{+}$. We have then $\mathrm{SU}(p,q) = \mathrm{KAN}$, a global Iwasawa decomposition of the Lie group $\mathrm{SU}(p,q)$. Now, $\mathfrak{m} = \mathfrak{g}_0 \cap \mathfrak{k}$ is given by

$$\mathfrak{m} = \left\{ \begin{bmatrix} \mathbf{i}\Lambda & 0 & 0 \\ 0 & \mathbf{i}\Lambda & 0 \\ 0 & 0 & Z \end{bmatrix} \middle| \ \begin{array}{c} \Lambda \text{ is diagonal with real entries,} \\ Z^* = -Z \end{array} \right\}.$$

The standard minimal parabolic subalgebra is given by

$$\mathfrak{p} = \mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n}^{+} = \left\{ \begin{bmatrix} \Lambda & B & Y^* \\ 0 & -\Lambda^* & 0 \\ 0 & Y & Z \end{bmatrix} \middle| \ \begin{array}{c} \Lambda \text{ is upper triangular,} \\ B^* = -B,\ Z^* = -Z \end{array} \right\}.$$

Given Θ a set of simple roots, we can describe the sets $\mathfrak{n}^{\pm}(\Theta)$. We will look only at the general case $p < q$, since the case $p = q$ is quite similar. We have several cases to study. Considering the simple system of roots

$$\Sigma = \{\lambda_1 - \lambda_2, \ldots, \lambda_{p-1} - \lambda_p, \lambda_p\},$$

we can take

1. $\Theta = \{\lambda_1 - \lambda_2, \ldots, \lambda_{p-1} - \lambda_p\}$, which gives

$$\langle\Theta\rangle^{+} = \{\lambda_i - \lambda_j,\ 1 \le i < j \le p\},$$

and

$$\mathfrak{n}^+(\Theta) = \left\{ \begin{bmatrix} A & 0 & 0 \\ 0 & -A^* & 0 \\ 0 & 0 & 0 \end{bmatrix} \middle| \ A \text{ is strictly upper triangular} \right\}.$$

2. Choosing $\Theta = \{\lambda_2 - \lambda_3, \ldots, \lambda_{p-1} - \lambda_p, \lambda_p\}$, we get

$$\begin{aligned} \langle\Theta\rangle^+ &= \{\lambda_i - \lambda_j,\ 2 \le i < j \le p\} \\ &\cup\{\lambda_i,\ 2\lambda_i,\ 2 \le i \le p\} \\ &\cup\{\lambda_i + \lambda_j,\ 2 \le i < j \le p\}, \end{aligned}$$

so

$$\mathfrak{n}^+(\Theta) = \left\{ \begin{bmatrix} A & B & Y^* \\ 0 & -A^* & 0 \\ 0 & Y & 0 \end{bmatrix} \middle| \begin{array}{c} A \text{ is strictly upper triangular, } a_{12} = 0 \\ \text{first row of } B \text{ equals to zero} \\ \text{first column of } Y \text{ and } B \text{ equals to zero} \end{array} \right\}.$$

3. In the case $\Theta = \{\lambda_p\}$, we have $\langle\Theta\rangle^+ = \{\lambda_p, 2\lambda_p\}$, giving us the set

$$\mathfrak{n}^+(\Theta) = \left\{ \begin{bmatrix} 0 & B & Y^* \\ 0 & 0 & 0 \\ 0 & Y & 0 \end{bmatrix} \middle| \begin{array}{c} b_{pp} \text{ is the only nonzero entry of } B, \text{ and} \\ Y \text{ has nonzero entries only in the } p\text{-th column} \end{array} \right\}.$$

Let us take a closer look at the Lie brackets in this algebra. If

$$M = \begin{bmatrix} 0 & A & X^* \\ 0 & 0 & 0 \\ 0 & X & 0 \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} 0 & B & Y^* \\ 0 & 0 & 0 \\ 0 & Y & 0 \end{bmatrix},$$

then

$$[M,N] = \begin{bmatrix} 0 & X^*Y - Y^*X & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and this implies that the algebra $\mathfrak{n}^+(\Theta)$ is isomorphic to the generalized Heisenberg Lie algebra of dimension $2(q-p)+1$.

5.2 The Lie algebra $\mathfrak{su}(1,2)$

The Hermitian forms $\mathcal{I}_3$ and $\mathcal{J}_3$ having matrices

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{and} \quad J_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

are unitarilly equivalent, and they have signature $(1,2)$. So the Lie algebra $\mathfrak{su}(1,2)$ can be defined using any of them. That is, $\mathfrak{su}(1,2)$ can be defined as being the Lie algebra

$$\mathfrak{su}_I(1,2) = \{X \in M_3(\mathbb{C}) \mid XI_3 + I_3X^* = 0,\ \mathrm{tr}(X) = 0\}$$

or

$$\mathfrak{su}_J(1,2) = \{X \in M_3(\mathbb{C}) \mid XJ_3 + J_3X^* = 0,\ \mathrm{tr}(X) = 0\}.$$

The Lie isomorphism between these two realizations is performed by

$$\begin{aligned} \varphi :\ \mathfrak{su}_I(1,2) &\longrightarrow \mathfrak{su}_J(1,2) \\ X &\longmapsto PXP^{-1}, \end{aligned}$$

where

$$P = \begin{bmatrix} 1/2 & 1/2 & 1/\sqrt{2} \\ 1/2 & 1/2 & -1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}.$$

5.2.1 Cartan and Iwasawa decompositions

In the realization $\mathfrak{su}_I(1,2)$ we have that a general matrix $X \in \mathfrak{su}_I(1,2)$ is of the form

$$X = \begin{bmatrix} \mathbf{i}a & b & \overline{x} \\ \overline{b} & \mathbf{i}d & -\overline{y} \\ x & y & \mathbf{i}z \end{bmatrix},$$

where $a, d, z \in \mathbb{R}$ and $a + d + z = 0$.

In this realization a Cartan decomposition $\mathfrak{su}_I(1,2) = \mathfrak{k} \oplus \mathfrak{s}$ is given by

$$\mathfrak{k} = \left\{ \begin{bmatrix} \mathbf{i}a & 0 & 0 \\ 0 & \mathbf{i}d & -\overline{y} \\ 0 & y & \mathbf{i}z \end{bmatrix} \middle| \; a = -d - z \in \mathbb{R} \right\}$$

and

$$\mathfrak{s} = \left\{ \begin{bmatrix} 0 & b_1 & \overline{b_2} \\ \overline{b_1} & 0 & 0 \\ b_2 & 0 & 0 \end{bmatrix} \middle| \; b_1, b_2 \in \mathbb{C} \right\}.$$

Note that the elements $X \in \mathfrak{k}$ can be written as

$$X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{d-z}{2}\mathbf{i} & -\overline{y} \\ 0 & y & \frac{z-d}{2}\mathbf{i} \end{bmatrix} + \begin{bmatrix} (-d-z)\mathbf{i} & 0 & 0 \\ 0 & \frac{d+z}{2}\mathbf{i} & 0 \\ 0 & 0 & \frac{d+z}{2}\mathbf{i} \end{bmatrix}.$$

And this implies that we can identify $\mathfrak{k} = \mathfrak{su}(2) \oplus \mathfrak{z} = \mathfrak{u}(2)$, where $\mathfrak{z}$ is identified with the one-dimensional center of the unitary Lie algebra $\mathfrak{u}(2)$. From now on we will work with the realization $\mathfrak{su}_J(1,2)$, and we should denote it simply by $\mathfrak{su}(1,2)$. Any mentions to the realization $\mathfrak{su}_I(1,2)$ will be specified. The conditions $XJ_3 + J_3X^*$ and $\mathrm{tr}(X) = 0$ imply that a general element $X \in \mathfrak{su}(1,2)$ must be of the form

$$X = \begin{bmatrix} a & b & \overline{y} \\ c & -\overline{a} & \overline{x} \\ x & y & -2\mathrm{Im}(a) \end{bmatrix},$$

where $b = -\overline{b}$ and $c = -\overline{c}$. A Cartan decomposition for $\mathfrak{su}(1,2)$ is $\mathfrak{su}(1,2) = \mathfrak{k} \oplus \mathfrak{s}$, where

$$\mathfrak{k} = \mathfrak{su}(1,2) \cap \mathfrak{su}(3) = \left\{ \begin{bmatrix} a\mathbf{i} & b\mathbf{i} & -\overline{x} \\ b\mathbf{i} & a\mathbf{i} & \overline{x} \\ x & -x & -2a\mathbf{i} \end{bmatrix} \in M_3(\mathbb{C}) \middle| \; a, b \in \mathbb{R} \right\},$$

$$\mathfrak{s} = \mathfrak{su}(1,2) \cap \mathbf{i}\mathfrak{su}(3) = \left\{ \begin{bmatrix} a & b\mathbf{i} & \overline{x} \\ -b\mathbf{i} & -a & \overline{x} \\ x & x & 0 \end{bmatrix} \in M_3(\mathbb{C}) \middle| \; a, b \in \mathbb{R} \right\}.$$

A maximal abelian subalgebra contained in $\mathfrak{s}$ is

$$\mathfrak{a} = \left\{ \begin{bmatrix} a & 0 & 0 \\ 0 & -a & 0 \\ 0 & 0 & 0 \end{bmatrix} \in M_3(\mathbb{C}) \middle| a \in \mathbb{R} \right\}.$$

A Cartan subalgebra in $\mathfrak{sl}(3,\mathbb{C})$ is the Lie subalgebra $\mathfrak{h}$ formed by the diagonal matrices in $\mathfrak{sl}(3,\mathbb{C})$. The root system associated with this Cartan subalgebra is

$$\Pi_{\mathbb{C}} = \{\pm(\lambda_1 - \lambda_2), \pm(\lambda_1 - \lambda_3), \pm(\lambda_2 - \lambda_3)\}.$$

The simple root system is

$$\Sigma = \{\lambda_1 - \lambda_2, \lambda_2 - \lambda_3\},$$

from which

$$\Pi^+_{\mathbb{C}} = \{\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, \lambda_1 - \lambda_3\}.$$

The set of restricted roots (to the split subalgebra $\mathfrak{a}$) is given by $\Pi = \{\pm\lambda_1, \pm 2\lambda_1\}$, and the simple root system is $\Sigma = \{\lambda_1\}$. The root spaces associated with the previous roots are:

1.

$$\begin{aligned} \mathfrak{g}_{\lambda_1} &= \{X \in \mathfrak{su}(1,2) \,|\, \mathrm{ad}(H)X = \lambda_1(X),\ \forall H \in \mathfrak{a}\} \\ &= \{X \in \mathfrak{su}(1,2) \,|\, [H,X] = \alpha X,\ \forall H = \mathrm{diag}(\alpha, -\alpha, 0) \in \mathfrak{a}\} \\ &= \mathrm{ger}\left\{ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -\mathbf{i} \\ 0 & 0 & 0 \\ 0 & \mathbf{i} & 0 \end{bmatrix} \right\}. \end{aligned}$$

2.

$$\begin{aligned} \mathfrak{g}_{-\lambda_1} &= \{X \in \mathfrak{su}(1,2) \,|\, [H,X] = -\alpha X,\ \forall H = \mathrm{diag}(\alpha, -\alpha, 0) \in \mathfrak{a}\} \\ &= \mathrm{ger}\left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathbf{i} \\ \mathbf{i} & 0 & 0 \end{bmatrix} \right\}. \end{aligned}$$

3.

$$\begin{aligned} \mathfrak{g}_{2\lambda_1} &= \{X \in \mathfrak{su}(1,2) \,|\, [H,X] = 2\alpha X,\ \forall H = \mathrm{diag}(\alpha,-\alpha,0) \in \mathfrak{a}\} \\ &= \mathrm{ger}\left\{\begin{bmatrix} 0 & \mathbf{i} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right\}. \end{aligned}$$

4.

$$\begin{aligned} \mathfrak{g}_{-2\lambda_1} &= \{X \in \mathfrak{su}(1,2) \,|\, [H,X] = -2\alpha X,\ \forall H = \mathrm{diag}(\alpha,-\alpha,0) \in \mathfrak{a}\} \\ &= \mathrm{ger}\left\{\begin{bmatrix} 0 & 0 & 0 \\ \mathbf{i} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right\}. \end{aligned}$$

5.

$$\begin{aligned} \mathfrak{g}_0 = \mathfrak{z}(\mathfrak{a}) &= \{X \in \mathfrak{su}(1,2) \,|\, [H,X] = 0,\ \forall H = \mathrm{diag}(\alpha,-\alpha,0) \in \mathfrak{a}\}\} \\ &= \mathrm{ger}\left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \mathbf{i} & 0 & 0 \\ 0 & \mathbf{i} & 0 \\ 0 & 0 & -2\mathbf{i} \end{bmatrix}\right\}. \end{aligned}$$

We have also $\mathfrak{n}^+ = \mathfrak{g}_{\lambda_1} \oplus \mathfrak{g}_{2\lambda_1}$, that is,

$$\mathfrak{n}^+ = \left\{\begin{bmatrix} 0 & \mathbf{i}b & \overline{x} \\ 0 & 0 & 0 \\ 0 & x & 0 \end{bmatrix} \,\middle|\, x \in \mathbb{C},\ b \in \mathbb{R}\right\}.$$

Consider the basis for $\mathfrak{n}^+$ formed by

$$A = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & -\mathbf{i} \\ 0 & 0 & 0 \\ 0 & \mathbf{i} & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & \mathbf{i} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that $[A,B] = C$, $[A,C] = 0$, $[B,C] = 0$. If we choose in the Heisenberg Lie algebra $Heis$ the basis formed by E_{12}, E_{23} and E_{13}, we can see that $\mathfrak{n}^+$ and $Heis$ are isomorphic via

$$A \longmapsto E_{12}, \quad B \longmapsto E_{23}, \quad C \longmapsto E_{13}.$$

Explicitly, the isomorphism $\varphi : \mathfrak{n}^+ \to Heis$ is given by

$$\varphi \begin{bmatrix} 0 & c\mathbf{i} & a - b\mathbf{i} \\ & 0 & 0 \\ 0 & a + b\mathbf{i} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{bmatrix}.$$

Remembering that $\mathfrak{k} = \mathfrak{u}(2)$, we obtain the Iwasawa decomposition for $\mathfrak{su}(1,2)$:

$$\mathfrak{su}(1,2) = \mathfrak{u}(2) \oplus \mathfrak{a} \oplus Heis.$$

5.2.2 The flag manifold of $\mathrm{SU}(1,2)$

The standard minimal parabolic subalgebra of $\mathfrak{su}(1,2)$ is given by $\mathfrak{p} = \mathfrak{m} \oplus \mathfrak{a} \oplus Heis$, where

$$\mathfrak{m} = \mathfrak{g}_0 \cap \mathfrak{k} = \left\{ \begin{bmatrix} a\mathbf{i} & 0 & 0 \\ 0 & a\mathbf{i} & 0 \\ 0 & 0 & -2a \end{bmatrix} \middle| \; a \in \mathbb{R} \right\}.$$

that is

$$\begin{aligned} \mathfrak{p} &= \left\{ \begin{bmatrix} x & b\mathbf{i} & \overline{y} \\ 0 & -\overline{x} & 0 \\ 0 & y & -2\mathrm{Im}(x) \end{bmatrix} \middle| \; b \in \mathbb{R}, \; x, y \in \mathbb{C} \right\} \\ &= \left\{ \begin{bmatrix} x & 0 & 0 \\ 0 & -\overline{x} & 0 \\ 0 & 0 & -2\mathrm{Im}(x) \end{bmatrix} \middle| \; x \in \mathbb{C} \right\} \oplus Heis. \end{aligned}$$

Since the only subalgebra of $\mathfrak{su}(1,2)$ containing $\mathfrak{p}$ is $\mathfrak{su}(1,2)$ itself, there are no parabolic subalgebras other than the standard minimal one. This implies that $\mathrm{SU}(1,2)$ has only one flag manifold , that we will characterize now.

Proposition 1.3.1 implies that $\mathbb{F}_\Theta = K/K_\Theta$, where the stabilizer is given by $K_\Theta = K \cap P_\Theta$. Viewing $\mathbb{F}_\Theta$ as an homogeneous space of K we can identify it as a K-orbit under the adjoint representation. Specifically for the group $SU(1,2)$ we have

Proposition 5.2.1. *The only flag manifold $\mathbb{F}$ of $SU(1,2)$ embeds in the component $\mathfrak{s}$ of the Cartan decomposition as the $\mathrm{Ad}(U(2))$-orbit of $H \in \mathfrak{a}^+$.*

Proof: The component $\mathfrak{s}$ of the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s} = \mathfrak{u}(2) \oplus \mathfrak{s}$ is invariant under the adjoint representation of $K = U(2)$. Since $\mathfrak{a}$ is a one dimensional subalgebra, a Weyl chamber $\mathfrak{a}^+ \subset \mathfrak{a}$ is just a ray starting at the origin. Choosing an element $H \in \mathfrak{a}^+$, the stabilizer of H under the adjoint action of K on $\mathfrak{s}$ is the centralizer K_H of H in K, which is given by $K \cap P$. It follows that the adjoint orbit $\mathrm{Ad}(U(2))H \subset \mathfrak{s}$ is identified with the coset space K/K_H, and this one is the same as the flag manifold $\mathbb{F} = G/P$. □

To get a better description of the flag manifold $\mathbb{F}$ of $SU(1,2)$ we take a look at the stabilizer of $H \in \mathfrak{a}^+$ under the adjoint action of K. If the Lie group $SU(1,2)$ is realized via the hermitian form $\mathcal{I}_3$ we get

$$\mathfrak{k} = \left\{ \begin{bmatrix} \mathbf{i}a & 0 & 0 \\ 0 & \mathbf{i}d & -\overline{y} \\ 0 & y & \mathbf{i}z \end{bmatrix} \middle| \; a = -d - z \in \mathbb{R} \right\} = \mathfrak{u}(2),$$

and thus the group K is of the form

$$K = \left\{ \begin{bmatrix} \det(u^*) & 0 \\ 0 & u \end{bmatrix} \middle| \; u \in U(2) \right\}.$$

Remember that $H \in \mathfrak{a}^+$ in this realization is of the form

$$H = \begin{bmatrix} 0 & a & 0 \\ a & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad a > 0.$$

Writing

$$x = \begin{bmatrix} \det(u^*) & 0 \\ 0 & u \end{bmatrix},$$

where

$$u = \begin{bmatrix} m & n \\ p & q \end{bmatrix},$$

we have

$$xH = Hx \iff \begin{bmatrix} 0 & a\det(u^*) & 0 \\ am & 0 & 0 \\ ap & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & am & an \\ a\det(u^*) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

from which

$$\begin{cases} a\det(u^*) = am \\ ap = an = 0 \end{cases} \Longrightarrow \begin{cases} p = n = 0 \\ m = \det(u^*) \end{cases} \quad \text{(since } a \neq 0).$$

Now we have $1 = \det(u^*)mq = mmq$, that is, $q = 1/m^2$. By the other hand,

$$1 = \det(u^*)mq = \overline{m}\,\overline{q}mq = |mq|^2 = \left|\frac{1}{m}\right|^2 \iff |m| = 1.$$

This means that $x \in K_H$ is of the form

$$x = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & 1/m^2 \end{bmatrix}, \quad \text{with } |m| = 1.$$

Hence, the stabilizer

$$K_H = \{x \in K \mid \mathrm{Ad}(x)H = H\} = \{x \in K \mid xHx^{-1} = H\}$$

can be seen as the circle group $U(1) = \{m \in \mathbb{C}\,;\,|m| = 1\}$. Now, $U(n)/U(n-1) = S^{2n-1}$ and this implies that

$$\mathbb{F} = \mathrm{Ad}(K)H = K/K_H = U(2)/U(1) = S^3,$$

that is, the only flag manifold of $SU(1,2)$ is the sphere S^3.

5.2.3 Cartan subalgebras and Jordan-Schur decompositions

There are only two conjugation classes of Cartan subalgebras in $\mathfrak{su}(1,2)$ (see Sugiura [33]). It is easy to see that

$$\mathfrak{h}_1 = \left\{ \begin{bmatrix} a+\mathbf{i}b & 0 & 0 \\ 0 & -a+\mathbf{i}b & 0 \\ 0 & 0 & -2\mathbf{i}b \end{bmatrix} \in M_3(\mathbb{C}) \;\middle|\; a,b \in \mathbb{R} \right\} = \mathfrak{a} + \mathfrak{h}_{\mathfrak{k}}$$

is a Cartan subalgebra in $\mathfrak{su}(1,2)$, where

$$\mathfrak{h}_{\mathfrak{k}} = \left\{ \begin{bmatrix} \mathbf{i}b & 0 & 0 \\ 0 & \mathbf{i}b & 0 \\ 0 & 0 & -2\mathbf{i}b \end{bmatrix} \in M_3(\mathbb{C}) \;\middle|\; b \in \mathbb{R} \right\}.$$

Now, we have that

$$\mathfrak{h}_2 = \left\{ \begin{bmatrix} \mathbf{i}b & \mathbf{i}c & 0 \\ \mathbf{i}c & \mathbf{i}b & 0 \\ 0 & 0 & -2\mathbf{i}b \end{bmatrix} \in M_3(\mathbb{C}) \;\middle|\; b,c \in \mathbb{R} \right\}.$$

is another Cartan subalgebra of $\mathfrak{su}(1,2)$, since it is abelian and $X \in N_{\mathfrak{g}}(\mathfrak{h}_2)$ if and only if $X \in \mathfrak{h}_2$. For the Cartan subalgebra $\mathfrak{h}_1$ its decomposition into toroidal part $\mathfrak{h}_1^+$ and vector part $\mathfrak{h}_1^-$ is given by $\mathfrak{h}_1^+ = \mathfrak{h}_{\mathfrak{k}}$ and $\mathfrak{h}_2^- = \mathfrak{a}$ (in fact, $\mathfrak{h}_1$ is a maximal abelian subalgebra in $\mathfrak{su}(1,2)$ containing $\mathfrak{a}$). For the Cartan subalgebra $\mathfrak{h}_2$ the toroidal part is given by $\mathfrak{h}_2^+ = \mathfrak{h}_2$ and the vector part $\mathfrak{h}_2^- = \{0\}$. For the lie algebra $\mathfrak{su}(p,q)$ we know that two Cartan subalgebras are conjugate under an inner automorphism if and only if their toroidal parts have the same dimension. This means that $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are not conjugate Cartan subalgebras and so we can choose them as representative Cartan subalgebras of their own conjugate classes. We now take a look at the Jordan-Schur decompositions of the elements belonging to $\mathfrak{h}_1$ and $\mathfrak{h}_2$. If $X_s \in \mathfrak{h}_1$, its Jordan-Schur decomposition is given by $X_s = X_h + X_e$, where the hyperbolic part $X_h \in \mathfrak{a}$ and the elliptic part $X_e \in \mathfrak{h}_{\mathfrak{k}}$. If $X_s \in \mathfrak{h}_2$ then X_s is an elliptic element, and so its Jordan-Schur decomposition is $X_s = X_h$. Since we are working on a semisimple Lie algebra (a simple one to be precise), the Cartan subalgebras can be expressed as being the maximal toral subalgebras (a subalgebra is said to be toral if it consists only of semisimple elements). Since toral subalgebras are abelian (cf. Humphreys, Chapter 8) we have that for every semisimple element X_s there is a Cartan subalgebra containing it. In fact, if there is no Cartan subalgebra containing it, we can take a Cartan subalgebra $\mathfrak{h}$ and consider $\tilde{\mathfrak{h}}$ the subalgebra generated by $\mathfrak{h}$ and X_s. Clearly $\tilde{\mathfrak{h}}$ is a toral subalgebra containing $\mathfrak{h}$, which contradicts the maximality of $\mathfrak{h}$. In this way, for an arbitrary element $X \in \mathfrak{su}(1,2)$ we have that the semisimple part X_s in its Jordan decomposition is conjugate to an element belonging to $\mathfrak{h}_1$ or belonging to $\mathfrak{h}_2$. Since the Jordan-Schur decomposition $X = X_h + X_e + X_n$ is unique for every $X \in \mathfrak{su}(1,2)$ and the components X_h, X_e and X_n commute with each other, we are lead to know all the possible Jordan-Schur decompositions for the elements $X \in \mathfrak{su}(1,2)$. Write $X = X_s + X_n$, where $X_s = X_h + X_e$. The possibilities are then the following

1. $X_s = 0$.
   - Then $X = X_n$, with X_n conjugate to an element in $\mathfrak{n}^+$ (which is isomorphic to the Heisenberg Lie algebra).
2. $X_s \neq 0$, $X_s \in \mathfrak{h}_2$.
   - Then $X = X_e$, with X_e conjugate to an element in $\mathfrak{h}_2$.
3. $X_s \neq 0$, $X_s \in \mathfrak{h}_1$. In this case the possibilities are
   - $X = X_e$, with X_e conjugate to an element in $\mathfrak{h}_1$.
   - $X = X_h$, with X_h conjugate to an element in $\mathfrak{h}_1$.
   - $X = X_e + X_h$, with $X_e + X_h$ conjugate to an element in $\mathfrak{h}_1$.
   - $X = X_h + X_n$, with X_h conjugate to an element in $\mathfrak{h}_1$ and X_n of the form

$$X_n = \begin{bmatrix} 0 & \mathbf{i}c & 0 \\ \mathbf{i}d & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

5.2.4 Elements of $\mathfrak{su}(1,2)$ as vector fields on the sphere S^3

The subspace formed by the imaginary quaternions $\mathrm{Im}\mathbb{H}$ endowed with the brackets $[z,w] = zw - wz$ is a Lie algebra, and this Lie algebra is isomorphic to $\mathfrak{su}(2)$. In fact, in

$\mathrm{Im}\mathbb{H}$ we have

$$[\mathbf{i},\mathbf{j}] = 2\mathbf{k}, \quad [\mathbf{j},\mathbf{k}] = 2\mathbf{i} \ \text{ and } \ [\mathbf{i},\mathbf{k}] = -2\mathbf{j},$$

while in $\mathfrak{su}(2)$ for the basis

$$A = \begin{bmatrix} 0 & \mathbf{i} \\ \mathbf{i} & 0 \end{bmatrix}, \ B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \ C = \begin{bmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{bmatrix},$$

we have

$$[A,B] = 2C \ \ [B,C] = 2A \ \text{ and } \ [A,C] = -2B.$$

This implies that $\mathrm{Im}\mathbb{H} = \mathfrak{su}(2)$. The Lie algebra $\mathrm{Im}\mathbb{H}$ is represented in $\mathbb{H}$ through left multiplication and also through right multiplication:

$$\begin{array}{rcl} E : \mathrm{Im}\mathbb{H} & \longrightarrow & \mathfrak{gl}(\mathbb{H}) \\ z & \longmapsto & E_z \end{array} \quad \text{and} \quad \begin{array}{rcl} D : \mathrm{Im}\mathbb{H} & \longrightarrow & \mathfrak{gl}(\mathbb{H}) \\ z & \longmapsto & D_z \end{array}$$

where $E_z, D_z : \mathbb{H} \longrightarrow \mathbb{H}$ are given respectively by $E_z(q) = zq$ and $D_z(q) = qz$, $q \in \mathbb{H}$. The representations E and D are faithful and commute, that is, for all $z \in \mathrm{Im}\mathbb{H}$ we have

$$E_z \circ D_z(q) = E_z(qz) = zqz = D_z(zq) = D_z \circ E_z(q), \quad \forall q \in \mathbb{H}$$

and if $E_z = D_z = 0$ then we must have $z = 0$, that is, $\ker(E) = \ker(D) = \{0\}$. This fact implies that the Lie algebra $\mathfrak{u}(2)$ is isomorphic to the Lie algebra of linear transformations

$$\{E_z + aD_{\mathbf{i}} \,|\, z \in \mathrm{Im}\mathbb{H}, \ a \in \mathbb{R}\},$$

where the center $\mathfrak{z}$ is given by the transformations $aD_{\mathbf{i}}$, $a \in \mathbb{R}$. The group of the unit quaternions, that is, the sphere

$$S^3 = \{q \in \mathbb{H}\,;\ \|q\|^2 = q\overline{q} = 1\}$$

is a Lie group isomorphic to $\mathrm{SU}(2)$. This can be seen via the identification

$$\mathbf{1} \Leftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \mathbf{i} \Leftrightarrow \begin{bmatrix} 0 & \mathbf{i} \\ \mathbf{i} & 0 \end{bmatrix}, \ \mathbf{j} \Leftrightarrow \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \ \mathbf{k} \Leftrightarrow \begin{bmatrix} -\mathbf{i} & 0 \\ 0 & \mathbf{i} \end{bmatrix}.$$

Remember that an infinitesimal action of $\mathfrak{su}(1,2)$ on S^3 is a homomorphism $\mathfrak{su}(1,2) \to \Gamma(TS^3)$, where $\Gamma(TS^3)$ stands for the Lie algebra of vector fields on S^3. By means of an infinitesimal action of $\mathfrak{su}(1,2)$ on S^3 one can see the Lie algebra $\mathfrak{su}(1,2)$ as a Lie algebra of vector fields on the sphere S^3. Considering the Cartan decomposition $\mathfrak{su}(1,2) = \mathfrak{k} \oplus \mathfrak{s} = \mathfrak{u}(2) \oplus \mathbb{H}$, the following propositions give us a better description of these vector fields.

Proposition 5.2.2. *The elements in the compact component $\mathfrak{k}$ under the infinitesimal action of $\mathfrak{su}(1,2)$ induced by the action of $\mathrm{SU}(1,2)$ on S^3 have the form*

$$X_{(0,z,a)}(x) = zx + ax\mathbf{i}, \qquad x \in S^3,$$

where $z \in \mathbf{i}\mathbb{H} = \mathfrak{su}(2)$ and $a \in \mathbb{R}$.

Proof: Consider the adjoint representation

$$\begin{aligned} \mathrm{Ad} : K &\to \mathrm{Gl}(\mathfrak{su}(1,2)) \\ k &\mapsto \mathrm{Ad}(k) = d(C_k)_1. \end{aligned}$$

We know that it is differentiable and gives rise to the action

$$K \times \mathfrak{s} \to \mathfrak{s}, \;\; (k,x) \mapsto \mathrm{Ad}(k)x.$$

Since the flag S^3 embeds in $\mathfrak{s}$ as an $\mathrm{Ad}(K)$-orbit, we can consider the restriction of the above action to S^3 (viewed as an $\mathrm{Ad}(K)$-orbit) to get an infinitesimal action of $\mathfrak{k}$ on S^3. Thus, for $X \in \mathfrak{k}$ the corresponding infinitesimal action on S^3 is given by

$$\tilde{X}(x) = d(\mathrm{Ad})_1(X)x = \mathrm{ad}(X)x,$$

that is, the induced vector field is given by the adjoint action of $\mathfrak{k}$ on $\mathfrak{s}$. Recall that $\mathfrak{k} = \mathfrak{u}(2) = \mathfrak{su}(2) \oplus \mathfrak{z}$. The adjoint representation of $\mathfrak{su}(2) \subset \mathfrak{k}$ in $\mathfrak{s}$ corresponds to left multiplication by immaginary quaternions in $\mathbb{H}$. We can check this as follows. For an element

$$Y = \begin{bmatrix} 0 & z_1 & \overline{z_2} \\ \overline{z_1} & 0 & 0 \\ z_2 & 0 & 0 \end{bmatrix} \in \mathfrak{s},$$

denoting $z_1 = p + q\mathbf{i}$ and $z_2 = r + s\mathbf{i}$ we can write

$$\begin{bmatrix} 0 & z_1 & \overline{z_2} \\ \overline{z_1} & 0 & 0 \\ z_2 & 0 & 0 \end{bmatrix} \in \mathfrak{s} \leftrightarrow (z_1, z_2) \in \mathbb{C}^2 \leftrightarrow (p, q, r, s) \in \mathbb{H}.$$

So, if we take

$$z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -a\mathbf{i} & -\overline{y} \\ 0 & y & a\mathbf{i} \end{bmatrix} \in \mathfrak{su}(2), \qquad y = c + d\mathbf{i},$$

we get

$$\mathrm{ad}(z)(Y) = \begin{bmatrix} 0 & a\mathbf{i}z_1 - \overline{z_2}y & z_1\overline{y} - a\mathbf{i}\overline{z_2} \\ -\overline{y}z_2 - a\mathbf{i}\overline{z_1} & 0 & 0 \\ \overline{z_1}y + a\mathbf{i}z_2 & 0 & 0 \end{bmatrix},$$

from which

$$\begin{bmatrix} 0 & a\mathbf{i}z_1 - \overline{z_2}y & z_1\overline{y} - a\mathbf{i}\overline{z_2} \\ -\overline{y}z_2 - a\mathbf{i}\overline{z_1} & 0 & 0 \\ \overline{z_1}y + a\mathbf{i}z_2 & 0 & 0 \end{bmatrix} \leftrightarrow (a\mathbf{i}z_1 - \overline{z_2}y, \overline{z_1}y + a\mathbf{i}z_2) \in \mathbb{C}^2$$

$$\leftrightarrow (-aq - cr - ds + (ap + cs - dr)\mathbf{i}, cp + dq - as + (ar + dp - cq)\mathbf{i}) \in \mathbb{C}^2.$$

$$\leftrightarrow (-aq - cr - ds, ap + cs - dr, cp + dq - as, ar + dp - cq) \in \mathbb{H}.$$

By the other hand, if we multiply $z = (0, a, c, d) \in \mathrm{Im}(\mathbb{H})$ on the left by $(p, q, r, s) \in \mathbb{H}$, we obtain

$$(0, a, c, d) \cdot (p, q, r, s) = (-aq - cr - ds, ap + cs - dr, cp + dq - as, ar + dp - cq),$$

that is, $\mathrm{ad}(z)$ corresponds to left multiplication by the immaginary quaternion z on $\mathbb{H}$, and the vector field induced by z on S^3 is given by $\tilde{X}_z(x) = zx$, $x \in S^3$. For the elements belonging to $\mathfrak{z}$, we describe the adjoint action in the following way. For

$$Z = \begin{bmatrix} 2a\mathbf{i} & 0 & 0 \\ 0 & -a\mathbf{i} & 0 \\ 0 & 0 & -a\mathbf{i} \end{bmatrix} \in \mathfrak{z}$$

we have

$$\mathrm{ad}(Z)(Y) = \begin{bmatrix} 0 & 3a\mathbf{i}z_1 & 3a\mathbf{i}\overline{z_2} \\ -3\mathbf{i}a\overline{z_1} & 0 & 0 \\ -3\mathbf{i}az_2 & 0 & 0 \end{bmatrix}$$

from which

$$\begin{bmatrix} 0 & 3a\mathbf{i}z_1 & 3a\mathbf{i}\overline{z_2} \\ -3\mathbf{i}a\overline{z_1} & 0 & 0 \\ -3\mathbf{i}az_2 & 0 & 0 \end{bmatrix} \leftrightarrow (3a\mathbf{i}z_1, -3a\mathbf{i}z_2) \leftrightarrow (-3aq, 3ap, 3as, -3ar).$$

By the other hand,

$$3a(p,q,r,s)\mathbf{i} = (-3aq, 3ap, 3as, -3ar),$$

that is, $\mathrm{ad}(Z)$ can be identified with right multiplication by $3a\mathbf{i}$ on $\mathbb{H}$, and the vector field induced by X on S^3 is thus given by $\tilde{X}_a(x) = 3ax\mathbf{i}$, $x \in S^3$. This implies that the induced vector field corresponding to a general element in $\mathfrak{k}$ has the form $\tilde{X}_z + \tilde{X}_a$.

For the $\mathfrak{s}$ component, there exists a K-invariant Riemannian metric such that for every $q \in \mathfrak{s} = \mathbb{H}$ the vector field $\tilde{X}_q$ induced by q on S^3 is the gradient of the height function $f_q(\cdot) = \langle q, \cdot \rangle$ with respect to this K-invariant metric (see [7] and [34] for details). □

In what concerns the elements in the $\mathfrak{s}$ component, there exists a K-invariant Riemannian metric such that for every $q \in \mathbb{H} = \mathfrak{s}$ the vector field $\tilde{X}_q$ induced by q on S^3 is the gradient of the height function $f_q(\cdot) = \langle q, \cdot \rangle$ with respect to this K-invariant metric (see [7] and [34] for details), which is called a Borel metric.

To get the expression for the vector fields $\tilde{X}_{(q,0,0)} = \tilde{X}_q$ induced by $q \in \mathfrak{s}$ in S^3 we will compare the expressions for the gradient of the height function f_q with respect to both the canonical Riemannian metric (immersion) and the Borel metric in S^3.

We begin considering S^3 endowed with the canonical metric. The height function f_q is linear on $\mathfrak{s}$, so its gradient vector field evaluated at $p \in S^3$ is obtained from the orthogonal projection of q over p. In fact, $(\mathrm{grad} f_q)_p = d(f_q)_p(v)$ is the cotangent vector ω such that $\omega(v) = \langle q, v \rangle$, that is, the cotangent vector ω such that $\langle \omega, v \rangle = \langle q, v \rangle$, for all

$v \in T_pS^3$. Since $\langle q - \langle q,p\rangle p, v\rangle = \langle q, v\rangle \ \ \forall v \in T_pS^3$, we get

$$(\mathrm{grad} f_q)_p = q - \langle q,p\rangle p.$$

The vector field $\tilde{X}_q$ is thus given by

$$\begin{aligned} \tilde{X}_q(p) &= q - \langle q,p\rangle p \\ &= q - \frac{1}{2}\left(q\overline{p} + p\overline{q}\right)p \\ &= q - \frac{1}{2}(q\overline{p}p + p\overline{q}p), \qquad p \in S^3. \end{aligned}$$

Since $\|p\| = p\overline{p} = 1$, we get

$$\tilde{X}_q(p) = \frac{1}{2}(q - p\overline{q}p).$$

Finally, a general element $X \in \mathfrak{su}(1,2)$ can be decomposed as

$$X = q + k = q + z + a\mathbf{i},$$

where $k = z + a\mathbf{i} \in \mathfrak{k} = \mathfrak{su}(2) + \mathfrak{z}$ ($z \in \mathfrak{su}(2)$ and $a\mathbf{i} \in \mathfrak{z}$) and $q \in \mathfrak{s}$. This implies that the vector field induced by X on S^3 has the form

$$\tilde{X}_{(q,z,a)}(x) = \frac{1}{2}(q - x\overline{q}x) + zx + ax\mathbf{i}, \quad x \in S^3.$$

Now we consider the Borel metric, which is defined by the inner product in the origin ($\mathfrak{su}(2) = \mathrm{Im}\mathbb{H}$) such that $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ is an orthogonal basis satisfying

$$\|\mathbf{j}\|_B = \|\mathbf{k}\|_B = 1 \quad \|\mathbf{i}\|_B^2 = 2.$$

This because the Borel metric arises from the fact that S^3 is a flag manifold of $\mathfrak{su}(1,2)$. The Borel metric is the K-invariant metric in the flag S^3 such that the a subspace $\mathfrak{k}_\alpha = \mathfrak{k} \cap (\mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha})$ is orthogonal with $\mathfrak{k}_\beta$ if $\alpha \neq \beta$ and such that $X \in \mathfrak{k}_\alpha$ has norm

$$\|X\|_B^2 = \alpha(H),$$

where $H \in \mathfrak{a}^+$ is such that $S^3 = \mathrm{Ad}(K)H$ (see Proposition 5.2.1).

For convenience we can rewrite the realization $\mathfrak{su}_I(1,2)$ as

$$\begin{bmatrix} \mathbf{i}a & z \\ \overline{z}^t & A \end{bmatrix}, \text{ where } a \in \mathbb{R},\ z \in \mathbb{C}^2 = \mathbb{H} = \mathbb{R}^4,\ A \in \mathfrak{u}(2).$$

The origin of the flag can be chosen as being H given by $a = 0$, $A = 0$ and $z = (1,0) \in \mathbb{C}^2$. The positive roots are α and 2α, $\alpha(H) = 1$, and their multiplicities are respectively 2 and 1. The root spaces in $\mathfrak{k}$ are then $\mathfrak{k}_\alpha = \mathrm{span}\{\mathbf{j},\mathbf{k}\}$ and $\mathfrak{k}_{2\alpha} = \mathrm{span}\{\mathbf{i}\}$, which justifies the above expression for the Borel metric.

Now we can compare what happens with the gradient vector fields in the canonical and the Borel metrics.

In general a left invariant metric $(\cdot,\cdot)$ on a Lie group G is given by an inner product in the origin. If $\{Z_1,\ldots,Z_n\}$ is an orthonormal basis with respect to this inner product, then the left invariant vector fields $\{Z_1^l,\ldots,Z_n^l\}$ form an orthonormal referential of the Riemannian metric. The gradient $\mathrm{grad} f$ of a given function f on G with respect to $(\cdot,\cdot)$ is thus

$$\mathrm{grad} f = (Z_1 f)Z_1 + \cdots + (Z_n f)Z_n,$$

where Z_f indicates the directional derivative $df(Z)$.

In the following we are going to denote by $\mathrm{grad}_C f$ the gradient with respect to the canonical metric on S^3 and by $\mathrm{grad}_B f$ the gradient with respect to the Borel metric. Let $\{Z_\mathbf{i}, Z_\mathbf{j}, Z_\mathbf{k}\}$ be the left invariant vector fields corresponding to the basis $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ of $\mathfrak{su}(2) = \mathrm{Im}\mathbb{H}$. Then $\{Z_\mathbf{i}, Z_\mathbf{j}, Z_\mathbf{k}\}$ is orthonormal with respect to the canonical metric while $\{\frac{1}{\sqrt{2}}Z_\mathbf{i}, Z_\mathbf{j}, Z_\mathbf{k}\}$ is orthonormal with respect to the Borel metric. Hence we have

$$\begin{aligned} \mathrm{grad}_B f &= \frac{1}{\sqrt{2}}\left(Z_\mathbf{i} f\right) Z_\mathbf{i} + \left(Z_\mathbf{j} f\right) Z_\mathbf{j} + \left(Z_\mathbf{k} f\right) Z_\mathbf{k} \\ &= \left(\frac{1}{\sqrt{2}} - 1\right)\left(Z_\mathbf{i} f\right) Z_\mathbf{i} + \left(Z_\mathbf{i} f\right) Z_\mathbf{i} + \left(Z_\mathbf{j} f\right) Z_\mathbf{j} + \left(Z_\mathbf{k} f\right) Z_\mathbf{k} \\ &= \left(\frac{1}{\sqrt{2}} - 1\right)\left(Z_\mathbf{i} f\right) Z_\mathbf{i} + \mathrm{grad}_C f. \end{aligned}$$

Since $Z_\mathbf{i} f = (\mathrm{grad}_C f, Z_\mathbf{i})$ the above formula becomes

$$\mathrm{grad}_B f = \mathrm{grad}_C f + \left(\frac{1}{\sqrt{2}} - 1\right)\left(\mathrm{grad}_C f, Z_\mathbf{i}\right) Z_\mathbf{i} = \mathrm{grad}_C f + \frac{\sqrt{2}-2}{2}\left(\mathrm{grad}_C f, Z_\mathbf{i}\right) Z_\mathbf{i}.$$

which correlates $\mathrm{grad}_B f$ and $\mathrm{grad}_C f$.

Regarding $\mathrm{SU}(2)$ as the unit sphere of $\mathbb{H}$ endowed with the quaternionic product, the left invariant vector fields are the linear vector fields given by right multiplication by imaginary quaternions (see [28], Section 5.2.1). In special,

$$Z_{\mathbf{i}}(p) = p\mathbf{i} \qquad Z_{\mathbf{j}}(p) = p\mathbf{j} \qquad Z_{\mathbf{k}}(p) = p\mathbf{k}.$$

And this allows us to rewrite the above gradients in $p \in S^3$ as

$$\mathrm{grad}_C f(p) = \frac{\partial_d f}{\partial \mathbf{i}}(p)\, p\mathbf{i} + \frac{\partial_d f}{\partial \mathbf{j}}(p)\, p\mathbf{j} + \frac{\partial_d f}{\partial \mathbf{k}}(p)\, p\mathbf{k}$$

where

$$\frac{\partial_d f}{\partial \mathbf{i}}(p) = \frac{d}{dt} f\left(pe^{\mathbf{i}t}\right)\Big|_{t=0} \qquad \frac{\partial_d f}{\partial \mathbf{j}}(p) = \frac{d}{dt} f\left(pe^{\mathbf{j}t}\right)\Big|_{t=0} \qquad \frac{\partial_d f}{\partial \mathbf{i}}(p) = \frac{d}{dt} f\left(pe^{\mathbf{k}t}\right)\Big|_{t=0}.$$

The gradient with respect to the Borel metric is then

$$\mathrm{grad}_B f(p) = \left(\frac{1}{\sqrt{2}} - 1\right) \frac{\partial_d f}{\partial \mathbf{i}}(p)\, p\mathbf{i} + \mathrm{grad}_C f(p).$$

For the height function $f_q(p) = \langle q, p \rangle$, $p \in S^3$ and $q \in \mathbb{H}$ fixed, we already have its expression in the canonical metric, that is,

$$\mathrm{grad}_C f_q(p) = \frac{1}{2}(q - p\overline{q}p).$$

However,

$$\frac{\partial_d f_q}{\partial \mathbf{i}}(p) = \frac{d}{dt} f_q\left(pe^{\mathbf{i}t}\right)\Big|_{t=0} = \frac{d}{dt}\langle q, pe^{\mathbf{i}t}\rangle\Big|_{t=0} = \langle q, p\mathbf{i}\rangle = \frac{1}{2}(-q\mathbf{i}\overline{p} + p\mathbf{i}\overline{q})$$

which leads us to

$$\begin{aligned}
\mathrm{grad}_B f_q(p) &= \frac{1}{2}(q - p\overline{q}p) + \frac{\sqrt{2}-2}{4}(-q\mathbf{i}\overline{p} + p\mathbf{i}\overline{q})\, p\mathbf{i} \\
&= \frac{1}{2}(q - p\overline{q}p) + \frac{\sqrt{2}-2}{4} p\mathbf{i}\overline{q}p\mathbf{i} + \frac{\sqrt{2}-2}{4} q.
\end{aligned}$$

For example, if $q = \mathbf{1}$ the corresponding vector field is given by

$$\mathrm{grad}_B f_{\mathbf{1}}(p) = \frac{1}{2}(\mathbf{1} - p^2) + \frac{\sqrt{2}-2}{4}\left((p\mathrm{i})^2 + \mathbf{1}\right).$$

By attempting to understand the behavior of this vector field, we can use the stereographic projection to visualize it as a vector field in a three-dimensional space.

To achieve this, let us momentarily consider $\mathbf{1} \in \mathbb{H}$ as the north pole of the sphere $S^3 = \{q \in \mathbb{H}\,;\, |q| = 1\}$. Moreover, let us identify the pure quaternions $\mathrm{Im}\mathbb{H}$ with the hyperplane

$$\mathrm{Im}\mathbb{H} = \{z \in \mathbb{H}\,;\, z = 0 + u\mathbf{i} + v\mathbf{j} + w\mathbf{k}\} = \{(u,v,w) \in \mathbb{R}^3\,;\, u,v,w \in \mathbb{R}\} = \mathbb{R}^3.$$

Let $\pi_{\mathbf{1}} : S^3\backslash\{\mathbf{1}\} \to \mathbb{R}^3$ be the stereographic projection from the north pole $\mathbf{1}$ onto $\mathrm{Im}\mathbb{H}$, defined as

$$\pi_{\mathbf{1}}(p) = z, \;\; p \in S^3, \;\; z \in \mathrm{Im}\mathbb{H},$$

where $z \in \mathrm{Im}\mathbb{H}$ is the unique point in the intersection $r_p^{\mathbf{1}} \cap \mathrm{Im}\mathbb{H}$, being $r_p^{\mathbf{1}}$ the straight line in $\mathbb{H}$ containing $\mathbf{1}$ and p.

An expression for $\pi_{\mathbf{1}}(p) = (u,v,w)$ in terms of $p = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in S^3$ can be written as follows. Let

$$r_p^{\mathbf{1}} = \{1 + t\mathbf{n}\,;\, \mathbf{n} = p - \mathbf{1} \text{ and } t \in \mathbb{R}\} = \{(1 + t(a-1)) + tb\mathbf{i} + tc\mathbf{j} + td\mathbf{k}\,;\, t \in \mathbb{R}\} \subset \mathbb{H}.$$

The straight line $r_p^{\mathbf{1}}$ intersects $\mathrm{Im}\mathbb{H}$ when $1 + t(a-1) = 0$, that is, when $t = {}^{1}/{}_{1-a}$. Therefore, if $p = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in S^3$, we obtain

$$\pi_p^{\mathbf{1}}(p) = u\mathbf{i} + v\mathbf{j} + w\mathbf{k} = \frac{b}{1-a}\mathbf{i} + \frac{c}{1-a}\mathbf{j} + \frac{d}{1-a}\mathbf{k} \in \mathrm{Im}\mathbb{H}.$$

Consequently,

$$D\pi_{\mathbf{1}} = \begin{bmatrix} \nabla u \\ \nabla v \\ \nabla w \end{bmatrix} = \begin{bmatrix} \frac{b}{(1-a)^2} & \frac{1}{1-a} & 0 & 0 \\ \frac{c}{(1-a)^2} & 0 & \frac{1}{1-a} & 0 \\ \frac{d}{(1-a)^2} & 0 & 0 & \frac{1}{1-a} \end{bmatrix}.$$

The vector field $\text{grad}_B f_{\mathbf{1}}$ can be expressed as

$$\begin{aligned}\text{grad}_B f_{\mathbf{1}}(p) &= \left(1 - a^2 + \frac{\sqrt{2}-2}{2}b^2\right) + \left(\frac{\sqrt{2}-2}{2} + 1\right) ab\mathbf{i} + \\ &- \left(\frac{\sqrt{2}-2}{2}bd + ac\right)\mathbf{j} + \left(\frac{\sqrt{2}-2}{2}bc - ad\right)\mathbf{k},\end{aligned}$$

where $p = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ is such that $a^2 + b^2 + c^2 + d^2 = 1$.

Then the expression for the projected vector field becomes

$$D\pi_1\left(\text{grad}_B f_{\mathbf{1}}(p)\right) = \begin{bmatrix} \frac{b}{(1-a)^2}\left(1 - a^2 + \frac{\sqrt{2}-2}{2}b^2\right) + \frac{\sqrt{2}-2}{2}\frac{ab}{1-a} \\ \frac{c}{(1-a)^2}\left(1 - a^2 + \frac{\sqrt{2}-2}{2}b^2\right) - \frac{1}{1-a}\left(\frac{\sqrt{2}-2}{2}bd + ac\right) \\ \frac{d}{(1-a)^2}\left(1 - a^2 + \frac{\sqrt{2}-2}{2}b^2\right) + \frac{1}{1-a}\left(\frac{\sqrt{2}-2}{2}bc - ad\right) \end{bmatrix}.$$

With this expression, we hope to be able to find a way to describe the vector field $\text{grad}_B f_{\mathbf{1}}$, its trajectories, and continue the study of controllability in this case. However, the expressions obtained for this specific vector field are already quite complicated, and for more general gradient vector fields, the expressions might become even more intricate. Consequently, we are compelled to seek alternate approaches for advancing this study.

REALIFICATION AND COMPLEXIFICATION OF VECTOR SPACES

This appendix is devoted to give a brief compilation of the basics concepts about realification and complexification of vector spaces. Many textbooks in Linear Algebra discuss such subjects in details, e.g., [23]. For an approach in the context of Lie Algebras and Lie Groups we refer [26] and [15].

A.1 Realification

A real vector space can be immediately obtained from a complex one simply by restriction of the scalars from $\mathbb{C}$ to $\mathbb{R}$. This procedure is called realification, and we shall discuss it now.

Definition A.1.1. *If V is a $\mathbb{C}$-vector space, the $\mathbb{R}$-vector space $V^{\mathbb{R}}$ obtained from V by restricting the scalars from $\mathbb{C}$ to $\mathbb{R}$ is called* ***realified*** *space of V and its dimension is twice that one of V, that is,*

$$\dim_{\mathbb{R}} V^{\mathbb{R}} = 2 \dim_{\mathbb{C}} V.$$

Note that if $\{e_1, \ldots, e_n\}$ is a $\mathbb{C}$-basis of V, then $\{e_1, \mathbf{i}e_1, \ldots, e_n, \mathbf{i}e_n\}$ is a $\mathbb{R}$-basis for $V^{\mathbb{R}}$, which justifies the equality above.

Definition A.1.2. *Let V be a $\mathbb{R}$-vector space. A linear operator $J : V \to V$ for which $J^2 = -\mathrm{Id}_V$ is said to be a* ***complex structure*** *on V.*

Given a complex structure J on a $\mathbb{R}$-vector space V, we can consider V as a $\mathbb{C}$-vector space with the complex scalar multiplication defined by

$$(a + b\mathbf{i}) \cdot v := a \cdot v + b \cdot J(v), \quad \text{where } v \in V, \ a + b\mathbf{i} \in \mathbb{C}.$$

Example A.1.3. *The most natural example of complex structure is given by the multiplication by* $\mathbf{i}$*, which defines a complex structure on* $V^{\mathbb{R}}$*. In other words, if* V *is a* $\mathbb{C}$*-vector space, then the linear operator* $J : V \to V$*,* $v \mapsto \mathbf{i} \cdot v$*, induces a complex structure on* $V^{\mathbb{R}}$ *in a very natural way, since both vector spaces have the same underlying set and* $J^2 = -\mathrm{Id}_V$*.*

This means that it is perfectly possible to get back the complex scalar multiplication in V *from* $V^{\mathbb{R}}$ *via the real linear operator* J*, as mentioned above.*

On the other hand, a complex structure J *on a* $\mathbb{R}$*-vector space* W *has minimal polynomial of the form* $\lambda^2 + 1$*, whose eigenvalues are* $\pm\mathbf{i}$*, and since its eigenvalues occurs in pairs,* $\dim W$ *is even. This means that we can use the complex structure* J *to define on* W *a complex scalar multiplication, and so we get a* $\mathbb{C}$*-vector space for which the realified space is exactly* W*.*

Example A.1.4. *Let on* $\mathbb{R}^2$ *the complex structure* J *having the matrix in the canonical basis given by*

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

The complex scalar product defined on $\mathbb{R}^2$ *from this complex structure is*

$$\begin{aligned} (a + \mathbf{i}b)(u, v) &= a(u, v) + bJ(u, v) \\ &= a(u, v) + b(-v, u) \\ &= (au - bv, av + bu) \\ &= (au - bv) + \mathbf{i}(av - bu), \quad u, v \in \mathbb{R}, \end{aligned}$$

which is the usual representation of $\mathbb{C}$*. However, if we choose* J *as above but in a different basis, we obtain a different complex scalar multiplication. This means that, in general, a given* $\mathbb{R}$*-vector space of even dimension admits many different complex structures that gives rise to different (although isomorphic)* $\mathbb{C}$*-vector spaces.*

If W is a vector subspace of the $\mathbb{C}$-vector space V, then its realified space $W^{\mathbb{R}}$ is a vector subspace of $V^{\mathbb{R}}$. However, not all vector subspaces of $V^{\mathbb{R}}$ are obtained by the realification of subspaces of V. For this to be true the vector subspace must be closed under the multiplication by $\mathbf{i}$, which is equivalent to say that a vector subspace of $V^{\mathbb{R}}$ comes from a complex subspace if and only if it is invariant under the natural complex structure J.

A.2 Complexification

Consider now a $\mathbb{R}$-vector space V. To obtain from V a $\mathbb{C}$-vector space naturally associated with V, we first must to give a meaning to the complex scalar multiplication $(a + b\mathbf{i}) \cdot v$, where $v \in V$. Moreover, one should expect for this multiplication to satisfy

$$(a + b\mathbf{i}) \cdot v = av + b\mathbf{i}v = av + \mathbf{i}bv.$$

Since V is a $\mathbb{R}$-vector space, the multiplication by $\mathbf{i}$ has no meaning in V, and thus we should regard $av + \mathbf{i}bv$ as a formal sum of av with bv, where the factor $\mathbf{i}$ keeps the addends apart. A more precise way to do it is considering $av + \mathbf{i}bv$ as being the ordered pair (av, bv).

Definition A.2.1. *The **complexification** of a $\mathbb{R}$-vector space V is defined as $V_{\mathbb{C}} = V \oplus V$, with the scalar multiplication defined by*

$$(a + b\mathbf{i})(v_1, v_2) = (av_1 - bv_2, bv_1 + av_2), \quad a, b \in \mathbb{R}.$$

The choice of this operation is justified if we think of the ordered pair $(v_1, v_2) \in V \oplus V$ as a formal sum $v_1 + \mathbf{i}v_2$:

$$(a + b\mathbf{i})(v_1 + \mathbf{i}v_2) = av_1 + a\mathbf{i}v_2 + b\mathbf{i}v_1 - bv_2 = (av_1 - bv_2) + \mathbf{i}(bv_1 + av_2).$$

Further, as

$$\mathbf{i}(v, 0) = (0, v),$$

it follows that

$$(v_1, v_2) = (v_1, 0) + (0, v_2) = (v_1, 0) + \mathbf{i}(v_2, 0).$$

And this tells us that we can formally think about the elements of $V_{\mathbb{C}}$ as elements of the form $v_1 + \mathbf{i}v_2$. However, it is worth noting that $\mathbf{i}v_2$ has no meaning while $\mathbf{i}(v_2, 0)$ is exactly $(0, v_2)$.

Example A.2.2. *Let $\mathbb{R}_{\mathbb{C}}$ be the set of ordered pairs (x, y), $x, y \in \mathbb{R}$, equipped with the scalar* *multiplication*

$$(a + \mathbf{i}b)(x, y) = (ax - by, bx + ay).$$

This makes it clear that $\mathbb{R}_{\mathbb{C}}$ *is isomorphic to* $\mathbb{C}$ *as* $\mathbb{C}$*-vector spaces, with isomorphism given by* $(x, y) \mapsto x + \mathbf{i}y$*.*

If $V \neq \{0\}$ is a $\mathbb{R}$-vector space and $\{e_i\}$ is a $\mathbb{R}$-basis for V, then $\{(e_i, 0)\}$ is a $\mathbb{C}$-basis for $V_{\mathbb{C}}$. In fact, given $(v_1, v_2) \in V_{\mathbb{C}}$ we can write v_1 and v_2 as $\mathbb{R}$-linear combinations of the elements e_i, and this shows that every element of $V_{\mathbb{C}}$ is a $\mathbb{R}$-linear combination of the elements $(e_i, 0)$ and $(0, e_i)$. The equality $(0, e_i) = \mathbf{i}(e_i, 0)$, allows us to pass to $\mathbb{C}$-linear combinations of the elements $(e_i, 0)$. It remains only to show that such set is linearly independent. But

$$(a_1 + \mathbf{i}b_1)(e_1, 0) + \cdots + (a_n + \mathbf{i}b_n)(e_n, 0) = (0, 0), \quad a_i, b_i \in \mathbb{R},$$

is equivalent to

$$(a_1e_1 + \cdots + a_ne_n, b_1e_1 + \cdots + b_ne_n) = (0, 0),$$

and the linear independence of e_i over $\mathbb{R}$ ensures that a_i and b_i are all zero. This also proves that $\dim_{\mathbb{C}}(V_{\mathbb{C}}) = \dim_{\mathbb{R}}(V)$.

Both the $\mathbb{R}$-subspaces $V \oplus \{0\}$ and $\{0\} \oplus V$ of $V_{\mathbb{C}}$ behave like V, since the sum is component-wise defined, $a(v, 0) = (av, 0)$ and $a(0, v) = (0, av)$, $a \in \mathbb{R}$. For this reason, the $\mathbb{R}$-linear transformation $v \mapsto (v, 0)$ is called **canonical embedding** of V into $V_{\mathbb{C}}$.

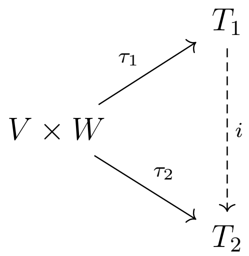
It is also possible to define the complexification of a real vector space V with an approach based on tensor products. Let us see how to proceed.

Definition A.2.3. *The tensor product between the real vector spaces* V *and* W *is a real vector space* $V \otimes_{\mathbb{R}} W$ *together with a* $\mathbb{R}$*-bilinear map* $\tau : V \times W \to V \otimes_{\mathbb{R}} W$ *such that for every* $\mathbb{R}$*-bilinear map* $\varphi : V \times W \to X$ *there exists only one linear map* $\Phi : V \otimes_{\mathbb{R}} W \to X$ *that commutes the diagram*

$$\begin{array}{ccc} V \times W & \xrightarrow{\tau} & V \otimes_{\mathbb{R}} W \\ \downarrow \varphi & \swarrow \Phi & \\ X & & \end{array}$$

that is, $\varphi = \Phi \circ \tau$*. The usual notation to the image* $\tau(v \times w)$ *of* $v \times w$ *under* τ *is* $v \otimes w$*, which is called* ***monomial tensor****.*

It is possible to prove that the tensor product between V and W always exist. Moreover, the tensor product is unique up to isomorphisms, that is, if $\tau_1 : V \times W \to T_1$ and $\tau_2 : V \times W \to T_2$ are two tensor products between V and W then there is an unique isomorphism $i : T_1 \to T_2$ that commutes the diagram



It is also possible to show that the monomial tensors $v \otimes w$ generate $V \otimes_{\mathbb{R}} W$ as a real vector space. In what concerns complexification, we are particularly interested in the tensor product of the form $V \otimes_{\mathbb{R}} \mathbb{C}$, where V is a real vector space.

If V has $\{e_1, \ldots, e_n\}$ as a basis, then the set

$$\{e_1 \otimes 1, \ldots, e_n \otimes 1, e_1 \otimes \mathbf{i}, \ldots e_n \otimes \mathbf{i}\}$$

is a basis for $V \otimes_{\mathbb{R}} \mathbb{C}$. Furthermore, as every element belonging to $V \otimes_{\mathbb{R}} \mathbb{C}$ can be uniquely expressed as $v_1 \otimes 1 + v_2 \otimes \mathbf{i}$, $v_1, v_2 \in V$, we can do the identification

$$V \otimes_{\mathbb{R}} \mathbb{C} \Leftrightarrow V \oplus \mathbf{i}V,$$

being $\mathbf{i}V$ a vector subspace isomorphic to V.

If $a+\mathbf{i}b \in \mathbb{C}$, then the multiplication by $a+\mathbf{i}b$ is a $\mathbb{R}$-linear map from $\mathbb{C}$ to $\mathbb{C}$. Denoting by $m(a+\mathbf{i}b)$ such multiplication, we have that $1 \otimes m(a+\mathbf{i}b)$ defines a $\mathbb{R}$-linear map from $V \otimes_{\mathbb{R}} \mathbb{C}$ to $V \otimes_{\mathbb{R}} \mathbb{C}$. This allows us to define the multiplication by the complex scalar $\alpha = a + \mathbf{i}b$ in $V \otimes_{\mathbb{R}} \mathbb{C}$ by setting

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

And so we define the **complexification** of V as being the complex vector space $V \otimes_{\mathbb{R}} \mathbb{C}$ together with the scalar multiplication above defined.

The vector space V is identified inside $V \otimes_{\mathbb{R}} \mathbb{C}$ with $V \otimes 1$, which is the canonical embedding of V into $V \otimes_{\mathbb{R}} \mathbb{C}$. It is possible to show that there exists a unique $\mathbb{C}$-isomorphism $f_V : V_{\mathbb{C}} \to V \otimes_{\mathbb{R}} \mathbb{C}$ that commutes the diagram

$$\begin{array}{ccc} & V & \\ \swarrow & & \searrow \\ V_{\mathbb{C}} & \xrightarrow{\quad f_V \quad} & V \otimes_{\mathbb{R}} \mathbb{C} \end{array}$$

where the downward arrows indicate the canonical embeddings. This isomorphism is defined by $f_V(v_1, v_2) = v_1 \otimes 1 + v_2 \otimes \mathbf{i}$.

Both $V_{\mathbb{C}}$ and $V \otimes_{\mathbb{R}} \mathbb{C}$ are direct sums of subspaces V and $\mathbf{i}V$ (via canonical embeddings), which suggests that a complex vector space W is the complexification of the real space V when $W = V + \mathbf{i}V$ and $V \cap \mathbf{i}V = \{0\}$. Note that this differs from the previous meaning given to the complexification (Definition A.2.1) by the reason that now we are considering a preexisting vector space as the complexification instead of starting with a real vector space V and constructing the complexification outside of it.

Definition A.2.4. *Let V be a $\mathbb{C}$-vector space. A conjugation in V is an antilinear transformation σ such that $\sigma^2 = 1$. In other words, a conjugation in V is a transformation $\sigma : V \to V$ satisfying*

1. $\sigma(v + v') = \sigma(v) + \sigma(v')$, $v, v' \in V$;
2. $\sigma(cv) = \bar{c}\sigma(v)$, $c \in \mathbb{C}$ *and* $v \in V$;
3. $\sigma(\sigma(v)) = v$, $v \in V$.

It is clear that every conjugation is $\mathbb{R}$-linear even if it is not $\mathbb{C}$-linear.

Given a $\mathbb{R}$-vector space V, the elements belonging to $V_{\mathbb{C}}$ can be identified with elements of the form $u + \mathbf{i}v$, $u, v \in V$ and this identification allows us to define a conjugation $\sigma : V_{\mathbb{C}} \to V_{\mathbb{C}}$ via

$$\sigma(u + \mathbf{i}v) = u - \mathbf{i}v.$$

It is immediate that $\sigma^2 = 1$ and that σ is antilinear in $V_{\mathbb{C}}$. Moreover,

$$V = \{w \in V_{\mathbb{C}} \mid \sigma(w) = w\}.$$

A $\mathbb{C}$-vector space W endowed with a conjugation can be understood as a complexified space of a $\mathbb{R}$-vector space. To see it, note that σ is a linear transformation from $W^{\mathbb{R}}$

on $W^{\mathbb{R}}$ having eigenvalues 1 and -1, and thus

$$W^{\mathbb{R}} = W_1 \oplus W_{-1},$$

being W_1 and W_{-1} the eigenspaces associated to the eigenvalues 1 and -1, respectively. So, if $v \in W_1$, then $\sigma(\mathbf{i}v) = -\mathbf{i}v$, that is, $\mathbf{i}v \in W_{-1}$.

In other words, if J is the standard complex structure of $W^{\mathbb{R}}$, we have $J(W_1) \subset W_{-1}$ and, similarly, $J(W_{-1}) \subset W_1$, which implies that W_1 and W_{-1} has the same dimension. It follows that W is the complexified space of W_1. Further,

$$W_1 = \{v \in W \mid \sigma(v) = v\},$$

where σ is the conjugation obtained from the decomposition $W = W_1 + W_{-1} \sim W_1 + \mathbf{i}W_{-1}$. It is worth to note that different conjugations give us different ways to see W as the complexified space of a real vector space.

There is a bijection between the conjugations of a $\mathbb{C}$-vector space V and its real subspaces that complexifies to V, and there is also a bijection between the $\mathbb{C}$-vector spaces and the $\mathbb{R}$-vector spaces that admit complex structures.

Finally, the realified of a complex vector space can be complexified. Let $(V^{\mathbb{R}})_{\mathbb{C}}$ be such complexification. We know that the natural complex structure of $V^{\mathbb{R}}$, J, is a $\mathbb{R}$-linear transformation of $V^{\mathbb{R}}$ with eigenvalues $\pm\mathbf{i}$, that is, J does not have eigenvectors in $V^{\mathbb{R}}$, but in its complexification.

Write $V_{\mathbf{i}}$ and $V_{-\mathbf{i}}$ for the generalized eigenspaces of J in $(V^{\mathbb{R}})_{\mathbb{C}}$ associated, respectively, to the eigenvalues $\pm\mathbf{i}$. Restricted to $V_{\mathbf{i}}$, the matrix of J has the form

$$\begin{bmatrix} \mathbf{i} & & * \\ & \ddots & \\ & & \mathbf{i} \end{bmatrix}.$$

By the fact that $J^2 = -1$ such matrix is diagonal, which allows us to conclude that $Jv = \mathbf{i}v$ for every $v \in V_{\mathbf{i}}$. Similarly, $Jw = -\mathbf{i}w$ for every $w \in V_{-\mathbf{i}}$.

Now, as J is real in $V^{\mathbb{R}}$, we have $V_{-\mathbf{i}} = \sigma(V_{\mathbf{i}})$, where σ is the conjugation of $(V^{\mathbb{R}})_{\mathbb{C}}$ corresponding to $V^{\mathbb{R}}$. This ensures that $V_{\mathbf{i}}$ and $V_{-\mathbf{i}}$ have the same dimension and, since

$(V^{\mathbb{R}})_{\mathbb{C}} = V_{\mathbf{i}} \oplus V_{-\mathbf{i}}$, then the dimension of $V_{\pm\mathbf{i}}$ is the half of that one of $V^{\mathbb{R}}$, which agrees with the dimension of V. As these spaces are complex, it follows that they are isomorphic to V.

This suggests that the complexification of a realified space can be seen as a quaternionic vector space with scalar multiplication given by

$$(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})w = aw + Ew + Fw + Gw, \quad a, b, c, d \in \mathbb{R}, \;\; w \in (V^{\mathbb{R}})_{\mathbb{C}},$$

where $E = J\sigma$, being σ the conjugation of $(V^{\mathbb{R}})_{\mathbb{C}}$, F is the scalar multiplication by $\mathbf{i}$ in $(V^{\mathbb{R}})_{\mathbb{C}}$ and $G = EF$.

BIBLIOGRAPHY

[1] AGRACHEV, A. A., AND SACHKOV, Y. L. *Control theory from the geometric viewpoint*, vol. 2. Springer Science & Business Media, 2004.

[2] ASSOUDI, R. E., AND GAUTHIER, J. P. Controllability of right invariant systems on real simple lie groups of type F_4, G_2, C_n, and B_n. *Mathematics of Control, Signals and Systems 1* (1988), 293–301.

[3] ASSOUDI, R. E., GAUTHIER, J. P., AND KUPKA, I. Controllability of right invariant systems on semi-simple lie groups. *Banach Center Publications 32*, 1 (1995), 199–208.

[4] ASSOUDI, R. E., GAUTHIER, J.-P., AND KUPKA, I. A. K. On subsemigroups of semisimple lie groups. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire 13*, 1 (1996), 117–133.

[5] AYALA, V., AND SAN MARTIN, L. A. B. Controllability properties of a class of control systems on lie groups. In *Nonlinear control in the year 2000*. Springer, 2007, pp. 83–92.

[6] CARMO, M. P., AND FLAHERTY FRANCIS, J. *Riemannian geometry*, vol. 6. Springer, 1992.

[7] DUISTERMAAT, J. J., KOLK, J. A. C., AND VARADARAJAN, V. S. Functions, flows and oscillatory integrals on flag manifolds and conjugacy classes in real semisimple lie groups. *Compositio Mathematica 49*, 3 (1983), 309–398.

[8] GAUTHIER, J. P., AND BORNARD, G. Contrôlabilité des systèmes bilinéaires. *SIAM Journal on Control and Optimization 20*, 3 (1982), 377–384.

[9] GAUTHIER, J. P., KUPKA, I., AND SALLET, G. Controllability of right invariant systems on real simple lie groups. *Systems & control letters 5*, 3 (1984), 187–190.

[10] HELGASON, S. *Differential geometry, Lie groups, and symmetric spaces*. Academic press, 1979.

[11] HILGERT, J., HOFMANN, K. H., AND LAWSON, J. D. Lie groups, convex cones, and semigroups. *(No Title)* (1989).

[12] JURDJEVIC, V., AND KUPKA, I. Control systems on semi-simple lie groups and their homogeneous spaces. *Annales de l'institut Fourier 31*, 4 (1981), 151–179.

[13] JURDJEVIC, V., AND KUPKA, I. Control systems subordinated to a group action: Accessibility. *Journal of differential equations 39*, 2 (1981), 186–211.

[14] JURDJEVIC, V., AND SUSSMANN, H. J. Control systems on lie groups. *Journal of Differential equations 12*, 2 (1972), 313–329.

[15] KNAPP, A. W. *Lie groups beyond an introduction*, vol. 140. Springer, 1996.

[16] LEE, J. M. *Smooth manifolds*. Springer, 2012.

[17] ONISHCHIK, A. L., AND VINBERG, E. B. *Lie groups and algebraic groups*. Springer Science & Business Media, 2012.

[18] RABELO, L., AND SAN MARTIN, L. A. B. Cellular homology of real flag manifolds. *Indagationes Mathematicae 30*, 5 (2019), 745–772.

[19] ROCIO, O. G. D., SAN MARTIN, L. A. B., AND SANTANA, A. J. Invariant cones and convex sets for bilinear control systems and parabolic type of semigroups. *Journal of dynamical and control systems 12*, 3 (2006), 419–432.

[20] ROCIO, O. G. D., SANTANA, A. J., AND VERDI, M. A. Semigroups of affine groups, controllability of affine systems and affine bilinear systems in $Sl(2,\mathbb{R}) \rtimes \mathbb{R}^2$. *SIAM Journal on Control and Optimization 48*, 2 (2009), 1080–1088.

[21] RODRIGUES, B. A., SAN MARTIN, L. A. B., AND SANTANA, A. J. Invariant control sets for symmetric vector fields induced by the infinitesimal action of $\mathfrak{so}(1,4)$ on S^3. (to appear).

[22] RODRIGUES, B. A., SAN MARTIN, L. A. B., AND SANTANA, A. J. Controllability and semigroups of invariant control systems on $Sl(n,\mathbb{H})$. *Mathematics of Control, Signals, and Systems 34*, 2 (2022), 393–404.

[23] ROMAN, S. *Advanced linear algebra*, vol. 135. Springer, 2007.

[24] SAN MARTIN, L. A. B. Invariant control sets on flag manifolds. *Mathematics of Control, Signals and Systems 6*, 1 (1993), 41–61.

[25] SAN MARTIN, L. A. B. Maximal semigroups in semi-simple lie groups. *Transactions of the American Mathematical Society 353*, 12 (2001), 5165–5184.

[26] SAN MARTIN, L. A. B. *Algebras de Lie*. Unicamp, 2010.

[27] SAN MARTIN, L. A. B. Flag type of semigroups: A survey. *Advances in Mathematics and Applications: Celebrating 50 years of the Institute of Mathematics, Statistics and Scientific Computing, University of Campinas* (2018), 351–372.

[28] SAN MARTIN, L. A. B. *Lie groups*. Springer, 2021.

[29] SAN MARTIN, L. A. B., AND TONELLI, P. A. Semigroup actions on homogeneous spaces. *Semigroup Forum 50* (1995), 59–88.

[30] SANTOS, A. L. D., AND SAN MARTIN, L. A. B. Controllability of control systems on complex simple lie groups and the topology of flag manifolds. *Journal of Dynamical and Control Systems 19*, 2 (2013), 157–171.

[31] SANTOS, A. L. D., AND SAN MARTIN, L. A. B. A method to find generators of a semi-simple lie group via the topology of its flag manifolds. *Semigroup Forum 95*, 2 (2017), 259–280.

[32] SILVA, F., AND CROUCH, P. E. Controllability on classical lie groups. *Mathematics of Control, Signals and Systems 1*, 1 (1988), 31–42.

[33] SUGIURA, M. Conjugate classes of cartan subalgebras in real semisimple lie algebras. *Journal of the Mathematical Society of Japan 11*, 4 (1959), 374–434.

[34] TAKEUCHI, M., AND KOBAYASHI, S. Minimal imbeddings of r-spaces. *Journal of Differential Geometry 2*, 2 (1968), 203–215.

[35] VARGA, R. S. Iterative analysis. *New Jersey 322* (1962).

[36] WARNER, G. *Harmonic analysis on semi-simple Lie groups I*, vol. 188. Springer Science & Business Media, 2012.