UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
(Mestrado)

EDUARDO CELSO VISCOVINI

CONTROLLABILITY OF CONTROL SYSTEMS[1]

Maringá-PR

2022

[1]O presente trabalho foi realizado com apoio do Ministério da Ciência, Tecnologia e Inovação - Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq - Brasil

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

CONTROLLABILITY OF CONTROL SYSTEMS

EDUARDO CELSO VISCOVINI

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Departamento de Matemática, Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito para obtenção do título de Mestre em Matemática.
Área de concentração: Geometria e Topologia.

Orientador: Prof. Dr. Alexandre José Santana.

Maringá-PR
2022

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

Viscovini, Eduardo Celso
V827c Controllability of control systems / Eduardo Celso Viscovini. -- Maringá, 2022.
111 f. : il.

Orientador: Prof. Dr. Alexandre José Santana
Dissertação (Mestrado) - Universidade Estadual de Maringá, Centro de Ciências Exatas, Programa de Pós-Graduação em Matemática - Área de Concentração: Geometria e Topologia, 2022.

1. Sistemas de controle. 2. Grupos de Lie. 3. Semigrupos. 4. Conjuntos de controle. 5. Variedades flag. 6. Control systems. 7. Controllability. 8. Lie groups. 9. Semigroups. 10. Control sets. 11. Flag manifolds. 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. 515.352

Edilson Damasio CRB9-1.123

EDUARDO CELSO VISCOVINI

CONTROLLABILITY OF CONTROL SYSTEMS

Dissertação 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 Mestre em Matemática tendo a Comissão Julgadora composta pelos membros:

COMISSÃO JULGADORA:

Prof. Dr. Alexandre José Santana - UEM (Presidente)
Prof. Dr. Fritz Colonius - Universität Augsburg - Alemanha
Prof. Dr. Luiz Antonio Barrera San Martin - UNICAMP
Prof. Dr. César Adolfo Hernandez Melo - UEM
Prof. Dr. Ryuichi Fukuoka - UEM

Aprovado em: 24 de fevereiro de 2022.
Local de defesa: Videoconferência – Google Meet (https://meet.google.com/pft-sown-gpx)

CONTROLLABILITY OF CONTROL SYSTEMS

Eduardo Celso Viscovini

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Departamento de Matemática, Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito para obtenção do título de Mestre em Matemática.

BANCA JULGADORA

Prof. Dr. Alexandre José Santana
Universidade Estadual de Maringá
(Orientador)

Prof. Dr. Fritz Colonius
Universität Augsburg

Prof. Dr. Ryuichi Fukuoka
Universidade Estadual de Maringá

Prof. Dr. Cesar Adolfo Hernandez Melo
Universidade Estadual de Maringá

Prof. Dr. Luiz Antonio Barrera San Martin
Universidade Estadual de Campinas

Maringá-PR

2022

To my mom, who taught me so much

Acknowledgments

First and foremost, I would like to thank God, for creating this amazing universe, for the strengths and weaknesses given to me, and for all the help throughout the journey.

I would like to thank my family, for always being on my side and always supporting me. For teaching me how to be a person, and encouraging my education whenever possible.

I would like to thank my teachers, for their excellence in the teaching of many math topics, and occasionally other areas as well. In special to professor Alexandre, whose help and orientation were essential to this thesis and also to my academic life in general.

I would like to thank my friends, who also supported and motivated me, and taught me many things.

And I would like to thank CNPq and the Brazilian people for the financial support.

"Mathematics is the most beautiful and most powerful creation of the human spirit."

Stefan Banach

Abstract

In this thesis we study the controllability problem for control systems: the question of whether any point in the space can reach any other point using the positive time trajectories of a given control system. We give special attention to bilinear and affine systems.

In chapter 1 we recall various known results from control theory and from Lie theory, which will be used in later chapters. In chapter 2 we show necessary and sufficient conditions for controllability in one class of affine system. In chapter 3 we construct the tangent system using curves originating an isotropy subgroup of an action, and use this idea to get partial results for a class of bilinear control systems. In chapter 4 we show an equivalence between the flag type of a connected semigroup in $\mathrm{Sl}(\mathbb{R}^d)$ and the existence of invariant cones for the action of this semigroup in exterior products.

Keywords: Control systems, Controllability, Lie groups, Semigroups, Control sets, Flag manifolds.

Resumo

Nesta dissertação estudamos o problema da controlabilidade para sistemas de controle: se de qualquer ponto no espaço é possível chegar à qualquer outro ponto utilizando as trajetórias em tempo positivo do sistema de controle. Damos atenção especial à sistemas bilineares e afins.

No capítulo 1 relembramos diversos resultados da teoria de controle e da teoria de Lie, que serão utilizados nos capítulos seguintes. No capítulo 2 mostramos condições necessárias e suficientes para a controlabilidade de uma classe de sistemas afins. No capítulo 3 construímos o sistema tangente utilizando curvas com origem em um subgrupo de isotropia de uma ação, e utilizamos essa ideia para obter resultados parciais para uma classe de sistemas bilineares. No capítulo 4 mostramos uma equivalência entre o tipo flag de um semigrupo conexo de $\mathrm{Sl}(\mathbb{R}^d)$ e a existência de cones invariantes pela ação desse semigrupo em espaços exteriores.

Palavras-chave: Sistemas de controle, Grupos de Lie, Semigrupos, Conjuntos de controle, Variedades flag.

CONTENTS

CHAPTER 1

INTRODUCTION

In this chapter we recall several concepts and results related to control systems and semigrupos of Lie groups. In order to have a self contained text we prove various of those known results. These results will be useful for our goal of studying global and local controllability for certain control systems. We will freely use the notations and concepts of the references [1], [8], [16], [17], [18], [19], [20].

1.1 Control systems

Control systems have many different definitions, depending on the context being studied. There are also many different types of control systems, such that it is hard to give a general definition including all of them (see e.g. [1], [8], [19]). In this section we introduce 2 definitions of control systems. First, we define a continuous control system using vector fields in differentiable manifolds. This first definition is better suited to present the specific control systems which will be attacked in this thesis, and also some of the examples. Later, we define a control system as a Lie semigroup acting in a manifold M. This second definition is better suited for some of the more general results in the later chapters. During this section we will also include observations on some of the alternative ways one could define a control system.

Whenever we say a function is differentiable we mean continuously differentiable, unless stated otherwise. Furthermore, in some occasions we will use the notation $x \to$

$f(x)$ for the function that takes x into $f(x)$, mostly to keep some observations short while avoiding ambiguity. Of course we also use $f : A \to B$ to mean f has domain in A and codomain in B.

Definition 1.1.1. *A continuous control system is an ordered set* $(M, \mathcal{F}, U, \mathcal{U})$ *where* M *is a differentiable manifold,* U *is a nonempty arbitrary set,* $\mathcal{F} : M \times U \to TM$ *is such that for any fixed* $c \in U$ *the function* $\mathcal{F}_c$ *defined by* $m \to \mathcal{F}(c, m)$ *is a vector field in* M*, and* $\mathcal{U}$ *is a set of functions with domain in* $\mathbb{R}$ *and codomain in* U*, satisfying the following properties:*

1. *For any* $c \in U$*, the constant function*

$$f : \mathbb{R} \to U$$

$$t \to c$$

   is in $\mathcal{U}$*.*

2. *For any* $u \in \mathcal{U}$ *and* $\alpha \in \mathbb{R}$ *the function*

$$\alpha u : \mathbb{R} \to U$$

$$t \to u(\alpha + t)$$

   is in $\mathcal{U}$*.*

3. *For any piecewise constant function* $f : \mathbb{R} \to \mathcal{U}$ *the function*

$$w : \mathbb{R} \to U$$

$$t \to f(t)(t)$$

   is in $\mathcal{U}$*.*

4. *For any* $x_0 \in M$ *and* $u \in \mathcal{U}$*, the initial valued problem*

$$\dot{x}(t) = \mathcal{F}(x(t), u(t))$$

$$x(0) = x_0$$

has unique and global solution such that, for a fixed u and t, the solution on time t depends differentiably on the initial condition x_0. Here a solution is assumed to be a continuous function satisfying the differential equation for almost all points.

The definition above is more complicated than it needs to be. For our purposes, a continuous control system could simply be defined as a set of complete vector fields in a manifold M without losing much, as explained later in this section. The reason we chose this definition is to better match the problems presented in the next chapters.

The functions in $\mathcal{U}$ are called controls. Conditions 1 and 3 imply that all piecewise constant functions $f : \mathbb{R} \to U$ are controls, as follows.

Proposition 1.1.2. *For a control system defined as previously, any piecewise constant function $f : \mathbb{R} \to U$ is contained in $\mathcal{U}$.*

Proof. Let $f : \mathbb{R} \to U$ a piecewise constant function, and define

$$g : \mathbb{R} \to \mathcal{U}$$

$$t \to g_t$$

where g_t is the constant function

$$g_t : \mathbb{R} \to U$$

$$s \to f(t).$$

Note that $g_t \in \mathcal{U}$ by the item 1, such that g is well defined. Furthermore, g is piecewise constant as it is constant in any interval where f is constant. Therefore, by condition 3, the function

$$w : \mathbb{R} \to U$$

$$t \to g(t)(t)$$

is contained in $\mathcal{U}$. But, for all $t \in \mathbb{R}$,

$$w(t) = g(t)(t) = g_t(t) = f(t)$$

such that $w = f$. Therefore $f \in \mathcal{U}$. □

On the other hand, if $\mathcal{U}$ is the set of piecewise constant functions $u : \mathbb{R} \to U$, then conditions 1,2 and 3 follow naturally, while condition 4 is true if $\mathcal{F}_c$ is a complete vector field with differentiable flow for all $c \in U$ (for example, if $\mathcal{F}_c$ are all complete and differentiable vector fields). Thus, assuming $\mathcal{F}$ satisfies the condition above, the smallest set of controls needed for a control system $(M, \mathcal{F}, U, \mathcal{U})$ is the set of all piecewise constant functions with domain in $\mathbb{R}$ and codomain in U. In general, the properties of a control system are very often preserved by restricting a set of controls $\mathcal{U}$ to only these piecewise constant functions, such that the set $\mathcal{U}$ could be instead be defined as the set of all piecewise constant functions $u : \mathbb{R} \to U$ without losing much.

Furthermore, the function $\mathcal{F}$ in the definition serves mostly a transition purpose, so that $\mathcal{F}_{u(t)}$ is a vector field in M for each $t \in \mathbb{R}$. One could instead define U to be the set of vector fields in M, and use the differential equation

$$\dot{x}(t) = u(t)(x(t))$$

to define the system, eliminating the need for the intermediate function $\mathcal{F}$.

From the uniqueness and globality of solution in item 4, it is possible to define the function

$$\phi : M \times \mathcal{U} \times \mathbb{R} \to M$$

$$(x_0, u, T) \to \phi(x_0, u, T)$$

where $\phi(x_0, u, T)$ is the solution to the initial valued problem

$$\dot{x}(t) = \mathcal{F}(x(t), u(t))$$

$$x(0) = x_0$$

on time $t = T$. Note that, by definition,

$$\phi(x, u, 0) = x.$$

The function ϕ is called the solution of the system, and has interesting properties. One of the most important of said properties is the cocycle property, which is presented in the following result.

Proposition 1.1.3. *Let $x \in M$, $u, v \in \mathcal{U}$ and $T_1, T_2 \in \mathbb{R}$. If T_1 and T_2 are both non negative, then*

$$\phi(\phi(x, u, T_1), v, T_2) = \phi(x, w, T_1 + T_2)$$

where $w \in \mathcal{U}$ is defined by

$$w(t) = \begin{cases} u(t); \text{ if } t < T_1 \\ v(t - T_1); \text{ if } t \geq T_1. \end{cases}$$

Furthermore, if T_1 and T_2 are both non positive, then

$$\phi(\phi(x, u, T_1), v, T_2) = \phi(x, w, T_1 + T_2)$$

where

$$w(t) = \begin{cases} u(t); \text{ if } t > T_1 \\ v(t - T_1); \text{ if } t \leq T_1. \end{cases}$$

Proof. We will prove the case $T_1, T_2 \geq 0$, the other case is analogous. First, note that the function $s \to v(s - T_1)$ is in $\mathcal{U}$ by 2. Define the piecewise constant function

$$f : \mathbb{R} \to \mathcal{U}$$

$$t \to \begin{cases} u; \text{ if } t < T_1 \\ (s \to v(s - T_1)); \text{ if } t \geq T_1 \end{cases}$$

then, for any $t \in \mathbb{R}$

$$f(t)(t) = \begin{cases} u(t); \text{ if } t < T_1 \\ v(t - T_1); \text{ if } t \geq T_1 \end{cases} = w(t),$$

and, by condition 3, $w \in \mathcal{U}$.

Now we show $\phi(\phi(x, u, T_1), v, T_2) = \phi(x, w, T_1 + T_2)$. Note that this equality is trivial if $T_1 = 0$ or $T_2 = 0$, by using the previously mentioned $\phi(y, z, 0) = y$ for any $y \in M, z \in \mathcal{U}$. We therefore assume $T_1, T_2 > 0$. Since $w(t) = u(t)$ for all $t < T_1$, then, for $t \in (0, T_1)$,

$$\frac{d}{dt}\phi(x, w, t) = \mathcal{F}(\phi(x, w, t), w(t)) = \mathcal{F}(\phi(x, w, t), u(t)).$$

Furthermore, $\phi(x, w, 0) = x$ by definition, therefore, $\phi(x, w, t)$ is solution to the differential equation

$$\dot{x}(t) = u(t)(x(t))$$

$$x(0) = x$$

in the interval $[0, T_1]$, therefore $\phi(x, u, T_1) = \phi(x, w, T_1)$. Then, the function

$$g : \mathbb{R} \to M$$

$$t \to \phi(x, w, t + T_1)$$

is such that $g(0) = \phi(x, w, T_1) = \phi(x, u, T_1)$ and

$$\frac{d}{dt} g(t) = \mathcal{F}(g(t), w(t + T_1)) = \mathcal{F}(g(t), v(t))$$

for all $t > 0$, therefore $g(t) = \phi(\phi(x, u, T_1), v, t)$ for all $t > 0$ and, in particular,

$$\phi(\phi(x, u, T_1), v, T_2) = g(T_2) = \phi(x, w, T_1 + T_2).$$

□

The cocycle property has a very interesting consequence. By fixing u and T we can define the application

$$\phi_u^T : M \to M$$

$$x \to \phi(x, u, T).$$

The cocycle property then implies that, for $T_1, T_2 \geq 0$ or $T_1, T_2 \leq 0$, $\phi_v^{T_1} \circ \phi_u^{T_2} = \phi_w^{T_1 + T_2}$ for some $w \in \mathcal{U}$. This means that the sets

$$S := \{\phi_u^T ; T \geq 0, u \in \mathcal{U}\}$$

$$S^{-1} := \{\phi_u^T ; T \leq 0, u \in \mathcal{U}\}$$

are closed for the composition of functions and are, therefore, semigroups with this operation. The notation S^{-1} is used because this semigroup is, in fact, the inverse of S. This is a consequence of the next property.

Proposition 1.1.4. *For $x \in M$, $T \in \mathbb{R}$ and $u \in \mathcal{U}$,*

$$\phi(\phi(x,u,T),v,T) = x$$

where v is defined by

$$v(t) = u(t+T)$$

Proof. We will show, in fact, that

$$\phi(\phi(x,u,T),v,t) = \phi(x,u,T+t).$$

for all $t \in \mathbb{R}$. Let

$$g : \mathbb{R} \to M$$

$$t \to \phi(x,u,T+t)$$

Then

$$g(0) = \phi(x,u,T)$$

and

$$\frac{d}{dt}g(t) = \frac{d}{dt}\phi(x,u,T+t) = \mathcal{F}(g(t),u(t+T)) = \mathcal{F}(g(t),v(t))$$

for all $t \in \mathbb{R}$, therefore

$$\phi(\phi(x,u,T),v,t) = g(t) = \phi(x,u,T+t).$$

In particular,

$$\phi(\phi(x,u,T),v,-T) = \phi(x,u,T-T) = x.$$

□

This means that ϕ_v^{-T} as defined in the proposition is the inverse of ϕ_u^T. Since an element ϕ_u^T is in S if, and only if, ϕ_v^{-T} is in S^{-1} and vice versa, we have that S^{-1} is the inverse of S. In particular, all of the applications ϕ_u^T are bijections in M, and it is similarly possible to define the following group:

$$G = \{\phi_{u_1}^{T_1}\phi_{u_2}^{T_2}...\phi_{u_n}^{T_n}; n \in \mathbb{N}, u_1, u_2, ..., u_n \in \mathcal{U}, T_1, T_2, ..., T_n \in \mathbb{R}\}.$$

Since the cocycle property requires T_1, T_2 to not have opposite signs, it is not always possible to write an element of G as a single ϕ_u^T for some $u \in \mathcal{U}, T \in \mathbb{R}$, thus the definition using finite compositions of these functions is required.

By item 4, the applications ϕ_u^T are differentiable. Since the inverse of an application ϕ_u^T is also written as ϕ_w^{-T} for some $w \in \mathcal{U}$ and $-T \in \mathbb{R}$, then these inverses are also differentiable, such that each ϕ_u^T is a diffeomorphism. Furthermore, since a concatenation of diffeomorphisms is also a diffeomorphism, all elements of G are diffeomorphisms.

The set S defined previously is called the semigroup of the system, and the set G is called the group of the system. The group G motivates the definition of group action:

Definition 1.1.5. *Let G be a group and M a set. An action of G in M is a function*

$$\rho : G \times M \to M$$

$$(g, x) \to gx$$

satisfying

$$Idx = x$$

$$g(hx) = (gh)x$$

for all $x \in M$ and $g, h \in G$.

If $S \subset G$ is a semigroup we also say that S acts on M.

Note that the group and the semigroup of a continuous control system act naturally in the respective manifold M by $sx = s(x)$ for all $s \in G, x \in M$.

A continuous control system is controllable in M if for any $x, y \in M$ there are $u \in \mathcal{U}$ and $T \geq 0$ such that $\phi(x, u, T) = y$. Equivalently, the system is controllable if for any $x, y \in M$ there is $s = \phi_u^T$ in the semigroup S of the system such that $sx = y$, or, yet, if

$$Sx = M \; \forall x \in M.$$

A semigroup acting on M and satisfying $Sx = M$ for all x is called transitive. This means that controllability can be studied simply from the semigroup of the system. In fact, from the controllability point of view a control system could also be defined as the action of a semigroup in a manifold. This definition has the advantage of including non continuous control systems. The control systems which we want to study in this thesis are continuous ones, mainly the bilinear control system and some other special cases of the affine control system in $\mathbb{R}^n$, both of which will be defined later. Nonetheless, the association of semigroups with control systems is used on many of the results which will be shown, both for simplicity and for generality. This motivates us to include the definition of control systems through semigroups. First, we define a Lie group.

Definition 1.1.6. *A Lie group is a smooth manifold G with a smooth group product*

$$p : G \times G \to G$$

$$(g, h) \to gh.$$

Some other relations between Lie theory and control systems will be presented in the next section. For now, we define a control system as a Lie semigroup acting on a manifold.

Definition 1.1.7. *A control system is an ordered set (M, G, S, ρ) where M is a manifold, G is a Lie group, $S \subset G$ is a nonempty semigroup, and*

$$\rho : G \times M \to M$$

$$(g, x) \to gx$$

is a differentiable action. In this case we also say that S is the semigroup of the system, and the semigroup S^{-1} is defined as the inverse semigroup of S.

As previously mentioned, there are other definitions for control system (e.g. [1], [8], [19]).

If there is no risk of confusion regarding the other elements, we will just say that the semigroup S is the control system, or that S is a control system in M

At a first glance it might seem restrictive to require for G to be a Lie group. However, any group G is a lie group with the discrete topology (although possibly not second countable). By choosing this topology to the group G of a continuous control system, the action

$$\rho : G \times M \to M$$

$$(g, x) \to gx = g(x)$$

becomes differentiable as it is differentiable in the second coordinate (each $g \in G$ is a differentiable application), such that, by taking $S \subset G$ as the semigroup of the system, (M, G, S, ρ) becomes a control system. This means any continuous control system can also be seen as a control system. However, this discretization of G is not necessarily the best approach for defining a Lie structure on it, as all of the geometry in this group is lost by doing so, and the resulting manifolds end up on most cases having uncountably many connected components, such that they are not second countable. Second countability is often a wanted property for manifolds, sometimes even required in their definition, as the lack of the second countable property leads to pathological behaviors in multiple scenarios. The next section in this chapter will explore Lie semigroups generated from invariant vector fields in G and the Lie-Palais Theorem which provides, under certain conditions, a much more interesting way of relating continuous control systems to Lie semigroups. On the other hand, not all control systems are continuous, as illustrated by the next example.

Example 1.1.8. *Let G be the group $(\mathbb{R}, +)$. Since $+$ is smooth, G is a Lie group with the canonical manifold structure of $\mathbb{R}$. Consider the action of G on itself defined by*

$$\rho : G \times G \to G$$

$$(g, h) \to g + h.$$

It is then possible to define a control system from the semigroup $S = \{1\} \cup (2, +\infty)$. This control system cannot be written as a continuous continuous control system, since the sets $Sx = \{x + 1\} \cup (x + 2, +\infty)$ are disconnected and, therefore, cannot be obtained from solutions in positive time of the differential equations in a continuous control system.

The previous definition of controllability can be generalized naturally to control systems. The semigroup S is associated with positive time, while S^{-1} is associated with negative time. In this case, the control system is defined to be controllable if for any $x, y \in M$ there is $s \in S$ such that $sx = y$, or, equivalently, if

$$Sx = M$$

for all $x \in M$, that is, if S is transitive in M. Remember that, for continuous control systems, this condition is equivalent to the previous definition of controllability, by considering $S = \{\phi_u^T; u \in \mathcal{U}, T \geq 0\}$.

Other important concepts regarding controllability of control systems are the positive and negative orbits. For a point $x \in M$ the positive orbit of x, denoted by $\mathcal{O}^+(x)$, is defined by Sx where S is the semigroup of the system. In continuous control systems, this set coincides with the set of all y such that $y = \phi(x, u, T)$ for some $u \in \mathcal{U}$ and $T \geq 0$. If $y \in \mathcal{O}^+(x)$ we say that y can be reached from x. The negative orbit is defined as $S^{-1}(x)$, and, similarly, for a continuous control system this set coincides with the set of all y such that $y = \phi(x, u, T)$ for some $u \in \mathcal{U}$ and $T \leq 0$.

By definition, if $y \in \mathcal{O}^+(x)$ then there is $s \in S$ such that $y = sx$. Since S^{-1} is the inverse of S, then $s^{-1} \in S^{-1}$ and $s^{-1}y = s^{-1}sx = x$, such that $x \in \mathcal{O}^-(y)$. Analogously, if $x \in \mathcal{O}^-(y)$ then $y \in \mathcal{O}^+(x)$, that is, $y \in \mathcal{O}^+(x)$ if, and only if, $x \in \mathcal{O}^-(y)$. Furthermore, if $y \in \mathcal{O}^+(x)$ and $z \in \mathcal{O}^+(y)$ then $z \in \mathcal{O}^+(x)$. This is because $y \in \mathcal{O}^+(x)$ and $z \in \mathcal{O}^+(y)$ implies there are $s, r \in S$ such that $sx = y$ and $ry = z$. Then $z = ry = (rs)x$, where $rs \in S$ since S is a semigroup, and, therefore $z \in \mathcal{O}^+(x)$. Analogously, if $y \in \mathcal{O}^-(x)$ and $z \in \mathcal{O}^-(y)$ then $z \in \mathcal{O}^-(x)$. By repeatedly applying this property, if there is a chain of points $x_1, x_2, ..., x_n$ such that each x_i is in the positive orbit of its predecessor x_{i-1} then $x_n \in \mathcal{O}^+(x_1)$, and analogously for negative orbits. The next result is a classical result on control theory and shows some equivalences regarding controllability control system and their positive and negative orbits (see [17]).

Proposition 1.1.9. *Let (M, G, S, ρ) be a control system with $M \neq \emptyset$. then the following are equivalent.*

1. *The system is controllable in M*
2. *$\mathcal{O}^+(x) = M$ for all $x \in M$*
3. *$\mathcal{O}^-(x) = M$ for all $x \in M$*
4. *$\mathcal{O}^+(x) = \mathcal{O}^-(x) = M$ for all $x \in M$*
5. *There exists $x \in M$ such that $\mathcal{O}^+(x) = \mathcal{O}^-(x) = M$.*

Proof. As previously mentioned, if the system is controllable then $Sx = M$ for all $x \in$

M. By definition of orbit, $\mathcal{O}^+(x) = Sx$, therefore $\mathcal{O}^+(x) = M$ for all $x \in M$, showing $1 \Rightarrow 2$.

For the implication $2 \Rightarrow 3$, let $x, y \in M$ be arbitrary points. Then $x \in \mathcal{O}^+(y)$ by item 2, and, therefore, $y \in \mathcal{O}^-(x)$. Since y is arbitrary, $\mathcal{O}^-(x) = M$ and since x is arbitrary this equality is true for all $x \in M$, showing item 3.

The implication $3 \Rightarrow 4$ is shown analogously. Let arbitrary $x, y \in M$, then $y \in \mathcal{O}^-(x)$. Therefore $x \in \mathcal{O}^+(y)$. Since x and y are both arbitrary then $\mathcal{O}^+(x) = M$ for all $x \in M$. The equality $\mathcal{O}^-(x) = M$ is already true by item 3 therefore $\mathcal{O}^+(x) = \mathcal{O}^-(x) = M$ for all $x \in M$.

The implication $4 \Rightarrow 5$ is direct by restriction to a single element of M.

Finally, to show $5 \Rightarrow 1$, let $x \in M$ as described in item 5 and let $y, z \in M$ arbitrary. Then $y \in \mathcal{O}^-(x)$ and $z \in \mathcal{O}^+(x)$, as $\mathcal{O}^-(x) = \mathcal{O}^+(x) = M$. Therefore $x \in \mathcal{O}^+(y)$, and we have the chain $x \in \mathcal{O}^+(y), z \in \mathcal{O}^+(x)$, which implies $z \in \mathcal{O}^+(y)$. Since y, z are arbitrary, the system is controllable. □

We then have some topological definitions regarding controllability of the system. In many cases it is easier to show that a point is contained in the topological closure of an orbit, rather than in the orbit itself. Motivated by this we have the definition of approximate controllability: If for any $x, y \in M$ and any open set V containing y there is $z \in V$ such that $z \in \mathcal{O}^+(x)$, then the system is said to be *approximately controllable*. This is equivalent to

$$\overline{Sx} = M$$

for all $x \in M$, where $\overline{Sx}$ denotes the topological closure of this set. One can also define *backward approximate controllability* by asking whether

$$\overline{S^{-1}x} = M$$

for all x. Interestingly, unlike the the previous case, approximate controllability is not equivalent to backward approximate controllability. This is illustrated by the following example.

Example 1.1.10. *Define the following vector fields in $\mathbb{R}^2$, where the tangent bundle of $\mathbb{R}^2$ is* *associated with* $\mathbb{R}^2$ *itself:*

$$X : \mathbb{R}^2 \to \mathbb{R}^2$$

$$(x, y) \to (-y^2, 0)$$

$$Y : \mathbb{R}^2 \to \mathbb{R}^2$$

$$(x, y) \to (1, 0)$$

$$Z : \mathbb{R}^2 \to \mathbb{R}^2$$

$$(x, y) \to \begin{cases} (0, 0);\ \textit{if}\ x \leq 0 \\ (0, x^2);\ \textit{if}\ x > 0 \end{cases}$$

$$W : \mathbb{R}^2 \to \mathbb{R}^2$$

$$(x, y) \to \begin{cases} (0, 0);\ \textit{if}\ x \leq 0 \\ (0, -x^2);\ \textit{if}\ x > 0 \end{cases}$$

All of these vector fields are differentiable and complete, such that their flows are differentiable and globally defined. In fact,

$$X_t(x, y) = (x - ty^2, y)$$

$$Y_t(x, y) = (x + ty, y)$$

$$Z_t(x, y) = \begin{cases} (x, y);\ \textit{if}\ x \leq 0 \\ (x, y + tx^2);\ \textit{if}\ x > 0 \end{cases}$$

$$W_t(x, y) = \begin{cases} (x, y);\ \textit{if}\ x \leq 0 \\ (x, y - tx^2);\ \textit{if}\ x \geq 0 \end{cases}$$

Let $U = \{X, Y, Z, W\}$ *and* $\mathcal{U}$ *the set of all piecewise constant function* $f : \mathbb{R} \to U$*, and consider the continuous control system* $(M, \mathcal{F}, U, \mathcal{U})$ *where* $\mathcal{F}$ *is defined by* $\mathcal{F}(x, c) = c(x)$ *for all* $c \in U$ *and* $M = \mathbb{R}^2$*. We note that the vector fields* X, Y *are horizontal while the vectors fields* Z, W *are vertical. Furthermore,* Y *is always a vector field to the right, while* X *is always to the left, except on* $y = 0$ *where it vanishes. This means that from a point* $(a, b) \in \mathbb{R}^2$*, it is possible, with these two vector fields, to reach any point in the horizontal line* $y = b$ *if* $b \neq 0$*, but only the points* (x, b) *where* $x > a$ *if* $b = 0$*. For the vertical controls, we note that both* Z *and* W *vanish if* $x \leq 0$*, and, otherwise,* Z *is always pointing up while* W *is always pointing down. Then, from a point* (a, b) *and using only these two vector fields it is possible to reach the entire vertical line* $x = a$ *if* $a > 0$*, but only the point* (a, b) *if* $a \leq 0$*.*

Let $r = \{(x,0) \in \mathbb{R}^2\}$ be the horizontal line $y = 0$. Then it is possible to show that, for any point $p \in \mathbb{R}^2$, $\mathbb{R}^2 \setminus r \subset Sp$, as follows. Let $q \in \mathbb{R}^2 \setminus r$ arbitrary. We first choose a sufficiently large t such that $Y_t p$ is in the set $x > 0$. Remember that this can be done to any point. Let $p_1 = Y_t p$. Since p_1 is in the set $x > 0$, it is possible to reach any point in it's vertical line using the vector fields Z and W. In particular, it's possible to reach any value for the coordinate y. Then, let p_2 be a point in this vertical line such that the y coordinate of p_2 matches the y coordinate of q, or, equivalently, such that p_2 and q are in the same horizontal line. Note that this line is not r, as we are assuming $q \notin r$. Therefore, it is possible to reach any point in this line from p_2, in particular, it is possible to reach q. Then we have

$$q \in Sp_2, p_2 \in Sp_1, p_1 \in Sp,$$

therefore

$$q \in Sp.$$

Since q is arbitrary in $\mathbb{R}^2 \setminus r$, we conclude that $\mathbb{R}^2 \setminus r \subset Sp$.

In particular, $\overline{Sp} = \mathbb{R}^2$ for all $p \in \mathbb{R}^2$, such that the system is approximately controllable. However, a point $(a,0)$ with $a \leq 0$ can only be reached from a point $(b,0)$ with $b \leq a$. Then, $S^{-1}(a,0)$ is a ray if $a < 0$, and, in particular, is not dense. The system is, therefore, not backward approximately controllable.

Note that controllability implies approximate controllability, because if $Sx = M$ then $\overline{Sx} = M$. Analogously, controllability implies backward approximate controllability. In particular, the control system shown in the previous example cannot be controllable, as it is not backward approximate controllable. Thus, the previous example is also an example of a control system which is approximately controllable but not controllable.

A control system is said to be *forward accessible* if Sx has nonempty interior for all $x \in M$, and is said to be *backward accessible* if $S^{-1}x$ has nonempty interior for all x. A system is said to be *accessible* if it is forward accessible and backwards accessible. Similarly to the previous case, if the system is controllable then it is forward and backward accessible, but they do not imply controllability. The next result shows that the combination of accessibility with approximate controllability is enough to prove controllability.

Proposition 1.1.11. *If a control system S on a manifold M is approximately controllable and backward accessible, then it is controllable. Alternatively, if S is backward approximately controllable and forward accessible then S is controllable.*

Proof. Assume S is approximately controllable and backward accessible. Then, for arbitrary $x, y \in M$, Sx is dense in M while $S^{-1}y$ has nonempty interior in M. Therefore, these two sets intersect each other. Let $z \in Sx \cap S^{-1}y$, then $z \in Sx$ and $y \in Sz$, therefore $y \in Sx$. Since x, y are arbitrary, the system is controllable. The other implication is analogous. □

Another important concept involving approximate controllablity is control sets.

Definition 1.1.12. *A control set is a set $C \subset M$ satisfying*

1. For all $x \in C$, $C \subset \overline{Sx}$.

2a. C has more than one element.

3. C is a maximal set satisfying 1.

For a continuous control system, condition $2a$ is sometimes replaced with the following:

2b. For all $x \in C$ there is $u \in \mathcal{U}$ such that $\phi(x, u, t) \in C$ for all $t > 0$.

In this case the previous condition 3 instead requires C to be a maximal set satisfying both 1 and 2b. 2b is not equivalent to 2a, such that these two definitions of control system differ from each other. In this thesis we use the version stated in definition 1.1.12 whenever talking about control systems, unless stated otherwise.

Remember that if $y \in Sx$ and $z \in Sy$ then $z \in Sx$. The closures of the orbits satisfy a similar relation: if $y \in \overline{Sx}$ and $z \in \overline{Sy}$ then $z \in \overline{Sx}$. This is due to the continuity of the action, as follows: for any open set V containing z there is $s \in S$ such that $sy \in V$. Since s is a diffeomorphism and, in particular, an homeomorphism, then $s^{-1}(V)$ is an open set containing y. Then, there is $r \in S$ such that $rx \in s^{-1}(V)$. Then $srx \in V$, where $sr \in S$. Since V is an arbitrary open set containing z, then $z \in \overline{Sx}$. Analogously for the closures of the negative orbits.

Control sets have an interesting property of no intersection: if two control sets C, D intersect each other, then $C = D$. The reason is as follows: assume C, D intersect, and let $x \in C, D$. Then $x \in \overline{Sy}$ for any y in C or D, such that $\overline{Sx} \subset \overline{Sy}$ for any

$y \in C \cup D$. Furthermore, $\overline{Sx}$ contains C and D, as x is a point contained in both control sets. This means $C \cup D \subset \overline{Sx} \subset \overline{Sy}$ for all $y \in C \cup D$. By the maximality of control sets, $C = D = C \cup D$.

We close this section with a very interesting result regarding control sets in compact manifolds, which makes good use of these properties.

Theorem 1.1.13. *Let S be a control system in a compact manifold M with dimension ≥ 1, and assume S is forward accessible. Then, for any $x_0 \in M$ there exists an invariant control set $C \subset \overline{Sx_0} \subset M$, that is, C is a control set such that $SC \subset C$. Furthermore, C has nonempty interior, and there is only a finite number of invariant control sets in M.*

Proof. Let D be the set

$$D = \{\overline{Sx}; x \in \overline{Sx_0}\}.$$

Note that all elements of D are contained in $\overline{Sx_0}$, as they are written as $\overline{Sx}$ with $x \in \overline{Sx_0}$. We order D as follows: given two sets $c, d \in D$ define $c \leq d$ if $d \subset c$. Note that this is the inverse of the inclusion order for sets. Now, for any totally ordered set $E \subset D$ let

$$e := \bigcap_{d \in E} d.$$

Note that the elements in D are all closed and nonempty, such that e is a decreasing intersection of closed and nonempty sets. Since M is compact, e is also nonempty. Let $x \in e$, then, by definition of e, x is contained in all $d \in E$. Remember that each $d \in E$ is written as $\overline{Sy}$, such that any z contained in $\overline{Sx}$ is also contained in $\overline{Sy}$ as $x \in d = \overline{Sy}$. This means $\overline{Sx}$ is contained in all $d \in E$, and, therefore, is an upper bound for the set E. By Zorn's lemma, there is a maximal element $C \in D$. By definition of D, C can be written as $\overline{Sx}$ for some $x \in \overline{Sx_0}$. Note that, for any $y \in C$, we have $y \in \overline{Sx}$ such that $\overline{Sy} \subset \overline{Sx} = C$ and, therefore $\overline{Sy} \geq C$. Since C is a maximal element of D and $\overline{Sy} \in D$ as $y \in \overline{Sx} \subset \overline{Sx_0}$, then $\overline{Sy} = C$. Since y is arbitrary, C satisfies the first condition of a control system, and is also invariant as $Sy \subset \overline{Sy} \subset C$ for any $y \in C$. Furthermore, since $Sx \subset C$ and the system is accessible, then C has nonempty interior and, in particular, is not unitary (this is why we ask dimension at least 1). Note that the same argument shows that any invariant set in M has nonempty interior. Finally, if D is another set satisfying the first 2 conditions of a control set and containing C, then $x \in D$ and, therefore, $D \subset \overline{Sx} = C$, such that C is maximal. Therefore C is an invariant control set with nonempty interior.

Now assume there are infinite invariant control sets in M. Then it is possible to create a sequence $(C_i)_{i\in\mathbb{N}}$ of these sets such that $C_i \neq C_j$ whenever $i \neq j$. These sets are all nonempty by condition 2 of control sets, then, for each $i \in \mathbb{N}$ let $x_i \in C_i$. Since M is compact, we can assume, without loss in generality, that the sequence x_i converges to a point $x \in M$. By the first part of the theorem, there is an invariant control set $C \subset \overline{Sx}$, and, by a previous observation, C has nonempty interior. In particular, the interior of C must intercept Sx, such that $sx \in Int(C)$ for some $s \in S$. Then, $sV \subset Int(C)$ for some open set V containing x, and, therefore, there is $n_0 \in \mathbb{N}$ such that $sx_i \in Int(C)$ for all $i > n_0$. But $sx_i \in C_i$, as the C_i are invariant. Therefore each C_i for $i > n_0$ intersect C. By the no intersection theorem, $C_i = C$ for all $i > n_0$, which contradicts the hypothesis that the C_i are all distinct. Therefore, there must be only a finite number of invariant control sets in M. □

1.2 Lie semigroups and control theory

In this section we recall some interesting properties regarding actions of Lie groups in manifolds, and their implications on control systems. As in the first section, we will consider control systems to be semigroups of Lie groups acting differentiably in a manifold.

For Lie group theory, in special Lie group actions we suggest [20], and, for semigroup actions see [11], [12], [16].

One thing to note is that for a semigroup $S \subset G$ to have any chance of bring transitive in a manifold M, first the group G itself must be transitive. Interestingly, the transitivity of G can be calculated, under some very general conditions, from the differential of the action at the identity. This is a consequence of the following local lemma

Lemma 1.2.1. *Let G be a Lie group acting in a manifold M by*

$$\phi : G \times M \to M$$

$$(g, m) \to gm.$$

Assume G is second countable. Then the following are equivalent for any $x \in M$, where ϕ_x

denotes the application $g \to gx$:

1. *The differential $D_{Id}\phi_x : T_{Id}G \to T_xM$ is surjective.*

2. *ϕ_x is open.*

3. *$x \in Int(Gx)$.*

4. *$Int(Gx) \neq \emptyset$*

Proof. If 1 is true, then, by the submersed manifold theorem, ϕ_x is locally surjective in Id, that is, for any open set V containing Id, $x \subset Vx$. Now, for any open set $W \subset G$ and $w \in W$, we have that $w^{-1}W$ is an open set containing Id, such that $x \in Int(w^{-1}Wx)$. Since ϕ_w is an homeomorphism in M, then $\phi_x w = wx \in Int(Wx) = Int(\phi_x(W))$. Since W and $w \in W$ are arbitrary, ϕ_x is open.

The implication $2 \Rightarrow 3$ can be obtained from the inclusion

$$x = Idx \subset Gx = (IntG)x = Int(Gx)$$

where the last equality is true if condition 2 is true.

Condition 3 implies 4 directly.

The implication $4 \Rightarrow 1$ is more complex, involving concepts which were not discussed in this thesis. As such, we will only provide a sketch of how it can be proven.

Consider the quotient G/H_x where H_x is the isotropy subgroup of x:

$$H_x = \{g \in G; gx = x\}.$$

It can be shown that this quotient admits a natural manifold structure, such that it is second countable if G is second countable, and such that the function

$$f : G/H_x \to M$$

$$gH_x \to gx$$

is an immersion satisfying $f(G/H_x) = Gx$. Furthermore, if G does not satisfy the rank condition in x then $dim(G/H_x) < dim(M)$. Sard's theorem can then be used to show that $f(G/H_x)$ has empty interior in M.

Then, 4 implies 1 by contrapositive. □

Proposition 1.2.2. *Let G be a Lie group acting in a connected manifold M by*

$$\phi : G \times M \to M$$

$$(g, m) \to gm.$$

Assume G second countable. Then the following are equivalent

1. *The differential $D_{Id}\phi_x : T_{Id}G \to T_xM$ is surjective for all $x \in M$*
2. *ϕ_x is open for any $x \in M$*
3. *$x \in Int(Gx)$ for any $x \in M$*
4. *G is accessible in M*
5. *G is transitive in M.*

Proof. The equivalences $1 \Leftrightarrow 2 \Leftrightarrow 3 \Leftrightarrow 4$ are direct consequences of the previous lemma (Remember that $G = G^{-1}$ such that G is accessible if, and only if, Gx has nonempty interior for all $x \in M$). Furthermore, we know that transitivity implies accessibility, such that 5 implies 4. To complete the proof, note that 2 implies Gx is open for all $x \in M$. However, the union

$$\bigcup_{x \in M} Gx$$

can be shown to be a partition of M. Since the Gx are all open and nonempty, and M is connected, then $Gx = M$ for all $x \in M$. □

Remember that a Lie group is second countable if, and only if, it has countable many components. From this point on, all Lie groups are assumed to be second countable, unless stated otherwise.

This proposition is very interesting as it not only gives a way to compute the controllability of Lie group, but also shows that controllability is equivalent to accessibility in this case. When a Lie group G satisfies the hypothesis of $D(\phi_x)_I d$ being surjective in a point $x \in M$ we say that G satisfies the rank condition in x, or that G has full rank in x. If G satisfies the rank condition in all $x \in M$, we say G satisfies the rank condition or has full rank in M.

When studying controllability of semigroups it is usually assumed that the associated Lie group satisfies the rank condition on the entire manifold, as, otherwise, the semigroup is sure to not be controllable. This assumption then implies that G is also accessible, as mentioned in the previous result.

It is also usual to ask the semigroup S to have nonempty interior in G. In some cases, if a semigroup S does have empty interior in G, it is possible to restrict the study to a subgroup H still containing S and such that S has nonempty interior in H. This is not always possible, and depends a lot on the type of semigroup being studied, such that in some occasions it's possible to ask the semigroup to have nonempty interior without loss in generality, while in others this is a restrictive condition. The advantage in asking the semigroup to have nonempty interior is that accessibility of the group then implies accessibility of the semigroup. This is because if G satisfies the rank condition then ϕ_x is open for all x such that $(Int(S))x = \phi_x(Int(S))$ is an open set contained in Sx for all $x \in M$, and $\phi_x(Int(S^{-1})) = \phi_x(Int(S)^{-1})$ is an open set contained in $S^{-1}x$ for all $x \in M$, such that S is accessible.

In particular, if S has non empty interior in G and G satisfies the rank condition in M then S is controllable if, and only if, it is approximately controllable.

A very important type of Lie semigroups are the semigroups generated by sets in the Lie algebra. Let G a Lie group and $C \subset \mathfrak{g}$ a nonempty subset. The semigroup generated by C is the semigroup S generated by all exponentials of elements in C at positive time:

$$S = \langle e^{tc}; t \geq 0, c \in C \rangle$$

$$= \{e^{t_1 c_1} e^{t_2 c_2} ... e^{t_k c_k}; k \in \mathbb{N}, t_1, t_2, ..., t_k \geq 0, c_1, c_2, ..., c_k \in C\}.$$

We will denote such a semigroup by $\langle C \rangle$. These semigroups have an interesting property that allows to compute whether their interior in G is empty.

Proposition 1.2.3. *Let $S \subset G$ be the semigroup generated by a set $C \subset \mathfrak{g}$, and denote by $\mathfrak{h}$ the smallest Lie subalgebra containing C. $Int(S)$ is nonempty in G if, and only if, $\mathfrak{h} = \mathfrak{g}$.*

If S is generated by a set $C \subset \mathfrak{g}$ that is not contained in any proper sub algebra of $\mathfrak{g}$, we say that S satisfies the rank condition in G, or that S has full rank in G. By the previous proposition, if S satisfies the rank condition on G then it has nonempty interior in G, and if, furthermore, G satisfies the rank condition in M, then S is accessible in M.

On the other hand, if S does not satisfy the rank condition in G, then there is a smallest sub algebra $\mathfrak{h}$ that contains C. Denoting by H the connected subgroup of G with Lie algebra $\mathfrak{h}$, we have that $S \subset H$, such that for these kinds of semigroups it is always possible to restrict the study to a subgroup where S has nonempty interior. Note that in the process of restricting the subgroup like this we might lose the rank condition on the manifold, as the application $D(\phi_x)_{Id}$ becomes restricted to $\mathfrak{h}$. Nonetheless, this have an interesting consequence. Let S be a semigroup generated by a set $C \subset \mathfrak{g}$, and let $D \subset \mathfrak{g}$ the smallest closed convex cone containing C. If $R = \langle D \rangle$ then it can be shown that $\overline{S} = \overline{R}$. This is because $e^{t(\alpha X)} = e^{(t\alpha)X}$ and $e^{t(X+Y)}$ can be arbitrarily approximated by the concatenations in the form

$$\left(e^{\frac{t}{k}X}e^{\frac{t}{k}Y}\right)^k$$

with $k \in \mathbb{N}$. This then allows the following result.

Proposition 1.2.4. *Let $S \subset G$ be a semigroup generated by a set $C \subset \mathfrak{g}$ and acting in a manifold M, and let D the smallest closed convex cone containing C and $R = \langle D \rangle$. Then S is controllable in M if, and only if, R is controllable in M.*

Proof. Let $\mathfrak{h}$ be the smallest sub algebra containing C, note that $D \subset \mathfrak{h}$ as $\mathfrak{h}$ is a subspace and, in particular, a closed convex cone containing C. If H denotes the connected subgroup generated by $\mathfrak{h}$ then both S and R are contained in H and satisfy the rank condition in H. Then, if H is not controllable in M, neither S nor R are controllable. Otherwise, controllability is equivalent to approximate controllability for S, and the same for R, such that it suffices to show that local controllability for these two subgroups is equivalent. In fact, we have $Sx \subset \overline{S}x \subset \overline{Sx}$ and $Rx \subset \overline{R}x \subset \overline{Rx}$ such that

$$\overline{Sx} = \overline{\overline{S}x} = \overline{\overline{R}x} = \overline{Rx}$$

showing that approximate controllability is equivalent for these two semigroups. □

Note that $S = \langle C \rangle$ coincides with the positive orbit from Id of the continuous control system $(G, \mathcal{F}, C, \mathcal{U})$ where $\mathcal{U}$ is the set of all piecewise constant functions $u : \mathbb{R} \to C$ and $\mathcal{F}$ is defined by $\mathcal{F}(g, X) = X^r(g)$. Here, X^r denotes the only right invariant vector field satisfying $X^r(Id) = X$. It can be calculated in a point $g \in G$ using the right translation

$$R_g : G \to G$$

$$h \to hg$$

by

$$X^r(g) = D(R_g)_{Id}(X).$$

S^{-1} coincides with the negative orbit from Id of this same system. This control system can be transported to the manifold M, by defining the function $\mathcal{F}_2$ as

$$\mathcal{F}_2(x, X) = \left.\frac{d}{dt}\right|_{t=0} e^{tX} x.$$

If ϕ_1, ϕ_2 denote the solutions of $(G, \mathcal{F}, C, \mathcal{U})$ and $(M, \mathcal{F}_2, C, \mathcal{U})$, respectively, it can be shown that $\phi_2(x, u, T) = \phi_1(Id, u, T)x$, such that the positive, negative orbits from the second system coincide with $Sx, S^{-1}x$, respectively, for all $x \in M$. As such, controllability properties of the continuous control system $(M, \mathcal{F}_2, C, \mathcal{U})$ are equivalent controllability properties of S.

Interestingly, if we instead define $\mathcal{U}$ as the larger set of integrable functions $u : \mathbb{R} \to C$, we still get the equality $\phi_2(x, u, T) = \phi_1(Id, u, T)x$. In this case, the positive orbit from Id in the system $(G, \mathcal{F}, C, \mathcal{U})$ can still be shown to be a semigroup R, such that controllability properties of $(M, \mathcal{F}_2, C, \mathcal{U})$ are equivalent to controllability properties of the semigroup R. Furthermore, it can be shown that $\overline{R}$ coincides with the closure of the previous semigroup S. This happens because integrable functions can be approximated by piece-wise constant function. Then, an argument similar to the previous proposition shows that controllability for R is equivalent to controllability for S. As such, not much in gained by adding these extra functions to $\mathcal{U}$.

One natural question is which continuous control systems can be obtained in a similar way as subsets of the Lie algebra of some Lie group acting on the respective manifold. This is equivalent to the question of whether given a set C of vector fields in M there is a Lie group G acting differentiably in M and a function $f : C \to \mathfrak{g}$ such that, for any $X \in C$ and $x \in M$,

$$X(x) = \frac{d}{dt} e^{tf(X)} x.$$

A very important result which answers this question is Lie-Palais theorem (see

[20]), which assures that such a group G exists if, and only if, the set C generates a Lie algebra of complete vector fields with finite dimension.

Theorem 1.2.5. *(Lie-Palais) Let $\mathfrak{h}$ a real Lie algebra of smooth vector fields in a manifold M. Assume all $X \in \mathfrak{g}$ are complete and that $\mathfrak{h}$ has finite dimension. Then there is a connected Lie group G acting in M by ϕ such that the function*

$$f : \mathfrak{g} \to \mathfrak{h}$$

$$X \to f(X)$$

where

$$f(X)(x) = D(\phi_x)_{Id}(v) = \frac{d}{dt}e^{tX}x$$

is an isomorphism of Lie algebras.

For the inverse implication, it can be shown that if G is a Lie group acting in a manifold M and f is as in the theorem, then $f(X)$ is complete for all $X \in \mathfrak{g}$ and f is a Lie homomorphism such that $f(\mathfrak{g})$ is a Lie algebra with finite dimension.

Remember that we ask the differential equations in a continuous control system to be global, such that the vector fields $\mathcal{F}_c$ are complete for each $c \in G$. Thus, a continuous control system defined from smooth vector fields can be viewed as the semigroup of a connected Lie group G acting differentiably on M if the set $\{\mathcal{F}_c; c \in U\}$ generates a Lie algebra of finite dimension.

Chapter 4 will use many results from Flag theory for semigroups in semissimple Lie groups. This is a very rich and deep theory, and is worth an entire study on its on. We talk more about it in section 4 itself.

CHAPTER 2

THE SYSTEM $\dot{x} = Ax + a + Bu$

2.1 Preliminaries

The control systems studied in this chapter are defined by families of differential equations in the form

$$\frac{d}{dt}x(t) = Ax(t) + a + Bu(t)$$

$$A \in M_n, B \in M_{n\times m}, a \in \mathbb{R}^n, u \in \mathcal{U},$$

where $\mathcal{U} = \{u : \mathbb{R} \to U; u \text{ is integrable}\}$, and U is a nonempty subset of $\mathbb{R}^m$. By integrable we mean that u is Riemann integrable in any interval of $\mathbb{R}$.

When a function u and an initial point $x(0) = x_0$ are fixed, the equation above becomes an ordinary differential equation with unique and global solution, such that the solution depends smoothly on the starting condition.

Using the notation for continuous control systems introduced in section 1.1, such a control system is defined by an ordered pair $(\mathbb{R}^n, \mathcal{F}, U, \mathcal{U})$ where $U, \mathcal{U}$ are as defined above and $\mathcal{F}$ is defined by

$$\mathcal{F} : \mathbb{R}^n \times U \to \mathbb{R}^n$$

$$(x, c) \to Ax + a + Bc.$$

Here, the tangent bundle of $\mathbb{R}^n$ is associated with the space itself.

Such control system is completely determined by A, a, B, U, therefore, we will denote it by $(A, a, B)_U$. If $U = \mathbb{R}^m$, we will also use the notation (A, a, B).

The control system $(A, a, B)_U$ can be shown to be equivalent to a semigroup S in the affine Lie group $\mathrm{Aff}(\mathbb{R}^n)$, generated by the set

$$\{(A, a + Bc); c \in U\} \subset \mathfrak{aff}(\mathbb{R}^n).$$

As previously mentioned, nothing is lost in terms of controllability by instead restricting $\mathcal{U}$ to piece-wise constant function and requiring U to be a convex set. In section 2.3 we also include a proof that is specific for the case considered in this chapter.

As usual, we denote the solution of the system by ϕ. Remember that the solution is a function $\phi : \mathbb{R}^n \times \mathcal{U} \times \mathbb{R} \to \mathbb{R}^n$ where $\phi(x_0, u, T)$ is defined as the solution of $\dot{x}(t) = \mathcal{F}(x(t), u(t))$ on time T.

A linear control system is defined by a family of differential equations in the form

$$\frac{d}{dt}x(t) = Ax(t) + Bu(t)$$

$$A \in M_n, B \in M_{n\times m}, u \in \mathcal{U},$$

where $\mathcal{U} = \{u : \mathbb{R} \to U; u \text{ is locally integrable}\}$, and $U \subset \mathbb{R}^m$ is nonempty.

This differential equation also has unique and global solution depending smoothly on the starting conditions, such that it also defines a control system $(\mathbb{R}^n, \mathcal{F}_l, U, \mathcal{U})$ where

$$\mathcal{F}_l(x, c) = Ax + Bc,$$

similar to the previous system. This system also satisfies the conditions of the Lie-Palais theorem, and can be associated with the semigroup in $\mathrm{Aff}(\mathbb{R}^n)$ generated by

$$\{A + Bc; c \in U\}.$$

As in the previous case, nothing is lost in terms of controllability if $\mathcal{U}$ is restricted to piece-wise constant functions or if U is required to be convex.

The linear system is completely determined by A, B, U, and we will denote such system by $(A, B)_U$, or, if $U = \mathbb{R}^m$, by (A, B). To avoid ambiguity, we will sometimes denote the linear system's solution by ϕ' if the symbol ϕ is already being used to denote the solution of the system $(A, a, B)_U$. This will be made clear beforehand in the cases where it is used.

The problem studied in this chapter is the one of finding conditions for the controllability or uncontrollability of the affine systems previously described.

The first result which will be shown is regarding the solution of those systems.

Proposition 2.1.1. *The solution of the affine system $(A, a, B)_U$ is given by:*

$$\phi(x_0, u, T) = e^{TA}x_0 + \int_0^T e^{(T-s)A}(Bu(s) + a)\, ds.$$

Proof. The affine solution for the class studied can be derived from the linear solution. For each control u, let u' be defined by $u'(t) = \begin{pmatrix} u(t) \\ 1 \end{pmatrix} \in \mathbb{R}^{m+1}$, where the elements of $\mathbb{R}^m$ and $\mathbb{R}^{m+1}$ are viewed as column-matrices, and let $B' = \begin{pmatrix} B & a \end{pmatrix} \in \mathbb{R}^{n\times(m+1)}$. Then

$$B'u'(t) = Bu(t) + a$$

for all $t \in \mathbb{R}$. Consequently, the ordinary differential equation associated with control u can be rewritten as a differential equation from the linear system:

$$\dot{x} = Ax(t) + a + Bu(t) = Ax(t) + B'u'(t).$$

Since the differential equations coincide, then ϕ is also solution to the system (A, B') with control u'. It's a known fact (see [2, 3] for details) that said solution is unique and is the function

$$e^{TA}x_0 + \int_0^T e^{(T-s)A}(B'u'(s))ds.$$

Therefore,

$$\phi(x_0, u, T) = e^{TA}x_0 + \int_0^T e^{(T-s)A}(B'u'(s))ds =$$

$$e^{TA}x_0 + \int_0^T e^{(T-s)A}(Bu(s) + a)ds$$

□

2.2 Unrestricted case

In this section we show necessary and sufficient conditions for the controllability of the unrestricted systems (A, a, B). A very useful construction is the following quotient. This idea was used by Willens in the Section 5 of [21].

Given V a subspace of $\mathbb{R}^n$ and a function $f : \mathbb{R}^n \to \mathbb{R}^n$, we say that f can be projected on $\mathbb{R}^n/V$ if $a + V = b + V$ implies $f(a) + V = f(b) + V$ for all $a, b \in \mathbb{R}^n$, and define the projection of f as the function:

$$\overline{f} : \mathbb{R}^n/V \to \mathbb{R}^n/V$$

$$x + V \to \overline{f}(x + V) := f(x) + V.$$

If f can be projected on $\mathbb{R}^n/V$ then it's projection is well defined.

In a similar way, given a control system defined by a flow $\phi : \mathbb{R}^n \times \mathcal{U} \times \mathbb{R} \to \mathbb{R}^n$, we say that it can be projected on $\mathbb{R}^n/V$ if, for all fixed u and T the function $x \to \phi(x, u, T)$ can be projected, and define the projected flow by

$$\overline{\phi}(x + V, u, T) = \phi(x, u, T) + V.$$

As in the previous case, if the control system can be projected then the projected flow is well defined. Furthermore, orbits in the original system project into orbits in the projected system. In particular, if the original system is controllable then the projected system is also controllable. For this reason projections will be very useful for showing non controllability of some systems: if we can project a system on a a system that is not controllable then the original system is also not controllable.

Natural examples for functions that can be projected are linear transformations on their invariant spaces, and also the exponentials of these transformations on those same spaces, since if a space is invariant under a linear transformation then it is also invariant under it's exponential. In any of those cases, the projected function is still linear. It is also possible to show that, if A is a linear transformation, V is one of it's invariant spaces, and $\overline{A}, \overline{e^{tA}}$ denote, respectively, the projections of A and e^{tA} on $\mathbb{R}^n/V$, then $e^{t\overline{A}} = \overline{e^{tA}}$.

An interesting invariant subspace shows up on linear systems: in an unrestricted system (A, B), the positive and negative orbits from the origin coincide as the same set, and are both A invariant subspaces. In fact, if $\mathcal{O}$ denotes the positive/negative orbit from the origin, then $\mathcal{O}$ is the image of the Kalman matrix

$$\begin{pmatrix} A^{n-1}B & A^{n-2}B & ... & AB & B \end{pmatrix}$$

(see [2, 3]) or, equivalently, $\mathcal{O}$ is the smallest A invariant subspace containing the image of B. In the next result we show that $(A, a, B)_U$ can be projected on $\mathbb{R}^n/\mathcal{O}$.

Lemma 2.2.1. *Consider the system $(A, a, B)_U$ and let $\mathcal{O}$ be the positive/negative orbit from the origin by the linear system (A, B). Then $(A, a, B)_U$ can be projected on $\mathbb{R}^n/\mathcal{O}$. Furthermore, the solution of the projected system is in given by*

$$\overline{\phi}(x + \mathcal{O}, u, T) = e^{TA}x + \int_0^T e^{sA}a\, ds + \mathcal{O}.$$

Proof. Let ϕ' denote the flow of the linear system (A, B), and let $x, y \in \mathbb{R}^n$ be such that $x - y \in \mathcal{O}$. Then:

$$\phi(x, u, T) - \phi(y, u, T) =$$

$$= e^{TA}x + \int_0^T e^{(T-s)A}(Bu(s) + a)\, ds - e^{TA}y - \int_0^T e^{(T-s)A}(Bu(s) + a)\, ds =$$

$$= e^{TA}(x - y) = \phi'(x - y, 0, T).$$

By hypotheses, $x - y \in \mathcal{O}$. If $T = 0$, then $\phi'(x - y, 0, T) = x - y \in \mathcal{O}$. If $T > 0$ then $\phi'(x - y, 0, T)$ is in the positive orbit from $x - y$, while $x - y$ is in the positive orbit from the origin. Therefore, $\phi'(x - y, 0, T)$ is in the positive orbit from the origin, that is, $\phi'(x - y, 0, T) \in \mathcal{O}$. Analogously, if $T < 0$ then $\phi'(x - y, 0, T)$ is in the negative orbit from the origin, and, therefore, is in $\mathcal{O}$.

Therefore, $\phi(x, u, T) + \mathcal{O} = \phi(y, u, T) + \mathcal{O}$, and the system can be projected in $\mathbb{R}^n/\mathcal{O}$.

For the solution, we have:

$$\overline{\phi}(x + \mathcal{O}, u, T) = \phi(x, u, T) + \mathcal{O} =$$

$$= e^{TA}x + \int_0^T e^{(T-s)A}(Bu(s) + a)\, ds + \mathcal{O} =$$

$$= e^{TA}x + \int_0^T e^{(T-s)A}Bu(s)\,ds + \int_0^T e^{(T-s)A}a\,ds + \mathcal{O} =$$

$$= e^{TA}x + \phi'(0,u,T) + \int_0^T e^{(T-s)A}a\,ds + \mathcal{O}$$

$$= e^{TA}x + \int_0^T e^{(T-s)A}a\,ds + \mathcal{O}\ ,$$

writing $k = T - s$ in the integral we get:

$$e^{TA}x + \int_0^T e^{(T-s)A}a\,ds + V = e^{TA}x + \int_T^0 -e^{kA}a\,dk + \mathcal{O} =$$

$$= e^{TA}x + \int_0^T e^{kA}a\,dk + V = e^{TA}x + \int_0^T e^{sA}a\,ds + \mathcal{O}\ .$$

Ending the proof. □

Note that the solution from the projected system is independent from the control such that the trajectory from a point will always be same, regardless of the control chosen. That is a good indication that these systems will never be controllable except for trivial cases, such as when $\mathbb{R}^n/\mathcal{O}$ is an unitary set. This is proven in the next results.

Lemma 2.2.2. *The projected system from lemma 2.2.1 can be projected once again on $Img(\overline{A})$, and the solution of this second projection is*

$$\overline{\overline{\phi}}(x + \mathcal{O} + Img(\overline{A}), u, T) = x + Ta + \mathcal{O} + Img(\overline{A}).$$

Proof. Let $x + \mathcal{O}$, $y + \mathcal{O}$ be such that $x + \mathcal{O} - y + \mathcal{O} = x - y + \mathcal{O} \in Img(\overline{A})$. Then

$$\overline{\phi}(x + \mathcal{O}\ , u, T) - \overline{\phi}(y + \mathcal{O}\ , u, T) =$$

$$= e^{TA}x + \int_0^T e^{sA}a\,ds - e^{TA}y - \int_0^T e^{sA}a\,ds + \mathcal{O} =$$

$$= e^{TA}(x - y) + \mathcal{O} = e^{T\overline{A}}(x - y + \mathcal{O}\ ).$$

Since $x - y + \mathcal{O} \in Img(\overline{A})$ and $Img(\overline{A})$ is $\overline{A}$ invariant, then $e^{T\overline{A}}(x - y + \mathcal{O}\ ) \in Img(A)$, showing that the system can be projected.

Now for the flow, note that if $\overline{\overline{A}}$ is the projection of $\overline{A}$ then $\overline{\overline{A}}$ is the null transformation, since $\overline{A}(x+\mathcal{O}\ ) \in Img(\overline{A})$ for all $x+\mathcal{O} \in \mathbb{R}^n/\mathcal{O}$, and, therefore, $e^{T\overline{\overline{A}}}$ is the identity transformation for any real T. Therefore,

$$e^{tA}x + \mathcal{O} + Img(\overline{A}) = e^{t\overline{\overline{A}}}(x + \mathcal{O} + Img(\overline{A})) = x + \mathcal{O} + Img(\overline{A}),$$

for any $t \in \mathbb{R}, x \in \mathbb{R}^n$, and, therefore:

$$\overline{\overline{\phi}}(x + \mathcal{O} + Img(\overline{A}), u, T) = e^{TA}x + \int_0^T e^{sA}a\, ds + \mathcal{O} + Img(\overline{A})$$

$$= x + \int_0^T a\, ds + \mathcal{O} + Img(\overline{A}) =$$

$$= x + Ta + \mathcal{O} + Img(\overline{A}).$$

□

Theorem 2.2.3. *The projected system*

$$\overline{\phi}(x + \mathcal{O}, u, T) = e^{TA}x + \int_0^T e^{sA}a\, ds + \mathcal{O}$$

is controllable if, and only if, $\mathbb{R}^n/\mathcal{O}$ *has dimension* 0*, or, equivalently,* $\mathcal{O} = \mathbb{R}^n$*.*

Proof. If $\mathbb{R}/\mathcal{O}$ has dimension 0 then it is an unitary set, and, therefore, the system is controllable. We have to show that the system is not controllable if the space has dimension greater or equal to one. For that, we consider two cases,

$$a + \mathcal{O} \in Img(\overline{A})$$

or

$$a + \mathcal{O} \notin Img(\overline{A}).$$

If $a + \mathcal{O} \in Img(\overline{A})$ then there exists $a' + \mathcal{O}$ such that $\overline{A}(-a' + \mathcal{O}) = a + \mathcal{O}$. Then the orbit of $-a' + \mathcal{O}$ is $\{a' + \mathcal{O}\}$, in fact, for all $T \in \mathbb{R}, u \in \mathcal{U}$:

$$\overline{\phi}(a' + \mathcal{O}, u, T) = -e^{TA}a' + \int_0^T e^{sA}a\, ds + \mathcal{O} =$$

$$= -e^{TA}a' + a' - a' + \int_0^T e^{sA}a\, ds + \mathcal{O} = -(e^{TA}a' - a') - a' + \int_0^T e^{sA}a\, ds + \mathcal{O}$$

$$= -\int_0^T e^{sA} Aa' \, ds - a' + \int_0^T e^{sA} a \, ds + \mathcal{O} \; = -a' + \int_0^T e^{sA}(a - Aa') \, ds + \mathcal{O} \; .$$

By definition of a' we have that $a - Aa' \in \mathcal{O}$, and, since $\mathcal{O}$ is A invariant, then $e^{sA}(a - Aa') \in \mathcal{O}$ for all $s \in \mathbb{R}$. Therefore the integral in the last term is contained in $\mathcal{O}$ and is null in the quotient. Then

$$\overline{\phi}(a' + \mathcal{O} \; , u, T) = a' + \mathcal{O} \; ,$$

and the orbit of $a' + \mathcal{O}$ is $\{a' + \mathcal{O} \; \}$. Since we're assuming that $\mathbb{R}^n/\mathcal{O}$ has dimension greater than 0, and, therefore, isn't unitary, then $\{a' + \mathcal{O} \; \}$ is a proper set, and the system is not controllable.

If $a + \mathcal{O} \; \notin Img(\overline{A})$, then $\overline{A}$ is not surjective in $\mathbb{R}^n/\mathcal{O}$. In particular, the quotient $(\mathbb{R}^n/\mathcal{O} \;)/\overline{A}$ has dimension greater than 0. By lemma 2.2.2 we can project the system in this quotient, and the projected system is given by

$$\overline{\overline{\phi}}(x + \mathcal{O} \; + Img(\overline{A}), u, T) = x + Ta + \mathcal{O} \; + Img(\overline{A}).$$

In particular, the positive orbit from the origin is the set $\{Ta + \mathcal{O} \; + Img(\overline{A}; T \geq 0\}$. That set is either a ray or a single point, and, since $(\mathbb{R}/\mathcal{O} \;)/Img(\overline{A})$ has dimension greater than 0, it is also proper. Therefore the projected system is not controllable. As mentioned before, this implies that the system from the theorem is also not controllable. □

Corollary 2.2.4. *If (A, B) is not controllable then (A, a, B) is not controllable.*

Proof. The controllability of (A, B) is equivalent to the condition $\mathcal{O} \; = \mathbb{R}^n$. If (A, B) is not controllable then $\mathcal{O} \; \neq \mathbb{R}^n$ is not controllable, which, by previous theorem, implies that (A, a, B) can be projected in a non controllable system and therefore is not controllable. □

The previous corollary gives a necessary condition for the controllability of the unrestricted systems (A, a, B) studied in this section. In the next part of this section we show that this condition is also sufficient for the controllability of these systems.

Lemma 2.2.5. *If the linear system (A, B) is controllable in $\mathbb{R}^n$ then any vector $x \in \mathbb{R}^n$ can be written as $Ax_A + Bx_B$ where $x_A \in \mathbb{R}^n, x_B \in \mathbb{R}^m$.*

Proof. Let

$$V := Img(A) + Img(B) = \{a + b; a \in Img(A), b \in Img(B)\}$$

Since $Img(A), Img(B)$ are both subspaces, then so is V. Furthermore, we have that $Img(B) \subset V$ and $A(V) \subset Img(A) \subset V$, such that V is an A-invariant subspace that contains $Img(B)$. Since $\mathcal{O}$ is the smallest subspace with these properties, then $\mathcal{O} \subset V$. If we assume that (A, B) is controllable, then $\mathcal{O} = \mathbb{R}^n$ such that $x \in \mathbb{R}^n = \mathcal{O} \subset V$. By definition of V, we have

$$x = a + b$$

where $a \in Img(A), b \in Img(B)$. Since a, b are in the respective images, then there are $x_a \in \mathbb{R}^n, x_b \in \mathbb{R}^m$ such that $a = Ax_a, b = Bx_B$, and, therefore,

$$x = Ax_A + Bx_B$$

. □

Theorem 2.2.6. *The system (A, a, B) is controllable if, and only if, the linear system (A, B) is controllable.*

Proof. It was already show in previous results that controllability for (A, a, B) implies controllability for (A, B) (corollary 2.2.4). What is left is to show the other implication.

Assume (A, B) is controllable. Then, by lemma 2.2.5, $a = Aa_A + Ba_B$ for some $a_A \in \mathbb{R}^n, a_B \in \mathbb{R}^m$. Denote by ϕ the flow of the affine system (A, a, B) and by ϕ' the flow of the linear system (A, B). Note that

$$\frac{d}{dt}(\phi'(x_0 + a_A, u + a_B, t) - a_A) =$$

$$A\phi'(x_0 + a_A, u + a_B, t) + B(u(t) + a_b) =$$

$$A(\phi'(x_0 + a_A, u + a_B, t) - a_A) + Aa_A + Ba_B + B(u(t)) =$$

$$A(\phi'(x_0 + a_A, u + a_B, t) - a_A) + a + Bu(t)$$

and

$$\phi'(x_0 + a_A, u + a_B, 0) - a_A = x_0 + a_A - a_A = x_0,$$

that is, $\phi'(x_0 + a_A, u + a_B, T) - a_A$ is solution to the differential equation

$$\dot{x} = Ax + a + Bu.$$

By uniqueness of solution,

$$\phi(x_0, u, T) = \phi'(x_0 + a_A, u + a_B, T) - a_A$$

for all $x_0 \in \mathbb{R}^n, u \in \mathcal{U}, T \in \mathbb{R}$. Now choose arbitrary $x, y \in \mathbb{R}^n$. Since (A, B) is controllable, then there exists $u \in \mathcal{U}$ and $T > 0$ such that

$$\phi'(x + a_A, u, T) = y + a_A.$$

Then:

$$\phi(x, u - a_B, T) = \phi'(x + a_A, u, T) - a_A = y.$$

□

2.3 Restricted case

In this section we study the controllability of the restricted systems $(A, a, B)_U$, assuming that U is a bounded subset of $\mathbb{R}^m$. Without loss in generality, we also assume that U is not contained in a proper affine subspace of $\mathbb{R}^m$. In fact, if it is the case that $U \subset V + b$ where V is a proper subspace of $\mathbb{R}^m$ and $b \in \mathbb{R}^m$, then the system $(A, a, B)_U$ can be shown equivalent to $(A, a + B(b), B|_V)_{U-b}$, in the sense that both systems are composed of the same vector fields and the first system is controllable if and only if the second one is. Therefore, if U is contained in a proper affine subspace of $\mathbb{R}^m$, then the problem can be simplified into a version where U is not contained in any proper affine subspace.

Our objective is to show conditions for the controllability of the systems $(A, a, B)_U$. The first such condition is derived from the previous section. Since $(A, a, B)_U$ is a restriction of (A, a, B), then all of it's trajectories will also be trajectories from the system (A, a, B). In particular, if $(A, a, B)_U$ is controllable then (A, a, B) is also controllable. That means that a necessary condition for the controllability in the restricted version is the controllability of the unrestricted version, or, equivalently, the controlability of (A, B).

This condition is quite useful because, as we will see in the next results, $(A, a, B)_U$ is equivalent to a translated restricted linear system whenever (A, B) is controllable.

Proposition 2.3.1. *Let $(A, a, B)_U$ be such that $a = Aa_A + Ba_B$ for some $a_A \in \mathbb{R}^n, a_B \in \mathbb{R}^n$, and let $W = U + a_B$. Then the system $(A, a, B)_U$ is controllable if, and only if, $(A, B)_W$ is controllable.*

Proof. Let ϕ denote the solution of $(A, a, B)_U$ and ϕ' the solution of $(A, B)_W$. Then, for $x \in \mathbb{R}^n$ and a control u from $(A, a, B)_U$,

$$\frac{d}{dt}\phi'(x + a_A, u + a_B, t) - a_A = A(\phi'(a + a_A, u + a_B, t)) + B(u(t) + a_B)$$

$$= A(\phi'(a + a_A, u + a_B, t) - a_B) + B(u(t)) + Aa_A + Ba_B$$

$$= A(\phi'(a + a_A, u + a_B, t) - a_B) + B(u(t)) + a,$$

and

$$\phi'(x + a_A, u + a_B, 0) - a_A = x + a_A - a_A = x,$$

therefore, $\phi'(x + a_A, u + a_B, T) - a_A$ satisfies the diferential equation of the system $(A, a, B)_U$, and, therefore $\phi'(x + a_A, u + a_B, T) - a_A = \phi(x, u, T)$. Furthermore, $(u - a_B)(t) \in W$ if, and only if, $u(t) \in U$. Therefore, the trajectories from one of the systems are the translated trajectories from the other one, and their controlabilities are equivalent. □

Corollary 2.3.2. *If (A, B) is controllable then there is $v \in \mathbb{R}^m$ and $W = U + v$ such that the controllability of $(A, a, B)_U$ is equivalent to the controllability of $(A, B)_W$.*

Remember that a_A, a_B as above exist whenever (A, B) is controllable, such that the corollary is a direct implication from the proposition.

There is a know result about the controllability of $(A, B)_U$: if $0 \in Int(U)$ and U is bounded, then $(A, B)_U$ is controllable if, and only if (A, B) is controllable and all the eigenvalues of A are 0 or purely imaginary [2, 3]. Note that this condition does not depend on the restriction U, and, therefore, the controllability of $(A, B)_U$ is equivalent for all restrictions U that are bounded and contain the origin in their interior. That motivates the following definition:

Definition 2.3.3. *We say that a linear system (A, B) is controllable restricted to the origin (C.R.O) if $(A, B)_U$ is controllable for some restriction U that is bounded and contains the origin in it's interior. Equivalently, (A, B) is C.R.O. if $(A, B)_U$ is controllable for all such restrictions, or, also equivalently, if (A, B) satisfies the Kalman rank condition and all eigenvalue of A are 0 or purely imaginary [2, 3].*

Proposition 2.3.4. *If U is bounded and $(A, B)_U$ is controllable, then (A, B) is C.R.O.*

Proof. Let $W = U \cup B(0, 1)$, where $B(0, 1)$ is the open ball in $\mathbb{R}^m$ centered on the origin. Note that W is still bounded and $0 \in Int(W)$. Furthermore, $U \subset W$, and then, the set of controls of $(A, B)_U$ is contained in the set of controls of $(A, B)_W$. Therefore, the controllability of $(A, B)_U$ implies the controllability of $(A, B)_W$. Since W is a bounded set that contains the origin in this interior, then (A, B) is C.R.O. □

Corollary 2.3.5. *If $(A, a, B)_U$, with bounded U, is controllable, then (A, B) is C.R.O.*

Proof. We saw that the controllability of $(A, a, B)_U$ implies the controllability of $(A, B)_W$ where $W = U+v$ for some $v \in \mathbb{R}^m$. Since U is bounded, then $W = U+v$ is also bounded, therefore, by the previous proposition, (A, B) is C.R.O. □

This corollary gives us another necessary condition for the controllability of $(A, a, B)_U$: (A, B) must be C.R.O.

For the next and last condition, we need some results about convex sets. Some of those results are topological ones. An useful fact connecting both is that if C is a convex set then $x \in Int(C)$ if, and only if, there exists a basis β of $\mathbb{R}^n$ such that $x + jb \in Int(C)$ for all $b \in \beta$ and $j \in \{-1, 1\}$.

A definition that is very useful when studying convex sets is the definition of convex closure.

Definition 2.3.6. *Let $C \subset \mathbb{R}^n$. The convex closure of C is the set of all convex sums of elements in C. We will denote it by $cv(C)$:*

$$cv(C) = \left\{ x \in \mathbb{R}^n; \exists\, (x_1, ..., x_k \in C;\ a_1, ..., a_k \in [0, 1]); \sum_{i=1}^{k} a_i x_i = x \text{ and } \sum_{i=1}^{k} a_i = 1 \right\}.$$

Note that the convex closure itself is a convex set. In fact, $cv(C)$ can be equivalently defined as the intersection of all convex set that contain C, or the smallest convex set containing C.

Lemma 2.3.7. *If $C \subset \mathbb{R}^n$ is a convex set and $0 \notin Int(C)$, then the set*

$$D := \overline{\{\alpha x; \alpha > 0 \text{ e } x \in C\}}$$

doesn't contain the origin in it's interior and is a convex cone, that is, if $x_1, x_2, ..., x_d \in D$ and $\alpha_1, \alpha_2, ..., \alpha_d$ are positive real numbers then $\sum_{i=1}^{d} \alpha_d x_d \in D$.

Proof. First we show that D is a convex cone. For that, consider the set

$$E := \{\alpha x; \alpha > 0 \text{ e } x \in C\}.$$

Note that $D = \overline{E}$. We will show that E is a convex cone. In fact, given $\alpha_1 c_1, \alpha_2 c_2, ..., \alpha_d c_d \in E$ and $\lambda_1, \lambda_2, ..., \lambda_d > 0$, let

$$M = \sum_{i=1}^{d} \lambda_i \alpha_i > 0.$$

Note that

$$\sum_{i=1}^{d} \lambda_i \alpha_i c_i = M \sum_{i=1}^{d} \frac{\lambda_i \alpha_i}{M} c_i,$$

where

$$\sum_{i=1}^{d} \frac{\lambda_i \alpha_i}{M} c_i$$

is a convex sum of elements of C, and, therefore, is in C. Then,

$$M \sum_{i=1}^{d} \frac{\lambda_i \alpha_i}{M} c_i \in E.$$

Therefore E is a convex cone. Now, to show that D is a convex cone, let $x_1, x_2, ..., x_d \in D$ and $\alpha_1, \alpha_2, ..., \alpha_d > 0$. Since $D = \overline{E}$ then, for each x_i there is a sequence $x_i^1, x_i^2, ... \in E$ converging for x_i. Then,

$$\sum_{i=1}^{d} \alpha_i x_i = \sum_{i=1}^{d} \alpha_i \lim_{j \to +\infty} x_i^j = \lim_{j \to +\infty} \sum_{i=1}^{d} \alpha_i x_i^j \in \overline{E} = D.$$

Now we show that $0 \notin Int(D)$. Assume, by contradiction, that $0 \in Int(D)$. Then, for some $\epsilon > 0$, $\epsilon e_1, \epsilon e_2, ..., \epsilon e_n, -\epsilon e_1, -\epsilon e_2, ..., -\epsilon e_n \in D$, where $e_1, e_2, ..., e_n$ denote the vectors in the canonical base. But D is a convex cone, therefore $\pm e_1, \pm e_2, ..., \pm e_n \in D$. Since $D = \overline{E}$, then, for each e_i there is a sequence $a_i^1, a_i^2, ...$ and $b_i^1, b_i^2, ...$ such that $\lim_{j\to+\infty} a_i^j = e_i$ and $\lim_{j\to+\infty} b_i^j = -e_i$. Then, for suficiently big k,

$$0 \in Int(cv(\{a_1^k, a_2^k, ..., a_n^k, b_1^k, b_2^k, ..., b_n^k\})) \subset Int(E)$$

$$\Rightarrow 0 \in Int(E).$$

Then

$$\pm e_1, \pm e_2, ..., \pm e_n \in E,$$

and, therefore,

$$\alpha_1 e_1, \alpha_2 e_2, ..., \alpha_n e_n, -\beta_1 e_1, -\beta_2 e_2, ..., -\beta_n e_n \in C$$

for some positive numbers $\alpha_1, \alpha_2, ..., \alpha_n, \beta_1, \beta_2, ..., \beta_n$. If

$$\epsilon = \min(\alpha_1, \alpha_2, ..., \alpha_n, \beta_1, \beta_2, ..., \beta_n) > 0,$$

then $\pm\epsilon e_1, \pm\epsilon e_2, ..., \pm\epsilon e_n \in C$, and, then, $0 \in Int(C)$, contradicting the hypothesis $0 \notin Int(C)$. □

Lemma 2.3.8. *If $C \subset \mathbb{R}^2$ is a convex set such that $0 \notin Int(C)$, then there exists a basis X, Y of $\mathbb{R}^2$ such that, for every $x = \alpha X + \beta Y \in C$, $\alpha \geq 0$*

Proof. Let $D := \overline{\{\alpha x; \alpha > 0, x \in C\}}$. By lema 2.3.7, D is a convex cone and $0 \notin Int(D)$. Note that if $D = \{0\}$ or if $D = \emptyset$, then $C = \{0\}$ or $C = \emptyset$, and the lemma is trivial. Otherwise, if δD denotes the boundary of D, then $\delta D \not\subset \{0\}$. Then, there is a nonzero $Y \in \delta D$. If $D \subset \langle Y \rangle$, then, for any X which is linearly independent to Y, X, Y is the desired basis, as any element in D and, in particular C, is in the form $0X + \beta Y$, where $0 \geq 0$. Otherwise, let $X \in D$ such that $X \notin \langle Y \rangle$. Then X, Y are linearly independent, and are a basis. We will show that this basis is as described in the lemma.

Assume, by contradiction, that there is $x = -\alpha X + \beta Y \in C$ such that $\alpha > 0$. Since $C \subset D$, then $x \in D$. Furthermore, $Y \in D$, since $Y \in \delta D$ and D is a closed set, and $X \in D$ by construction. D is closed for positive linear sums, since it's a convex cone.

We will show that $Y - \epsilon_1 X \in D$, for some $\epsilon_1 > 0$. In fact, if $\beta < 1$, then $-\beta + 1 > 0$, and

$$(-\beta + 1)Y + (-\alpha X + \beta Y) = Y - \alpha X \in D,$$

and, if $\beta \geq 1$, then $\frac{1}{\beta} > 0$, and:

$$\frac{1}{\beta}(-\alpha X + \beta Y) = Y - \frac{\alpha}{\beta} X \in D.$$

On both cases, there exists $\epsilon_1 > 0$ such that $Y - \epsilon_1 X \in D$. Now, since D is convex, then the segment from Y to $Y - \epsilon_1 X$ is contained in D, and, therefore $Y - \epsilon X \in D$ for all $\epsilon \leq \epsilon_1$. Let $\epsilon = \min\{\epsilon_1, \frac{1}{2}\}$. Then $1 - \epsilon > 0$, and, using again that D is closed for positive linear sums, we get

$$(1 + \epsilon)Y = Y + \epsilon Y \in D;$$

$$(1 - \epsilon)Y = Y - \epsilon Y \in D;$$

$$Y + \epsilon X \in D.$$

Therefore, $Y \pm \epsilon X, Y \pm \epsilon Y \in D$. Since X, Y is a base of $\mathbb{R}^2$, and D is convex, then $Y \in Int(D)$, which is a contradiction with $Y \in \delta(D)$. □

Proposition 2.3.9. *If C is a convex subset of $\mathbb{R}^n$ with $n > 0$, and $0 \notin Int(C)$, there there exists a basis $\{X, Y_1, Y_2, ...Y_{n-1}\}$ of $\mathbb{R}^n$ such that, for any $x = \alpha X + \sum_{i=1}^{n-1} \beta_i Y_i$, if $x \in C$ then $\alpha \geq 0$.*

Proof. The proposition is trivial for $\mathbb{R}^1$. We will show using induction that it holds for $\mathbb{R}^n$.

Let $C \subset \mathbb{R}^n$ a convex set such that $0 \notin Int(C)$. Then, for some vector e_k from the canonical basis, $\alpha e_k \notin C$ for all $\alpha > 0$ or $\alpha e_k \notin C$ for all $\alpha < 0$. Let V be any $n - 1$ dimension subspace containing e_k. Note that 0 is not in the interior of $V \cap C$ on the subspace topology of V, since $\alpha e_k \in (V - (V \cap C))$, for all $\alpha > 0$ or for all $\alpha < 0$. Then, by the induction hypothesis, there is a basis $Z, Y_1, ..., Y_{n-2}$ of V such that

$$x = \alpha Z + \sum_{i=1}^{n-2} \beta_i Y_i,$$

with $\alpha \geq 0$ for all $x \in C \cap V$. Let $W = \langle Y_1, Y_2, ..., Y_{n-2}\rangle$. Note that W has dimension $n-2$, therefore $\mathbb{R}^n/W$ has dimension 2. Furthermore, if $\pi : \mathbb{R}^n \to \mathbb{R}^2/W$ is the canonical projection, then $\alpha Z + W \notin \pi(C)$ for all $\alpha < 0$, since the elements of $\pi^{-1}(\alpha Z + W)$ are contained on V and are in the form

$$\alpha Z + \sum_{i=1}^{n-2} \beta_i Y_i.$$

where $\alpha < 0$. Therefore, $0 + W \notin Int(\pi(C))$. W has dimension 2 and $\pi(C)$ is convex since it is the image of a convex set by a linear transformation, therefore, the lema 2.3.8 holds for $\pi(C)$, and there is a basis $X + W, Y + W$ of $\mathbb{R}^n/W$ such that $\alpha \geq 0$ whenever $\alpha X + W + \beta Y + W \in \pi(C)$. Then, writing $Y_{n-1} = Y$, we have, for any $x = \alpha X + \sum\limits_{i=1}^{n-1} \beta_i Y_i \in C$, $\pi(x) = \alpha X + \beta_{n-1} Y_{n-1} \in \pi(C)$, therefore $\alpha \geq 0$. □

The above property allows us to show a last necessary condition for controllabillity

Proposition 2.3.10. *If $(A, B)_U$ is controllable, then*

$$0 + V \in Int(\pi(B(cv(U)))),$$

where $V = Img(A)$ and $\pi : \mathbb{R}^n \to \mathbb{R}^n/V$ is the canonical projection

Proof. If $V = \mathbb{R}^n$ then $\mathbb{R}^n/V$ is null, and the proposition is trivial, since $0 + V \in Int(\pi(B(cv(U))))$ will be true for any nonempty set U. We will assume otherwise, and show that the proposition hold by contrapositive. If $0 + V \notin Int(\pi(B(cv(C))))$ then, by proposition 2.3.9, there exists a basis $\beta = \{X, Y_1, ..., Y_d\}$ of $\mathbb{R}^n/V$ such that $\alpha \geq 0$ whenever $\alpha X + \sum\limits_{i=1}^{d} \beta_i Y_i \in \pi(B(cv(U)))$.

Let $x \in \mathcal{O}^+(0)$ arbitrary. Then $x = \phi(0, u, T)$ for some control u and positive time T. Then:

$$\pi(x) = \pi\left(e^{TA}0 + \int_0^T e^{(T-s)A}B(u(s))\,ds\right) = \int_0^T e^{(T-s)A}B(u(s)) + Img(A)\,ds.$$

We recall a property which was used previously:

$$e^{tA}x + Img(A) = x + A\left(\sum_{i=1}^{+\infty} \frac{t^i A^{i-1}}{i!} x\right) + Img(A) = x + Img(A),$$

for all $t \in \mathbb{R}$, therefore,

$$\pi(x) = \int_0^T B(u(s)) + Img(A)\, ds = \int_0^T \pi(B(u(s)))\, ds.$$

Note that $\pi(B(u(s))) \in \pi(B(U)) \subset \pi(B(cv(U)))$, such that the integral above is an integral of elements in $\pi(B(cv(U)))$. But every element in this set have a non negative first coordinate on the basis β, therefore, the numerical result of the integral also has a non negative first coordinate on that basis. Then, x is an element of the set

$$S := \left\{ \alpha X + \sum_{i=1}^{d} \beta Y ; \alpha \geq 0 \right\}.$$

Since $x \in \mathcal{O}^+(0)$ is arbitrary, we have $\pi(\mathcal{O}^+(0)) \subset S$. But S is a proper subset of $\mathbb{R}^n/V$ and π is surjective, therefore, $oo^+(0) \neq \mathbb{R}^n$ and the system is not controllable. □

Corollary 2.3.11. *If $(A, a, B)_U$ is controllable, then*

$$0 + V \in Int(\pi(B(cv(U)) + a)),$$

where $V = Img(A)$ and $\pi : \mathbb{R}^n \to \mathbb{R}^n/V$ is the canonical projection.

Proof. If $(A, a, B)_U$ is controllable then (A, B) is controllable and then

$$a = Aa_A + Ba_B$$

For some $a_A \in \mathbb{R}^n$ and $a_B \in \mathbb{R}^m$. From the previous proposition and the equivalence of the systems $(A, a, B)_U$ and $(A, B)_{U+a_B}$ with respect to controllability, we have that the controllability of $(A, a, B)_U$ implies

$$0 + V \in Int(\pi(B(cv(U + a_B)))).$$

The convex closure of a translated set is the translation of the convex closure, then we have

$$B(cv(U + a_B)) = B(cv(U) + a_B) = B(cv(U)) + Ba_B.$$

Furthermore, $Aa_A \in Img(A)$ by definition, therefore $\pi(Aa_A) = 0 + Img(A)$. Then,

$$\pi(B(cv(U)) + Ba_B) = \pi(B(cv(U)) + Ba_B + Aa_A) = \pi(B(cv(U)) + a).$$

Therefore,

$$0 + V \in Int(\pi(B(cv(U)) + a)).$$

□

So far we have that, if U is bounded, the following conditions are necessary for controllability of $(A, B)_U$:

- (A, B) must be C.R.O. Note that this condition also is also implying the controllability of (A, B). Equivalently, the Kalman matrix must have full rank and all of the eigenvalues of A must have real part equal to zero.
- $0+V \in Int(\pi(B(cv(U))))$, where $V = Img(A)$ and $\pi : \mathbb{R}^n \to \mathbb{R}^n/V$ is the canonical projection.

We can also make a similar list for the controlability of $(A, a, B)_U$. It is very similar to the previous one, due to the relations between those systems:

- (A, B) must be C.R.O. Note that this condition also is also implying the controllability of (A, B). Equivalently, the Kalman matrix must have full rank and all of the eigenvalues of A must have real part equal to zero.
- $0 + V \in Int(\pi(B(cv(U)) + a))$, where $V = Img(A)$ and $\pi : \mathbb{R}^n \to \mathbb{R}^n/V$ is the canonical projection.

In the remaining of this section we show that those conditions are also sufficient for the controllability of the respective systems. The first step is to show that the controllability of $(A, B)_U$ is equivalent to the controllability of $(A, B)_{cv(U)}$. An idea that will be used for that is the idea of approximated controllability. Remember that a control system is approximatelly controllable if the positive and negative orbits from any point x are both dense in the manifold where the system is being considered.

Remember as well that the controllability of a system implies the approximated controllability for the same system, and that the inverse implication is true if the system is equivalent to a Lie semigroup generated by exponentials, which includes the systems studied here. In the following results we show a less general proof of this, specific for the systems considered in this chapter.

Lemma 2.3.12. *Denoting by $\mathcal{O}^-_{(A,B)_U}, \mathcal{O}^+_{(-A,-B)_U}$ the orbits by $(A,B)_U, (-A,-B)_U$, respectively, then*

$$\mathcal{O}^-_{(A,B)_U}(x) = \mathcal{O}^+_{(-A,-B)_U}(x)$$

for all $x \in \mathbb{R}^n$.

Proof. Denote by ϕ the solution of the system $(A,B)_U$ and by ϕ' the solution of the system $(-A,-B)_U$. We will show that $\phi(x,u,T) = \phi'(x,v,-T)$, where v is defined by $v(t) = u(-t)$. We do that by showing that $\phi'(x,v,-t)$ satisfies the initial value problem from the first system:

$$\frac{d}{dt}\phi'(x,v,-t) = \left(-A\phi'(x,v,-t) - B(v(-t))\right)(-1) = A\phi'(x,v,-t) + B(u(t))$$

$$\phi'(x,v,0) = x.$$

Then $\phi(x,u,T) = \phi'(x,v,-T)$, since this problem has unique solution. Therefore, $y \in \mathcal{O}^-_{(A,B)_C}(x)$ if and only if there is a control u and a time $T < 0$ such that $y = \phi(x,u,T) = \phi'(x,-u,-T)$, which happens if, and only if, $y \in \mathcal{O}^+_{(-A,-B)_C}(x)$. Therefore the two sets are equal. □

Lemma 2.3.13. *Let $C \subset \mathbb{R}^n$ be a set, $A : \mathbb{R}^n \to \mathbb{R}^n$ a linear transformation and I any open interval in $\mathbb{R}$. If V is the smallest A invariant subspace containing C, then the set $D := \{e^{tA}c; t \in I, c \in C\}$ spans V.*

Proof. First we show that the lemma holds when $0 \in I$.
Let W be the space spanned by D. We have that $W \subset V$, since any $e^{tA}c$ is in V. Assume, by contradiction, that $W \neq V$. Since $0 \in I$,then $e^0 c = c \in W$ for all $c \in C$, then $C \subset W$. Since $W \subsetneq V$ and V is the smallest A invariant subset containing C, then W is not A invariant. Then, there is $x \in W$ such that $Ax \notin W$. Write

$$x = \sum_{i=1}^{n} e^{t_i A} c_i,$$

and note that

$$e^{tA}x = \sum_{i=1}^{n} e^{(t+t_i)A} c_i.$$

Since I is open, then $e^{tA}x \in W$ for sufficiently small t. But

$$\left.\frac{d}{dt}\right|_{t=0} e^{tA}x = Ax \notin W,$$

then $e^{tA} \notin W$ for sufficiently small t, which is a contradiction on the hypothesis that the $e^{t+t_iA}c_i$ are all in W.
Now, for the general case, let $\alpha \in I$. Note that

$$\{e^{tA}x; t \in I, x \in C\} = \{e^{\alpha A}e^{tA}; t \in (I-\alpha), x \in C\} = e^{\alpha A}\{e^{tA}; t \in (I-\alpha), x \in C\}.$$

Since $\alpha \in I$, then $0 \in (I-\alpha)$, then, by what was shown previously, the set $\{e^{tA}; t \in (I-\alpha), x \in C\}$ spans V. Since $e^{\alpha A}$ is an isomorphism, and V is a finite dimension subspace invariant by $e^{\alpha A}$, then the restriction of $e^{\alpha A}$ to V is still an isomorphism, therefore $e^{\alpha A}\{e^{tA}; t \in (I-\alpha)\}$ spans V. □

Lemma 2.3.14. *Given C, A, I, V as in the previous lemma, it is possible to choose $t_1, t_2, ..., t_n \in I, x_1, x_2, ..., x_n \in C$ such that $e^{t_1}x_1, e^{t_2}x_2, ..., e^{t_n}x_n$ spans V and $t_1, t_2, ..., t_n$ are two by two distinct.*

Proof. The previous lemma assures the existence of $e^{t_1A}x_1, e^{t_2A}x_2..., e^{t_nA}x_n$ spanning V. A spanning set for a finite dimension space will still span the space when perturbed (assuming said perturbation is contained withing the space itself). Then, for sufficiently small perturbations of $t_1, t_2, ..., t_n$, the vectors $e^{t_1A}x_1, e^{t_2A}x_2..., e^{t_nA}x_n$ will still span $\mathbb{R}^n$. I is an open set, by hypothesis, then those values will remain in I under small perturbations. Therefore, we can change $t_1, t_2, ..., t_n$ slightly to make them distinct without losing the properties of the lemma. □

Lemma 2.3.15. *If (A, B) is controllable and U is not contained in a proper affine subspace, then the positive and negative orbits from 0 of $(A, B)_U$ have nonempty interior.*

Proof. Since (A, B) is controllable, then the smallest A invariant subspace containing the image of B is $\mathbb{R}^n$. Let $D := \{x - y; x, y \in C\}$, since C is not contained in any proper affine subspace, then D spans $\mathbb{R}$, and, therefore, $B(D)$ spans $Img(B)$. Consequently, the smallest A invariant subspace containing $B(D)$ is still $\mathbb{R}^n$. By lemma 8 there are distinct $t_1, t_2, ..., t_n > 0$ and $z_1, z_2, ..., z_n \in B(D)$ such that $e^{t_1}z_1, e^{t_2}z_2, ..., e^{t_n}z_n$ span $\mathbb{R}^n$.

Without losing generality, assume that $t_1 < t_2 < ... < t_n$. Then, writing $s_0 = 0$, there are $s_1, s_2, ..., s_n$ such that $0 = s_0 < t_1 < s_2 < t_2 < s_2 < ... < t_n < s_n$. Define:

$$\alpha_i = s_i - s_{i-1},$$

$$\beta_i = s_i - t_i,$$

$$\lambda_i = t_i - s_{i-1} = \alpha_i - \beta_i,$$

write $z_i = B(x_i) - B(y_i)$, with $x_i, y_i \in C$, and, for each $i \in \{1, 2, ..., n\}$ and $t \in \mathbb{R}$, define the control

$$u_{it} : \mathbb{R} \to C$$

$$s \to \begin{cases} y_i, \text{ if } t < s \\ x_i \text{ if } s \leq t \end{cases}$$

and the functions

$$f_i : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

$$(t, v) \to f_i^t(v) := \phi(v, u_{it}, \alpha_i).$$

Note that f_i is of class C^1 on the set $S_i := \{(t, x) \in \mathbb{R} \times \mathbb{R}^n; 0 < t < \alpha_i\}$. This is verified by the continuity of it's partial derivatives:

$$\frac{df_i}{dt} = \frac{d}{dt}\left(e^{\alpha_i A}v + \int_0^{\alpha_i} e^{(\alpha_i - s)A}B(u_{it}(s))\, ds\right) =$$

$$= \frac{d}{dt}\left(\int_0^t e^{(\alpha_i - s)A}B(x_i)\, ds - \int_{\alpha_i}^t e^{(\alpha_i - s)A}B(y_i)\, ds\right) =$$

$$= e^{(\alpha_i - t)A}B(x_i - y_i) = e^{(\alpha_i - t)A}z_i,$$

$$\frac{df_i}{dv} = \frac{d}{dv}\left(e^{\alpha_i A}v + \int_0^{\alpha_i} e^{(\alpha_i - s)A}B(u_{it}(s))\, ds\right) = e^{\alpha_i A}v.$$

since both functions above are continuous in S_i, then f_i is C^1 in this set. Note that when $t = \beta_i$:

$$\left.\frac{df_i}{dt}\right|_{t=\beta_i} = e^{(\alpha_i - \beta_i)A}z_i = e^{\lambda_i A}z_i,$$

furthermore, by the definition of f_i, we have that if $t > 0$ then $f_i^t(v)$ is in the positive orbit from v. Define the function

$$F : \mathbb{R}^n \to \mathbb{R}^n$$

$$(t_1, t_2, ..., t_n) \to f_1^{t_1} f_2^{t_2} ... f_n^{t_n}(0).$$

F is C^1 in $(0, \alpha_1) \times (0, \alpha_2) \times ... \times (0, \alpha_n)$, since it's a composition of C_1 functions. Furthermore, if e_i is the i-th vector of the canonical basis in $\mathbb{R}^n$, then:

$$\left.\frac{dF}{de_i}\right|_{(\beta_1,\beta_2,...,\beta_n)} = Df_1 Df_2 ... Df_{i-1} \left(\left.\frac{df_i}{dt}\right|_{\beta_i} \right) = e^{\alpha_1 A} e^{\alpha_2 A} ... e^{\alpha_{i-1} A} e^{\lambda_i A} z_i =$$

$$e^{\left(\lambda_i + \sum_{j=1}^{i-1} \alpha_j\right) A} z_i = e^{(\lambda_i + s_i - 1)A} z_i = e^{t_i A} z_i.$$

Where Df_j denote the respective spacial derivatives. Since the vectors $e^{t_i A} z_i$ span $\mathbb{R}^n$, then the derivative of F in $(\beta_1, \beta_2, ..., \beta_n)$ is bijective, therefore, by the inverse function theorem, F is locally invertible in this point. In particular, $F((0, \alpha_1) \times (0, \alpha_2) \times ... \times (0, \alpha_n))$ has nonempty interior. As was previously mentioned, this set is contained in $\mathcal{O}^+(0)$, therefore $\mathcal{O}^+(0)$ has nonempty interior.

To show the same for $\mathcal{O}^-(0)$, note that the Kalman criteria is equivalent for (A, B) and $(-A, -B)$. Then, since (A, B) is controllable, $(-A, -B)$ must also be. Denoting by $\mathcal{O}^+_{(-A,-B)}(0)$ the positive orbit in the system $(-A, -B)$, the previous argument assures that $Int(\mathcal{O}^+_{(-A,-B)}) \neq \emptyset$. But, by lemma 2.3.12, $\mathcal{O}^-(0) = \mathcal{O}^+_{(-A,-B)}(0)$, therefore $Int(\mathcal{O}^-(0)) \neq \emptyset$. □

Corollary 2.3.16. *If U is not contained in any proper affine subspace and $(A, B)_U$ is approximately controllable then $\mathcal{O}^+(0)$ and $\mathcal{O}^-(0)$ have nonempty interior.*

Proof. If $(A, B)_U$ is approximately controllable then so is (A, B), since the system (A, B) contains all of the orbits from $(A, B)_U$. However, the positive and negative orbits from the origin for the system (A, B) coincide with the image of the Kalman matrix ([2, 3]), which is either maximal or not dense. Since we're assuming that this system is approximately controllable, then the orbit from the origin must be maximal, which implies that the Kalman matrix has full rank and the system is controllable. Then, the previous lemma ensures that $\mathcal{O}^+(0), \mathcal{O}^-(0)$ in the system $(A, B)_U$ have nonempty interior. □

From the Lie theory and semigroup theory point of view, the hypothesis of controllability for (A, B) can be shown to imply full rank for the semigroup associated to the system $(A, B)_U$ in some subgroup H which is transitive in $\mathbb{R}^n$.

Proposition 2.3.17. *If U is not contained in a proper affine subspace, then $(A, B)_U$ is controllable if, and only if, it is approximately controllable.*

Proof. If $(A, B)_U$ is controllable then it is also approximately controllable. We will show the other implication.

Assume $(A, B)_U$ to be approximately controllable. Then for any $x, y \in \mathbb{R}^n$, $\mathcal{O}^-(y)$ and $\mathcal{O}^+(x)$ are dense. Furthermore, by the corollary 2.3.16, $\mathcal{O}^+(0)$ and $\mathcal{O}^-(0)$ have nonempty interior. Then, there are $z \in \mathcal{O}^+(x) \cap \mathcal{O}^-(0)$ and $w \in \mathcal{O}^-(y) \cap \mathcal{O}^+(0)$. We have $y \in \mathcal{O}^+(w), w \in \mathcal{O}^+(0), 0 \in \mathcal{O}^+(z), z \in \mathcal{O}^+(x)$, therefore $y \in \mathcal{O}^+(x)$. Since x, y are arbitrary, the system is controllable. □

On the next results we fall back to the notation $\mathcal{O}^+_\Sigma(x), \mathcal{O}^-_\Sigma(x)$ for the the positive, negative orbits of x in the system Σ, respectively.

Proposition 2.3.18. *For a control system $(A, B)_U$ with U bounded and not contained in an affine proper subspace, the following are equivalent.*

- *$(A, B)_U$ is approximately controllable.*
- *$(A, B)_{cv(U)}$ is approximately controllable.*

Proof. We will show that $\overline{\mathcal{O}^+_{(A,B)_U}(x)} = \overline{\mathcal{O}^+_{(A,B)_{cv(U)}}(x)}$ for all $x \in \mathbb{R}^n$. Note that $\overline{\mathcal{O}^+_{(A,B)_U}(x)} \subset \overline{\mathcal{O}^+_{(A,B)_{cv(U)}}(x)}$, since $U \subset cv(U)$. to show the other inclusion, let $\epsilon > 0$, $x \in \mathbb{R}^n$, $y \in \overline{\mathcal{O}^+_{(A,B)_{cv(U)}}(x)}$ be arbitrary. Then there is a control u and a time $T > 0$ such that

$$\|\phi(x, u, T) - y\| < \frac{\epsilon}{3}. \tag{2.3-1}$$

Recall that

$$\phi(x, u, T) = e^{TA}x + \int_0^T e^{(T-s)A}B(u(s))\, ds.$$

Since $B(U)$ is bounded, there is M such that $\|v\| < M$ for all $v \in B(U)$. Since $e^{(T-s)A}$ is a continuous function, and $[0, T]$ is compact, there is $\delta > 0$ such that $\|e^{(T-t_1)A} - e^{(T-t_2)A}\| <$

$\frac{\epsilon}{3MT}$ whenever $|t_1 - t_2| < \delta$. Take a partition $\mathcal{P} = \{0 = \alpha_1, \alpha_1, ..., \alpha_{d+1} = T\}$ with intervals smaller than δ such that:

$$\left\| \int_0^T e^{(T-s)A} B(u(s))\, ds - \sum_{i=1}^{d} (\alpha_{i+1} - \alpha_i) e^{(T-\alpha_i)A} B v_i \right\| < \frac{\epsilon}{3}$$

$$\Rightarrow \left\| \phi(x, u, T) - \left(e^{TA} x + \sum_{i=1}^{d} (\alpha_{i+1} - \alpha_i) e^{(T-\alpha_i)A} B v_i \right) \right\| < \frac{\epsilon}{3}, \tag{2.3-2}$$

where $v_i = u(\alpha_i)$. Since $v_i \in cv(U)$, there exists $w_{i1}, w_{i2}, ..., w_{il_i} \in U$ and positive numbers $\beta_{i1}, \beta_{i2}, ..., \beta_{il_i}$ such that:

$$v_i = \sum_{j=1}^{l_i} \beta_{ij} w_{ij},$$

$$\sum_{j=1}^{l_i} \beta_{ij} = 1.$$

Let

$$\gamma_{ij} := \alpha_i + (\alpha_{i+1} - \alpha_i) \sum_{k=1}^{j} \beta_{ij},$$

for $i = 1, ..., d$, $j = 0, 1, ..., l_i$. Note that $\gamma_{il_i} = \gamma_{(i+1)0} = \alpha_{i+1}$. Define the function:

$$u_2 : (0, T] \to U$$

$$t \to w_{ij}, \text{ se } t \in \left(\gamma_{i(j-1)}, \gamma_{ij}\right].$$

Note that:

$$\left\| \phi(x, u_2, T) - \left(e^{TA} x + \sum_{i=1}^{d} (\alpha_{i+1} - \alpha_i) e^{(T-\alpha_i)A} B v_i \right) \right\| =$$

$$\left\| \left(\sum_{i=1}^{d} \sum_{j=1}^{l_i} \int_{\gamma_{i(j-1)}}^{\gamma_{ij}} e^{(T-s)A} B w_{ij} \right) - \left(\sum_{i=1}^{d} \sum_{j=1}^{l_i} \beta_{ij} (\alpha_{i+1} - \alpha_i) e^{(T-\alpha_i)A} B w_{ij} \right) \right\| =$$

$$\left\| \sum_{i=1}^{d} \sum_{j=1}^{l_i} \int_{\gamma_{i(j-1)}}^{\gamma_{ij}} (e^{(T-s)A} - e^{(T-\alpha_i)A}) B w_{ij} \right\| \leq$$

$$\sum_{i=1}^{d} \sum_{j=1}^{l_i} \int_{\gamma_{i(j-1)}}^{\gamma_{ij}} \left\| (e^{(T-s)A} - e^{(T-\alpha_i)A}) B w_{ij} \right\| < \sum_{i=1}^{d} \sum_{j=1}^{l_i} \int_{\gamma_{i(j-1)}}^{\gamma_{ij}} \frac{\epsilon}{3MT} M \leq$$

$$= \sum_{i=1}^{d} \sum_{j=1}^{l_1} (\alpha_{i+1} - \alpha_i) \beta_{ij} \frac{\epsilon}{3T} = \sum_{i=1}^{d} (\alpha_{i+1} - \alpha_i) \frac{\epsilon}{3T} = \frac{\epsilon}{3}$$

$$\Rightarrow \left\| \phi(x, u_2, T) - \left(e^{TA}x + \sum_{i=1}^{d} (\alpha_{i+1} - \alpha_i) e^{(T-\alpha_i)A} B v_i \right) \right\| < \frac{\epsilon}{3}. \tag{2.3-3}$$

From 2.3-1, 2.3-2, 2.3-3 and the triangular inequality we have

$$\|\phi(x, u_2, T) - y\| < \epsilon,$$

showing that $y \in \overline{\mathcal{O}^+_{(A,B)_U}(x)}$.

The argument above, together with lemma 2.3.12, also assures the equality of the closures for the negatie orbits:

$$\overline{\mathcal{O}^-_{(A,B)_U}} = \overline{\mathcal{O}^+_{(-A,-B)_U}} = \overline{\mathcal{O}^+_{(-A,-B)_{cv(U)}}} = \overline{\mathcal{O}^-_{(A,B)_{cv(U)}}}.$$

Then, the approximated controllability for the systems $(A, B)_U$ and $(A, B)_{cv(U)}$ is equivalent. □

Theorem 2.3.19. *If U is bounded and not contained in a proper affine subspace, then the following are equivalent:*

- *$(A, B)_U$ is controllable*
- *$(A, B)_{cv(U)}$ is controllable*

Proof. By propositions 2.3.17, 2.3.18, $(A, B)_U$ is controllable if, and only if, $(A, B)_U$ is approximately controllable, if and only if $(A, B)_{cv(C)}$ is approximately controllable if and only if $(A, B)_{cv(C)}$ is approximately controllable. □

Corollary 2.3.20. *If U is bounded and not contained in a proper affine subspace, then the following are equivalent:*

$(A, a, B)_U$ is controllable

$(A, a, B)_{cv(U)}$ is controllable

Proof. It's a direct consequence of the previous lemma and the equivalence between the two systems. □

For the final theorem we will use a few more properties of convex sets.

Lemma 2.3.21. *If $C \subset \mathbb{R}^n$ is convex then $Int(C) \neq \emptyset$ if, and only if, C is not contained in any proper affine subspace.*

Proof. Any propper affine subspace has empty interior, therefore if C is contained in one then $Int(C) = \emptyset$. For the other implication, assume that C is not contained in any proper affine subspace. Then there is a basis $y_1 - x, y_2 - x, ..., y_n - x$ such that $y_1, y_2, ..., y_n, x \in C$. Since C is convex, then

$$\left\{\left(1 - \sum_{i=1}^{n} \alpha_n\right) x + \alpha_1 y_1 + \alpha_2 y_2 + ... + \alpha_n y_n; 0 < \alpha_i < \frac{1}{n}\right\} \subset C \Rightarrow$$

$$\left\{x + \alpha_1(y_1 - x) + \alpha_2(y_2 - x) + ... + \alpha_n(y_n - x); 0 < \alpha_i < \frac{1}{n}\right\} \subset C.$$

This set is open since $y_1 - x, y_2 - x, ..., y_n - x$ is a basis. Therefore, $Int(C) \neq \emptyset$. □

Lemma 2.3.22. *If $C \subset \mathbb{R}^n$ is a convex set with nonempty interior, $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and $x \in Int(T(C))$, then there is $y \in Int(C)$ such that $T(y) = x$.*

Proof. Let $v \in Int(C)$, and let $e_1, ..., e_n$ the canonical basis of $\mathbb{R}^n$. Then there is $\alpha > 0$ such that $v + \alpha e_i \in C$ and $v - \alpha e_i \in C$, for all $i \in \{1, ..., k\}$. Given $y \in Int(T(C))$, there is sufficiently small $\beta > 0$ such that $y + \beta(y - T(v)) \in T(C)$, and, then, $T(w) = y + \beta(y - T(v))$, for some $w \in C$. Let $x = \frac{\beta}{1+\beta}v + \frac{1}{1+\beta}w \in C$. Note that:

$$T(x) = \frac{\beta}{1+\beta}T(v) + \frac{1}{1+\beta}T(w) = \frac{\beta}{1+\beta}T(v) + \frac{1}{1+\beta}(y + \beta(y - T(v))) =$$

$$= \frac{\beta}{1+\beta}T(v) + \frac{1+\beta}{1+\beta}y - \frac{\beta}{1+\beta}T(v) = y.$$

Furthermore, writing $\gamma = \frac{\beta}{1+\beta}\alpha > 0$, then, for all $i \in \{1, ..., n\}$, $j \in \{-1, 1\}$:

$$\frac{\beta}{1+\beta}(v + j\alpha e_i) + \frac{1}{1+\beta}w = \frac{\beta}{1+\beta}v + \frac{1}{1+\beta}w + j\frac{\beta}{1+\beta}\alpha e_i = x + j\gamma e_i \in C,$$

that is, $x + \gamma e_i \in C$ and $x - \gamma e_i \in C$ for all $i \in \{1, ..., n\}$, therefore, by the convexity of C, $x \in Int(C)$. □

Finally, we show that the necessary conditions found previously are also sufficient.

Theorem 2.3.23. *A system $(A, B)_U$ with U bounded and not contained in any proper affine subspace is controllable if, and only if, both of the following conditions are true:*

- *(A, B) is C.R.O.. Equivalently, the Kalman matrix of (A, B) has full rank, and all of the eigenvalues of A have their real component equal to zero.*

- $0 + V \in Int(\pi(B(cv(U))))$, *where* $V \subset \mathbb{R}^n$ *is the image of* A *and* $\pi : \mathbb{R}^n \to \mathbb{R}^n/V$ *is the canonical projection.*

Proof. The necessity of those two items for controllability was already shown in the first part of the section. We will now assume that $(A, B)_U$ satisfies both conditions and show that it is controllable.

Note that $cv(U)$ is a convex set which is not contained in an affine subspace, therefore, by lemma 2.3.21, $cv(U)$ has nonempty interior. Then, since $0+V \in Int(\pi(B(cv(U))))$, and $\pi \circ B$ is a linear application, by lemma 2.3.22 there is $v \in Int(cv(U))$ such that $\pi(B(v)) = 0 + V$. Equivalently, there is $v \in Int(cv(U))$ such that $\pi(B(v)) \in Img(A)$, therefore, there exists $w \in \mathbb{R}^n$ such that $Aw = Bv \iff Aw - Bv = 0$. Now, the system $(A, B)_{cv(U)}$ is equivalent to $(A, 0, B)_{cv(U)}$, which is equivalent to the system $(A, B)_{cv(U)-v}$, by proposition 2.3.1. Since $\in Int(cv(U))$, then $0 \in Int(cv(U) - v)$. Then, since (A, B) is C.R.O., we have that $(A, B)_{cv(U)-v}$ is controllable, and, therefore, $(A, B)_{cv(U)}$ is controllable. By theorem 2.3.19, $(A, B)_U$ is also controllable.

□

Corollary 2.3.24. *An affine system* $(A, a, B)_U$ *with* U *bounded and not contained in any proper affine subspace of* $\mathbb{R}^m$ *is controllable if, and only if, both of the following conditions are true:*

- (A, B) *is C.R.O.. Equivalently, the Kalman matrix of the system* (A, B) *has full rank and all the eigenvalues of* A *have positive component equal to zero.*

- $0 + V \in Int(\pi(B(cv(U)) + a))$, *where* $V \subset \mathbb{R}^n$ *is the image of* A *and* $\pi : \mathbb{R}^n \to \mathbb{R}^n/V$.

Proof. Once again, the necessity of those two conditions was shown in the first part of the section. For the other implication, assume that both conditions are true. Then, from the controllability of (A, B), a can be written as

$$a = Aa_A + Ba_B,$$

such that the controllability of $(A, a, B)_U$ is equivalent to the controllability of $(A, B)_{U+a_B}$. From the previous theorem, the system $(A, B)_{U+a_B}$ is controllable if, and only if, (A, B) is $C.R.O.$ and

$$0 + V \in Int(\pi(B(cv(U))) + a_B).$$

The first condition is already true by hypothesis, and, as mentioned previously, the equality $a = Aa_A + Ba_B$ implies

$$Int(\pi(B(cv(U))) + a_B) = Int(\pi(B(cv(U)) + a)).$$

Therefore the second condition is also true, so $(A, B)_{U+a_B}$, and, therefore $(A, B)_U$, are controllable. □

CHAPTER 3

TANGENT CONTROL SYSTEM

3.1 Definition of a tangent control system

In this chapter we define the tangent control system, which is an useful tool for studying local controllability. The idea is to consider curves originating in the isotropy subgroup of a point $x \in M$ and contained in the semigroup of the system for positive time, and differentiate the action of these curves in x in time 0. Interesting results can be obtained from this construction.

Let G be a Lie group and $R \subset G$ a semi group. For this construction R is not assumed to have nonempty interior. Define

$$\mathcal{C}_R = \{\phi : \mathbb{R} \to G : \phi \text{ is } C^1 \text{ and } \phi(0) \in R\}$$

the set of all diferentiable curves in G originating in R. We study local properties of these curves, so $\mathcal{C}_R$ could alternatively be defined as a set of germs without losing much.

Note that $\mathcal{C}_R$ is a semigroup with the product $\phi\varphi : \mathbb{R} \to G$ defined by

$$(\phi\varphi)(t) = \phi(t)\varphi(t).$$

If R is a subgroup then $\mathcal{C}_R$ is group with inverse defined by $\phi^*(t) = (\phi(t))^{-1}$. We use the notation ϕ^* in this case as to not be mistaken with the notation for inverse function, which is usually denoted by ϕ^{-1}. Moreover, if R is not a subgroup, then $\mathcal{C}_R \subset \mathcal{C}_G$, such that $\mathcal{C}_R$ is contained in a group. Now fix a semigroup $S \subset G$. Given a subsemigroup R contained in $\overline{S}$, the topological closure of S, we denote

$$S_R = \{\phi \in \mathcal{C}_R : \phi(t) \in S \text{ for all } t > 0\},$$

the set of curves in $\mathcal{C}_R$ that stay in S in positive time, and by S^*_R the set of curves in $\mathcal{C}_{R^{-1}}$ that stay in S^{-1} in positive time. It can be shown that $S^*_R = (S_R)^{-1}$ and if S_R and S^*_R are nonempty then both are semigroups.

Now suppose that G acts on the manifold M, let $v \in M$ and denote by $H_v = \{g \in G; gv = v\}$ the isotropy subgroup of v. For simplicity, denote $\mathcal{C}_{H_v}$ as $\mathcal{C}_v$. Suppose also that a subsemigroup $R \subset \overline{S}$ is contained in H_v, then $S_R \subset \mathcal{C}_v$. Observe that in this case the curves $f(t)v$ are in M and $f(0)v = v$ for all $f \in \mathcal{C}_v$. Moreover, if $f \in S_R$, then the curve $f(t)v$ is also in $\mathcal{O}^+(v)$ for all positive t, since $f(t) \in S$ for all $t \geq 0$. Now for every $g \in G$ denote by ϕ_g the diffeomophism

$$\phi_g : M \to M$$

$$\phi_g(m) = g(m).$$

When $g \in H_v$, $\phi_g(v) = v$ and therefore $D_v\phi_g : T_vM \to T_vM$ is an automorphism. In particular, if $f \in \mathcal{C}_v$ then $f(0) \in H_v$ and hence $D_v\phi_{f(0)} \in \mathrm{Gl}(T_vM)$. Also define the map

$$F : \mathcal{C}_v \to T_vM, F(f) = \left.\frac{d}{dt}\right|_{t=0} f(t)v.$$

These maps were defined in order to represent C_v in the following affine group:

$$\mathrm{Aff}(T_vM) = \mathrm{Gl}(T_vM) \rtimes T_vM.$$

Recall that the affine group operation is defined by $(g, v) \cdot (h, w) = (gh, v + gw)$ for all $(g, v), (h, w) \in \mathrm{Gl}(T_vM) \rtimes T_vM$. We call affine action the natural action of $\mathrm{Gl}(T_vM) \rtimes T_vM$ on T_vM given by $(g, v) \cdot w = gw + v$ with $(g, v) \in \mathrm{Gl}(T_vM) \rtimes T_vM$ and $w \in T_vM$.

Define

$$\rho : \mathcal{C}_v \to \text{Aff}(T_vM)$$

by $\rho(f) = (D_v\phi_{f(0)}, F(f))$. It is not difficult to see that this is a group homomorphism. The image $\rho(S_R)$ is, therefore, a semigroup of $Aff(T_vM)$, and defines a control system in T_vM by the affine action. In the next results we show that the controllability of $\rho(S_R)$ is closely related with the local controllability of S in v. We call $\rho(S_R)$ the tangent semigroup and the system associated to it the tangent system.

Note that

$$\left.\frac{d}{dt}f(t)v\right|_{t=0} = F(f) = \rho(f)(0).$$

We first show that the controllability of $\rho(S_R)$ implies local controllability for S in v. To show this we need the lemma 3.1.1, presented next. The proof of this lemma is rather large and diverges a bit from the other results in this section, so we decided to included it in section 3.3.

Lemma 3.1.1. *Let M a finite dimensional differentiable manifold, $F : \mathbb{R}^n \to M$ a differentiable map and its derivative*

$$D_0F : \mathbb{R}^n \to T_{F(0)}M.$$

If $C \subset \mathbb{R}^n$ is a generating cone (that is, $Int(C) \neq 0$) with $D_0F(C) = T_{F(0)}M$, then $F(0) \in Int(F(Int(C)))$.

Proof. See section 3.3 □

Theorem 3.1.2. *If $(\rho(S_R))(0) = T_vM$, then $v \in Int(S(v))$. If $(\rho(S_R^*))(0) = T_vM$, then $v \in Int(S^{-1}(v))$.*

Proof. Suppose that $(\rho(S_R))(0) = T_vM$ and that $b_1, b_2, \ldots, b_k \in T_vM$ generate positively T_vM. Now we define the curves $f_1, f_2, \ldots, f_k \in S_R$ by recurrence. As $(\rho(S_R))(0) = T_vM$, there exists f_1 such that $\rho(f_1)(0) = b_1$. For $i = 2, \ldots, k$, suppose that $f_1, f_2, \ldots, f_{i-1}$ are defined. Let

$$T_i := ((D_v\phi_{f_1(0)})(D_v\phi_{f_2(0)}) \cdots (D_v\phi_{f_{i-1}(0)}))^{-1} \in \text{Gl}(T_vM)$$

and let f_i such that $\rho(f_i)(0) = T_i(b_i) \in T_vM$.

Define the map

$$A : \mathbb{R}^k \to G$$

$$(t_1, t_2, \ldots, t_k) \to f_1(t_1)f_2(t_2)\cdots f_k(t_k).$$

If Q denotes the positive orthant of $\mathbb{R}^k$, i.e.,

$$Q = \{(t_1, t_2, \ldots, t_k) \in \mathbb{R}^k; t_1, t_2, \ldots, t_k \geq 0\},$$

then $Int(Q) = \{(t_1, t_2, ..., t_k) \in \mathbb{R}^k; t_1, t_2, ..., t_k > 0\}$ and $A(Int(Q)) \subset S$, since $f_i(t_i) \in S$ for $t_i > 0$. Note that $A(0) \in R \subset H_v$. Now take

$$B : \mathbb{R}^k \to M$$

$$(t_1, t_2, \ldots, t_k) \to A(t_1, t_2, \ldots, t_k)v = f_1(t_1)f_2(t_2)\cdots f_k(t_k)v.$$

As $A(Int(Q)) \subset S$, then $B(Int(Q)) = A(Int(Q))v \subset \mathcal{O}^+(v)$, and knowing that $A(0) \in H_v$, we have $B(0) = v$. To compute the partial derivatives of B note that for every $i \in \{1, 2, \ldots, k\}$, we have

$$B(0, 0, \ldots, t_i, \ldots, 0) = f_1(0)f_2(0)\cdots f_i(t_i)\cdots f_k(0)v = \phi_{f_1(0)}\phi_{f_2(0)}\cdots\phi_{f_{i-1}(0)}f_i(t_i)v$$

hence,

$$\left.\frac{d}{de_i}\right|_0 B = \left.\frac{d}{dt}\right|_0 (\phi_{f_1(0)}\phi_{f_2(0)}\cdots\phi_{f_{i-1}(0)}f_i(t_i)v) =$$

$$(D_v\phi_{f_1(0)})(D_v\phi_{f_2(0)})\cdots(D_v\phi_{f_{i-1}(0)})\left(\left.\frac{d}{dt}\right|_{t=0} f_i(t_i)v\right) =$$

$$(D_v\phi_{f_1(0)})(D_v\phi_{f_2(0)})\cdots(D_v\phi_{f_{i-1}(0)})T_ib_i = b_i$$

Since D_0B is a linear map and Q is a cone generated by $e_1, e_2, \ldots, e_k$, then $D_0B(Q)$ is a cone generated by $D_0B(e_1), D_0B(e_2), \ldots, D_0B(e_k)$. As we see in the above equality, these vector coincide with $b_1, b_2, \ldots, b_k$, that generate T_vM positively by definition. Hence, $D_0B(Q) = T_vM$. By Lemma 3.1.1, we have $v = B(0) \in Int(B(Q)) \subset Int(S(v))$. Similarly we prove that $v \in Int(S^{-1}(v))$ if $(\rho(S_R^*))(0) = T_vM$. □

As a consequence we have the main result of this section

Corollary 3.1.3. *If the tangent system is controllable then S is locally controllable in v.*

Proof. If $\rho(\mathcal{S}_R)$ is controllable then

$$(\rho(\mathcal{S}_R))(0) = (\rho(\mathcal{S}_R))^{-1}(0) = (\rho(\mathcal{S}_R^*))(0) = T_vM.$$

Hence by above result $v \in Int(S(v)) \cap Int(S^{-1}(v))$, therefore S is locally controllable in v. □

From definition of the tangent semigroup we have the following property

Proposition 3.1.4. *If $(T, v) \in \rho(\mathcal{S}_R)$ then $(T, \alpha v) \in \rho(\mathcal{S}_R)$ for all $\alpha > 0$.*

To see this, take a curve f such that $\rho(f) = (T, v)$ and define $g(t) = f(\alpha t)$, hence $\rho(g) = (T, \alpha v)$. This property implies that the positive and negative orbits from the origin are cones (not necessarily convex). Knowing that a cone is the entire space if and only if the origin is in its interior we have that these orbits are maximal if and only if the origin is in their interiors. Then we have the following result.

Proposition 3.1.5. *The tangent system is controllable if and only if it is locally controllable in the origin.*

Therefore the previous relation between the tangent system and local controllability can also be characterized as follows.

Corollary 3.1.6. *If the tangent system is locally controllable in the origin then S is locally controllable in v.*

The other implication is true under some additional hypotheses. One natural thing to ask is for R to coincide with $\overline{S} \cap H_v$. Some hypothesis must also be required from S to ensure enough curves in S_R. Here, we ask for $R \cap \overline{Int(S)}$ to be nonempty and for the group G to be second countable. In reality, we want G to satisfy the rank condition in v, which comes as a consequence of the orbit of v having nonempty interior if G is second countable. Thus, by assuming G to be second countable, local controllability in v then implies the rank condition in v.

Note that the rank condition implies the following property: for any $x \in T_vM$ there is $X \in T_{Id}G$ such that $D_{Id}\phi_v(X) = x$, and then there is a curve $f : \mathbb{R} \to G$ such that $f(0) = Id$ and $f'(0) = X$, and, consequently, $\frac{d}{dt}\big|_{t=0} f(t)v = x$

The next theorem will also make use of the following observation: given a curve $f : \mathbb{R} \to G$ such that $f(0) = g$ and $f'(0) = x \in T_g G$, and an open set V containing g, it is possible to construct a curve $f_2 : \mathbb{R} \to G$ such that $f_2(0) = f(0) = g$, $f_2'(0) = f'(0) = x$ and $f_2(\mathbb{R}) \subset V$. In fact, let $I = (-\epsilon, \epsilon)$ be a sufficiently small interval such that $f(I) \subset V$, we first construct a diffeomorfism $\psi : \mathbb{R} \to (-\epsilon, \epsilon)$ such that $\psi(0) = 0$ and $\psi'(0) = 1$. One example is

$$\psi(t) = \frac{\epsilon}{\pi} arctan\left(\frac{\pi}{\epsilon} t\right)$$

Then f_2 can be defined as $f \circ \psi$. Furthermore, if G acts on a manifold M, then $\frac{d}{dt}\big|_{t=0} f(t)m = \frac{d}{dt}\big|_{t=0} f_2(t)m$ for any $m \in M$. This means no derivatives are lost by confining the curves to V.

Theorem 3.1.7. *If G is second countable, $R = \overline{S} \cap H_v$ and $R \cap \overline{Int(S)}$ is nonempty then the following are equivalent:*

1. S is locally controllable in v
2. $\rho(S_R)$ is controllable in $T_v M$
3. $\rho(S_R)$ is locally controllable in the origin in $T_v M$

Proof. The equivalence between 2 and 3 and the implication $2 \Rightarrow 1$ were already shown in previous results. We must show $1 \Rightarrow 2$. Assume S locally controllable in v and let W be an open set containing v such that $W \subset S(x)$ for all $x \in W$. The hypothesis $R \cap \overline{Int(S)} \neq \emptyset$ implies there is $h \in Int(S)$ such that $h^{-1}(v) \in W$. Then there is $g \in S$ such that $g(v) = h^{-1}(v)$, and, therefore, $hg \in R \cap Int(S)$. Let $V \subset S$ an open set containing hg. Then $Vg^{-1}h^{-1}$ is an open set containing the Identity. Remember that local controllability in v implies full rank for the group G in v, as G is second countable by hypothesis. Then, for any $x \in T_v M$, there is a curve $f : \mathbb{R} \to G$ such that $f(0) = Id$, $\frac{d}{dt}\big|_{t=0} f(t)v = x$ and $f(\mathbb{R}) \subset Vg^{-1}h^{-1}$. Then $f(t)hg$ is contained in $V \subset S$ for all t and satisfies $f(0)hg = hg \in R$, therefore $(t \to f(t)hg) \in S_R$. Furthermore

$$\rho(t \to f(t)hg)0 = \frac{d}{dt}\bigg|_{t=0} f(t)hgv = \frac{d}{dt}\bigg|_{t=0} f(t)v = x.$$

Since x is arbitrary in $T_v M$, we have $\rho(S_R)(0) = T_v M$.

For the negative orbit, note that S^{-1} is also locally controllable in v. In fact, for any $n, m \in W$ there is an element $g \in S$ such that $gm = n$, and, therefore, $g^{-1}n = m$.

Furthermore, $R \cap \overline{Int(S^{-1})} = (R \cap \overline{Int(S)})^{-1} \neq \emptyset$. Then, the same arguments show that $\rho(S_R^*)(0) = T_v M$, and, therefore, $\rho(S_R)$ is controllable in $T_v M$. $\square$

3.2 An application in Bilinear Control System

In this section we show an application of the tangent system in bilinear control systems:

$$\dot{x} = Ax + uBx, x \in \mathbb{R}^d \setminus \{0\}, u : \mathbb{R} \to \mathbb{R}, \tag{3.2-1}$$

where A and B are $d \times d$-matrices and u is piecewise constant. In the notation of the first chapter, this is the continuous control system $(\mathbb{R}^n, \mathcal{F}, \mathbb{R}, \mathcal{U})$ where

$$\mathcal{F}(x, r) = Ax + rBx$$

Here we ask the controls to be piecewise constant as nothing is lost in terms of controllability and this makes it easier to define the semigroup of the system: the bilinear control system is equivalent to the semigroup in $\mathrm{Gl}(\mathbb{R}^n)$ generated by exponentials of the set

$$C = \{A + rB; r \in \mathbb{R}\} \subset \mathfrak{gl}(\mathbb{R}^n).$$

(see e.g Colonius and Kliemann [1] and Elliot [8]).

It is often usefull to instead consider the semigroup generated by the set

$$D = \{A, B, -B\} \subset \mathfrak{gl}(\mathbb{R}^n).$$

Note that C and D generate the same closed convex cone in $\mathfrak{gl}(\mathbb{R}^n)$, such that their two subgroups are equivalent controllability wise. The set D has the advantage of being a discrete set containing only 3 elements. As such, in many contexts, including this chapter, the semigroup used for the study of the bilinear control system is the semigroup generated by D. We will denote this semigroup by S. Note that

$$S = \{e^{r_1 B} e^{s_1 A} ... e^{r_k B} e^{s_k A}; k \in \mathbb{N}; r_1, r_2, ..., r_k \in \mathbb{R}; s_1, s_2, ..., s_k \in (0, +\infty)\}$$

A very useful tool in studying the controllablility in $\mathbb{R}^n - \{0\}$ of bilinear control systems and semigroups of matrices in general is the projective space, defined by

$$P(\mathbb{R}^n) = \{V \subset \mathbb{R}^n : V \text{ is a one dimentional subspace}\}.$$

An element $V \in P(\mathbb{R}^n)$ is often denoted by $[x]$ for a vector $x \in V$ different from 0.

A linear automorphism $g \in \mathrm{Gl}(\mathbb{R}^n)$ takes one dimensional subspaces into one dimensional subspaces, and acts in $P(\mathbb{R}^n)$ by defining

$$g : P(\mathbb{R}^n) \to P(\mathbb{R}^n)$$

$$[v] \to g[v] := g([v]) = [gv]$$

where $g([v])$ denotes the set of images by g from every element of $[v]$. This induces a control system in $P(\mathbb{R}^n)$, by acting the semigroup S defined previously on it.

It can be shown that the bilinear control system is controllable in $\mathbb{R}^n \setminus \{0\}$ if, and only if, it is controllable in $P(\mathbb{R}^n)$ and $\mathbb{R}^*_+ x \subset Sx$ for some $x \in \mathbb{R}^n \setminus \{0\}$. In particular, if B has at least one eigenvalue with nonzero real part and the system is controllable in $P(\mathbb{R}^n)$ and accessible in $\mathbb{R}^n \setminus \{0\}$ then the second condition can be shown to also be true, such that the system is controllable in $\mathbb{R}^n$. This means that the set of pairs (A, B) which make the system controllable in $P(\mathbb{R}^n)$ but not in $\mathbb{R}^n$ has measure zero in $\mathfrak{gl}(\mathbb{R}^n) \times \mathfrak{gl}(\mathbb{R}^n)$. This is a particular property of the unrestricted bilinear system. For other types of linear semigroups of $\mathrm{Gl}(\mathbb{R}^n)$, while their action in the projective space is strongly related with their action in $\mathbb{R}^n \setminus \{0\}$ and controllability in $\mathbb{R}^n \setminus \{0\}$ implies controllability in $P(\mathbb{R}^n)$, in general the set of the semigroups which fail the inverse implication isn't always of measure zero. Alternatively, the reciprocal holds with a lot more generality when semigroups are considered in $\mathrm{Sl}(\mathbb{R}^n)$.

The set $P(\mathbb{R}^n)$ has a natural manifold structure such that the function

$$f : S^{n-1} \to P(\mathbb{R}^n)$$

$$x \to [x]$$

defines a covering of $P(\mathbb{R}^n)$. In particular, the tangent space in a point $[v] \in P(\mathbb{R}^n)$ is isomorphic to the space $v^\perp$, the tangent of $\frac{v}{\|v\|}$ in S^{n-1}.

It can be shown that the function

$$\rho : \mathrm{Gl}(\mathbb{R}^n) \times P(\mathbb{R}^n) \to P(\mathbb{R}^n)$$

$$(g, [v]) \to [gv]$$

is well defined and is a differentiable action.

We say that B has a real maximum Eigenvalue if B has a real eigenvalue α such the Eigenspace associated with α has dimension 1 and any eigenvalue λ distinct from α has real part smaller than α. In this case we say that α is the real maximum Eigenvalue of B. If $[v]$ is the Eigenspace associated with α, then there is a proper subspace $V \subset \mathbb{R}^n$ of possible exceptions such that for all $w \in \mathbb{R}^n \setminus V$,

$$\lim_{t \to +\infty} e^{tB}[w] = [v].$$

By rewriting this equation one also has

$$\lim_{t \to +\infty} e^{(-t)(-B)}[w] = [v].$$

This means that $[v] \in \overline{S[w]} \cap \overline{S^{-}[w]}$ for all w not in V. This has interesting consequences regarding controllability, as will be shown in the following lemma.

Lemma 3.2.1. *If $[v]$ and V are as described above, then S is controllable in $P(\mathbb{R}^n)$ if, and only if, S is locally controllable in $[v]$ and there is no nontrivial subspace simultaneously invariant by both A and B.*

Proof. If S is controllable then it is also locally controllable in any point, and in particular $[v]$. Furthermore, note that if W is a subspace distinct from $\{0\}$ and $\mathbb{R}^n$ and invariant by both A and B, then it is also invariant by their exponentials, and therefore by S. That means orbits from W are stuck in W, and therefore S can't be controllable. Therefore controllability of S implies no nontrivial space is simultaneously invariant by both A and B.

For the other implication we first show that the inclusion $[v] \in \overline{S[x]}$ holds for all $[x] \in P(\mathbb{R}^n)$. We already had the inclusion for all $[x] \notin V$. For $[x] \in [V]$, we first show that there is $[y] \in S[x]$ such that $[y] \notin [V]$. To do this assume, by contradiction, that

$S[x] \subset [V]$. Let W be the subspace generated by Sx, then W is nontrivial since $W \subset V$ which is proper in $\mathbb{R}^n$ and $W \neq \{0\}$ since W contains $x \neq 0$. Furthermore, W is S invariant. This is because any element $w \in W$ can be written as

$$w = \sum_{i=1}^{k} \alpha_i s_i x$$

where $\alpha_i \in \mathbb{R}$ and $s_i \in S$. Then, for any $r \in S$:

$$rw = r \sum_{i=1}^{k} \alpha_i s_i x = \sum_{i=1}^{k} \alpha_i (r s_i) x \in W.$$

By hypothesis, since W is nontrivial, there is $w \in W$ such that $Aw \notin W$ or $Bw \notin W$. If $Aw \notin W$ then $e^{\epsilon A} w \notin W$ for sufficiently small $\epsilon > 0$, which contradicts W being invariant. Similarly, if $Bw \notin W$, then $e^{\epsilon B} w \notin W$ for sufficiently small $\epsilon > 0$.

Either case leads to a contradiction, therefore $S[x] \not\subset [V]$. Then, there is $[y] \in S[x]$ such that $[y] \notin V$. By the previous argument,

$$[v] \in \overline{S[y]} \subset \overline{S[x]}.$$

Therefore, $[v] \subset \overline{S[x]}$ for all $[x] \in P(\mathbb{R}^n)$.

It can be shown analogously that

$$[v] \in \overline{S^{-1}[x]}$$

for all $[x] \in P(\mathbb{R}^n)$.

Now let arbitrary $[x], [y] \in P(\mathbb{R}^n)$, and a controllable open set U containing v. By the previous argument, there are $[z] \in S[x] \cap U$ and $[w] \cap S^{-1}[y] \cap U$. Since U is controllable and both $[z], [w]$ are in U, then $[w] \in S[z]$. We then have the chain $[z] \in S[x], [w] \in S[z], [y] \in S[w]$, therefore $[y] \in S[x]$, showing the controllability of S in $P(\mathbb{R}^n)$. □

By the previous section, local controllability of S in $[v]$ can be studied from it's tangent system in $[v]$. Unfortunately it is not easy to calculate $\overline{S} \cap H_v$, which would be required for the equivalence shown in that section, but we can study subsemigroups $R \subset \overline{S} \cap H_v$ for one way conditions that guarantee local controllability if true. For this, we first show how to compute the application $\rho : \mathbb{C}_{[v]} \to \mathrm{Aff}(T_{[v]} P(\mathbb{R}^n))$ described in the previous chapter for this particular case.

3.2.1 Computing the tangent application for bilinear control systems.

Let

$$\rho : \mathcal{C}_{[v]} \to \mathrm{Aff}(T_{[v]}P(\mathbb{R}^n))$$

$$f \to (D\phi_{f(0)}, \frac{d}{dt}\Big|_{t=0} f(t)[v])$$

as in the previous section, where $\phi_{f(0)}$ denotes the application $[x] \to f(0)[x]$. Let $f \in \mathcal{C}_{[v]}$ and $M = f(0) \in \mathrm{Gl}(\mathbb{R}^n)$, $N = f'(0) \in \mathfrak{gl}(\mathbb{R}^n) = M_n(\mathbb{R})$. We can assume, without loss in generality, that $\|v\| = 1$, as $[v] = \left[\frac{v}{\|v\|}\right]$. Let $\beta_1 = \{v, w_2, w_3, ..., w_n\}$ be an orthonormal basis for $\mathbb{R}^n$. Note that $M[v] = f(0)[v] = [v]$, and, therefore, the matrix representation of M in β_1 is written as

$$[M]^{\beta_1}_{\beta_1} = \begin{pmatrix} m & M_1 \\ 0 & M_2 \end{pmatrix}$$

for some $m \in \mathbb{R}$, $M_1 \in M_{1\times(n-1)}$, $M_2 \in M_{(n-1)\times(n-1)}$. Note that $m \neq 0$ and $\det(M_2) \neq 0$ since $M = f(0) \in G$ must be invertible. Let

$$\pi : S^1 \to P(\mathbb{R}^n)$$

$$x \to [x]$$

be the natural covering of $P(\mathbb{R}^n)$. The tangent T_vS^{n-1} can be associated with the space $v^\perp$, which is spanned by the basis $\beta_2 := \{w_2, w_3, ..., w_n\}$. Since $P(\mathbb{R}^n)$ is a local diffeomorphism, the space $T_{[v]}P(\mathbb{R}^n)$ is equal to $D\pi_v(T_vS^{n-1})$, and, by our previous association, $T_{[v]}P(\mathbb{R}^n)$ is the space spanned by the basis $\beta_3 := \{D_v\pi(w_2), D_v\pi(w_3), ..., D_v\pi(w_n)\}$. Note that, by definition, $D_v\pi$ sends the elements of β_2 into the elements of β_3, preserving order, and therefore

$$[D_v\pi]^{\beta_2}_{\beta_3} = Id_{n-1}.$$

An invertible linear transformation g acts in S^{n-1} by

$$\psi_g : S^{n-1} \to S^{n-1}$$

$$x \to \frac{gx}{\|gx\|}.$$

This action is equivariant by π to the action in $P(\mathbb{R}^n)$:

$$\phi_g \circ \pi = \pi \circ \psi_g.$$

Differentiating both sides we get

$$D\phi_g \circ D\pi = D\pi \circ D\psi_g.$$

Assuming $g[v] = [v]$ and restricting this equality to $T_{[v]}G$, we have that $D\pi_v$ is an invertible linear transformation, and, therefore

$$D_{[v]}\phi_g = D_v\pi \circ D_v\psi_g \circ D_{[v]}\pi^{-1}.$$

We calculate $D\phi_M$ by replacing g with M. First, define the normalization application:

$$\eta : \mathbb{R}^n - \{0\} \to S^{n-1}$$
$$x \to \frac{x}{\|x\|}.$$

Note that $\psi_g = \eta \circ g|_{S^{n-1}}$ for arbitrary g, and, therefore,

$$D_v\psi_g = D_{g(v)}\eta \circ g|_{TS^{n-1}}.$$

Let $u, w \in \mathbb{R}^n$, $u \neq 0$. Remember that

$$\frac{d}{dt}\bigg|_{t=0} \|u + tw\| = \frac{d}{dt}\bigg|_{t=0} \sqrt{\langle u + tw, u + tw\rangle} = \frac{1}{\|u\|}\langle u, w\rangle.$$

Then,

$$D_u\eta(w) = \frac{d}{dt}\bigg|_{t=0} \frac{u + tw}{\|u + tw\|} = \frac{w\|u\| - u\frac{1}{\|u\|}\langle u, w\rangle}{\|u\|^2} =$$
$$= \frac{1}{\|u\|}\left(w - \left\langle \frac{u}{\|u\|}, w\right\rangle \frac{u}{\|u\|}\right)$$

Note that the application $w \to w - \left\langle \frac{u}{\|u\|}, w\right\rangle \frac{u}{\|u\|}$ is the orthogonal projection on the plane $u^\perp$. Then, $D_u\eta$ is the composition of this projection with $\frac{1}{\|u\|}Id$. In particular, for $g = M$

and $u = mv = M(v)$, its matrix from β_1 to β_2 is

$$[D_{mv}\eta]_{\beta_2}^{\beta_1} = \begin{pmatrix} 0_{(n-1)\times 1} & \frac{1}{m}Id_{n-1} \end{pmatrix}$$

composing with M we get

$$[D_{[v]}\phi_M]_{\beta_2}^{\beta_2} = \begin{pmatrix} 0_{(n-1)\times 1} & \frac{1}{m}Id_{n-1} \end{pmatrix} [M|_{v^\perp}]_{\beta_1}^{\beta_2}.$$

Remember that

$$[M]_{\beta_1}^{\beta_1} = \begin{pmatrix} m & M_1 \\ 0 & M_2 \end{pmatrix}$$

and, therefore, the restriction $M|_{v^\perp}$ is represented by the matrix

$$[M|_{v^\perp}]_{\beta_1}^{\beta_2} = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}.$$

Then,

$$[D_v\psi_M]_{\beta_2}^{\beta_2} = \begin{pmatrix} 0_{(n-1)\times 1(n-1)\times 1} & \frac{1}{m}Id_{n-1} \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{m}M_2 \end{pmatrix}.$$

That is, in order to get the matrix $[D_v\psi_M]_{\beta_2}^{\beta_2}$ we just have to remove the first line and the first column of $[M]_{\beta_1}^{\beta_1}$ and multiply the resulting $(n-1)\times(n-1)$ matrix by $\frac{1}{m}$, where m is entry $1,1$ of $[M]_{\beta_1}^{\beta_1}$. The differential $D_{[v]}\phi_M$ will have this same matrix representation in the basis β_3. In fact, since $D_v\pi$ takes β_2 in β_3, preserving order, then its matrix on these basis is

$$[D_v\pi]_{\beta_3}^{\beta_2} = Id_{n-1}.$$

Furthermore,

$$[D_v\pi^{-1}]_{\beta_2}^{\beta_3} = Id_{n-1}^{-1} = Id_{n-1}.$$

Therefore,

$$[D_{[v]}\phi_M]_{\beta_3}^{\beta_3} = Id_{n-1}\frac{1}{m}M_2Id_{n-1} = \frac{1}{m}M_2$$

Note that this is the first entry of the application ρ. Therefore, it's first coordinate can be calculated from an orthonormal basis $\beta_1 = \{v, w_2, ..., w_n\}$ by removing the first line and the first column of the matrix $[f(0)]_{\beta_1}^{\beta_1}$ and dividing the remaining $(n-1)\times(n-1)$ matrix by the entry $1,1$ of $[f(0)]_{\beta_1}^{\beta_1}$.

For the second coordinate we once again use the action in S^{n-1} to calculate the action in $P(\mathbb{R}^n)$. By the equivariance between these systems we have

$$\pi \circ \psi_{f(t)} = \phi_{f(t)} \circ \pi$$

for all t in the domain of f, and, therefore,

$$D_v\pi \left(\frac{d}{dt}\bigg|_{t=0} \psi_{f(t)}v \right) = \left(\frac{d}{dt}\bigg|_{t=0} \phi_{f(t)}\pi(v) \right)$$

$$\left(\frac{d}{dt}\bigg|_{t=0} f(t)[v] \right) = D_v\pi \left(\frac{d}{dt}\bigg|_{t=0} \eta(f(t)v) \right) = D_v\pi \circ D_{f(0)v}\eta \circ f(0)v.$$

Remember that

$$[D_{f(0)v}\eta]_{\beta_2}^{\beta_1} = [D_{mv}\eta]_{\beta_2}^{\beta_1} = \begin{pmatrix} 0_{(n-1)\times 1} & \frac{1}{m}Id_{n-1} \end{pmatrix} = \begin{pmatrix} 0_{(n-1)\times 1} & Id_{n-1} \end{pmatrix}.$$

Then, writing $[N]_{\beta_1}^{\beta_1}$ as

$$[N]_{\beta_1}^{\beta_1} = \begin{pmatrix} n & N_1 \\ N_2 & N_3 \end{pmatrix}$$

where $n \in \mathbb{R}$, and N_1, N_2, N_3 are $1 \times (n-1), (n-1) \times 1, (n-1) \times (n-1)$, respectively, we have

$$[D_v\pi \circ D_{f(0)v}\eta \circ f(0)v]_{\beta_3} = Id \begin{pmatrix} 0_{(n-1)\times 1} & \frac{1}{m}Id_{n-1} \end{pmatrix} \begin{pmatrix} n & N_1 \\ N_2 & N_3 \end{pmatrix} \begin{pmatrix} 1 \\ 0_{(n-1)\times 1} \end{pmatrix} =$$

$$\begin{pmatrix} 0_{(n-1)\times 1} & \frac{1}{m}Id_{n-1} \end{pmatrix} \circ \begin{pmatrix} n \\ N_2 \end{pmatrix} = \frac{1}{m}N_2$$

That is, the second coordinate of ρ can be calculated from the first column of $f'(0) = N$ by removing it's first coordinate and dividing by the coordinate $1,1$ of $f(0)$.

3.2.2 Real maximal eigenvalue

We now study the case when B has a maximal real eigenvalue associated with a one dimensional subspace $[v]$. As previously mentioned, denoting $H_{[v]} \cap S$ by R, it is usually not easy to calculate S_R. Instead, we study the controllability of smaller subgroups of S_R for a one way only condition. The subsemigroup we chose to study is the subsemigroup generated by the constant curves

$$C := \{f : \mathbb{R} \to \mathrm{Gl}(\mathbb{R}^n), t \to e^{sB} : s \in \mathbb{R}\}$$

and the curves

$$D := \{f : \mathbb{R} \to \mathrm{Gl}(\mathbb{R}^n), t \to e^{rtA} : r \in [0, +\infty)\}.$$

Denote by $\mathcal{R}$ the subsemigroup of S_R generated by $C \cup D$. Let $\beta = \{v, w_1, w_2, ..., w_n\}$ be an orthonormal basis of $\mathbb{R}^n$, and write

$$B = \begin{pmatrix} b & B_1 \\ 0_{(n-1\times 1)} & B_2 \end{pmatrix}$$

$$A = \begin{pmatrix} \alpha & A_1 \\ a & A_2 \end{pmatrix}$$

where $\alpha, b \in \mathbb{R}$, a is $(n-1) \times 1$, A_1, B_1 are $1 \times (n-1)$ and A_2, B_2 are $(n-1) \times (n-1)$. Note that

$$e^{sB} = \begin{pmatrix} e^{sb} & C \\ 0_{(n-1\times 1)} & e^{sB_2} \end{pmatrix}$$

for some $1 \times (n-1)$ matrix C, and, therefore, the image of the elements $f : t \to e^{sB}$ in C is

$$\left(\frac{e^{sB_2}}{e^{sb}}, 0 \right) = e^{(sB_2 - bId, 0)},$$

that is, $\rho(C)$ is the one parameter group in $Aff(T_{[v]}P(\mathbb{R}^n)$ generated by

$$(B_2 - bId, 0),$$

or, equivalently, the semigroup generated by

$$\{(B_2 - bId, 0), (-(B_2 - bId), 0)\}.$$

Furthermore, the elements $f : t \to e^{rtA}$ in D satisfy $f'(0) = rA$ and $f(0) = Id$, therefore,

$$\rho(f) = (Id, ra) = e^{r(0,a)},$$

and $\rho(D)$ is the semigroup generated by $(0,a)$. Since $\mathcal{R}$ is generated by $C \cup D$, then $\rho(\mathcal{R})$ is the semigroup generated by the set

$$\{(B_2 - bId, 0), -(B_2 - bId, 0), (0, a)\}.$$

Denoting $B_2 - bId$ by $\overline{B}$, the semigroup generated by the above set is the semigroup

$$\{e^{t_1 a} e^{s_1 \overline{B}} ... e^{t_k a} e^{s_k \overline{B}} : t_i, s_i \in \mathbb{R}, t_i \geq 0, k \in \mathbb{N}\},$$

where $e^{t_i a}$ denotes the application $x \to x + t_i a$. It is associated to the control system defined by the differential equation

$$\dot{x}(t) = a + u(t)\overline{B}x(t),$$

with the set of controls $\mathcal{U}$ including all piecewise constant functions $u : \mathbb{R} \to \mathbb{R}$. We denote this system by $\{a, \overline{B}, -\overline{B}\}$. It is a particular case of the biaffine system

$$\dot{x}(t) = a + Ax(t) + u(t)(b + Bx(t))$$

by taking A and b as 0. We study these systems in the next subsection.

3.2.3 The system $\{a, B, -B\}$

In the next results we study controllability of the system $\{a, B, -B\}$. For this entire subsection, whenever we mention cones it is implicit that we are also assuming convexity, unless stated otherwise.

Proposition 3.2.2. *Let $t_1, t_2, ..., t_k, s_1, s_2, ..., s_k \in \mathbb{R}^k$, then*

$$e^{t_k B} e^{s_k a} ... e^{t_2 B} e^{s_2 a} e^{t_1 B} e^{s_1 a}(0) = \sum_{i=1}^{k} e^{\left(\sum\limits_{j=i}^{k} t_j\right) B} s_i a$$

Proof. The proof follows by induction. The case $k = 1$ is trivial. Assuming the equality holds for $k = n - 1$ we have

$$e^{t_k B} e^{s_k a} ... e^{t_1 B} e^{s_1 a}(0) = e^{t_k B}(e^{t_{k-1} B} e^{s_{k-1} a} ... e^{t_1 B} e^{s_1 a}(0) + s_k a) =$$

$$= e^{t_k B}\left(\sum_{i=1}^{k-1} e^{\left(\sum\limits_{j=i}^{k-1} t_i\right)B} s_i a\right) + e^{t_k B} s_k a =$$

$$= \left(\sum_{i=1}^{k-1} e^{\left(\sum\limits_{j=i}^{k} t_j\right)B} s_i a\right) + e^{\left(\sum\limits_{j=k}^{k} t_j\right)B} s_k a =$$

$$= \left(\sum_{i=1}^{k} e^{\left(\sum\limits_{j=i}^{k} t_j\right)B} s_i a\right)$$

□

For a linear transformation $B : \mathbb{R}^n \to \mathbb{R}^n$ we say that a set C is invariant by the flow of B if $e^{tB}C \subset C$ for all $t \geq 0$. Note that if C is a subspace then this is equivalent to B invariance, however this equivalence is not true for all kinds of sets. This definition is used in the following corollary

Corollary 3.2.3. *Denote by $\mathcal{O}^+, \mathcal{O}^-$ the positive and negative orbits of 0 in the system $\{a, B, -B\}$, respectively. Then*

$$\mathcal{O}^+ = \left\{\sum_{i=1}^{k} \alpha_i e^{t_i B} a; k \in \mathbb{N}, \alpha_i > 0, t_i \in \mathbb{R}, \forall i \in \{1, ..., k\}\right\}$$

$$\mathcal{O}^- = \left\{\sum_{i=1}^{k} -\alpha_i e^{t_i B} a; k \in \mathbb{N}, \alpha_i > 0, t_i \in \mathbb{R}, \forall i \in \{1, ..., k\}\right\} = -\mathcal{O}^+$$

In other words, $\mathcal{O}^+$ is the cone generated by $\{e^{tB}a; t \in \mathbb{R}\}$, and $\mathcal{O}^-$ is the cone generated by $\{-e^{tB}a; t \in \mathbb{R}\}$. $\mathcal{O}^+$ coincides, also, with the smallest cone invariant by the flow of B that contains a. As a consequence, the following are equivalent:

- *The system is controllable.*
- *The system is locally controllable in the origin*
- *$\mathcal{O}^+ = \mathbb{R}^n$*
- *$\mathbb{R}^n$ is the smallest cone invariant by the flow of B that contains a.*

Proposition 3.2.4. *If B has a real eigenvalue then the system is not controllable.*

Proof. Let $\beta = \{b_1, b_2, ..., b_n\}$ be a basis of $\mathbb{R}^n$ such that

$$[B]_\beta^\beta := \begin{pmatrix} J_1 & 0 \\ 0 & M \end{pmatrix},$$

where J_1 is the Jordan block associated to the real eigenvalue, let's say, λ:

$$J_1 = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Denote by k the number of lines/columns of J_1, and define the linear functional

$$p : \mathbb{R}^n \to \mathbb{R}$$

$$(a_1, a_2, ..., a_k, ..., a_n) \to a_k.$$

Then $p(e^{tB}v) = e^\lambda p(v)$ for any $t \in \mathbb{R}, v \in \mathbb{R}^n$. In particular, $p^{-1}((0, +\infty)), p^{-1}((-\infty, 0)), p^{-1}(0)$ are cones invariant by the flow of B such that their union covers all of $\mathbb{R}^n$. Therefore, $\mathcal{O}^+ \subset p^{-1}((0, +\infty))$, or $\mathcal{O}^+ \subset p^{-1}((-\infty, 0))$, or $\mathcal{O}^+ \subset p^{-1}(0)$, depending on which of these sets contain a. In any case, $\mathcal{O}^+ \neq \mathbb{R}^n$, and the system is not controllable. □

That means that for $\{a, B, -B\}$ to be controllable B must have only complex eigenvalue. Interestingly, the only cones invariant by such a matrix are subspaces. This is shown in the next result

Proposition 3.2.5. *Let $B : \mathbb{R}^n \to \mathbb{R}^n$ a real linear transformation with no real eigenvalues. If $V \neq \emptyset$ is a cone invariant by the flow of B then V is a subspace.*

Proof. We will first prove that if $Int(V) \neq 0$ then $V = \mathbb{R}^n$. The proof of this will follow by induction on the dimension n of the space.

The claim is trivial is trivial in $\mathbb{R}^0$. We will prove it in $\mathbb{R}^n, n > 0$ assuming it holds true for dimensions smaller than n.

Since B has no real eigenvalues and we are assuming $n > 0$ then B must have at least one pair of conjugated complex eigenvalues. Then, there must be a B invariant space W of dimension 2 such that the restriction B_W is represented, in some basis of W, by the matrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

with $a, b \in \mathbb{R}, b \neq 0$. Then,

$$e^{tB}_W = e^{tB_W} = e^{ta} R_{tb}$$

where R_{tb} denotes the rotation by tb in the previously mentioned basis of W. The only cones in W invariant by the flow of B_W are $\{0\}$ and V. Now let $\pi : \mathbb{R}^n \to \mathbb{R}^n/V$ be the canonical projection. Since W is B invariant, the linear transformation B projects into a linear transformation

$$\hat{B} : \mathbb{R}^n/W \to \mathbb{R}^n/W$$

$$x + W \to Bx + W$$

such that $\pi \circ B = \hat{B} \circ \pi$ and $\pi \circ e^{tB} = e^{t\hat{B}} \circ \pi$ for all $t \in \mathbb{R}$. The linear transformation $\hat{B}$ has no real eigenvalues. In fact, if $\hat{B}$ has a real eigenvalue then it would have a real eigenvector $z + W \neq 0 + W$ associated to it, such that $\langle z + W \rangle$ would be $\hat{B}$ invariant. But then the subspace $Z := W + \langle z \rangle = \pi^{-1}(\langle z + W \rangle)$ is a 3 dimensional subspace invariant by B. A real linear transformation in an odd dimension space must always have a real eigenvalue, such that B must have a real eigenvector in W. But that contradicts the hypothesis of the theorem. Then, $\hat{B}$ must have no real eigenvalues.

The cone V projects into a cone $\pi(V)$. π is an open application and we are assuming V has nonempty interior, therefore $Int(\pi(V)) \neq \emptyset$. Furthermore, $\pi(V)$ is invariant by the flow of $\hat{B}$, as $e^{t\hat{B}} \circ \pi(V) = \pi \circ e^{tB}(V) \subset \pi V$ for $t > 0$. Note that $dim(\mathbb{R}^n/W) = n - 2 < n$. Then, by the induction hypothesis, $\pi(V) = \mathbb{R}^n/W$.

In particular, $0+W \in Int(\pi(V))$, then, $Int(V)$ intersects $\pi^{-1}(0+W) = W$ (see lemma 2.3.22). Note that V, W are both invariant by the flow of B such that $V \cap W$ must also be invariant by the flow of B. Since $V \cap W \subset W$, this intersection is invariant by the flow of B if, and only if, it is invariant by the flow of B_W. Then, $V \cap W$ is either $\{0\}$ or W. Since $Int(V)$ must intersect W, then $V \cap W$ cannot equal $\{0\}$, therefore $V \cap W = W$, that is, $W \subset V$.

We then have $W \subset V$ and $\pi(V) = \mathbb{R}^n/W$. This is enough to prove that $V = \mathbb{R}^n$. In fact, since V is a convex cone, then it is closed for sums of it's elements. Let $x \in \mathbb{R}^n$ arbitrary, then $x + W \in \pi(V)$ and there must be $y \in V$ such that $\pi(y) = x + W$, that is, $x - y \in W$. But $W \subset V$ such that $x - y \in V$, and, therefore, $x = y + (x - y) \in V$. Since $x \in \mathbb{R}^n$ is arbitrary, then $\mathbb{R}^n \subset V$.

Now for the theorem itself, let V as in the hypothesis and let W be the subspace generated by V. Since V is a convex cone, then V has nonempty interior in W. Furthermore, since W is generated by a set the is invariant by the flow of B, then W itself is invariant by the flow of B, as the e^{tB} are all linear application. Since W is a subspace then it is also invariant by B, and it is possible to consider the restriction B_W. Since B has no real eigenvalues then neither does B_W, and, since V is invariant by the flow of B then it is also invariant by the flow of B_W. Then, the previous argument assures that $V = W$, showing that V is a subspace. □

Theorem 3.2.6. *The system $\{a, B, -B\}$ is controllable if, and only if, both of the following are true:*

1. *B has no real eigenvalues*
2. *a is not contained in a proper B invariant subspace.*

Proof. As shown in proposition 3.2.2, the system is controllable if, and only if, a is contained in a proper cone that is invariant by the flow of B. Since any subspace is in particular a cone, a must not be contained in any proper B invariant subspace, and, by proposition 3.2.4, B must also not have any real eigenvalue. This shows that controllability implies 1 and 2. For the other implication, assume B has no real eigenvalues. By proposition 3.2.5, all of the cones invariant by the flow of B are subspaces. Then, if a is not contained in any B invariant proper subspace the system is controllable. □

It is possible to calculate the conditions of the previous theorem as follows: condition 1 can be calculated from the characteristic polynomial of B and condition 2 can be calculated from the B cyclic space of a, that is, condition 2 is true if the vectors $a, Ba, B^2a, ..., B^{n-1}a$ span $\mathbb{R}^n$.

3.3 Appendix

In this section we prove the lemma 3.1.1 from section 3.1.

Lemma 3.1.1. *Let*

$$F : \mathbb{R}^n \to M$$

be a continuously differentiable function such that $F(0) = v$. *If* $V \subset \mathbb{R}^n$ *is a closed convex cone with nonempty interior such that* $DF_0(V) = T_vM$ *then* $v \in Int(F(Int(V)))$.

Proof. DF_0 must be a surjective linear function, since $DF_0(V) = T_vM$.

Let

$$\phi : T_vM \to \mathbb{R}^m$$

be a chart such that $\phi(v) = 0$ and

$$\hat{F} := \phi \circ F : \mathbb{R}^n \to \mathbb{R}^m.$$

The lemma is equivalent to the inclusion

$$0 \in Int(\hat{F}(Int(V))).$$

Let $X = \ker(D\hat{F}_0)$ and $Y = X^\perp$. Then

$$\mathbb{R}^n = X \oplus Y.$$

Furthermore, both DF_0 and $D\phi_v$ are surjective, therefore $D\hat{F}_0$ is surjective and is an isomorphism from Y to $\mathbb{R}^m$. Since V is convex with nonempty interior and $0 \in Int(\mathbb{R}^n) = Int(D\hat{F}_0(V))$, then there is $x_0 \in Int(V)$ such that $D\hat{F}_0(x_0) = 0$ (see lemma 2.3.22). If $0 = x_0 \in Int(V)$ then $V = \mathbb{R}^n$ and

$$0 \in Int(\hat{F}(\mathbb{R}^n)) = Int(\hat{F}(Int(\mathbb{R}^n))) = Int(\hat{F}(Int(V)))$$

by the submersion theorem. Assume $x_0 \neq 0$. To simplify the calculations, choose an inner product $\langle \cdot, \cdot \rangle$ in $\mathbb{R}^n$ and the respective norm and metric such that $\|x_0\| = 1$. Let $\epsilon > 0$ such that $B(x_0, \epsilon) \subset V$. Note that $B(\alpha x_0, \alpha\epsilon) \subset V$ for all $\alpha > 0$ since V is a cone.

Define

$$\tilde{F} : X \oplus Y = \mathbb{R}^n \to \mathbb{R}^m \oplus X$$

$$(x, y) \to (\hat{F}(x, y), x).$$

Note that $\tilde{F}$ is a differentiable function and $D\tilde{F}_0$ is an isomorphism. Let $\langle \cdot, \cdot \rangle_2$ be the inner product in $\mathbb{R}^m \oplus X$ induced from $\langle \cdot, \cdot \rangle$ by $D\tilde{F}_0$, that is, the product defined by

$$\langle u, w \rangle_2 = \langle (D\tilde{F}_0)^{-1} u, (D\tilde{F}_0)^{-1} w \rangle.$$

We choose this inner product as it is the only product of $\mathbb{R}^m \oplus X$ which makes $D\tilde{F}_0$ an isometry from the product $\langle \cdot, \cdot \rangle$ in $\mathbb{R}^n$. This further simplifies some of the calculations.

Since $\tilde{F}$ is differentiable and has invertible differential in 0, then it has a local differentiable inverse $\tilde{F}^{-1}$. Let $U \subset \mathbb{R}^m \oplus X$ be an open set containing 0 such that U is contained in the domain of $\tilde{F}^{-1}$ and

$$\frac{1}{\|u\|_2} \|\tilde{F}^{-1}(u) - D(\tilde{F}^{-1})_0(u)\| < \epsilon$$

for all $u \in U$. Let $\alpha > 0$ sufficiently small such that

$$(0, \alpha x_0) \in U$$

Note that

$$\|(0, \alpha x_0)\|_2 = \|D\tilde{F}_0(\alpha x_0)\|_2 = \alpha.$$

Therefore

$$\epsilon > \frac{1}{\alpha} \|\tilde{F}^{-1}(0, \alpha x_0) - D(\tilde{F}^{-1})_0(0, \alpha x_0)\| = \frac{1}{\alpha} \|\tilde{F}^{-1}(0, \alpha x_0) - \alpha x_0\|.$$

In particular, if $y_0 = \tilde{F}^{-1}(0, \alpha x_0)$, then $y_0 \in B(\alpha x_0, \alpha\epsilon) \subset Int(V)$. Also, since $(0, \alpha x_0) \in U$, which is contained in the domain of $\tilde{F}^{-1}$, then $\tilde{F}$ is still a local diffeomorphism in y_0. In particular, $\tilde{F}$ is open in y_0. Since $y_0 \in Int(V)$ and $\tilde{F}$ is open in y_0, then $(0, \alpha x) = \tilde{F}(y) \in Int(\tilde{F}(Int(V)))$. Finally, note that $\hat{F} = \pi \circ \tilde{F}$ where

$$\pi : \mathbb{R}^m \oplus X \to \mathbb{R}^m$$

$$(a, b) \to a$$

is the natural projection and, in particular, an open function. Therefore,

$$0 = \pi(0, \alpha x) \in \pi(Int(\tilde{F}(Int(V)))) = Int(\hat{F}(Int(V)))$$

completing the proof. □

CHAPTER 4

INVARIANT CONES FOR SEMIGROUPS OF $Sl(\mathbb{R}^n)$

The results presented in this chapter are a joint work with Emerson V. Castelani, João A. N. Cossich and Alexandre J. Santana, and we are very grateful to Luiz A. B. San Martin for suggesting this problem and many of the ideas used. We deal with invariant cones for semigroup actions to study controllability of control systems. In our context this question is related with the flag type of the semigroup (in particular semigroup of the control system) and hence with the control sets of the semigroup (or of the control system). Note that it is far from achieving global results on controllability of bilinear control systems, that is, to find sufficient conditions for controllability is a long term and still incomplete area of research (see e.g. Elliot [8]). But, in the last few decades, several papers have been published showing that the Lie theory, especially the theory of semigroups of semisimple Lie groups, provides tools to study controllability (see e.g. Do Rocio, San Martin and Santana [5], Do Rocio, Santana and Verdi [6], Dos Santos and San Martin [7] and San Martin [12]). As an example, we recall the bilinear control system presented in the previous chapter:

$$\dot{x} = Ax + uBx, x \in \mathbb{R}^d \setminus \{0\}, u \in \mathbb{R}, \tag{4.0-1}$$

where A and B are $d \times d$-matrices, we have that the semigroup S of the system is given by the concatenations of solutions:

$$S = \{e^{t_k(A+u_kB)}e^{t_{k-1}(A+u_{k-1}B)} \dots e^{t_1(A+u_1B)}, t_i \geq 0, k \in \mathbb{N}\}$$

and the group system has a similar definition just changing the positive times t_i by real times (see e.g Colonius and Kliemann [1] and Elliot [8]). And if we consider A and B generating a semisimple Lie algebra $\mathfrak{g}$ we have the possibility to use the semisimple Lie theory to study controllability of the system, for example in case of $\mathfrak{g} = \mathfrak{sl}(\mathbb{R}^d)$ we have that this system is controllable in $\mathbb{R}^d \setminus \{0\}$ ($Sx = \mathbb{R}^d \setminus \{0\}$ for all $x \in \mathbb{R}^d \setminus \{0\}$) if and only if $S = \mathrm{Sl}(\mathbb{R}^d)$ (see [5] and [16]).

One of the most interesting ways to prove that the above system is not controllable is to show the existence of some S-invariant proper subset of $\mathbb{R}^d$, a trap of the system. This problem was addressed in [10], by Sachkov, but in [5] the authors searched these invariant sets among the convex cones, since if a set C is invariant by the system then the convex closure of C is also invariant. In this chapter we follow a similar approach to improve and generalize the results contained in [5] and in particular to give a necessary and sufficient condition for controllability of the above system when $A, B \in \mathfrak{sl}(\mathbb{R}^d)$. More specifically, we prove that the system is controllable if and only if it does not have an invariant proper cone in the k-fold exterior product of $\mathbb{R}^d$, $\bigwedge^k \mathbb{R}^d$, for all $k \in \{1, \dots, d-1\}$. In fact, this is a consequence of our following transitivity result: Let $S \subset \mathrm{Sl}(\mathbb{R}^d)$ be a connected semigroup with nonempty interior. Then $S = \mathrm{Sl}(\mathbb{R}^d)$ if and only if there are no S-invariant and proper cones in $\bigwedge^k \mathbb{R}^d$, for all $k \in \{1, \cdots, d-1\}$. These two results are built from the theory of flag type of a semigroup.

We briefly recall the main concept or tool of this chapter. Consider $S \subset \mathrm{Sl}(\mathbb{R}^d)$ a semigroup with nonempty interior. Denote by $\mathbb{F}_\Theta$ the flag manifold of all flags $(V_1 \subset \cdots \subset V_k)$ of subspaces $V_i \subset \mathbb{R}^d$ with $\dim V_i = r_i$, $i = 1, \dots, k$ and $\Theta = \{r_1, \dots, r_k\}$. Take the canonical projection $\pi^{\Theta}_{\Theta_1} : \mathbb{F}_\Theta \to \mathbb{F}_{\Theta_1}$ with $\Theta_1 \subset \Theta$ and denote by $\mathbb{F}$ the full flag manifold with the sequence $\Theta_M = \{1, 2, \dots, d-1\}$. There is a natural (transitive) action of $\mathrm{Sl}(\mathbb{R}^d)$ in these flag manifolds. Recall that an invariant control set is closed and its interior is dense on it, as S is assumed to have nonempty interior and, therefore, be accessible. One important result is that in each flag manifold $\mathbb{F}_\Theta$ there exists just one S-invariant control set. Moreover, there exist $\Theta \subset \Theta_M$ such that $\pi^{-1}_{\Theta}(C_\Theta) = C$ where

$\pi_\Theta : \mathbb{F} \to \mathbb{F}_\Theta$ is the canonical projection, and C_Θ, C are the invariant control sets in $\mathbb{F}, \mathbb{F}_\Theta$ respectively. In addition, among these flag manifolds there is exactly one, denoted by $\mathbb{F}_{\Theta(S)}$, which is minimal (see [12]). The flag manifold $\mathbb{F}_{\Theta(S)}$ (or $\Theta(S)$) is called the flag (or parabolic) type of S (for details see San Martin [11] and San Martin and Tonelli [16]). We note that once we know the invariant control set $C_{\Theta(S)}$ in the flag type $\mathbb{F}_{\Theta(S)}$ then every invariant control set is described because for any Θ we have $C_\Theta = \pi_\Theta(C)$ and $C = \pi^{-1}_{\Theta(S)}\left(C_{\Theta(S)}\right)$. Given $\Theta = \{r_1, \ldots, r_n\}$ with $0 < r_1 < \cdots < r_n < d$ define $\Theta^* = \{d - r_n, \ldots, d - r_1\}$. The flag manifold $\mathbb{F}_{\Theta^*}$ is said to be dual of $\mathbb{F}_\Theta$. With this we have that the flag type of S^{-1} is given by the flag manifold $\mathbb{F}_{\Theta(S)^*}$ dual to the flag type of S (see [13]).

From this semigroup theoretical development, considering S a connected semigroup with nonempty interior and taking $\Theta(S)$ its flag type, we prove our main result: there exists a non-trivial S-invariant cone $W \subset \bigwedge^k \mathbb{R}^d$ if and only if $k \in \Theta(S)$. Hence, as a consequence we show the controllability and transitivity results mentioned above.

4.1 Preliminaries

Recall that the flag manifolds $\mathbb{F}_\Theta$ are compact and the **minimal flag manifolds** are the Grassmannians $\mathbb{F}_\Theta = \mathbb{G}_k(d)$, where $\Theta = \{k\}$. A particular case, when $k = 1$, is the projective space $\mathbb{P}^{d-1} = \mathbb{G}_1(d)$.

From now on, in this section we discuss the special case $\mathbb{G}_k(d)$, $1 \leq k \leq d-1$. In this work it is convenient represent $\mathbb{G}_k(d)$ in the following algebraic way. Let $B_k(d)$ be the set of $d \times k$ matrices of rank k. Define in $B_k(d)$ the following equivalence relation: $p \sim q$ if exists $a \in \mathrm{Gl}(\mathbb{R}^k)$ with $q = pa$. In other words, $p \sim q$ if, and only if, the columns of p and q generate the same subspace of $\mathbb{R}^d$. Then we can see $\mathbb{G}_k(d)$ as $B_k(d)/\sim$. Denote the elements of $\mathbb{G}_k(d)$ by $[p]$. There is a natural action ρ_k of the Lie group $\mathrm{Sl}(\mathbb{R}^d)$ on $\mathbb{G}_k(d)$, which is given by $\rho_k(g, [p]) = [gp]$.

Now take an arbitrary basis $\mathcal{B}$ of $\mathbb{R}^d$ and $N_\mathcal{B}$ the nilpotent group of lower triangular matrices (with respect to $\mathcal{B}$) with ones on the main diagonal. The decomposition of $\mathbb{G}_k(d)$ into $N_\mathcal{B}$-orbits is called **Bruhat decomposition** of $\mathbb{G}_k(d)$, moreover if we change the basis the decomposition also changes. There is just a finite number of these orbits, $N_\mathcal{B}[p]$ with $[p] \in \mathbb{G}_k(d)$. It is well known that exists only one open and dense orbit,

$N_{\mathcal{B}}[p_0]$, where $[p_0]$ is the subspace spanned by the first k basic vectors (see [14]). We have that $N_{\mathcal{B}}[p_0]$ can be written as

$$\begin{bmatrix} I_k \\ X \end{bmatrix}$$

with I_k the $k \times k$ identity and X an arbitrary $(d-k) \times k$ matrix. Taking

$$\eta = \begin{bmatrix} A_1 & 0 \\ Y & A_2 \end{bmatrix} \in N_{\mathcal{B}}$$

with A_1 and A_2 invertible, it follows that

$$\rho_k(\eta, [p_0]) = [\eta p_0] = \begin{bmatrix} A_1 & 0 \\ Y & A_2 \end{bmatrix} \begin{bmatrix} I_k \\ 0 \end{bmatrix} = \begin{bmatrix} A_1 \\ Y \end{bmatrix} = \begin{bmatrix} I_k \\ YA_1^{-1} \end{bmatrix}.$$

Note that this orbit is diffeomorphic to euclidean spaces.

Another important concept here is the **split regular** or just **regular** element, that is the $h \in \mathrm{Sl}(\mathbb{R}^d)$ with positive and distinct eigenvalues, where in some basis (denoted by $\mathcal{B}(h)$), $h = \mathrm{diag}\{\lambda_1, \ldots, \lambda_d\}$ with $\lambda_1 > \cdots > \lambda_d > 0$. Considering the action on $\mathbb{G}_k(d)$, the fixed points for h are the subspaces spanned by k basic vectors. Moreover, these fixed points are hyperbolic and with respect to $\mathcal{B}(h)$, the stable manifolds are the $N_{\mathcal{B}}$-orbits. One interesting dynamical property is that the stable manifold of the subspace $[p_0]$ is open and dense, and, if $[p_0]$ is the space generated by the first k vectors of $\mathcal{B}(h)$ then it is the unique attractor for h, such that $h^m[q] \to [p_0]$ for generic $[q]$. Now taking h^{-1} instead of h and reverting the order of the basis, it follows that h has also just one repeller, and it is the subspace spanned by the last k basic vectors $\{e_{d-k+1}, \ldots, e_d\}$ of $\mathcal{B}(h)$.

We recall other dynamical facts. Let $S \subset \mathrm{Sl}(\mathbb{R}^d)$ be a semigroup with nonempty interior and denote by reg(S) the set of regular elements in $\mathrm{int}S$. As before take C_k the S-invariant control set in $\mathbb{G}_k(d)$, its uniqueness implies that

$$C_k = \bigcap_{[p] \in \mathbb{G}_k(d)} \mathrm{cl}(S[p]).$$

According to the above comments, for $h \in$ reg(S) we have that $b_{\{k\}}(h) = [p_0]$ and $C_k \subset N_{\mathcal{B}(h)}[p_0]$ if $k \in \Theta(S)$. The **set of transitivity** of an invariant control set C_k is the set C_k^0 of the fixed points which are the attractors for elements in reg(S) (see [12]). Specifically, we have that for any $[p] \in C_k^0$, there exists a basis $\mathcal{B}(h) = \{e_1, \ldots, e_d\}$ of $\mathbb{R}^d$ and $h = \mathrm{diag}\{\lambda_1, \ldots, \lambda_d\}$ with $\lambda_1 > \cdots > \lambda_d > 0$ (in this basis), such that $h \in \mathrm{int}S$ and $[p] = \langle e_1, \ldots, e_k \rangle$, i.e., $[p]$ is the attractor of h. From this fact it follows that the set of attractors of elements in reg(S) coincides with C_k^0 and this set is dense in C_k. Hence reg(S) is dense in $\mathrm{int}S$ and C_k is formed, in some sense, by attractors for these regular elements. This is a kind of converse to the fact that $[p] \in C_k$ if $[p]$ is the attractor of a element $h \in$ reg(S). Therefore C_k is contained in the open Bruhat component corresponding to $\mathcal{B}(h)$. Another interesting result in this context is that $C_k = \mathbb{G}_k(d)$ for some k if and only if S is transitive on $\mathbb{G}_k(d)$. On the other hand, we have that if S is a proper semigroup of $\mathrm{Sl}(\mathbb{R}^d)$, then $C_k \neq \mathbb{G}_k(d)$ for any $k \in \{1, \ldots, d-1\}$ and S is not transitive on $\mathbb{G}_k(d)$ (see [14], [15] and [16] for more details).

We finish this section recalling some necessary facts about tensorial product and Grassmanianns.

For $k \in \{1, \ldots, d\}$, denote by $\bigwedge^k \mathbb{R}^d$ the k-fold exterior product of $\mathbb{R}^d$ and let $\mathcal{F}_k(d)$ be the set of all k multi-index $I = \{i_1, \ldots, i_k\} \subset \{1, \ldots, d\}$ with $1 \leq i_1 < \cdots < i_k \leq d$. It is well known that if we fix a basis $\mathcal{B} = \{e_1, \ldots, e_d\}$, then $\{e_I := e_{i_1} \wedge \cdots \wedge e_{i_k}; I = \{i_1, \ldots, i_k\} \in \mathcal{F}_k(d)\}$ is a basis of $\bigwedge^k \mathbb{R}^d$. Along the text, we use the notation $\mathcal{D}$ to designate the set of all decomposable elements of $\bigwedge^k \mathbb{R}^d$, that is, the set of elements that can be written as $u_1 \wedge \cdots \wedge u_k$ with $u_i \in \mathbb{R}^d$.

The manifold $\mathbb{G}_k(d)$, $k \in \{1, \ldots, d-1\}$, can be seen as a compact submanifold of the projective space $\mathbb{P}\left(\bigwedge^k \mathbb{R}^d\right)$ of $\bigwedge^k \mathbb{R}^d$ via Plücker embedding $\varphi : \mathbb{G}_k(d) \to \mathbb{P}\left(\bigwedge^k \mathbb{R}^d\right)$, $\varphi([p]) = [u_1 \wedge \cdots \wedge u_k]$, where $p = [u_1 \ \ldots \ u_k]$ is a $d \times k$ matrix and $[u_1 \wedge \cdots \wedge u_k] \in \mathbb{P}\left(\bigwedge^k \mathbb{R}^d\right)$ denotes the class of all non-zero multiples of $u_1 \wedge \cdots \wedge u_k \in \bigwedge^k \mathbb{R}^d$.

Identifying the Grassmaniann $\mathbb{G}_k(d)$ as a subset of $\mathbb{P}\left(\bigwedge^k \mathbb{R}^d\right)$, we can write the action ρ_k of $\mathrm{Sl}(\mathbb{R}^d)$ on $\mathbb{G}_k(d)$ as

$$\rho_k(g, [u_1 \wedge \cdots \wedge u_k]) = [gu_1 \wedge \cdots \wedge gu_k]$$

and denote $\rho_k(g, [p])$ simply by $g[p]$.

In the next sections $\pi : \left(\bigwedge^k \mathbb{R}^d\right) \backslash \{0\} \to \mathbb{P}\left(\bigwedge^k \mathbb{R}^d\right)$ represents the canonical projection.

4.2 Cones in k-fold exterior product

From now on we consider a connected semigroup $S \subset \mathrm{Sl}(\mathbb{R}^d)$ with nonempty interior. In this work a **cone** means a closed convex cone in a finite dimensional vector space V and if not otherwise specified the cones are proper and non-trivial. Remember that a cone W is **pointed** if $W \cap -W = \{0\}$ and **generating** if int$W \neq \emptyset$. Our main interest is to study the S-invariance of this kind of cones in $\bigwedge^k \mathbb{R}^d$, with $1 \leq k \leq d-1$.

In this section we present some technical and useful results about cones. In particular, we show that any S-invariant cone $W \subset \bigwedge^k \mathbb{R}^d$ contains a decomposable element and it is pointed and generating (see the following Propositions 4.2.4 and 4.2.5). To obtain these results we need some lemmas.

Lemma 4.2.1. *Let $F : V_1 \to V_2$ be an analytic map where V_1 and V_2 are finite dimensional vector spaces. Assume that for a nonempty open set $U \subset V_1$ there is a subspace $V \subset V_2$ such that $F(U) \subset V$. Then $F(V_1) \subset V$.*

Proof. The canonical projection $p : V_2 \to V_2/V$ is linear and then analytic. Therefore, $p \circ F$ is an analytic function, and $p \circ F(U) = 0 + V$, since $p(x) \in V$ for all $x \in U$. Therefore, $p \circ F(x) = 0 + V$, for every $x \in V_1$, because the unique analytic map between finite dimensional vector spaces which vanishes on an open subset of the domain is the null map. Hence, $F(x) \in V$, for all $x \in V_1$. □

Lemma 4.2.2. *Let V be a d-dimensional vector space and take the cone $W \subset V$. Then W is generating if, and only if, W is not contained in any proper subspace of V.*

Proof. If W has nonempty interior, then W is not contained in a proper subspace of V. For the converse, observe that convex cones spanned by any basis of V have nonempty interior. In fact, let $\{e_1, \ldots, e_d\}$ be a basis of V. The convex cone spanned by this basis is the set

$$\left\{ \sum_{i=1}^{d} \alpha_i e_i;\ \alpha_1, \ldots, \alpha_d \geq 0 \right\}.$$

Then the interior of this set is nonempty. Now, assuming that W is not contained in a proper subspace of V, we have that W contains a basis $\mathcal{B}$. Since W is a convex cone it follows that W also contains the convex cone spanned by $\mathcal{B}$. Therefore, int$W \neq \emptyset$. □

Lemma 4.2.3. *If U is open in the set of decomposable elements $\mathcal{D}$ (in the relative topology), then U contains a basis of $\bigwedge^k \mathbb{R}^d$.*

Proof. First note that we can write the set of decomposable elements as the image of a polynomial function. In fact, let $F : (\mathbb{R}^d)^k \to \bigwedge^k \mathbb{R}^d$ be the map given by $F(u_1, u_2, \cdots, u_k) = u_1 \wedge u_2 \wedge \cdots \wedge u_k$. Clearly, $F((\mathbb{R}^d)^k) = \mathcal{D}$. On the other hand, F is polynomial due to the multi-linearity of the wedges, hence it is analytic.

Since F is continuous, the set $F^{-1}(U)$ is open in $(\mathbb{R}^d)^k$. In fact, since U is open in $\mathcal{D}$, there is an open $U' \subset \bigwedge^k \mathbb{R}^d$ with $U = U' \cap \mathcal{D}$. So $F^{-1}(U) = F^{-1}(U') \cap (\mathbb{R}^d)^k = F^{-1}(U')$.

Now, suppose that U does not contain a basis of $\bigwedge^k \mathbb{R}^d$, then S is contained in a proper subspace Z, so $F(F^{-1}(U)) \subset U \subset Z$, and by Lemma 4.2.1, $F((\mathbb{R}^d)^k) \subset Z$. Hence $\mathcal{D}$ is contained in a proper subspace of $\bigwedge^k \mathbb{R}^d$. This is a contradiction, because $\mathcal{D}$ spans $\bigwedge^k \mathbb{R}^d$. Therefore U contains a basis of $\bigwedge^k \mathbb{R}^d$. □

In the next two propositions we consider the representation $\delta : \mathrm{Sl}(\mathbb{R}^d) \to \mathrm{Gl}(\bigwedge^k \mathbb{R}^d)$ where

$$\delta(g)(u_1 \wedge \cdots \wedge u_k) := gu_1 \wedge \cdots \wedge gu_k.$$

To abbreviate, we denote $\delta(g)$ simply by g.

Now we can prove that an invariant cone contains a decomposable element.

Proposition 4.2.4. *Take $S \subset \mathrm{Sl}(\mathbb{R}^d)$ a semigroup with nonempty interior. Let $\{0\} \neq W \subset \bigwedge^k \mathbb{R}^d$ be an S-invariant cone. Then W intercepts a non-null decomposable element of $\bigwedge^k \mathbb{R}^d$.*

Proof. Take $h \in \mathrm{reg}(S)$ and consider as before the basis $\mathcal{B} = \{e_1, \cdots, e_d\}$ of $\mathbb{R}^d$ such that $h = \mathrm{diag}(\lambda_1, \cdots, \lambda_d)$ with $\lambda_1 > \cdots > \lambda_d > 0$. Note that for $I = \{i_1, \ldots, i_k\} \in \mathcal{F}_k(d)$, the vectors $e_I = e_{i_1} \wedge \cdots \wedge e_{i_k} \in \bigwedge^k \mathbb{R}^d$ are eigenvectors of h, with eigenvalues $\lambda_{i_1} \cdots \lambda_{i_k}$. Moreover, they form a basis of $\bigwedge^k \mathbb{R}^d$.

Define the following order relation on $\mathcal{F}_k(d)$: given $I = \{i_1, \ldots, i_k\}$ and $J = \{j_1, \ldots, j_k\}$ in $\mathcal{F}_k(d)$,

$$I \prec J \text{ if } \lambda_{i_1} \cdots \lambda_{i_k} < \lambda_{j_1} \cdots \lambda_{j_k}.$$

If necessary, we take a perturbation of $h = \mathrm{diag}(\lambda_1, \ldots, \lambda_d) \in \mathrm{int}S$ such that $\prec$ become a total order. Consider $0 \neq v \in W$ with $v = \sum_{I \in \mathcal{F}_k(d)} \alpha_I e_I$ and define

$$J_0 = \{j_1, \ldots, j_k\} = \max\{I \in \mathcal{F}_k(d);\ \alpha_I \neq 0\}.$$

As W is S-invariant, we have that W is invariant under h. Hence,

$$\left(\frac{h^m(v)}{(\lambda_{j_1}\cdots\lambda_{j_k})^m}\right)_{m\in\mathbb{N}}$$

is a sequence in W and

$$\frac{h^m(v)}{(\lambda_{j_1}\cdots\lambda_{j_k})^m} = \sum_{I\in\mathcal{F}_k(d)} \alpha_I \frac{h^m(e_I)}{(\lambda_{j_1}\cdots\lambda_{j_k})^m} = \sum_{I\in\mathcal{F}_k(d)} \alpha_I \frac{(\lambda_{i_1}\cdots\lambda_{i_k})^m}{(\lambda_{j_1}\cdots\lambda_{j_k})^m} e_I,$$

for all $m \in \mathbb{N}$. Note that if $I = \{i_1, \cdots, i_k\} \notin \{I \in \mathcal{F}_k(d);\ \alpha_I \neq 0\}$, then $\alpha_I \frac{(\lambda_{i_1}\cdots\lambda_{i_k})^m}{(\lambda_{j_1}\cdots\lambda_{j_k})^m} = 0$, for all $m \in \mathbb{N}$. Moreover $\lambda_{j_1}\cdots\lambda_{j_k} > \lambda_{i_1}\cdots\lambda_{i_k}$ for all $\{i_1, \ldots, i_k\}$ in $\{I \in \mathcal{F}_k(d);\ \alpha_I \neq 0\}\backslash\{J_0\}$. Hence,

$$\lim_{m\to\infty} \frac{(\lambda_{i_1}\cdots\lambda_{i_k})^m}{(\lambda_{j_1}\cdots\lambda_{j_k})^m} = 0.$$

Therefore

$$\lim_{m\to\infty} \frac{h^m(v)}{(\lambda_{j_1}\cdots\lambda_{j_k})^m} = \lim_{m\to\infty} \sum_{I\in\mathcal{F}_k(d)} \alpha_I \frac{(\lambda_{i_1}\cdots\lambda_{i_k})^m}{(\lambda_{j_1}\cdots\lambda_{j_k})^m} e_I = \alpha_{J_0} e_{J_0}$$

The closeness of W implies that the decomposable element $\alpha_{J_0} e_{J_0}$ belongs to W, and moreover, this element is non-null. □

Hence we have the main result of this section.

Proposition 4.2.5. *Let $S \subset \mathrm{Sl}(\mathbb{R}^d)$ be a semigroup with non empty interior. If $\{0\} \neq W \subset \bigwedge^k \mathbb{R}^d$ is a S-invariant cone, then W is pointed and generating.*

Proof. First recall that the representation of $\mathrm{Sl}(\mathbb{R}^d)$ on $\bigwedge^k \mathbb{R}^d$ is irreducible.

Now, define $H = W \cap -W$. Then H is an S-invariant vector subspace. We have also that H is S^{-1}-invariant, because if $g \in S$, then $gH \subset H$. Since g is invertible, gH is a subspace of H with $\dim gH = \dim H$, i.e., $gH = H$. Consequently, $H = g^{-1}H$. The fact that int$S \neq \emptyset$ implies that $\mathrm{Sl}(\mathbb{R}^d)$ is generated by $S \cup S^{-1}$. Hence H is $\mathrm{Sl}(\mathbb{R}^d)$-invariant, now knowing that W is proper and $\mathrm{Sl}(\mathbb{R}^d)$ is irreducible we have that $H = \{0\}$. Hence W is pointed.

Finally, assume that int$W = \emptyset$. By Lemma 4.2.2, $W \cup -W$ is contained in a proper subspace V of $\bigwedge^k \mathbb{R}^d$. Consider a decomposable element $x \in W$ and take $\rho_k^q : \mathrm{Sl}(\mathbb{R}^d) \to \mathbb{G}_k(d)$ the open map $\rho_k^q(g) = [gq]$ where $[q] := \varphi^{-1}(\pi(x))$ and φ is the Plücker embedding defined in the second section. Then $\varphi(\rho_k^q(\text{int}S))$ is open in $\varphi(\mathbb{G}_k(d))$, that is, there exists an open set $B \subset \mathbb{P}\left(\bigwedge^k \mathbb{R}^d\right)$ such that

$$\varphi(\rho_k^q(\text{int}S)) = B \cap \varphi(\mathbb{G}_k(d)).$$

Knowing that

$$\pi^{-1}(\varphi(\phi(\text{int}S))) = \pi^{-1}(B \cap \varphi(\mathbb{G}_k(d))) = \pi^{-1}(B) \cap \mathcal{D},$$

we have that $\pi^{-1}(\varphi(\phi(\text{int}S)))$ is open in $\mathcal{D}$. By Lemma 4.2.3, $\pi^{-1}(\varphi(\phi(\text{int}S)))$ contains a basis of $\bigwedge^k \mathbb{R}^d$. Note also that

$$\pi^{-1}(\varphi(\phi(\text{int}S))) = \pi^{-1}(\varphi((\text{int}S)[q])) = \pi^{-1}(\pi((\text{int}S)x)) = \pi^{-1}((\text{int}S)\pi(x)).$$

So, if $y \in \pi^{-1}(\varphi(\phi(\text{int}S)))$, then $\pi(y) \in (\text{int}S)\pi(x)$, hence there is $g \in \text{int}S$ with $\pi(y) = g\pi(x) = \pi(gx)$, that is, $y = \alpha gx$ for some $\alpha \neq 0$. If $\alpha > 0$, then $y \in \alpha Sx \subset \alpha W = W$ and if $\alpha < 0$, then $y \in \alpha Sx \subset \alpha W = -W$. Anyway $y \in W \cup -W$ and we conclude that $\pi^{-1}(\varphi(\phi(\text{int}S))) \subset W \cup -W$. But it is a contradiction, because $\pi^{-1}(\varphi(\phi(\text{int}S)))$ is contained in the proper subspace V of $\bigwedge^k \mathbb{R}^d$ and contains a basis of $\bigwedge^k \mathbb{R}^d$. Therefore, $\text{int}W \neq \emptyset$. □

4.3 Cones, flag type and controllability

In this section we prove that there exists an S-invariant cone in $\bigwedge^k \mathbb{R}^d$ if and only if the flag type of S contains k. Consequently we have the main result of this section, Theorem 4.3.5, that gives a necessary and sufficient condition for the equality $S = \text{Sl}(\mathbb{R}^d)$ in terms of the existence of S-invariant cones in the spaces $\bigwedge^k \mathbb{R}^d$, $k \in \{1, \dots, d-1\}$.

Theorem 4.3.1. *Let $S \subset \text{Sl}(\mathbb{R}^d)$ be a connected semigroup with flag type given by $\Theta(S)$. If $k \in \Theta(S)$, then there exists an S-invariant cone $\{0\} \neq W \subset \bigwedge^k \mathbb{R}^d$.*

Proof. Take $h \in \text{reg}(S)$ and consider $\mathcal{B}(h) = \{e_1, \dots, e_d\}$ the special basis of $\mathbb{R}^d$. We saw in the second section that $b_{\{k\}}(h) = (\text{span}\{e_1, \ldots, e_k\})$ and the orbit

$$N_{\mathcal{B}(h)} b_{\{k\}}(h) = \left\{ \begin{bmatrix} I_k \\ X \end{bmatrix} ;\ X \in \mathbb{R}^{(d-k)\times k} \right\}$$

contains C_k. Note that $\varphi(N_{\mathcal{B}(h)} b_{\{k\}}(h)) \subset \pi(M)$, where M is the affine subspace

$$M = \left\{(1, x_2, \cdots, x_{\binom{d}{k}});\ x_2, \cdots, x_{\binom{d}{k}} \in \mathbb{R}\right\} \subset \bigwedge^k \mathbb{R}^d$$

in the basis $\{e_I;\ I \in \mathcal{F}_k(d)\}$. Since the invariant control set $C_k \subset \mathbb{G}_k(d)$ is contained in $N_{\mathcal{B}(h)} b_{\{k\}}(h)$, we have

$$\varphi(C_k) \subset \varphi(N_{\mathcal{B}(h)} b_{\{k\}}(h)) \subset \pi(M).$$

Define $M_1 := \pi^{-1}(\varphi(C_k)) \cap M$. Let W be the cone generated by M_1, W is clearly non-null.

Now, we show that W is S-invariant. Since C_k is S-invariant, it follows that $\varphi(C_k)$ is S-invariant. We claim that $(\mathbb{R}\backslash\{0\})M_1$ is S-invariant. In fact, given $\alpha \in \mathbb{R}\backslash\{0\}$, $u_1 \wedge \cdots \wedge u_k \in M_1$ and $g \in S$, we have that $\pi(g(\alpha\, u_1 \wedge \cdots \wedge u_k)) = \pi(g(u_1 \wedge \cdots \wedge u_k))$ is contained in $\pi(gM_1) = g\pi(M_1) = g\varphi(C_k) \subset \varphi(C_k)$, due to the equality $\pi(M_1) = \varphi(C_k)$ and the S-invariance of $\varphi(C_k)$. Hence knowing that $\pi|_M$ is injective, we conclude the claim. As S is connected this implies that C_k, $\varphi(C_k)$ and M_1 are connected.

Furthermore, since for every $x \in \left(\bigwedge^k \mathbb{R}^d\right) \backslash\{0\}$ the mapping $g \in S \mapsto gx \in \bigwedge^k \mathbb{R}^d$ is continuous, we conclude that S leaves invariant the connected components of $(\mathbb{R}\backslash\{0\})M_1$. As $(\mathbb{R}^+)M_1$ is one of these components, $(\mathbb{R}^+)M_1$ is invariant, implying that its convex closure W is S-invariant. □

Remark 4.3.2. *Our result generalizes Theorem 4.2 in [5] and also improves its hypotheses in the sense that we do not need to have the identity in* $\text{cl}S$. *In [5] the authors assume* $1 \in S$ *to guarantee that* S *leaves invariant the connected components of* $(\mathbb{R}\backslash\{0\})M_1$, *but we can show that this is not necessary. In fact, let* $g \in S$, *then* g *leaves* $(\mathbb{R}\backslash\{0\})M_1$ *invariant. So* g *is a bijection between the connected components of* $(\mathbb{R}\backslash\{0\})M_1$. *Denote by* $M_1^+ = (\mathbb{R}^+)M_1$ *and* $M_1^- = (\mathbb{R}^-)M_1$ *these connected components. Suppose that there is an element* $g \in S$ *which does not leave* M_1^+ *invariant. Then* $g(M_1^+) = M_1^-$ *and* $g(M_1^-) = M_1^+$. *Hence we have another element in* S, g^2, *that leaves invariant the components, but this contradicts the connectedness* *of S.*

We also note that by Proposition 4.2.5, the cone W, in the above theorem is pointed and generating.

The following results prove that the existence of a pointed invariant cone in $\bigwedge^k \mathbb{R}^d$ implies that the flag type of the semigroup contains k.

Lemma 4.3.3. *Assume that $k \notin \Theta(S)$. Let C_k be the invariant control set for the action of S on $\mathbb{G}_k(d)$. Then there is a two-dimensional subspace $V \subset \bigwedge^k \mathbb{R}^d$ such that $\pi(V) \subset \varphi(C_k)$.*

Proof. Denote by $\pi_k : \mathbb{F} \to \mathbb{G}_k(d)$ the natural projection and consider $[p] \in C_k$. Let f be an element of the invariant control set C of the full flag $\mathbb{F}$ with $\pi_k(f) = [p]$. Such element exists because $C_k = \pi_k(C)$. Let $\Theta(S) = \{r_1, \dots, r_n\}$ be the flag type of S and observe that $\pi^{-1}_{\Theta(S)}(\pi_{\Theta(S)}(f))$ is a subset of C, where $\pi_{\Theta(S)} : \mathbb{F} \to \mathbb{F}_{\Theta(S)}$. Therefore, $\pi_k(\pi^{-1}_{\Theta(S)}(\pi_{\Theta(S)}(f))) \subset C_k$. Since $k \notin \Theta(S)$ then $\pi_{\Theta(S)}(f) = (V_1 \subset \cdots \subset V_n)$ with $\dim V_i = r_i$, $1 \leq i \leq n$. We have the following cases:

Case 1: Assume that $r_1 < k < r_n$. In this case, there exists $l \in \{1, \dots, n-1\}$ such that the elements of $\pi_k(\pi^{-1}_{\Theta(S)}(\pi_{\Theta(S)}(f)))$ are the k-subspaces that contain V_l and are contained in V_{l+1}. Let $\{v_1, \cdots, v_{r_l}\}$ be a basis of V_l, and complete it to an ordered basis $\{v_1, \cdots, v_{r_l}, v_{r_l+1}, \cdots, v_{r_{l+1}}\}$ of V_{l+1}. Since $r_l < k$ and $r_{l+1} > k$, consider the element v_k in this basis of V_{l+1} and, moreover, there is a basic element v_j with $k < j \leq r_{l+1}$. In this way, define the subspace

$$V = \{v_1 \wedge \cdots \wedge v_{r_l} \wedge \cdots \wedge v_{k-1} \wedge (\alpha v_k + \beta v_j);\ \alpha, \beta \in \mathbb{R}\}.$$

Case 2: Now, suppose that $k < r_1$. Here, the elements of $\pi_k(\pi^{-1}_{\Theta(S)}(\pi_{\Theta(S)}(f)))$ are the k-subspaces contained in V_1. Since $k \geq 1$, then $r_1 \geq 2$. Hence, given an ordered basis $\{v_1, \cdots, v_{r_1}\}$ of V_1, we can find $v_k, v_j \in \{v_1, \cdots, v_{r_1}\}$ where j satisfies $k < j \leq r_1$. Consider the subspace

$$V = \{v_1 \wedge \cdots \wedge v_{k-1} \wedge (\alpha v_k + \beta v_j);\ \alpha, \beta \in \mathbb{R}\}. \tag{4.3-2}$$

Case 3: Finally, assume $k > r_n$. Hence, $\pi_k(\pi^{-1}_{\Theta(S)}(\pi_{\Theta(S)}(f)))$ is the set formed by the

k-subspaces which contains V_n. Since $k \leq d-1$, we can consider a basis $\{v_1, \ldots, v_{r_n}\}$ of V_{r_n} and complete it to obtain the ordered basis $\{v_1, \ldots, v_{r_n}, v_{r_n+1}, \ldots, v_d\}$ of $\mathbb{R}^d$. In this case, we can also take v_k and v_j in this basis, with $k < j \leq d$ and consider the subspace defined as in (4.3-2).

In the three cases, the subspace $V \subset \bigwedge^k \mathbb{R}^d$ is two-dimensional and satisfies

$$\pi(V) \subset \varphi(\pi_k(\pi^{-1}_{\Theta(S)}(\pi_{\Theta(S)}(f)))) \subset \varphi(C_k).$$

□

The following theorem is a reciprocal of Theorem 4.3.1.

Theorem 4.3.4. *If* $\{0\} \neq W \subset \bigwedge^k \mathbb{R}^d$ *is an* S*-invariant cone, then* $k \in \Theta(S)$*.*

Proof. Assume that $k \notin \Theta(S)$ and denote by L the intersection of W with the set $\mathcal{D}$ of the decomposable elements of $\bigwedge^k \mathbb{R}^d$. By Proposition 4.2.4 we have that L is nonempty. Moreover, L is S-invariant, since the set of decomposable elements is also S-invariant. Therefore, $\varphi^{-1}(\pi(L))$ is also invariant. As W is a closed set then L is closed in $\mathcal{D}$ and hence $\varphi^{-1}(\pi(L))$ is a closed set in $\mathbb{G}_k(d)$. Since $\mathbb{G}_k(d)$ is compact, $\varphi^{-1}(\pi(L))$ is also compact, then there is an invariant control set contained in $\varphi^{-1}(\pi(L))$. But there is only one invariant control set $C_k \subset \mathbb{G}_k(d)$ implying that $C_k \subset \varphi^{-1}(\pi(L))$ and hence $\pi^{-1}(\varphi(C_k)) \subset L \subset W$. As proved in Lemma 4.3.3, there is a two-dimensional subspace V such that $\pi(V) \subset \varphi(C_k)$. But this means that $V \subset \pi^{-1}(\varphi(C_k)) \subset W$, which is a contradiction because W is pointed (see Proposition 4.2.5). □

Recall that if $S \subset \mathrm{Sl}(\mathbb{R}^d)$ is a nonempty semigroup, then S is transitive on $\mathbb{R}^d \backslash \{0\}$ if and only if $S = \mathrm{Sl}(\mathbb{R}^d)$ (see [5]). In this context, the next theorem gives a necessary and sufficient condition in terms of the existence of invariant cones.

Theorem 4.3.5. *Let* $S \subset \mathrm{Sl}(\mathbb{R}^d)$ *be a connected semigroup with nonempty interior. Then* $S = \mathrm{Sl}(\mathbb{R}^d)$ *if and only if there are no* S*-invariant cones in* $\bigwedge^k \mathbb{R}^d$*, for all* $k \in \{1, \ldots, d-1\}$*.*

Proof. Let $W \subset \bigwedge^k \mathbb{R}^d$ be a proper S-invariant cone, for some $k \in \{1, \ldots, d-1\}$. Note that W does not contain $\mathcal{D}$, otherwise the convexity of W would imply that the convex closure of $\mathcal{D}$, $\bigwedge^k \mathbb{R}^d$, would be contained in W, which would contradicts the fact that W is proper.

By Proposition 4.2.4 we can consider an element $v_1 \in W \cap \mathcal{D}$. Take $v_2 \in \mathcal{D}\backslash W$. If $S = \mathrm{Sl}(\mathbb{R}^d)$ and knowing that $\mathcal{D}$ is S-invariant then there exists $g \in S$ such that $gv_1 = v_2 \notin W$, but this contradicts the S-invariance of W. Hence $S \neq \mathrm{Sl}(\mathbb{R}^d)$.

On the other hand, assume that $S \subset \mathrm{Sl}(\mathbb{R}^d)$ is proper. Then $\Theta(S) \neq \emptyset$, hence there exists $k \in \Theta(S)$, for some $k \in \{1, \ldots, d-1\}$. Therefore, Theorem 4.3.1 implies the existence of a such cone. □

Remark 4.3.6. *This theorem complement and improve Section 7 of [5].*

The next example shows that, as we commented before, the connectedness of S is fundamental in the previous results.

Example 4.3.7. *Let $S^+ \subset \mathrm{Sl}(\mathbb{R}^2)$ be the set of matrices with positive entries. It is not difficult to show that that S^+ is a proper semigroup with nonempty interior in $\mathrm{Sl}(\mathbb{R}^2)$, the positive orthant $Q^+ = \{(a,b) \in \mathbb{R};a,b \geq 0\}$ is S^+-invariant and S^+ is a open set. Now take the following proper semigroup*

$$S = S^+ \cup (-S^+) = (-1)^{\mathbb{Z}} S^+ = \{(-1)^k A; k \in \mathbb{Z}, A \in S^+\}.$$

Note that S has nonempty interior. Moreover, S is not transitive on $\mathbb{R}^2$ because it leaves invariant the double cone $Q^+ \cup -Q^+ = (-1)^{\mathbb{N}} Q^+$:

$$S((-1)^{\mathbb{N}} Q^+) = (-1)^{\mathbb{N}} S^+ (-1)^{\mathbb{N}} Q^+ = (-1)^{\mathbb{N}+\mathbb{N}} S^+ Q^+ = (-1)^{\mathbb{N}} Q^+.$$

However, S does not leave invariant proper cones in $\mathbb{R}^2 = \bigwedge^1 \mathbb{R}^2$. In fact, we have that $-I \in S$, therefore, if C is a proper S invariant cone then $-I(C) = -C \subset C$. This implies that C is a subspace, which is a contradiction.

As a consequence of the above results, we get a necessary and sufficient condition for controllability of

$$\dot{x} = Ax + uBx, x \in \mathbb{R}^d \setminus \{0\}, u \in \mathbb{R},$$

with $A, B \in \mathfrak{sl}(\mathbb{R}^d)$.

Recall that the system semigroup

$$S = \{e^{t_1(A+u_1 B)} \cdots e^{t_n(A+u_n B)};\ t_1, \ldots, t_n \geq 0,\ u_1, \ldots, u_n \in \mathbb{R},\ n \in \mathbb{N}\}$$

is a semigroup of $\mathrm{Sl}(\mathbb{R}^d)$. Moreover, if the Lie algebra, generated by A and B, coincides with $\mathfrak{sl}(\mathbb{R}^d)$, then int$S \neq \emptyset$. Furthermore, S is path connected. It is well know that this system is controllable if, and only if, $S = \mathrm{Sl}(\mathbb{R}^d)$ (see e.g. [5]). Hence, as a result of Theorem 4.3.5 we have the necessary and sufficient condition for controllability of this bilinear system.

Theorem 4.3.8. *The above system is controllable if and only if it does not leave invariant a cone in $\bigwedge^k \mathbb{R}^d$, for all $k \in \{1, \ldots, d-1\}$.*

4.4 Flag type and invariance of convex sets

In this section, we generalize the previous one. Or rather, instead of proper cones, we study the existence of proper convex sets in $\bigwedge^k \mathbb{R}$ which are invariant by the action of a semigroup $S \subset \mathrm{Sl}(\mathbb{R}^d)$. We also relate the existence of this convex sets with the flag type $\Theta(S)$ of S.

Initially, given $h \in \mathrm{reg}(S)$, take as before the basis $\mathcal{B}(h) = \{e_1, \ldots, e_d\}$ of $\mathbb{R}^d$. Since $1 = \det(h) = \lambda_1 \cdots \lambda_d$, then for all $k \in \{1, \ldots, d-1\}$, we can prove that $\lambda_1 \cdots \lambda_k > 1$.

The following lemma gives an expression for the closed convex cone generated by a convex set in $\bigwedge^k \mathbb{R}^d$.

Lemma 4.4.1. *If the set $K \subset \bigwedge^k \mathbb{R}^d$ is convex, then the closed convex cone W generated by K is*

$$W := \mathrm{cl}(\bigcup_{\alpha > 0} \alpha K).$$

Proof. Let $\{W_l\}_{l \in \Lambda}$ be the family of all closed cones that contains K and consider $V := \bigcap_{l \in \Lambda} W_l$ the closed convex cone generated by K.

Note that W is a closed cone which contains K. To show that W is convex, take $x, y \in W$. There are sequences $(\gamma_n x_n), (\delta_n y_n)$ in $\bigcup_{\alpha > 0} \alpha K$ with $\gamma_n, \delta_n > 0$ and $x_n, y_n \in K$ (for all $n \in \mathbb{N}$) converging to x and y respectively. Take $t \in [0,1]$ and define

$$z_n = \left(\frac{(1-t)\gamma_n}{(1-t)\gamma_n + t\delta_n} \right) x_n + \left(\frac{t\delta_n}{(1-t)\gamma_n + t\delta_n} \right) y_n, \ n \in \mathbb{N}.$$

Note that (z_n) is a sequence in K, then $(((1-t)\gamma_n + t\delta_n)\, z_n)$ is a sequence in $\bigcup_{\alpha > 0} \alpha K$, since $(1-t)\gamma_n + t\delta_n > 0$. But $((1-t)\gamma_n + t\delta_n)\, z_n = (1-t)\gamma_n x_n + t\delta_n y_n$ converges to

$(1-t)x+ty$, hence $(1-t)x+ty \in W$. Therefore $V \subset W$.

On the other hand, for each $\gamma > 0$ we have $\gamma K \subset W_l$, for all $l \in \Lambda$, then $\bigcup_{\gamma>0} \gamma K \subset W_l$, for all $l \in \Lambda$. Hence the closeness of each W_l implies that $W \subset W_l$, for all $l \in \Lambda$, so $W \subset V$. □

Proposition 4.4.2. *Let $K \subset \bigwedge^k \mathbb{R}^d$ be a proper S-invariant convex set. Then the closed cone generated by K is S-invariant.*

Proof. Denote by W the closed cone generated by K. Since K is S-invariant, for each $g \in S$ it holds that $gK \subset K$. Hence

$$gW = g\left(\mathrm{cl}(\bigcup_{\alpha>0} \alpha K)\right) \subset \mathrm{cl}(g\left(\bigcup_{\alpha>0} \alpha K\right)) = \mathrm{cl}(\bigcup_{\alpha>0} \alpha g K) \subset \mathrm{cl}(\bigcup_{\alpha>0} \alpha K) = W,$$

that is, W is S-invariant. □

Proposition 4.4.3. *If $K \subset \bigwedge^k \mathbb{R}^d$ is a proper S-invariant convex set, then $0 \notin \mathrm{int}K$.*

Proof. For each $h \in \mathrm{reg}(S)$ denote by $b_k(h)$ the attractor of h in G_k. The set of transitivity of C_k, C_k^0, satisfies

$$C_k^0 = \{b_k(h); h \in \mathrm{reg}(S)\},$$

and has nonempty interior. In particular, there is an open set $V \subset C_k^0 = \{b_k(h); h \in \mathrm{reg}(S)\}$. As a consequence, $\phi(V)$ is an open set in $\mathcal{D}$, and therefore, by Lemma 4.2.3, $\pi^{-1}(\phi(V))$ contains a basis $\{b_1, b_2, \ldots, b_n\}$ of the exterior space. Since $b_i \in \pi^{-1}(\phi(V))$ and V is a subset of $C_k^0 = \{b_k(h); h \in \mathrm{reg}(S)\}$, then, for each b_i exists $h_i \in \mathrm{reg}(S)$ such that $b_i \in \pi^{-1}(\phi(b_k(h_i)))$ or, equivalently, there is a basis $\{e_1(h_i), e_2(h_i), \ldots, e_d(h_i)\}$ of $\mathbb{R}^d$ where h_i is written as $\mathrm{diag}(\lambda_{1i}, \lambda_{2i}, \ldots, \lambda_{di})$ and $b_i = e_1(h_i) \wedge e_2(h_i) \wedge \cdots \wedge e_k(h_i) = e_I(h_i)$ with $I = \{1, \ldots, k\}$. So, if we suppose that $0 \in \mathrm{int}K$, then there are $\alpha \neq 0$ and $h_1, \ldots, h_r \in \mathrm{reg}(S)$ with $r = \binom{d}{k}$, such that $\alpha e_I(h_i)$ is a basis of $\bigwedge^k \mathbb{R}^d$ with $\pm\alpha e_I(h_i)$ contained in $\mathrm{int}K$, $i = 1, \ldots, r$.

But for all $m \in \mathbb{N}$ and $i \in \{1, \ldots, r\}$ we have $h_i^m(\pm\alpha e_I(h_i)) \in K$ due to S-invariance of K. Moreover,

$$\|h_i^m(\pm\alpha e_I(h_i))\| = |\alpha|(\lambda_{1i} \cdots \lambda_{ki})^m \|e_I(h_i)\| \to +\infty,$$

then the convexity of K implies that $K = \bigwedge^k \mathbb{R}^d$. □

The above proposition has the following consequence.

Corollary 4.4.4. *Let $K \subset \bigwedge^k \mathbb{R}^d$ be an S-invariant convex set and denote by W the closed cone generated by K. The following statements are equivalents:*

i) W is proper;

ii) K is proper.

iii) $0 \notin \mathrm{int}K$.

Proof. The implication $(i) \Rightarrow (ii)$ holds because $K \subset W$. Moreover, $(ii) \Rightarrow (iii)$ follows by Proposition 4.4.3. Finally, to prove that $(iii) \Rightarrow (i)$ we first note that if $W = \bigwedge^k \mathbb{R}^d$ then int$K \neq \emptyset$. In fact, if int$K = \emptyset$ then K is contained in a proper affine subspace $V + u_0$, where $V \subset \bigwedge^k \mathbb{R}^d$ is a proper vector subspace and $u_0 \in \bigwedge^k \mathbb{R}^d$. Hence

$$\begin{aligned}\bigwedge^k \mathbb{R}^d &= W = \mathrm{cl}(\bigcup_{\alpha>0} \alpha K) \subset \mathrm{cl}(\bigcup_{\alpha>0} \alpha(V + u_0)) = \mathrm{cl}(\bigcup_{\alpha>0}(V + \alpha u_0)) \\ &= V + [0, +\infty)u_0 \neq \bigwedge^k \mathbb{R}^d.\end{aligned}$$

which is a contradiction. Hence, given the open set $-$int$K$, there are $\alpha > 0$ and $k \in K$ with $\alpha k \in -$int$K$, that is, $-\alpha k \in$ int$K$. Since K is convex, the line $[-\alpha k, k) := \{(t - 1)\alpha k + tk;\ t \in [0, 1)\}$ is contained in int$K$, therefore $0 \in$ int$K$. □

The next result presents a synthesis of this section, the relation among invariant convex set, invariant cone and flag type.

Theorem 4.4.5. *Let $S \subset \mathrm{Sl}(\mathbb{R}^d)$ a semigroup with nonempty interior. Then the following statements are equivalents:*

i) There exists an S-invariant proper convex set in $\bigwedge^k \mathbb{R}^d$;

ii) There exists an S-invariant proper closed cone in $\bigwedge^k \mathbb{R}^d$;

iii) $k \in \Theta(S)$.

Proof. By Proposition 4.4.2 and Corollary 4.4.4 we have that $(i) \Rightarrow (ii)$. By Theorem 4.3.4 it follows that (ii) implies (iii). Moreover, since a cone is a convex set, the implication $(iii) \Rightarrow (i)$ follows by Theorem 4.3.1. □

4.5 Examples

In order to present examples to illustrate our results, we create a computational implementation in Julia Language [4] called `LieAlgebraRankCondition.jl`[1]. The basic idea of this implementation is the following: given the bilinear control system

$$\dot{x} = Ax + uBx, x \in \mathbb{R}^4 \setminus \{0\}, u \in \mathbb{R} \text{ and } A, B \in \mathfrak{sl}(\mathbb{R}^4)$$

put the Lie brackets in a convenient way and analyse all the possibilities until get, if possible, a linearly independent (L.I.) set for $\mathfrak{sl}(\mathbb{R}^4)$. In the following we describe a conceptual algorithm.

[1]Available in `https://github.com/evcastelani/LieAlgebraRankCondition.jl`

Algorithm 1: Lie Algebra Rank Condition Algorithm.

Data: A: Array, B: Array, `dim`: dimension of $\mathfrak{sl}(\mathbb{R}^4)$

Result: `True`: a set of L. I. arrays were found; `False`: Does not exists an L. I. set of arrays.

```
C ← {A, B, [A, B]};
if C is L. I. then
    k ← 3;
else
    return False;
end
while k ≤ dim do
    j ← k − 1;
    C_trial ← C_j;
    while (C ∪ [C_trial, C_k] is not L.I) and (j > 3) do
        j ← j − 1 ;
        C_trial ← C_j;
    end
    if j=3 then
        remove C_k from C ;
        k ← k − 1;
    else
        add [C_trial, C_k] to C;
        k ← k + 1;
    end
    if k=3 then
        return False;
    end
end
return True;
```

Remark 4.5.1. *The parameter* `dim` *can be changed in order to find solutions for higher order spaces.*

Example 4.5.2. *Consider the bilinear system*

$$(\Sigma)\quad \dot{x} = Ax + uBx,\ \textit{with } x \in \mathbb{R}^4 \setminus \{0\}, u \in \mathbb{R},$$

$$A = \begin{bmatrix} 0 & 2 & 0 & -1 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ -1 & 0 & 2 & 0 \end{bmatrix} \textit{ and } B = \mathrm{diag}(4, 1, -2, -3) \in \mathfrak{sl}(\mathbb{R}^4).$$

The matrix A has the distinct eigenvalues, $3, 2, -2, -3$, with the following eigenvectors $v_1 = (1, 2, 2, 1)$, $v_2 = (-2, -1, 1, 2)$, $v_3 = (2, -1, -1, 2)$ and $v_4 = (-1, 2, -2, 1)$, respectively. Let S be the semigroup of (Σ), that is,

$$S = \{e^{t_1(A+u_1B)} \cdots e^{t_k(A+u_nB)}; t_1, \ldots, t_n \geq 0,\ n \in \mathbb{N}\}.$$

Using the implementation of Algorithm 1, we can show that this system satisfies the Lie algebra rank condition, hence S has nonempty interior in $\mathrm{Sl}(\mathbb{R}^4)$. Moreover, S is a proper semigroup. In fact, by [15, Proposition 2], we have

$$A + uB \in \mathcal{L}(S_2) = \{X \in \mathfrak{sl}(\mathbb{R}^4);\ \exp(X) \in S_2\},$$

where $S_2 = \{g \in \mathrm{Sl}(\mathbb{R}^4);\ g\mathcal{O}_2 \subset \mathcal{O}_2\}$ is the the compression semigroup of the positive orthant $\mathcal{O}_2 = \left\{ \sum_{I=\{i_1<i_2\}\subset\{1,2,3,4\}} \alpha_I e_I;\ \alpha_I \geq 0 \right\} \subset \bigwedge^2 \mathbb{R}^4$. This semigroup coincides with the set of all matrix in $\mathrm{Sl}(\mathbb{R}^4)$ such that the minors of order 2 have non-negative determinant. Note that $S \subset S_2$. Since S_2 leaves invariant the cone $\mathcal{O}_2$, then $S\mathcal{O}_2 \subset \mathcal{O}_2$. Hence (Σ) is not controllable and therefore S is proper, in particular S leaves invariant the positive orthant of $\bigwedge^2 \mathbb{R}^4$.

On the other hand, neither $\pm A$ nor $\pm(A + uB)$ leave invariant an orthant of $\mathbb{R}^4$. In fact, by [10, Lemma 1], a matrix $X = (x_{ij})$ leaves invariant the orthant with signs $(\sigma_1, \ldots, \sigma_d)$ if and only if $\sigma_i\sigma_j x_{ij} > 0$. Applying this condition to $\pm A$, $\pm(A + uB)$, we get the contradictory fact that $\sigma_1\sigma_4$ must be simultaneously 1 and -1, so that there are no invariant orthants in $\mathbb{R}^4 = \bigwedge^1 \mathbb{R}^4$. The system (Σ) is a counter-example for the following conjecture proposed by Sachkov in [10]. Consider a bilinear control system with A symmetric and $B = \mathrm{diag}(b_1, \ldots, b_n)$ where $b_i \neq b_j$ for $i \neq j$. Is it true that if this system has no invariant orthants and everywhere satisfies the necessary Lie algebra rank controllability condition, then it is controllable in $\mathbb{R}^d \setminus \{0\}$?

Now we prove that, although (Σ) *is not controllable, there are no* S*-invariant cones in* $\mathbb{R}^4 = \bigwedge^1 \mathbb{R}^4$ *neither in* $\bigwedge^3 \mathbb{R}^4$*. Suppose that* $W \subset \mathbb{R}^4$ *is an* S*-invariant cone. Then* W *has nonempty interior and it is not contained in the plane generated by* $\{e_2, e_3, e_4\}$*. Therefore there is a vector* $w = (w_1, w_2, w_3, w_4) \in W$ *such that* $w_1 \neq 0$*. Since* $\mathrm{e}^{tB} \in \mathrm{cl(S)}$ *for all* $t \in \mathbb{R}$*, then if* $w_1 > 0$ *we have that*

$$\lim_{t\to+\infty} \frac{e^{tB}w}{\|e^{tB}w\|} = e_1 \in W.$$

If $w_1 < 0$ *then*

$$\lim_{t\to+\infty} \frac{e^{tB}w}{\|e^{tB}w\|} = -e_1 \in W.$$

Without loss of generality, assume that $e_1 \in W$*. Knowing that* v_1 *is the attractor eigenvalue of* A *and considering the basis* $\{v_1, v_2, v_3, v_4\}$*, a similar argument assures that either* $v_1 \in V$ *or* $-v_1 \in V$*.*

Let $H := \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_4 < 0\}$*, then, for all* $x \in H$*,*

$$\lim_{t\to+\infty} \frac{e^{t(-B)}x}{||e^{t(-B)}x||} = -e_4.$$

In particular, note that if $W \cap H \neq \emptyset$*, then* $-e_4 \in W$*. Now we show that* $W \cap H \neq \emptyset$*. Since the inner product between* Ae_1 *and* e_4 *is negative, then the curve* $t \mapsto e^{tA}e_1$ *intersects* H *for* $t > 0$*. By* S*-invariance and knowing that* $e_1 \in W$*, we have* $\mathrm{e}^{\mathbb{R}_+ A} e_1 \subset W$ *, then* $W \cap H \neq \emptyset$*.*

As stated early, either $v_1 \in W$ *or* $-v_1 \in W$*. As* v_1 *has a positive fourth coordinate, then*

$$\lim_{t\to+\infty} \frac{e^{t(-B)}v_1}{||e^{t(-B)}v_1||} = e_4,$$

and as $-v_1$ *has a negative first coordinate, we have*

$$\lim_{t\to+\infty} \frac{e^{tB}(-v_1)}{||e^{tB}(-v_1)||} = -e_1.$$

Hence if $v_1 \in W$ *then* $e_4 \in W$*. But* $-e_4$ *is also in* W*, then* W *is not pointed. On the other hand, if* $-v_1 \in W$*, then* $-e_1 \in W$*. Analogously, since* e_1 *is also in* W*, then* W *is not pointed also in his case. Anyway* W *is not pointed, but this contradicts Proposition 4.2.5.*

Since W *is arbitrary, we conclude that* (Σ) *does not have invariant cones in* $\mathbb{R}^4 = \bigwedge^1 \mathbb{R}$*.*

Now in the case of $\bigwedge^3 \mathbb{R}^4$*, we recall that* S *has invariant cones in* $\bigwedge^3 \mathbb{R}^4$ *if, and only if,* S^{-1} *has invariant cones in* $\mathbb{R}^4$*, and the linear isomorphism from* $\mathbb{R}^4$ *to* $\bigwedge^3 \mathbb{R}^4$ *(that preserves basis)* *is also a one to one correspondence between the respective invariant cones (see e.g. [15]).*

Therefore, it is enough to prove that S^{-1} *does not leave invariant cones in* $\mathbb{R}^4$*. Since* S *is generated by the exponential of the elements of* $\{A + uB;\ u \in \mathbb{R}\} \subset \mathfrak{sl}(\mathbb{R}^4)$*, then* S^{-1} *is generated by the exponential of the elements* $-A + uB$ *with* $u \in \mathbb{R}\}$*. Then* S^{-1} *is also the semigroup of the bilinear control system* $\dot{x} = -Ax + uBx$ *with* $x \in \mathbb{R}^4 \setminus \{0\}$ *and* $u \in \mathbb{R}$*.*

Let $W \neq \{0\}$ *be an* S^{-1}*-invariant cone. Note that* S^{-1} *has nonempty interior in* $\mathrm{Sl}(\mathbb{R}^4)$*. Therefore,* $e_1 \in W$ *or* $-e_1 \in W$*. Without loss of generality, we assume* $e_1 \in W$*. Since the highest eigenvalue of* $-A$ *is* 3 *and the corresponding eigenvector is* v_4*, then* $v_4 \in W$ *or* $-v_4 \in W$*. Furthermore, the inner product between* $-Ae_1$ *and* e_4 *is positive, and, therefore,* $e_4 \in W$*. If* $v_4 \in W$*, then* $\lim\limits_{t\to+\infty} \dfrac{e^{tB}v_4}{||e^{tB}v_4||} = -e_1 \in W$ *and* W *is not pointed, because* $e_1, -e_1 \in W$*. Otherwise, if* $-v_4 \in W$*, then* $\lim\limits_{t\to+\infty} \dfrac{e^{t(-B)}(-v_4)}{||e^{t(-B)}(-v_4)||} = -e_4$ *and* W *is still not pointed, because* $e_4, -e_4 \in W$*. Since* W *is not pointed in both cases, by Proposition 4.2.5 we have a contradiction. We conclude that the proper semigroup* S *does not leave invariant a proper cone in* $\bigwedge^1 \mathbb{R}^4$ *neither in* $\bigwedge^3 \mathbb{R}^4$ *but* S *has an invariant cone in* $\bigwedge^2 \mathbb{R}^4$ *(in fact, we showed that it leaves invariant the positive orthant of that space). Then by Theorem 4.3.5, the system* (Σ) *is not controllable. Moreover, Theorem 4.3.4 implies that* S *has parabolic type* $\Theta(S) = \{2\}$*, in other words,* $\mathbb{F}_{\Theta(S)} = \mathbb{G}_2(4)$*.*

Example 4.5.3. *Consider the above bilinear control system, but with*

$$A := \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & -1 \end{bmatrix}, \quad B := \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -\frac{3}{2} & -\frac{1}{10} & 0 \\ 0 & \frac{1}{10} & -\frac{3}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and denote the system semigroup by S*. Using again the implementation of Algorithm 1, we can see that* S *satisfies the Lie algebra rank condition, so* $\mathrm{int}S \neq \emptyset$*. Now we show that* S *does not have invariant cones in* $\bigwedge^1 \mathbb{R}^4$*,* $\bigwedge^2 \mathbb{R}^4$ *or* $\bigwedge^3 \mathbb{R}^4$ *and therefore* $S = \mathrm{Sl}(\mathbb{R}^4)$*.*

First note that

$$e^{\frac{\pi}{2}A} = \begin{bmatrix} 0 & d & 0 & 0 \\ -d & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{d}\frac{\sqrt{2}}{2} & \frac{1}{d}\frac{\sqrt{2}}{2} \\ 0 & 0 & -\frac{1}{d}\frac{\sqrt{2}}{2} & \frac{1}{d}\frac{\sqrt{2}}{2} \end{bmatrix}$$

with $e^{\frac{\pi}{2}} = d$.

Now we compute $e^{\frac{\pi}{2}A}$ *in the canonical basis of* $\bigwedge^3 \mathbb{R}^4$.

$$e^{\frac{\pi}{2}A}(e_1 \wedge e_2 \wedge e_3) = d\frac{\sqrt{2}}{2}e_1 \wedge e_2 \wedge e_3 - d\frac{\sqrt{2}}{2}e_1 \wedge e_2 \wedge e_4,$$

$$e^{\frac{\pi}{2}A}(e_1 \wedge e_2 \wedge e_4) = d\frac{\sqrt{2}}{2}e_1 \wedge e_2 \wedge e_3 + d\frac{\sqrt{2}}{2}e_1 \wedge e_2 \wedge e_4,$$

$$e^{\frac{\pi}{2}A}(e_1 \wedge e_3 \wedge e_4) = -\frac{1}{d}e_2 \wedge e_3 \wedge e_4$$

and

$$e^{\frac{\pi}{2}A}(e_2 \wedge e_3 \wedge e_4) = \frac{1}{d}e_1 \wedge e_3 \wedge e_4.$$

Then $e^{\frac{\pi}{2}A}$ *can be written, with respect to the canonical basis of* $\bigwedge^3 \mathbb{R}^4$*, as*

$$\begin{bmatrix} d\frac{\sqrt{2}}{2} & d\frac{\sqrt{2}}{2} & 0 & 0 \\ -d\frac{\sqrt{2}}{2} & d\frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{d} \\ 0 & 0 & -\frac{1}{d} & 0 \end{bmatrix} = \begin{bmatrix} dI & 0 \\ 0 & \frac{1}{d}I \end{bmatrix} \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}$$

with R_1, R_2 *rotations by angles different from* 0 *and* π.

In the next lemma we prove that the cones in $\mathbb{R}^4$*, which are invariant by above matrix, are subspaces.*

Lemma 4.5.4. *Let* $T \in \mathrm{Sl}(\mathbb{R}^4)$ *be the matrix*

$$T = \begin{bmatrix} dI & 0 \\ 0 & \frac{1}{d}I \end{bmatrix} \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}$$

where $R_1, R_2 \in \mathrm{SO}(2, \mathbb{R}) \setminus \{I, -I\}$*,* I *is* (2×2)*-identity matrix and* $d \in \mathbb{R} \setminus \{0\}$*. If* W *is a* T*-invariant cone in* $\mathbb{R}^4$ *then* W *is a subspace.*

Proof. Note that $\langle e_1, e_2 \rangle$ and $\langle e_3, e_4 \rangle$ are T-invariant spaces, and the restrictions of T to these spaces are αR where $\alpha > 0$ and R is the rotation different from I and $-I$. The only cones in a two-dimensional space that are invariant by these maps are $(0, 0)$ or the whole space, hence if $W \subset \langle e_1, e_2 \rangle$ then $W = \{0\}$ or $W = \langle e_1, e_2 \rangle$. If $W \subset \langle e_3, e_4 \rangle$ then $W = \{0\}$ or $W = \langle e_3, e_4 \rangle$. Suppose that W is not contained in these spaces. Then there exists $v \in W$ such that $v \neq \langle e_1, e_2 \rangle$ and $v \neq \langle e_3, e_4 \rangle$. As $\mathbb{R}^4$ is a direct sum of these two spaces, then v has the unique decomposition $v = u + w$, with $0 \neq u \in \langle e_1, e_2\rangle$ and $0 \neq w \in \langle e_3, e_4\rangle$. Knowing the eigenvalues of the restriction of A to $\langle e_1, e_2\rangle$ we can show that

$$\|T^n u\| = \|dR_1^n u\| \to +\infty \text{ and } \|T^n w| = \|(1/d)R_2^n w\| \to 0.$$

In particular, the distance of $\frac{T^n v}{\|T^n v\|}$ to $\langle e_1, e_2\rangle$ converges to zero, this sequence is contained in a compact set and has a subsequence that converges to p. Note that $p \in \langle e_1, e_2\rangle$ and $\|p\| = 1$. As W is a T-invariant cone then $p \in \mathrm{cl}(W) = W$.

We have also that $W \cap \langle e_1, e_2\rangle$ is a T-invariant cone which contains p. Then $W \cap \langle e_1, e_2\rangle = \langle e_1, e_2\rangle$ and so $\langle e_1, e_2\rangle \subset W$. It implies that $-u \in W$, then $w = v + (-u) \in W$ and therefore W has a non-null element of $\langle e_3, e_4\rangle$. In a similar way we can see that $\langle e_3, e_4\rangle \subset W$. Hence W contains $\langle e_1, e_2\rangle$ and $\langle e_3, e_4\rangle$, that is, $W = \mathbb{R}^4$. In all cases, W is a subspace of $\mathbb{R}^4$. □

By the above lemma, any $e^{\frac{\pi}{2}A}$-invariant cone in $\bigwedge^1 \mathbb{R}^4$ or in $\bigwedge^3 \mathbb{R}^4$, is a subspace. Therefore there are no S-invariant cones in $\bigwedge^1 \mathbb{R}^4$ neither in $\bigwedge^3 \mathbb{R}^4$.

Now it remains to prove that in $\bigwedge^2 \mathbb{R}^4$ there are no S-invariant cones. First note that the following submatrix of B,

$$B_2 = \begin{bmatrix} -\frac{3}{2} & -\frac{1}{10} \\ \frac{1}{10} & -\frac{3}{2} \end{bmatrix}$$

satisfies $\lim_{t\to+\infty} e^{tB_2} = 0$ implying that $\lim_{t\to+\infty} e^{tB}v = 0$ for all $v \in \langle e_2, e_3\rangle$. Moreover $e^{tB}e_1 = e^{2t}e_1$ and $e^{tB}e_4 = e^t e_4$.

Note that when $t \to +\infty$ we have that

$$\frac{e^{tB}(e_1 \wedge e_4)}{e^{2t}e^t} = \frac{e^{2t}e_1 \wedge e^t e_4}{e^{2t}e^t} = e_1 \wedge e_4 \to e_1 \wedge e_4$$

and moreover

$$\frac{e^{tB}(e_i \wedge e_j)}{e^{2t}e^t} \to 0 \text{ for } (i, j) \neq (1, 4).$$

Hence, for any vector $v \in \bigwedge^2 \mathbb{R}^4$ we have

$$v = \alpha_1 e_1 \wedge e_4 + \alpha_2 e_1 \wedge e_2 + \alpha_3 e_1 \wedge e_3 + \alpha_4 e_4 \wedge e_2 + \alpha_5 e_4 \wedge e_3 + \alpha_6 e_2 \wedge e_3, \tag{4.5-3}$$

for some $v_1, \ldots, v_4 \in \mathbb{R}$ and we have $\lim_{t\to+\infty} \frac{e^{tB}(v)}{e^{2t}e^t} = \alpha_1 e_1 \wedge e_4$. Now, suppose that exists an

S-invariant cone W. Then, there is $v \in W$ of the form (4.5-3) such that

$$\lim_{t \to +\infty} \frac{e^{tB}(v)}{e^{2t}e^{t}} = \alpha e_1 \wedge e_4$$

with $\alpha \neq 0$, because $\mathrm{int}W \neq \emptyset$.

As $\alpha e_1 \wedge e_4 \in W$ and

$$e^{2\pi A} = \begin{bmatrix} d^4 & 0 & 0 & 0 \\ 0 & d^4 & 0 & 0 \\ 0 & 0 & -\frac{1}{d^4} & 0 \\ 0 & 0 & 0 & -\frac{1}{d^4} \end{bmatrix}$$

we have that $e^{2\pi A}(\alpha e_1 \wedge e_4) = \alpha e_1 \wedge -e_4 = -\alpha e_1 \wedge e_4$. As W is invariant by the $e^{2\pi A}$-action, then $-\alpha e_1 \wedge e_4 \in W$, hence any straight line generated by $\alpha e_1 \wedge e_4$ is contained in W, that is, W is not pointed. Consequently, $\bigwedge^2 \mathbb{R}^4$ does not have S-invariant cones. Therefore, by Theorem 4.3.5, $S = \mathrm{Sl}(\mathbb{R}^4)$, that is, the system is controllable.

BIBLIOGRAPHY

[1] F. COLONIUS AND W. KLIEMANN, *The Dynamics of Control*, Birkhaüser, 2000.

[2] E. D. SONTAG, *Mathematical control theory: deterministic finite dimensional systems*, Vol. 6. Springer Science & Business Media, 2013.

[3] J. ZABCZYK, *Mathematical control theory*, Springer International Publishing, 2020.

[4] J. BEZANSON, A.EDELMAN, S. KARPINSKI AND V. B. SHAH, *Julia: A fresh approach to numerical computing*, SIAM review, 59(1), 65-98, 2017.

[5] O. G. DO ROCIO, L. A. B. SAN MARTIN, AND A. J. SANTANA, *Invariant Cones and Convex Sets For Bilinear Control Sistems and Parabolic Type of Semigroups*, Journal of Dynamical and Control Systems, 12(3), 419—432, 2006.

[6] O. DO ROCIO, A.J. SANTANA, AND M. VERDI, *Semigroups of affine groups, controllability of affine systems and affine bilinear systems in* $\mathrm{Sl}(2,\mathbb{R}) \rtimes \mathbb{R}^2$, SIAM J. Control Optim. 48(2), 1080-1088, 2009.

[7] A. L. DOS SANTOS AND L.A.B SAN MARTIN, *Controllability of Control Systems on Complex Simple Groups and the Topology of Flag Manifolds.* J. Dyn. Control Syst., 19, 157-171, 2013.

[8] D.L. ELLIOTT, *Bilinear Control Systems, Matrices in Action*, Kluwer Academic Publishers, 2008.

[9] V.JURDJEVIC AND I. KUPKA, *Control systems subordinate to a group action: accessibility*, J. Differ. Equ., 39 (1981), 186-211.

[10] YU. L. SACHKOV, *On invariant orthants of bilinear systems.*, J. Dynam. Control Systems, 4(1), 137-147, 1998.

[11] L.A.B. SAN MARTIN, *Flag Type of Semigroups: A Survey*. In: Lavor C., Gomes F. (eds) Advances in Mathematics and Applications. Springer Cham., 351-372, 2018.

[12] L.A.B. SAN MARTIN, *Invariant control sets on flag manifolds*. Mathematics of Control, Signals, and Systems, 6, 41-61, 1993.

[13] L.A.B. SAN MARTIN, *Maximal semigroups in semi-simple Lie groups*. Trans. Amer. Math. Soc., 353, 5165-5184, 2001.

[14] L. A. B. SAN MARTIN, *On global controllability of discrete-time control systems*, Mathematics of Control, Signals and Systems, 8, 279–297, 1995.

[15] L. A. B. SAN MARTIN, *A family of maximal noncontrollable Lie wedges with empty interior.*, Systems Control Lett., 43, 53–57, 2001.

[16] L.A.B. SAN MARTIN AND P.A. TONELLI, *Semigroup actions on homogeneous spaces*. Semigroup Forum, 50, 59-88, 1995.

[17] E. D. SONTAG, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Springer; 2nd edition, 1998.

[18] A. A. AGRACHEV, Y. L. SACHKOV, *Control Theory from the Geometric Viewpoint*. Springer, 2004

[19] V. JURDJEVIC, *Geometric Control Theory (Cambridge Studies in Advanced Mathematics)*. Cambridge: Cambridge University Press. 1996.

[20] L.A.B. SAN MARTIN, *Lie Groups*. Springer, 2021.

[21] J. C. WILLEMS, *Topological Classification and Structural Stability of Linear Systems*. Journal of Differential Equations, 35, pp 306-318 (1980)