New Mathematical Tools in Reach Control Theory

by

Melkior Ornik

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Electrical & Computer Engineering
University of Toronto

Abstract

New Mathematical Tools in Reach Control Theory

Melkior Ornik

Doctor of Philosophy

Graduate Department of Electrical & Computer Engineering

University of Toronto

2017

Reach control theory is an approach to satisfying complex control objectives on a constrained state space. It relies on triangulating the state space into simplices, and devising a separate controller on each simplex to satisfy the control specifications. A fundamental element of reach control theory is the Reach Control Problem (RCP). The goal of the RCP is to drive system trajectories of an affine control system on a simplex to leave this simplex through a predetermined facet. This thesis discusses a number of issues pertaining to reach control theory and the RCP. Its central part is a discussion of the solvability of the RCP. We identify strong necessary conditions for the solvability of the RCP by affine feedback and continuous state feedback, and provide elegant characterizations of these conditions using methods from linear algebra and algebraic topology. Additionally, using the theory of positive systems, Z-matrices, and graph theory, we obtain several new interpretations of the currently known set of necessary and sufficient conditions for the solvability of the RCP by affine feedback. The thesis also provides a rigorous foundation for the notion of exiting a simplex through a facet, and discusses uniqueness and existence of trajectories in reach control with discontinuous feedback. Finally, building on previous results, the thesis includes novel applications of reach control theory to parallel parking and adaptive cruise control. These applications serve to motivate new directions of theoretical research in reach control.

Acknowledgements

I wish to thank my supervisor, Mireille Broucke, for giving me the opportunity to pursue Ph.D. studies as her student. The Ph.D. theses almost always (rightfully) start with the same words of acknowledgements for supervisor's patience, wisdom, and vision. I want to take this opportunity to emphasize something else. As I stated in private many times before, Mireille brought an amazingly human element to my studies. An element that prompted me to never feel like a graduate student, but instead, to feel like *Mireille's* student. Without a doubt, this, more than anything else, served to make my experience an exceptionally valuable one.

I thank the entire Systems Control Group for their support throughout my studies. In particular, I am very grateful to Luca Scardovi and Lacra Pavel for serving on my thesis progress committee. I thank Raymond Kwong for serving on my qualifying examination. Additionally, I want to thank Angela Schoellig and Brian Ingalls for participating in my final defence, and Josh Taylor for heading my departmental defence committee. I imagine that duties such as these are never on the top of a scientist's list of excitements, and I am grateful for receiving such constant and consistent support.

I thank my collaborators of the past four years: Gabriele D'Eleuterio, Abdol-Reza Mansouri, Miad Moarref, Mateus Moura, Alexander Peplowski, and Adam Sniderman. You have helped turn a naturally solitary player such as myself into an unabashed supporter of collaboration. I also thank Mario Vukosavljev and Zachary Kroeze. We spent a lot of times discussing things around that big rectangular table in our office. These discussions helped us converge to a common philosophy about reach control, and a great deal of that philosophy is present in this work.

I thank Per-Olof Gutman for being an absolutely wonderful host to me during my visit to the Technion – Israel Institute of Technology in the spring of 2016. I truly cherish the time that I spent in Haifa. That is in no small part due to Per's research. To a much greater part it is due to Per's open, incredibly direct, and immensely caring personality that I can imagine he is well-known for.

Finally, I thank all the wonderful educators throughout my earlier university career. In particular I want to thank Andrew Lewis at Queen's University, whose incredible speed, mathematical rigour, and personal informality helped mould my scientific worldview.

Declaration of Previous Publication

With very minor exceptions, the entirety of the results presented in Chapters 4–9 of this thesis, and the majority of the accompanying text, is contained in previous manuscripts [89, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108] coauthored by the author of this thesis and his collaborators during Ph.D. studies. With the exception of the results of Section 6.1, these results have been published in peer-reviewed journal and conference venues, or are currently in the submission process.

The author of this thesis is the primary author of the work in Chapter 4, Chapter 5, Section 6.1, Chapter 7, and Chapter 8. This work comes from [99, 100, 101, 102, 103, 104, 105, 106]. The results and text of Sections 6.2-6.4 and Section 9.1 were produced in collaboration with Miad Moarref, and come from [89, 107]. The results and text of Section 9.2 were produced in collaboration with Mateus S. Moura and Alexander Peplowski, and are covered by [108].

Contents

List of Figures

Index of Basic Symbols and Acronyms

Some standard symbols used in the thesis:

$\emptyset$	empty set
$\mathcal{X} \subset \mathcal{Y}$	set $\mathcal{X}$ is a subset of set $\mathcal{Y}$
$\mathcal{X} \supset \mathcal{Y}$	set $\mathcal{X}$ is a superset of set $\mathcal{Y}$
$\mathcal{X} \cap \mathcal{Y}$	intersection of sets $\mathcal{X}$ and $\mathcal{Y}$
$\mathcal{X} \cup \mathcal{Y}$	union of sets $\mathcal{X}$ and $\mathcal{Y}$
$\mathbb{N}$	set of all strictly positive integers
$\mathbb{R}^n$	set of all real n-tuples
$\mathbb{B}^n$	n-dimensional unit ball
$\mathbb{S}^n$	n-dimensional unit sphere
$\mathbb{B}^n(x, r)$	closed ball in $\mathbb{R}^n$ with the center at x and radius r
$\mathbb{R}^{m \times n}$	set of all real $m \times n$ matrices
$\mathbb{1}_n$	column vector consisting of n 1's
e_i	column vector consisting of only 0's, except at a 1 at the i-th coordinate
$x \cdot y$	standard dot product of vectors x and y
$\|x\|$	standard norm of vector x
A^T	transpose of matrix A
$\mathrm{int}(\mathcal{X})$	(relative) interior of set $\mathcal{X}$
$\partial(\mathcal{X})$	(relative) boundary of set $\mathcal{X}$
$\mathrm{co}\{v_1, \ldots, v_n\}$	convex hull of points $v_1, \ldots, v_n$
$\mathrm{sp}\{v_1, \ldots, v_n\}$	vector space spanned by points $v_1, \ldots, v_n$
$\mathrm{aff}\{v_1, \ldots, v_n\}$	affine hull of points $v_1, \ldots, v_n$

Basic symbols used throughout the thesis:

$\mathcal{B}$	set of all vectors Bu with B from (3.1)
I	$\{1, 2, \ldots, n\}$, where n is the dimension of $\mathcal{S}$, defined above (3.3)
$I(x)$	smallest set of indices such that $x \in \mathcal{S}$ is contained in the convex hull of corresponding vertices in $\mathcal{S}$, defined defined above (3.3)
$\mathcal{C}(x)$	cone of velocity vectors associated with $x \in \mathcal{S}$, defined in (3.3)
$\mathcal{O}_{\mathcal{S}}$	set of possible closed-loop equilibria in the RCP, defined in (5.1)
$\mathcal{E}_{\mathcal{S}}$	set of open-loop equilibria in the RCP, defined in (7.4)

Acronyms used throughout the thesis:

RCP	Reach Control Problem
ORCP	Output Reach Control Problem
DRCP	Discontinuous Reach Control Problem
ODE	Ordinary differential equation
PWA	Piecewise affine
CPWA	Continuous piecewise affine
DPWA	Discontinuous piecewise affine
KKM	Knaster-Kuratowski-Mazurkiewicz
AE	Absolute extensor
AR	Absolute retract

Chapter 1

Introduction

There are two main fundamental pieces of motivation for this thesis. The first, presented in Section 1.1, is perhaps more extrinsic, objective, and, in one version or another, appears in nearly all works related to the Reach Control Problem. The second, presented in Section 1.2, is more abstract.

1.1 Control on a Constrained State Space

Safety, liveness, existence of physical obstacles and other constraints play a role in a large number of control scenarios. In other words, the natural state space of a system is not necessarily a full Euclidean space, but a subset. For instance, if a differential equation $\dot{x} = f(x, u)$ describes the movement of an autonomous vehicle in a room, a standard task is to design a control function u which allows the vehicle to reach a certain area of the room. If the room is boundless and does not contain any constraints on the vehicle's movement, this is a simple problem in controllability [66]. However, the set of possible vehicle positions is constrained by existence of walls and obstacles. In order for the vehicle not to hit those walls and obstacles, these constraints need to be taken into account.

Classical methods for geometric control theory, which focus on system stability, controllability and optimality, generally ignore state space constraints. Thus, a significant research effort arose from attempting to solve control problems in settings with constrained state spaces, or constrained control inputs. Arguably the most substantial investigation in this direction is Model Predictive Control (for comprehensive surveys of this technique, see [27, 84]). Additional methods include vertex control [46], interpolating control [96], reference governors [42], and anti-windup schemes (see [12] for a plethora of references).

The above methods largely pertain to systems with a convex state space and discrete time. Clearly,

there is a need to investigate constrained systems in a more general setting. *Reach control theory* is an approach to solving complex control objectives on a possibly nonconvex state space, and in continuous time. The method of reach control relies on triangulating the state space into simplices, or polytopes, and designing an appropriate controller on each simplex to satisfy the control objectives: either stabilize to an equilibrium in that simplex, or transition to a neighbouring simplex. See Figure 1.1 for an example taken from [107] and motivated by [134].

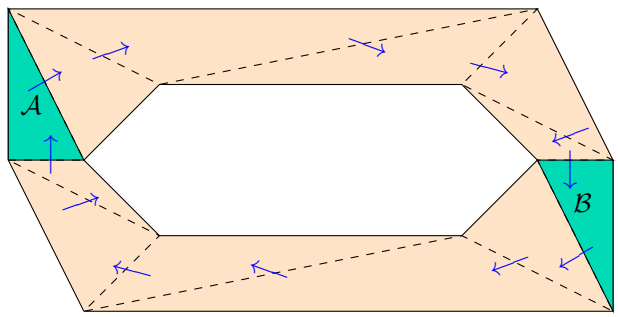

Figure 1.1: An example of a reach control approach to meeting a complex control objective. The system state, originally located in simplex $\mathcal{A}$, is required to visit simplex $\mathcal{B}$, return to $\mathcal{A}$, and continue repeating the procedure. In order to achieve that, the state space is triangulated into simplices, and a desired sequence of simplices (marked by blue arrows) is given.

On an abstract level, reach control theory is an extremely elegant concept for dealing with a constrained state space and complex control objectives. It aims to reduce often difficult computations on the entire non-convex state space to a series of computations on a simplex, with the idea that work on simplices is easy, and a wide class of control laws (e.g., affine laws) can be easily designed and computed on a simplicial structure. Due to its use of triangulations, reach control relies on the state space constraints being linear, but, given that nonlinear constraints can always be underapproximated or overapproximated by linear ones, this is not a major loss of generality.

In practice, however, reach control is *hard*. To outline the challenges and difficulties in reach control, we note that its theory can be divided into three layers:

- **Low-level** issues deal with designing a controller to guarantee the desired behaviour at each simplex in the state space triangulation.

- **Mid-level** issues deal with the switching mechanism between the control laws at two adjacent simplices.

- **High-level** issues deal with the discrete behaviour imposed by the triangulation of the state space.

Continuing with the example of Figure 1.1, low-level issues consist on designing a controller on each simplex that drives the system trajectories to exit the simplex through the desired edge, marked by an arrow. Mid-level issues deal with ensuring that leaving one simplex through the appropriate edge indeed leads to the system trajectory entering the next simplex. In other words, we need to ensure that the trajectory does not "escape" from the entire state space, instead of entering the next simplex, and that it does not start switching between two simplices infinitely many times. A fundamental high-level issue is determining a desired sequence of simplices that need to be visited. Another one is to determine the underlying triangulation of the state space in the first place. Let us describe these three layers in some more detail.

1.1.1 Low-level Reach Control

In the setting of reach control, the desired behaviour of system trajectories on a simplex $\mathcal{S}$ involves one of the following options:

(i) converging to an equilibrium $e \in \mathcal{S}$,

(ii) exiting through a predetermined facet $\mathcal{F} \subset \mathcal{S}$, without previously exiting $\mathcal{S}$ through any other facet.

The goal is to determine whether there exists a control law which enables such a behaviour.

Before we proceed, we note that control specifications which require combinations of these behaviours are also a possibility. For example, a part of Section 9.2.3 explores the control specification where trajectories are allowed both to converge to an equilibrium or, alternatively, to exit through a facet. However, this is largely a technical issue, and does not substantially differ from the two behaviours above.

Due to its difficulty, the bulk of research is done on case (ii) above. In the case of affine control systems $\dot{x} = Ax + Bu + a$, this is the *Reach Control Problem (RCP)*, which has been explored at length in the previous decade. The RCP is the central topic of this thesis, and Section 1.3 gives a somewhat substantial review of previous work on it. However, in the interest of readability, we briefly outline the current state of the art in the RCP research.

Although the RCP was attacked from a number of angles, the results on its solvability remain limited. In particular, in the case of continuous state feedback, no sufficient and necessary conditions to determine whether the trajectories of an affine control system can be controlled to leave a simplex through a particular facet are known. In the case of affine feedback, a set of sufficient and necessary conditions has been identified. However, this set of conditions is challenging for implementation purposes.

The initial thrust of this thesis work was largely driven by the following idea: if necessary and sufficient conditions for solvability are so hard to find, let us identify a necessary condition stronger than the ones previously known (even if it is still not sufficient), and determine whether we can find a computationally feasible and mathematically elegant characterization of it. If the answer is affirmative, then we found an easy necessary condition for the solvability of the RCP. This is a substantial addition to reach control, even if applying such a condition might result in some false positives.

The new necessary condition discussed above has been identified in terms of an affine/topological obstruction in the RCP. Its characterization is the main contribution of this thesis, given in Chapter 5 and Chapter 6. We showed that the existence of affine and topological obstructions can be characterized elegantly in terms of verifying whether a union of convex cones covers an entire Euclidean space. This result comes with a caveat: such a verification is generally NP-hard to do.

1.1.2 Mid-level Reach Control

As mentioned, the issue of transition from one simplex to another is largely technical. However, it has not been rigorously illuminated by previous research in reach control. Its core is in the observation that the low-level question (the RCP) concerns only exiting a particular simplex. Indeed, a control law which solves the RCP will drive a system trajectory to leave the simplex through a correct exit facet. However, this does not need to imply that a trajectory will automatically enter the adjacent simplex. For two examples of such pathological behaviour see Figure 1.2. The goal, thus, is to determine additional conditions, in terms of the velocity vectors of the trajectory or smoothness of the trajectory, which ensure that such a behaviour does not occur. We note that in the left image of Figure 1.2, the velocity vector at the point of intersection of the two simplices points outside both simplices, while the trajectory in the right image of Figure 1.2 is not analytic.

The second issue in the middle layer of the RCP is in switching. The idea of reach control is to design a separate simple controller on each simplex of the state space, and then concatenate these controllers into a controller on the whole space. Various classes of admissible controllers have been previously examined, among them discontinuous piecewise affine feedback (DPWA). In this set-up, the controller on each simplex is affine, but the concatenation of the vector fields on each simplex is not necessarily continuous. This leads to problems with the existence and uniqueness of a system trajectory, as well as in choosing which control law should be applied for a trajectory which starts at the intersection of the two simplices. Chapter 4 is dedicated to the two issues above.

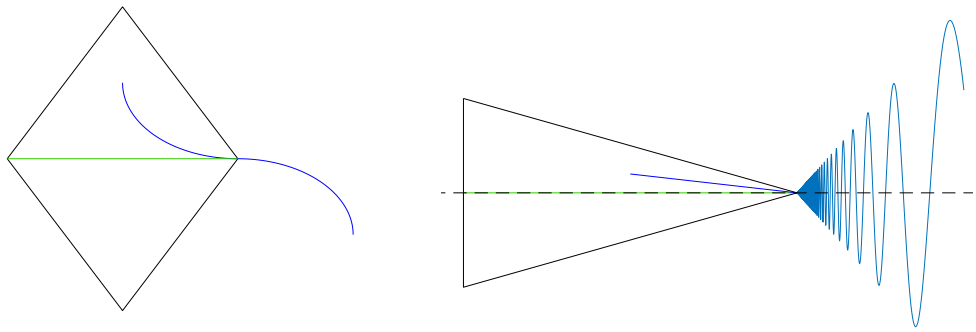

Figure 1.2: Two examples of pathological behaviour in the RCP. In the left image, the trajectory leaves the top simplex through the exit facet (marked in green), in accordance with the RCP specifications. However, that does not lead it to enter the bottom simplex. In the right image, the trajectory leaves the simplex through the exit facet. However, not only does it fail to enter the bottom simplex, it also continuously oscillates between the upper half-space, which contains the original simplex, and the lower half-space, which contains the simplex we wish to enter.

1.1.3 High-level Reach Control

The previous description of reach control as a mechanism founded on a triangulation of the state space hides the notorious fact that constructing an appropriate triangulation is a difficult task. While determining a triangulation of a state space is of course easy, there is no guarantee that such a triangulation will contain a *feasible path of simplices*, i.e., a sequence of simplices on which it is possible to construct a controller which guarantees that the control objective will be satisfied. Additionally, determining whether there exists a feasible path for a given triangulation is, at best, computationally difficult: it requires determining the solvability of the RCP at multiple simplices. This is, as we discussed in Section 1.1.1, itself problematic.

The primary problem in this layer is to determine a "good" triangulation of the state space, in the sense that it satisfies at least one of the following criteria:

(i) it contains a feasible path of simplices, or no other triangulation contains a feasible path either, or

(ii) it has geometric properties that make the solvability of the instances of the RCPs on the contained simplices easy to determine.

Unfortunately, no concerted effort has been taken to deal with this question. There has been some progress (e.g., [5, 23]), in the direction of (ii), where the faces of the simplices are chosen in a special way so that the solvability of the RCP is easier to determine, but this theory is far from complete. While

this thesis is not focused on the high-level issues in reach control, it takes a step towards finding an algorithmic way of determining an optimal triangulation, by determining a triangulation as a function of the parameters of the state space for a particular application in Section 9.2. However, this is again merely a first step.

The high-level outlook of reach control contains a number of additional questions that have not yet been explored at all. For instance, the state space in our discussion above is always assumed to be static. That is not necessarily the case in actual applications: it is realistic that obstacles move over time. Reach control theory has not been formed at all in a time-varying state space. Again, while the main focus of this thesis is on the matters described in Section 1.1.1, a first step towards this theory is presented in Chapter 9, in the setting of adaptive cruise control.

1.2 Novel Mathematical Techniques in Control

A second piece of motivation for this thesis takes a slightly more abstract outlook than the standard motivation presented above. In short, it is the author's belief that control theory often tries to run away from learning new mathematics, instead preferring to remain by its linear algebraic and dynamical systems roots. On the other hand, reach control offers an amazing opportunity to bring results from a number of fields into control theory. It has a rich discrete high-level structure which is amenable to graph theory analysis and temporal logic specifications. On the middle level it deals with fundamental questions of switched systems, including Zeno phenomena and uniqueness of solutions to discontinuous differential equations, and the low-level analysis in particular contains wide room for the use of interesting mathematics. In previous work, a Lyapunov style set-up was explored in [56], and optimization was used in [54]. In this thesis we present a number of additional approaches to low-level behaviour in the RCP.

To illustrate the richness of the mathematical structure of the RCP, let us list the major mathematical topics used in this thesis:

- Chapter 4 uses the theory of existence and uniqueness of differential equations, and explores the theory of sub-analytic differential equations,
- Chapter 5 uses the machinery of algebraic topology, in particular retraction theory, homotopy theory, obstruction theory and nerve theory, as well as Sperner's lemma and the KKM lemma,
- Chapter 6 uses linear algebra and optimization,
- Chapter 7 uses the theory of positive systems, Z- and M-matrices, and graph theory,

- Chapter 8 uses linear algebra,
- Chapter 9 explores the theory of nonlinear systems, linearizations around a non-equilibrium point, hybrid systems, and time-varying discrete behaviour of reach control.

Given this wealth of mathematical structure, a secondary objective of this thesis is, thus, to motivate the use of deeper mathematical results and tools in a wider variety of control problems.

1.3 Previous Work

As mentioned previously, there is a substantial body of work on reach control and the Reach Control Problem. Motivated by previous work on hybrid systems, the RCP was first introduced in [49]. In its current version, it was simultaneously stated in [47, 122]. As this thesis deals with several different facets of reach control and the RCP, the works that pertain the most to each individual topic are discussed more closely in the corresponding chapters. In this section we give a general overview of the state of the art in reach control by recounting the milestone papers in the field. As one can observe, the research effort on the RCP consists mainly of a primary thrust in the RCP with affine feedback, followed by subsequent efforts to define and investigate the use of more general control laws in the case when the RCP is not solvable by affine feedback. For an additional overview of the topic, see [53]; this literature review partly overlaps with the one contained therein.

Habets and van Schuppen, 2001 [49] This paper introduces reach control theory in the context of the hybrid systems methodology, and poses a variant of the Reach Control Problem on simplices and rectangles, using continuous piecewise affine feedback. However, this variant is significantly more restrictive than the current version of the problem. In it, the system trajectories are required to leave the simplex at the first instance that they reach the predetermined exit facet, and are required to do so in a way that is transversal to the exit facet (i.e., $h_0 \cdot x(t) > 0$, where h_0 is the outward-pointing normal to the exit facet $\mathcal{F}_0 \subset \mathcal{S}$). Sufficient and necessary conditions for the solvability of this version of the RCP is given, but such a set is no longer valid in the current version of the RCP.

Habets and van Schuppen, 2004 [50] The set-up and arguments of [49] are reiterated in [50]. The primary contribution of this paper is to generalize the results of [49] to the case of a variant of the RCP on general polytopes. In that set-up, sufficient and necessary conditions from [49] are no longer valid. Instead, there are now separate sets of sufficient and of necessary conditions. If the sufficient conditions are satisfied, this paper provides a computationally efficient algorithm for constructing a continuous

piecewise affine feedback solving the RCP on a polytope, by triangulating the polytope into a number of simplices. This algorithm will form the basis of computational work to date. We use it, for instance, in Chapter 9.

Habets, Collins and van Schuppen, 2006 [47] This paper provides (as Problem 4.1) the current version of the Reach Control Problem. In particular, the requirements that a system trajectory leaves the simplex at the first time of contact with the exit facet is removed. The requirement that a trajectory leaves the simplex transversally to the exit facet is removed as well. The paper also establishes a set of sufficient and necessary conditions for solvability of the RCP using affine feedback. These conditions, which we elaborate more on in Chapter 3, form the foundation of this thesis, and a major element of nearly all research in the RCP to date. As they are not given in a form amenable to control design, the goal is to explore them further to reach a computationally reasonable set of sufficient and/or necessary conditions for the solvability of the RCP.

Roszak and Broucke, 2006 [122] Published in parallel with [47], this paper also contains the new definition of the RCP, and introduces it using the notation that is prevalent in current research. The above set of sufficient and necessary conditions for the solvability of the RCP using affine feedback is also present in this paper, as Theorem 8. An additional investigation of these conditions is performed in this paper, and they are transformed into a feasibility problem with bilinear inequalities. However, while such a restatement is mathematically interesting and relates the RCP to the body of work in optimization, bilinear feasibility problems are generally NP-hard to solve. We build on this work in Chapter 6 and Chapter 7 of this thesis.

Roszak and Broucke, 2007 [123] This paper replaces the bilinear feasibility problem of [122] by a series of simple linear feasibility problems. However, this is only done in the special case where the system has $n-1$ inputs. As we will show in Section 5.3, the RCP has a very specific geometric structure in such a case, and these results are not readily generalizable to a general number of inputs.

Broucke, 2010 [23] Moving away from piecewise affine feedback, this paper is the first one to deal with RCP with general continuous state feedback. In the paper it is shown that the solvability of the RCP under affine feedback and under continuous state feedback are equivalent under very specific assumptions on the location of equilibria in the simplex. However, in general, continuous and affine feedback are not equivalent in the RCP, which was shown four years later in [55] and is an impetus for the exploration of different feedback classes in later papers. In terms of technical details, this paper

marks the first appearance of the use of Sperner's lemma to show the existence of equilibria in the RCP. We use a similar technique in Chapter 5.

Habets, Collins, and van Schuppen, 2012 [48] This paper is interested in the version of the RCP with output dynamics. In other words, instead of looking at affine or continuous feedback control based on the full system state trajectory, the feedback is based on the output, which can only partially observe the full state. This problem is interesting from a engineering perspective, and the authors create a machinery of *output polytopes* to be able to control the behaviour of the system in the original simplex. A version of this is later also discussed in [69]. However, this set-up makes the problem of finding feasible necessary and sufficient conditions for the solvability of the RCP more difficult. As that question is the primary focus of this thesis, there is no significant intersection between the content of [48] and this thesis.

Ashford and Broucke, 2013 [5] This paper continues the effort towards looking at more general classes of control laws to solve the RCP. We previously mentioned that the initial papers on the topic were devoted to affine and piecewise affine feedback. In 2010, this was expanded to continuous state feedback by [23]. In this paper, the authors consider the situation when not even continuous state feedback is sufficient to solve the RCP, and discuss time-varying feedback laws.

Helwa and Broucke, 2013 [54] This paper takes a different approach from the bulk of the previous work. Instead of generalizing the class of appropriate feedback controls for the RCP to be solvable, it slightly changes the definition of the RCP, requiring that the trajectories exit a simplex in a particular way. More precisely, it requires that there exists ξ such that $\xi \cdot (Ax + Bu(x) + a) < 0$ for all x. This is clearly a restrictive condition that does not exist in the definition of the RCP from [47, 122], although it may be taken as a generalization of the condition from [50]. With such a condition, the modified *monotonic RCP* is more easily solvable, and an easily verifiable set of sufficient and necessary conditions for solvability by continuous state feedback is developed in [54]. However, due to the restrictiveness of the above additional requirement, this does not immediately help with determining the solvability of the RCP by continuous state feedback.

Semsar-Kazerooni and Broucke, 2014 [125] This paper returns to the topic of solving the RCP under affine feedback. Instead of directly devising sufficient and necessary conditions for the solvability of the RCP, it studies the geometry of equilibria in the simplex in the case that an affine feedback does not solve the RCP. As non-existence of equilibria is a necessary condition for a feedback control to solve the RCP, this is useful as more than a mere mathematical exercise. In particular, locating the equilibria

can have two advantages: first, the simplex can potentially be cut in such a way that such equilibria are now outside of the simplex and hence no longer obstruct the solvability of the RCP. Secondly, observing that the equilibria are on the boundary of the simplex can help develop an intuition for an update to the affine feedback which will serve to "push" these equilibria outside of the simplex. This is the motivation behind the notion of reach controllability, which is introduced in [125], and further developed in Chapter 8.

Broucke and Ganness, 2014 [24] By discussing the solvability of the RCP in two cases: using discontinuous piecewise affine (DPWA) feedback and using general open-loop controls, this paper completes the set of currently investigated feedback controls in the RCP. It is shown that the RCP is solvable by DPWA feedback if and only if it is solvable by open-loop controls. In other words, it is meaningless to investigate classes of control other than DPWA. Additionally, [24] introduces the notion of reach control indices, which serve to encode the structure of equilibria in the RCP under continuous state feedback. We briefly revisit this notion in Chapter 7.

Unfortunately, [24] contains a small technical gap in the proof that the RCP can be solved by discontinuous piecewise affine feedback if and only if it can be solved by open-loop controls. While the idea is correct, this gap was identified and corrected in Chapter 4 of this thesis.

Helwa and Broucke, 2015 [56] This paper seeks to provide a Lyapunov-like approach to the solvability of the RCP. It does so by introducing the notion of a flow function, which needs to decrease along the closed-loop trajectories in a simplex until the trajectories finally leave the simplex. In analogy with the work on Lyapunov functions, it is shown that the existence of such a function is sufficient and necessary for the solvability of the RCP by continuous state feedback.

Helwa, Lin, and Broucke, 2016 [57] This paper discusses the solvability of the RCP by open-loop controls on a polytope. In that case, invariance conditions that are necessary for solvability of the RCP by continuous state feedback on a simplex (and that we discuss in Chapter 3) are no longer necessary on a polytope. While we discuss the use of open-loop controls in the RCP in a small part of Chapter 4, this paper is largely disconnected from the material of this thesis.

Wu and Shen, 2016 [136] This paper represents the first step in an entirely new and exciting direction of the Reach Control Problem. Unlike all previous versions of the RCP, which deal with an affine differential equation on a simplex, the work in [136] deals with linear differential inclusions on a simplex. In other words, the system dynamics are given by $\dot{x} \in \mathcal{A}(x) + Bu + a$, where $\mathcal{A}$ is a set-valued

function, and the RCP condition now comes in two flavours: in the weak version, one trajectory from every starting point x_0 is required to exit a simplex through a predetermined facet, while in the strong version, every trajectory is required to do so. The results in [136] solve the RCP for two particular classes of differential inclusions, both in the weak and strong version.

1.4 Thesis Organization

This thesis is primarily a compilation of work published in a number of papers over the past several years. While those papers are all within the relatively wide purview of reach control theory, their topics differ greatly. Hence, there are three general areas that the contributions of these thesis belong to: Chapter 4 deals with the switching mechanism between two adjecent simplices in reach control, and also provides a formal treatise of the existence and uniqueness of system trajectories. Chapters 5–8 deal with the solvability of the RCP and, related to it, with the structure of equilibria in affine systems. Finally, Chapter 9 uses the previous work on the RCP to develop two attractive novel applications for reach control: automated parallel parking and adaptive cruise control. We now list the main themes in each chapter. We note that the below chapter summaries are, in some cases, modified versions of the abstracts of articles contained in the corresponding chapters.

Chapter 2 covers basic mathematical notions used in the remainder of the thesis. We note that it is not comprehensive, and the reader is directed to standard reference books in mathematical analysis, geometry, and linear algebra for more details on the concepts used in this thesis. Additionally, this chapter introduces more advanced notions from topology and matrix theory that will be used primarily in Chapter 5 and Chapter 7.

Chapter 3 is the preliminary chapter containing the formal statement of the RCP and an exposition of previous results on its solvability. Apart from some largely notational details, it contains no novel work. The presentation of Chapter 2 and Chapter 3 is largely compiled from the papers that the thesis draws on in future chapters: [89, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108].

Chapter 4, based on [99], explores the notion of leaving a polytope through a predescribed facet. There are currently two subtly different notions available in the RCP literature. In one, the only condition is that at the last time instance when a trajectory is inside the polytope, it must also be inside the exit facet. In the other, the trajectory is required to also cross from the polytope into the outer open half-space bounded by the exit facet. In this chapter, we provide an example showing that these notions are not equivalent for general continuous or smooth state feedback. On the other hand, we prove that analyticity of the feedback control is a sufficient condition for equivalence of these definitions.

We generalize this result to several other classes of feedback control previously investigated in the RCP literature, most notably piecewise affine feedback. Additionally, we clarify or complete a number of previous results on the exit behavior of trajectories in the RCP.

Chapter 5, based on [101, 102, 103, 105, 106], studies a topological obstruction to solving the RCP by continuous state feedback. The existence of a topological obstruction is shown to be a sufficient condition for the RCP *not* to be solvable. The chapter has four major parts: in the first part, the problem of a topological obstruction is solved using retraction theory for two- and three-dimensional simplices. In the following sections the problem is distilled as one of continuously extending a function that maps into a sphere from the boundary of a simplex to its interior. As such, we employ techniques from the extension problem of algebraic topology. The problem of a topological obstruction is then solved for the RCP on two-input systems, giving a characterization based on null-homotopic functions on a one-dimensional sphere. Section 5.5 removes any restrictions on the dimension of the simplex, the number of inputs of the system, and the particular geometry of the subset of the state space where the obstruction arises. Thus, the results of this section represent the culmination of our efforts to characterize the topological obstruction. Finally, a small gap between the sufficient and necessary conditions for the existence of a topological obstruction identified in Section 5.5 is removed in Section 5.5.4 using the machinery of nerve theory.

In parallel with the previous chapter, Chapter 6 (based on [89, 102]) studies an obstruction to solvability of the RCP using affine feedback. It consists of two parts: the first part solves the problem of an affine obstruction in two- and three-dimensional systems using basic linear algebraic arguments. The second part provides necessary and sufficient conditions for occurrence of such an obstruction in the case of general state space dimensions. For two-input systems, these conditions are formulated in terms of scalar linear inequalities. Finally, computationally efficient necessary conditions are proposed for checking the obstruction for multi-input systems as feasibility programs in terms of linear inequalities.

Chapter 7, based on [100], takes a novel approach to the RCP, transforming it into a problem in positive system theory. Using the notions of Z-matrices, M-matrices and graph theory, this results in a number of new necessary and sufficient conditions for the set of open-loop equilibria in the RCP to be empty, and, consequently, a number of new necessary and sufficient conditions for the solvability of the RCP.

Chapter 8, based on [104], studies the geometric structure of open-loop equilibria in the RCP. Using a triangulation in which the set of all possible equilibria under closed-loop feedback intersects the interior of the simplex, we prove that the equilibrium set contains at most one point, both in the single-input and multi-input case. We additionally improve on the currently available results on reach controllability

to characterize when the closed-loop equilibria can be pushed off the simplex using affine feedback.

Chapter 9, based on [107, 108], consists of two applications. In the first part, we propose a novel method of parallel parking using reach control theory. We design a state space consisting of eight polytopes, and we use an automated procedure to construct a continuous piecewise affine controller for each polytope. Extensive simulations demonstrate the robustness of the approach: a vehicle starting from an acceptable initial position performs the maneuver safely and comes to a stop in the desired parking area. In the second part, we investigate a correct-by-construction synthesis of piecewise affine feedback controllers designed to satisfy the strict safety specifications set forth by the adaptive cruise control (ACC) problem. Our design methodology is based on the formulation of the ACC problem as a reach control problem on a polytope in a two-dimensional state space. The boundaries of this polytope, expressed as linear constraints on the states, arise from the headway and velocity safety requirements imposed by the ACC problem statement. We propose a model for the ACC problem, develop a controller that satisfies the ACC requirements, and produce simulations for the closed-loop system.

1.5 Main Contributions

1.5.1 Chapter 4

- Lemma 4.7 guarantees that chattering cannot occur in the RCP with analytic feedback,
- Lemma 4.12 guarantees that chattering cannot occur in the RCP with continuous piecewise affine feedback,
- Lemma 4.14 proves that, in the case of continuous piecewise affine feedback on a polytope, if invariance conditions in the RCP hold, a trajectory exiting a simplex necessarily crosses an exit facet,
- Theorem 4.30 corrects and generalizes an analogous result from [24]. It shows that if the RCP is solvable by open-loop controls, it is also solvable by discontinuous piecewise affine feedback. It also proves that, when using the proposed discontinuous piecewise affine feedback control to solve the RCP, chattering does not occur.

1.5.2 Chapter 5

- Theorem 5.19 provides a characterization of the existence of a topological obstruction in the RCP in two- and three-dimensional simplices,

- Theorem 5.28 provides a characterization of the existence of a topological obstruction in the RCP with two control inputs,
- Theorem 5.40 provides a necessary condition for the existence of a topological obstruction in the RCP with a general number of system states and inputs,
- Theorem 5.48 provides a sufficient condition for the existence of a topological obstruction in the RCP with a general number of system states and inputs,
- Theorem 5.49 combines Theorem 5.40 and Theorem 5.48 into a characterization of the existence of a topological obstruction in the RCP with a general number of system states and inputs.

1.5.3 Chapter 6

- Theorem 6.2 provides a characterization of the existence of a affine obstruction in the RCP in two- and three-dimensional simplices, and shows that, in that case, an affine obstruction exists if and only if a topological obstruction exists,
- Corollary 6.9 characterizes the existence of an affine obstruction in the RCP with a general number of system states and inputs,
- Algorithm 6.13 provides a computationally feasible method of checking the existence of an affine obstruction in the RCP with two inputs.

1.5.4 Chapter 7

- Lemma 7.3 proves that the RCP posed on any simplex can be transformed into an RCP on the standard orthogonal simplex,
- Theorem 7.7 characterizes the solvability of the RCP in terms of existence of positive solutions of linear equations,
- Corollary 7.10 characterizes the solvability of the RCP in terms of properties of a related directed graph.
- Proposition 7.11 characterizes the existence of open-loop equilibria in the RCP in terms of M-matrices, under the assumption that the set of all possible closed-loop equilibria does not intersect the exit facet.

1.5.5 Chapter 8

- Lemma 8.7 generalizes the work in [55] and proves that, under certain assumptions on the geometry of the equilibrium set in the RCP, the RCP contains only a single equilibrium,
- Theorem 8.14 generalizes the work in [125] and shows that, under related assumptions, the RCP is solvable if and only if the system matrices are reach controllable.

1.5.6 Chapter 9

- Section 9.1.3 provides a methodology for solving the problem of automated parallel parking using reach control,
- Section 9.2.3 provides a correct-by-construction strategy for solving the Adaptive Cruise Control problem with constant vehicle speed,
- Section 9.2.4 provides a methodology for solving Adaptive Cruise Control problem with varying vehicle speed.

Chapter 2

Mathematical Fundamentals

In this chapter we recount some fundamental geometric and topological notions that will be making an appearance throughout the thesis. For the sake of brevity and understanding, we only identify those concepts that are not already well-known in control-theoretic and basic mathematical literature. It is assumed that the reader is familiar with standard notions of calculus, linear algebra and ordinary differential equations, such as continuous functions, matrices, or uniqueness and existence of solutions of an ODE. For a more thorough threatise, we particularly recommend a classic textbook [127] which covers multivariate calculus, a well-known treatment [120] of geometry and convex analysis, the following books on matrix analysis: [11, 45, 62], the following books on topology: [20, 72, 94, 130], as well as the following classical works on differential equations: [40, 51]. As mentioned in Chapter 1, for the sake of readability, we defer discussing the introductory notions that will be used in specific later chapters, as opposed to the entire thesis, to those chapters.

2.1 Notation on Maps, Matrices, and Polytope Skeletons

If $\mathcal{A}$, $\mathcal{X}$ and $\mathcal{Y}$ are sets such that $\mathcal{A} \subset \mathcal{X}$, and $f : \mathcal{X} \to \mathcal{Y}$ is a map, then $f|_{\mathcal{A}} : \mathcal{A} \to \mathcal{Y}$ denotes the restriction of f to $\mathcal{A}$. $id_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}$ denotes an identity map on $\mathcal{X}$.

If A is a matrix, then $[A]_{ij}$ is the (i, j)-th element of A. If A is a block-matrix, then A_{ij} is the (i, j)-th block of A. If P is a matrix, P is *nonnegative* (denoted by $P \geq 0$) if $[P]_{ij} \geq 0$ for all i, j. It is *semipositive* (denoted by $P > 0$) if $P \geq 0$ and $P \neq 0$, and it is *positive* (denoted by $P \gg 0$) if $[P]_{ij} > 0$ for all i, j. If $a \in \mathbb{R}^n$ is a block-vector with p blocks, $\text{supp}(a) = \{i \in \{1, \ldots, p\} \mid a_i \neq 0\}$. If $v_1, v_2 \in \mathbb{R}^n$ are vectors, their scalar (dot) product is denoted by $v_1 \cdot v_2$ and equals $v_1^T v_2$.

Let $\mathcal{P}$ be a polytope with $\dim(\mathcal{P}) = r$. The *k-th skeleton* of $\mathcal{P}$, denoted by $\partial^{(k)}\mathcal{P}$, is the union of all

k-dimensional faces of $\mathcal{P}$. In particular, $\partial^{(0)}\mathcal{P}$ consists of the vertices of $\mathcal{P}$. In addition, $\partial^{(r)}\mathcal{P} = \mathcal{P}$. If $\mathcal{X}$ is a closed space, and $f : \mathcal{X} \to \mathcal{Y}$ is a map, then $\partial f = f|_{\partial\mathcal{X}}$.

2.2 Simplices

Definition 2.1. *A set* $\{v_0, v_1, \ldots, v_k\} \subset \mathbb{R}^n$ *is* affinely independent *if the set* $\{v_1 - v_0, \ldots, v_k - v_0\}$ *is linearly independent.*

Definition 2.2. *If* $\mathcal{V} \subset \mathbb{R}^n$ *is a vector space and* $x_0 \in \mathbb{R}^n$*, then* $x_0 + \mathcal{V} = \{x_0 + v \mid v \in \mathcal{V}\}$ *is an* affine space*. The* dimension *of an affine space* $\mathcal{A} = x_0 + \mathcal{V}$ *equals* $\dim(\mathcal{V})$*.*

Definition 2.3. *A set* $\mathcal{S} \subset \mathbb{R}^n$ *is a* k-dimensional simplex *if* $\mathcal{S} = \mathrm{co}\{v_0, \ldots, v_k\}$ *for some* $v_0, \ldots, v_k \in \mathbb{R}^n$*, and* $\{v_0, v_1, \ldots, v_k\}$ *is affinely independent.*

Definition 2.4. *A set* $\mathcal{P} \subset \mathbb{R}^n$ *is a* convex polytope *if* $\mathcal{P} = \mathrm{co}\{v_0, \ldots, v_k\}$ *for some* $v_0, \ldots, v_k \in \mathbb{R}^n$*.*

Definition 2.5. *A set* $\mathcal{P} \subset \mathbb{R}^n$ *is a* polytope *if it is a finite union of convex polytopes.*

Definition 2.6. *The* dimension *of a polytope* $\mathcal{P}$ *equals the dimension of the smallest affine space* $\mathcal{A}$ *such that* $\mathcal{P} \subset \mathcal{A}$*.*

2.3 Affine Functions

Definition 2.7. *A map* $f : \mathbb{R}^m \to \mathbb{R}^n$ *is an* affine function *if* $f(x) = Kx + g$ *for some* $K \in \mathbb{R}^{n\times m}$*,* $g \in \mathbb{R}^n$*.*

The following proposition is known (e.g., as Proposition 5.38 in [72]). However, defining affine functions by their values on simplex vertices is of crucial importance in reach control, so we provide its complete proof.

Proposition 2.8. *Let* $\{v_0, \ldots, v_m\} \subset \mathbb{R}^m$ *be an affinely independent set, and let* $y_0, \ldots, y_m \in \mathbb{R}^n$*. There exists exactly one affine function* $f : \mathbb{R}^m \to \mathbb{R}^n$ *such that* $f(v_i) = y_i$ *for all* $i \in \{0, \ldots, m\}$*.*

Proof. First, let us show that there exists such a function f. Since $\{v_0, \ldots, v_m\} \subset \mathbb{R}^m$ is an affinely independent set, $\{v_1 - v_0, \ldots, v_m - v_0\} \subset \mathbb{R}^m$ is a linearly independent set of m points in $\mathbb{R}^m$. Hence, $\{v_1 - v_0, \ldots, v_m - v_0\}$ is a basis for $\mathbb{R}^m$. Thus, we can define a matrix $K \in \mathbb{R}^{n\times m}$ such that $K(v_i - v_0) = y_i - y_0$ for all $i \in \{1, \ldots, m\}$. For this matrix K, let $g = y_0 - Kv_0$.

Let $f(x) = Kx + g$. We note that $f(v_i) = Kv_i + g = Kv_i + y_0 - Kv_0 = K(v_i - v_0) + y_0 = y_i$ for all $i \in \{0, \ldots, m\}$, where the last equality holds from the definition of K. Hence, we constructed a function f satisfying the desired property.

Now, assume that $f_1(x) = K_1x + g_1$ and $f_2(x) = K_2x + g_2$ are affine functions which satisfy $f_1(v_i) = f_2(v_i) = y_i$ for all $i \in \{0, \ldots, m\}$. Then, $f_1 - f_2$ is an affine function which satisfies $(f_1 - f_2)(v_i) = 0$ for all $i \in \{0, \ldots, m\}$. Thus,

$$(K_1 - K_2)v_i = g_2 - g_1 \tag{2.1}$$

for all $i \in \{0, \ldots, m\}$, i.e., $(K_1 - K_2)(v_i - v_0) = 0$ for all $i \in \{0, \ldots, m\}$. Since $\{v_1 - v_0, \ldots, v_m - v_0\}$ is a basis for $\mathbb{R}^m$, $K_1 - K_2$ necessarily equals 0. Thus, $K_1 = K_2$. By (2.1), we also obtain $g_1 = g_2$. Thus, $f_1 = f_2$. □

2.4 Differentiability

Definition 2.9. *Let $\mathcal{I} \subset \mathbb{R}^m$ be an open set. $f : \mathcal{I} \to \mathbb{R}^n$ is* differentiable *if at every point $x \in \mathcal{I}$ there exists a matrix $J_x \in \mathbb{R}^{n\times m}$ such that*

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - J_x h\|}{\|h\|} = 0.$$

Proposition 2.10 ([127])**.** *Let $\mathcal{I} \subset \mathbb{R}^m$ be an open set. Let $f = (f_1, \ldots, f_n) : \mathcal{I} \to \mathbb{R}^n$. Assume that all the partial derivatives $\partial f_j/\partial x_i$ exist and are continuous. Then, f is differentiable.*

Definition 2.11. *Let $\mathcal{I} \subset \mathbb{R}^m$ be an open set. $f = (f_1, \ldots, f_n) : \mathcal{I} \to \mathbb{R}^n$ is* infinitely diferentiable *(denoted by $f \in C^\infty$) if all partial derivatives*

$$\frac{\partial^k f_j}{\partial x_1^{k_1} \cdots \partial x_m^{k_m}}$$

exist and are continuous.

Definition 2.12. *Let $\mathcal{I} \subset \mathbb{R}^m$ be an open set. Let $f = (f_1, \ldots, f_n) : \mathcal{I} \to \mathbb{R}^n$ be infinitely differentiable. Map f is* analytic *at $x_0 = (x_{01}, \ldots, x_{0m}) \in \mathcal{I}$ (denoted by $f \in C^\omega$) if*

$$f_j(x_1, \ldots, x_m) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_m=0}^{\infty} \frac{\partial^{k_1+\ldots+k_m} f}{\partial x_1^{k_1} \cdots x_m^{k_m}}(x_{01}, \ldots, x_{0m})(x_1 - x_{01})^{k_1} \cdots (x_m - x_{0m})^{k_m}$$

for all $j \in \{1, \ldots, n\}$.

2.5 Topological and Metric Spaces

Definition 2.13. *A* topological space *is an ordered pair* $(\mathcal{X},\mathcal{U})$*, where* $\mathcal{X}$ *is a set and* $\mathcal{U} \subset 2^{\mathcal{X}}$*, such that the following holds:*

(i) $\emptyset, \mathcal{X} \in \mathcal{U}$*,*

(ii) If $U_i \in \mathcal{U}$ *for all* $i \in I$*, then*

$$\bigcup_{i \in I} U_i \in \mathcal{U},$$

(iii) If I *is a finite set, and* $U_i \in \mathcal{U}$ *for all* $i \in I$*, then*

$$\bigcap_{i \in I} U_i \in \mathcal{U}.$$

Sets $\mathcal{U}$ *are* open sets *in* $\mathcal{X}$*.*

Remark 2.14. *In the Euclidean space* $\mathbb{R}^n$*, the standard open sets are sets* U *such that for all* $x \in U$ *there exists* $\varepsilon > 0$ *such that* $\mathbb{B}^n(x, \varepsilon) \subset U$*. Whenever there is a standard notion of open sets in a topological space* $(\mathcal{X},\mathcal{U})$*, the notation is usually abused, and it is said that* $\mathcal{X}$ is a topological space*.*

Definition 2.15. *If* $(\mathcal{X},\mathcal{U})$ *is a topological space and* $\mathcal{Y} \subset \mathcal{X}$*, then* $\mathcal{U}' = \{U \cap \mathcal{Y} \mid U \in \mathcal{U}\}$ *satisfies the conditions in Definition 2.13. It is said that* $\mathcal{U}'$ *defines a* subspace topology *on* $\mathcal{Y}$*.*

Definition 2.16. *A* metric space *is an ordered pair* $(\mathcal{X}, d)$*, with* $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ *such that the following holds:*

(i) $d(x, y) \geq 0$ *for all* $x, y \in \mathcal{X}$*,*

(ii) $d(x, y) = 0$ *if and only if* $x = y$*,*

(iii) $d(x, y) = d(y, x)$ *for all* $x, y \in \mathcal{X}$*,*

(iv) $d(x, z) \leq d(x, y) + d(y, z)$ *for all* $x, y, z \in \mathcal{X}$*.*

Map d *is said to be the* metric *on* $\mathcal{X}$*.*

Remark 2.17. *In the Euclidean space* $\mathbb{R}^n$*, the standard metric is given by* $d(x, y) = \|y - x\|$*. Whenever there is a standard metric in a metric space* $(\mathcal{X}, d)$*, the notation is usually abused, and it is said that* $\mathcal{X}$ is a metric space*.*

The following proposition is trivial to verify.

Proposition 2.18. *Let $(\mathcal{X}, d)$ be a metric space. Let $\mathbb{B}_d(x, \varepsilon) = \{y \in \mathcal{X} \mid d(x, y) \leq \varepsilon\}$. Then, $\mathcal{U} = \{U \subset \mathcal{X} \mid$ for all $x \in U$, there exists $\varepsilon > 0$ such that $\mathbb{B}_d(x, \varepsilon) \subset U\}$ satisfies the conditions in Definition 2.13.*

Definition 2.19. *Let $(\mathcal{X}, d)$ be a metric space. $(\mathcal{X}, d)$ is* compact *if for any sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{X}$ there exists a convergent subsequence $(x_{n_p})_{p\in\mathbb{N}} \subset \mathcal{X}$, i.e., a sequence $(x_{n_p})_{p\in\mathbb{N}} \subset \mathcal{X}$ and $x \in \mathcal{X}$ such that for any $\varepsilon > 0$, $d(x, x_{n_p}) < \varepsilon$ for all large enough p.*

Theorem 2.20 (Heine-Borel)**.** *If $\mathcal{X} \subset \mathbb{R}^n$ is a metric space, then $\mathcal{X}$ is compact if and only if it is closed and bounded.*

Definition 2.21. *Let $\mathcal{Y} \subset \mathbb{R}^n$.* Relative interior *of $\mathcal{Y}$, denoted by* $\operatorname{int}(\mathcal{Y})$*, is defined by* $\operatorname{int}(\mathcal{X}) = \{y \in \mathcal{X} \mid$ *there exists $\varepsilon > 0$ such that $\mathbb{B}^n(x, \varepsilon) \cap \operatorname{aff}(\mathcal{Y}) \subset \mathcal{Y}\}$.*

The relative interior of $\mathcal{Y}$ is the analogue of interior of $\mathcal{Y}$, in the topology of $\operatorname{aff}(\mathcal{Y})$. The relative boundary is defined analogously.

Having provided a very general overview of the material that will be used throughout the thesis, we now move on to more advanced notions that will be used in Chapter 5 and Chapter 7.

2.6 Basic Topology

We say two topological spaces $\mathcal{X}$ and $\mathcal{Y}$ are *homeomorphic* (notation $\mathcal{X} \cong \mathcal{Y}$) if there exists a continuous function $f : \mathcal{X} \to \mathcal{Y}$ which is bijective and has a continuous inverse. Such an f is called a *homeomorphism.*

Theorem 2.22 ([20])**.** *Let $\mathcal{P}$ be a convex κ-dimensional polytope. Then $\mathcal{P} \cong \mathbb{B}^\kappa$, and $\partial\mathcal{P} \cong \mathbb{S}^{\kappa-1} = \partial\mathbb{B}^\kappa$ by the same homeomorphism.*

Theorem 2.23 ([90])**.** *Let $\mathcal{X} \subset \mathbb{R}^m$ be convex, compact and non-empty. Then, $\mathcal{X} \cong \mathbb{B}^\rho$ for some $\rho \leq m$.*

If $\mathcal{A}$ is a subspace of $\mathcal{X}$, a continuous map $r : \mathcal{X} \to \mathcal{A}$ is a *retraction* if $r|_{\mathcal{A}} \equiv id_{\mathcal{A}}$. The following claim is easy to prove.

Lemma 2.24. *Let $\mathcal{X} \cong \tilde{\mathcal{X}}$ by homeomorphism h, and $\mathcal{A} \subset \mathcal{X}$ and $\tilde{\mathcal{A}} \subset \tilde{\mathcal{X}}$ homeomorphic with the same homeomorphism. Then there exists a retraction $r : \mathcal{X} \to \mathcal{A}$ if and only if there exists a retraction $\tilde{r} : \tilde{\mathcal{X}} \to \tilde{\mathcal{A}}$.*

Proof. Take $\tilde{r} = h \circ r \circ h^{-1}$. This function is obviously continuous, and for each $\tilde{a} \in \tilde{\mathcal{A}}$, $\tilde{r}(\tilde{a}) = (h \circ r \circ h^{-1})(\tilde{a}) = (h \circ r)(a)$, where $a \in \mathcal{A}$ is such that $h(a) = \tilde{a}$. Since r is a retraction, $(h \circ r)(a) = h(a) = \tilde{a}$. We are done. Analogously, if we know $\tilde{r}$ exists, we can take $r = h^{-1} \circ \tilde{r} \circ h$. □

Two homeomorphic sets are considered to be essentially equal in a topological sense. Let us now give a generalization of the homeomorphism equivalence relation.

Let $f, g : \mathcal{X} \to \mathcal{Y}$ be continuous maps. We say f is *homotopic* to g, denoted by $f \simeq g$, if there exists a continuous function $F : \mathcal{X} \times [0,1] \to \mathcal{Y}$ such that $F(\cdot, 0) \equiv f$ and $F(\cdot, 1) \equiv g$. Topological spaces $\mathcal{X}$ and $\mathcal{Y}$ are *homotopy equivalent*, denoted $\mathcal{X} \simeq \mathcal{Y}$, if there exist continuous maps $f : \mathcal{X} \to \mathcal{Y}$, $g : \mathcal{Y} \to \mathcal{X}$ such that $f \circ g$ and $g \circ f$ are homotopic to $id_{\mathcal{Y}}$ and $id_{\mathcal{X}}$, respectively. A continuous map $f : \mathcal{X} \to \mathcal{Y}$ is said to be *inessential* or *null-homotopic* if f is homotopic to a constant map $c(x) = y_0$, where $y_0 \in \mathcal{Y}$. A topological space $\mathcal{X}$ is *contractible* if the identity map $id_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}$ is null-homotopic.

Lemma 2.25 ([130])**.** *If $f : \mathcal{X} \to \mathbb{S}^n$ is a continuous map such that $f(\mathcal{X}) \neq \mathbb{S}^n$, that is, f is not surjective, then f is null-homotopic.*

Lemma 2.26 ([94])**.** *If $f : \mathcal{X} \to \mathcal{Y}$ is a continuous map and $\mathcal{Y}$ is contractible, then f is null-homotopic.*

Lemma 2.27 ([94])**.** *Every continuous map $f : \mathbb{S}^n \to \mathbb{S}^1$ with $n \geq 2$ is null-homotopic.*

Example 2.28. *Any convex subset X of $\mathbb{R}^n$ is contractible. To see this, define $F : X \times [0,1] \to X$ by $F_p(x,t) = F(tp + (1-t)x)$, where p is any point in X. This function is a homotopy between the identity map id_X and a constant map.*

A *path* is defined as a continuous function $f : [0,1] \to \mathcal{Y}$, with endpoints at $f(0)$ and $f(1)$. The *reverse path* of path f is denoted by $\bar{f}$ and is defined by $\bar{f}(t) = f(1-t)$. A *concatenation* of two paths f and g with $f(1) = g(0)$ is denoted by fg and formally defined by

$$fg(t) = \begin{cases} f(2t) & \text{if } t \leq 1/2, \\ g(2t-1) & \text{if } t > 1/2. \end{cases}$$

A *loop* is a closed path, i.e., a path f with $f(0) = f(1)$. We say $f(0)$ is the *basepoint* of the loop f. Equivalently, a loop is a continuous map $f : \mathbb{S}^1 \to \mathcal{Y}$.

For paths, we have the following notion of path-homotopy. Let $f, g : [0,1] \to \mathcal{Y}$ be paths with the same endpoints, i.e. $f(0) = g(0)$ and $f(1) = g(1)$. We say f is *path-homotopic* to g (denoted, with slight abuse of notation, also by $f \simeq g$), if there exists a continuous function $F : [0,1] \times [0,1] \to \mathcal{Y}$ such that $F(\cdot, 0) \equiv f$, $F(\cdot, 1) \equiv g$, $f(0, \cdot) \equiv f(0)$, $F(1, \cdot) \equiv f(1)$. A loop f is *path-null-homotopic* if it is path-homotopic to a constant map.

We note that the definition of path-(null-)homotopy of paths contains a stronger condition that for homotopy of general maps.

Lemma 2.29 ([72])**.** *Let $f, g : [0,1] \to Y$ be paths with the same endpoints. Then, f is path-homotopic to g if and only if $f\bar{g}$ is null-path-homotopic. In particular, $f\bar{f}$ is null-path-homotopic.*

Returning back to contractibility, a topological space $\mathcal{X}$ is *locally contractible* if for every $x \in \mathcal{X}$ and every open neighborhood $\mathcal{V} \subset \mathcal{X}$ of x, there exists an open neighborhood $\mathcal{W} \subset \mathcal{V}$ of x which is contractible in the subspace topology from $\mathcal{V}$.

Lemma 2.30 ([3, 17])**.** *A finite union of convex sets is locally contractible.*

The following definition of absolute retract, which also resembles the definition of absolute extensor, is taken from [3]. While the definition is somewhat unorthodox, it may be noted that in metrizable spaces, the notions of absolute retract and absolute extensor (AE) are in fact equivalent [116].

Definition 2.31. *A metrizable space $\mathcal{X}$ is an* absolute retract *(AR) if for every metrizable space $\mathcal{Y}$ and every closed $\mathcal{V} \subset \mathcal{Y}$, each continuous map $f : \mathcal{V} \to \mathcal{X}$ is extendable to a continuous map $F : \mathcal{Y} \to \mathcal{X}$.*

Definition 2.32 ([117])**.** *A compact metric space is* finite-dimensional *if it is homeomorphic to a subset of $\mathbb{R}^k$ for some $k \in \mathbb{N}$.*

Theorem 2.33 ([3, 17])**.** *If $\mathcal{X}$ is a compact, contractible and locally contractible finite-dimensional metric space, it is AR.*

Finally, we will require a major result from extension theory. The *extension problem* regards the following question: given a continuous map $\partial f : \partial\mathcal{X} \to \mathcal{Y}$ defined on the boundary of a space $\mathcal{X}$, we would like to know if there exists a continuous extension of $f : \mathcal{X} \to \mathcal{Y}$ such that $f|_{\partial\mathcal{X}} \equiv \partial f$.

The main result relating null-homotopic maps and the extension problem is the Extension Theorem (see [94]). Because of its importance, we also include a proof of this result.

Theorem 2.34. *A continuous map $f : \mathbb{S}^n \to \mathcal{Y}$ is null-homotopic if and only if f extends to a map $F : \mathbb{B}^{n+1} \to \mathcal{Y}$.*

Proof. Suppose $f : \mathbb{S}^n \to \mathcal{Y}$ is inessential and let $H : \mathbb{S}^n \times [0,1] \to \mathcal{Y}$ be a homotopy between $f(x) = H(x,0)$ and a constant map $c(x) = H(x,1)$. Note that every point $y \in \mathbb{B}^{n+1}\backslash\{0\}$ can be written uniquely as tx, where $t \in (0,1]$ and $x \in \mathbb{S}^n$. Define $F : \mathbb{B}^{n+1} \to \mathcal{Y}$ by $F(tx) = H(x, 1-t)$ for $x \in \mathbb{S}^n$ and $t \in (0,1]$. Also define $F(0) = c(0)$. It is easy to check that F is continuous.

Conversely, suppose that $f : \mathbb{S}^n \to \mathcal{Y}$ extends to a map $F : \mathbb{B}^{n+1} \to \mathcal{Y}$. Define $H(x,t) : \mathbb{S}^n \times [0,1] \to \mathcal{Y}$ by $H(x,t) := F(tx)$. Then H is a homotopy between f and a constant map $c(x) = F(0)$. $\square$

2.7 Simplicial Complexes

Let J be a finite index set. $\mathcal{D} \subset 2^J \backslash \{\emptyset\}$ is a (finite) *abstract simplicial complex* if

$$(L \subset K \wedge K \in \mathcal{D}) \Rightarrow L \in \mathcal{D}, \quad K, L \in 2^J \backslash \{\emptyset\}. \tag{2.2}$$

Geometrically, an abstract simplicial complex $\mathcal{D}$ can be realized with the singletons in J playing the role of vertices, and with k-subsets of J being the k-dimensional faces of $\mathcal{D}$. It is well-known (see, e.g., Remark 1.3.4 in [98]) that we can indeed represent $\mathcal{D}$ as a usual geometric simplicial complex instead of an abstract one, without any loss of functionality. For an illustration of a geometric realization of an abstract simplicial complex see Figure 2.1.

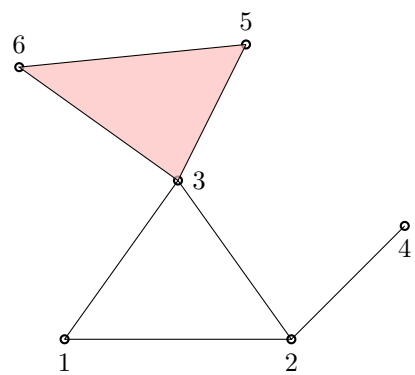

Figure 2.1: An example of a simplicial complex. In abstract form, this simplicial complex is given by $\mathcal{D} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1,2\}, \{1,3\}, \{2,3\}, \{2,4\}, \{3,5\}, \{3,6\}, \{5,6\}, \{3,5,6\}\}$.

Let $\mathcal{D}$ be a finite simplicial complex with vertices denoted by $1, 2, \ldots, l$. The *(first) barycentric subdivision* of $\mathcal{D}$ divides each k-dimensional simplex $\mathcal{P} \subset \mathcal{D}$ into $(k+1)!$ smaller simplices, $k = 1, \ldots, l-1$. Each of these smaller simplices consists of one vertex $x_0 \in \{1, 2, \ldots, l\}$ of the original simplex, one vertex in the middle of some edge $\mathrm{co}\{x_0, x_1\}$, one vertex in the middle of some two-dimensional simplex $\mathrm{co}\{x_0, x_1, x_2\}$, etc. The *barycentric center* of $\mathcal{P}$ is the vertex of its barycentric subdivision contained in the interior of $\mathcal{P}$. The *barycentric star* $\mathrm{bst}(j)$ in $\mathcal{D}$ consists of all simplices in the first barycentric subdivision of $\mathcal{D}$ which contain $\{j\}$. See Figure 2.2 for an illustration.

Finally, we give a technical lemma that will be crucial in our main result of Chapter 5.

Lemma 2.35 (Lemma 4.3, [8])**.** *Let Δ be a simplex, and let $f : \Delta \to \Delta$ be a continuous map such that*

$$f(\mathcal{P}) \subset \mathcal{P}$$

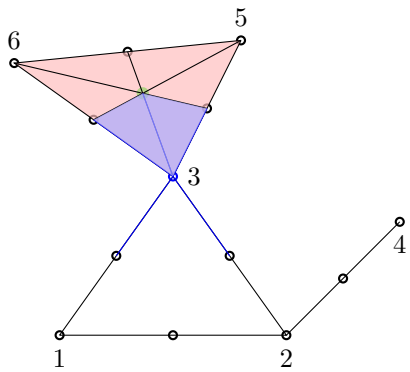

Figure 2.2: The barycentric subdivision of the simplicial complex from Figure 2.1. Each 1-dimensional edge has been split into two 1-dimensional edges, and the 2-dimensional face has been split into 6 smaller triangles. The barycentric star of vertex 3 is marked in blue, and the barycentric center of the two-dimensional simplex with vertices $3, 5, 6$ is marked in green.

holds for each face $\mathcal{P} \subset \Delta$. Then, f is surjective.

Corollary 2.36. *Let Δ be a simplex. There does not exist a continuous map $f : \Delta \to \partial\Delta$ such $f(\mathcal{P}) \subset \mathcal{P}$ holds for each face $\mathcal{P} \subset \Delta$.*

Proof. Assume otherwise. Then, by taking the codomain of f to be Δ instead of $\partial\Delta$, such a map f satisfies the conditions of Lemma 2.35 with $\mathrm{Im}(f) \subset \partial\Delta$. However, by Lemma 2.35, $\mathrm{Im}(f) = \Delta$. □

2.8 A Short Tutorial on Nerve Theory

The notion of a *set cover* is a central notion of topology. To remind the reader, a cover of a set $\mathcal{X}$ is a set $\{\mathcal{X}_j \mid j \in J\}$ such that $\cup_{j \in J} \mathcal{X}_j = \mathcal{X}$. If all the sets $\mathcal{X}_j$ are open (in a topology on $\mathcal{X}$), $\{\mathcal{X}_j \mid j \in J\}$ is an *open cover*. If all $\mathcal{X}_j$ are closed, $\{\mathcal{X}_j \mid j \in J\}$ is a *closed cover*. If the index set J is finite, the corresponding cover is a *finite cover*.

The properties of a set cover reveal a great deal of information about the underlying space. For instance, a commonly used definition of a compact space is that it is a space where every open cover has a finite subcover. As another example, the *covering dimension* of a topological space $\mathcal{X}$ is defined as the smallest number n such that every open cover can be refined (i.e., split into smaller open sets) in such a way that every point in $\mathcal{X}$ is contained in no more than $n+1$ elements of that refinement. It can be shown, for instance, that the covering dimension of the space $\mathbb{R}^n$ equals its vector space dimension $\dim(\mathbb{R}^n) = n$. For more on covers, see [91].

The latter example motivates the need to examine the structure of the intersections of elements of

a set cover. This is an impetus for a variety of results in geometry and topology, e.g., Helly's theorem [35], Knaster-Kuratowski-Mazurkiewicz (KKM) lemma [93], and Sperner's lemma [16]. We use the KKM lemma in Section 5.3 and Sperner's lemma in Section 5.5.3, and these tools were also used in other work in control and dynamics [65, 109]. Nerve theory seeks to provide a more comprehensive methodology for exploring the properties of an intersection of cover elements. In fact, it is possible to prove a generalization of Sperner's lemma using nerve theory [78].

Nominally, a nerve of a cover $\{\mathcal{X}_j \mid j \in J\}$ of a space is just a list of all subsets of the family $\{\mathcal{X}_j \mid j \in J\}$ such that all sets in the subset have a nonempty intersection. In formal terms, let J be a finite index set, and let $\{\mathcal{X}_j \mid j \in J\}$ be a family of subsets of a topological space $\mathcal{X}$. The *nerve* $\mathcal{N}$ *of* $\{\mathcal{X}_j \mid j \in J\}$ *with respect to* $\mathcal{X}$ is the set of non-empty subsets K of J given by

$$K \in \mathcal{N} \Leftrightarrow \bigcap_{k \in K} \mathcal{X}_k \neq \emptyset. \tag{2.3}$$

While other related objects can be described and investigated within the purview of nerve theory (e.g., the *nerve graph* introduced and used in [78]), they are not in the focus of our thesis, and we omit such a discussion. We refer the reader to [78], [95], and other references regarding nerve theory contained in the bibliography.

The structure of a nerve has remarkable geometric properties. First of all, by (2.2), $\mathcal{N}$ is an abstract simplicial complex: if $\cap_{j \in K} \mathcal{X}_j \neq \emptyset$, then clearly $\cap_{j \in L} \mathcal{X}_j \neq \emptyset$ for all $\emptyset \neq L \subset K$. Thus, $\mathcal{N}$ has a natural geometric interpretation: see Figure 2.3 for an example.

A far more remarkable result concerning nerves is the well-known nerve theorem. In informal terms, it states that, under certain conditions, a space and a nerve of its cover have the same homotopy structure, namely that they are homotopy equivalent. This is remarkable because a given space can clearly admit a variety of different covers, which are then shown to share the same homotopic properties.

We note that nerve $\mathcal{N}$ depends on $\{\mathcal{X}_j \mid j \in J\}$ and $\mathcal{X}$. In the remainder of this chapter, however, we will not explicitly note this. This is for two reasons. First, we will be using nerve theory exclusively to investigate a nerve of a single cover on one particular space. Secondly, the nerve theorem exactly states that the homotopic properties of $\mathcal{N}$ are preserved regardless of the underlying cover.

Although the statement of the nerve theory can be proved under slightly different conditions, the form given in [13] and [95] is most applicable to our work, and this is the one we present below.

Theorem 2.37 (Theorem 3.3, [95])**.** *Let $\mathcal{N}$ be the nerve of a closed finite cover $\{\mathcal{X}_j \mid j \in J\}$ with respect to $\mathcal{X}$. Assume that the following holds: $\cap_{j \in K} \mathcal{X}_j$ is AR for all $\emptyset \neq K \subset J$. Then, $\mathcal{X} \simeq \mathcal{N}$.*

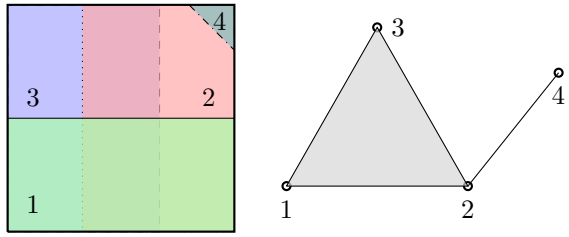

Figure 2.3: The figure on the left gives an example of a closed cover of a unit square $\mathcal{X} = [0,1] \times [0,1]$. The square is covered by four sets: one (green, denoted by 1, and bounded by a thin solid line) covering the bottom half, one (red, denoted by 2, and bounded by a dotted line) covering the right two thirds of the square, another (blue, denoted by 3, bounded by a dashed line) covering the left two thirds of the square, and the fourth one (gray, denoted by 4, bounded by a dash-dotted line) covering the upper right corner of the square. The corresponding nerve $\mathcal{N}$ is given by $\mathcal{N} = \{\{1\},\{2\},\{3\},\{4\},\{1,2\},\{1,3\},\{2,3\},\{2,4\},\{1,2,3\}\}$. The figure on the right gives a geometric realization of $\mathcal{N}$. We note that this nerve is indeed homotopy equivalent to the unit square, confirming Theorem 2.37.

The assumption that all $\cap_{j\in K}\mathcal{X}_j$ are AR stipulates that all $\mathcal{X}_j$ and all their intersections are reasonably "nice". By Theorem 2.33, this is satisfied, for instance, if all the intersections are compact, contractible and locally contractible. If a closed cover $\{\mathcal{X}_j \mid j \in J\}$ satisfies this assumption of Theorem 2.37, we say that it is *regular*.

2.9 Graph Theory and Matrices

We say that $P \in \mathbb{R}^{n\times n}$ is a *Z-matrix* if

$$[P]_{ij} \leq 0, \quad \text{for all } i,j \in \{1,\ldots,n\},\ i \neq j.$$

The set of all $n \times n$ Z-matrices is denoted by $\mathcal{Z}_n$. By the Perron-Frobenius theorem (e.g., [62]), Z-matrices have at least one real eigenvalue. If P is a Z-matrix, we denote its smallest real eigenvalue by $l(P)$. Matrix P is an *M-matrix* if it is a Z-matrix with

$$l(P) \geq 0. \tag{2.4}$$

For a detailed survey of Z- and M-matrices, see [11]. In particular, we note that there are dozens of equivalent definitions for M-matrices, which makes them particularly easy to work with.

A matrix is *irreducible* if its rows and columns cannot be simultaneously permuted to obtain a lower triangular matrix. By simultaneously permuting its rows and columns, any matrix P can be written in a (lower triangular) Frobenius normal form

$$P = \begin{bmatrix} P_{11} & 0 & \cdots & 0 \\ P_{21} & P_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ P_{p1} & P_{p2} & \cdots & P_{pp} \end{bmatrix},$$

where each block $P_{11}, \ldots, P_{pp}$ is either irreducible or a 1×1 zero-matrix. For more background, see [7].

In the following paragraph we use the notation from [58, 59]. Assume that P is in Frobenius normal form, with p^2 blocks. Let $R(P) = (V, E)$ be a directed graph with the vertex set $V = \{1, \ldots, p\}$ and edges

$$E = \{(i, j) \in V \times V \mid P_{ij} \neq 0,\, i \neq j\}.$$

Definition 2.38. *We say that i* accesses *j (alternatively, j* is accessed by *i) if $i = j$, or there is a path in $R(P)$ going from i to j. We denote this by $i \longrightarrow j$.*

Definition 2.39. *If $W \subset \{1, \ldots, p\}$, we say that W accesses j (denoted by $W \longrightarrow j$) if $i \longrightarrow j$ for some $i \in W$.*

Define

$$\text{above}(W) = \{j \in \{1, \ldots, p\} \mid W \longrightarrow j\}. \tag{2.5}$$

Let P be a Z-matrix in Frobenius normal form. Then, all blocks P_{ii} are Z-matrices as well. We define

$$\begin{aligned} S &= \{i \in \{1, \ldots, p\} \mid P_{ii} \text{ is singular}\}, \\ T &= \{i \in \{1, \ldots, p\} \mid l(P_{ii}) < 0\}. \end{aligned} \tag{2.6}$$

The following two propositions are known, and will be the key for obtaining a graph-theoretic characterization of the solvability of the RCP in Chapter 7:

Proposition 2.40 (Corollary 5.13, [58])**.** *Suppose P is a Z-matrix in Frobenius normal form. The linear equation $Py = 0$ has a solution $y > 0$ if and only if $S \backslash \text{above}(T) \neq \emptyset$.*

Proposition 2.41 (Theorem 3.11, [59])**.** *Suppose P is a Z-matrix in Frobenius normal form, and $a \geq 0$. The linear equation $Py = a$ has a solution $y \geq 0$ if and only if $\text{supp}(a) \cap \text{above}(S \cup T) = \emptyset$.*

The above definitions and results are illustrated by Example 2.42.

Example 2.42. *Let $P \in \mathbb{R}^{7\times 7}$ be a Z-matrix in Frobenius normal form given by*

$$P = \left(\begin{array}{cc|c|c|c|c|c} -2 & -2 & 0 & 0 & 0 & 0 & 0 \\ -2 & -2 & 0 & 0 & 0 & 0 & 0 \\ \hline -3 & -4 & 4 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ \hline -1 & -1 & 0 & -6 & 0 & 0 & 0 \\ \hline 0 & 0 & -7 & 0 & 0 & 5 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -4 & -3 \end{array}\right), \tag{2.7}$$

where the blocks are separated by vertical and horizontal lines.

We can easily verify that $p = 6$, $S = \{1, 4\}$, $T = \{1, 6\}$, and the edges between vertices generated by P can be graphically represented as in Figure 2.4.

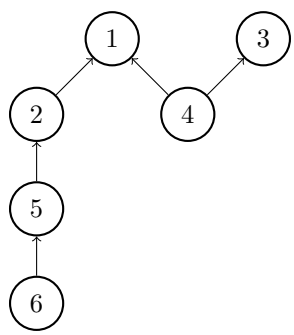

Figure 2.4: Graph $R(P)$ generated by matrix P from (2.7).

From Figure 2.4 we note that $\text{above}(T) = \{1, 2, 5, 6\}$ *and* $\text{above}(S \cup T) = \{1, 2, 3, 4, 5, 6\}$. *Let us revisit the results of Proposition 2.40 and Proposition 2.41. Since* $S \backslash \text{above}(T) = \{4\} \neq \emptyset$, *Proposition 2.40 guarantees that $Py = 0$ has a solution $y > 0$. This can be easily verified: for instance, $y = e_5$ satisfies $Py = 0$.*

Let $a \geq 0$. Since $\text{above}(S \cup T)$ *equals the entire set* $\{1, 2, 3, 4, 5, 6\}$, *Proposition 2.41 guarantees that the linear equation $Py = a$ will have a solution $y \geq 0$ if and only if* $\text{supp}(a) = \emptyset$, *i.e., $a = 0$.* □

Chapter 3

The Reach Control Problem

This chapter formally defines the Reach Control Problem (RCP) on a simplex. The RCP is the fundamental problem that this thesis is dealing with. We also introduce the related geometric notions to be used in the remainder of the thesis, and provide a brief recount of the crucial results on the solvability of the RCP.

3.1 Problem Statement

Let $\mathcal{S} := \mathrm{co}\{v_0, v_1, \ldots, v_n\} \subset \mathbb{R}^n$ be an n-dimensional simplex with vertices $v_0, \ldots, v_n$. Its facets shall be denoted by $\mathcal{F}_0, \ldots, \mathcal{F}_n$, where each facet is indexed by the vertex it does not contain. Furthermore, for each $i \in \{0, \ldots, n\}$ let h_i be the unit normal vector to the facet $\mathcal{F}_i$ pointing outside the simplex. For a figure illustrating the above notation, see Figure 3.1, adapted from [89] and [106].

We consider the system

$$\dot{x} = Ax + Bu + a, \quad x \in \mathcal{S}, \tag{3.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $a \in \mathbb{R}^n$. Define $\mathcal{B} = \mathrm{Im}(B)$.

As mentioned in the introduction, the Reach Control Problem is to find a closed-loop control feedback which results in every trajectory of (3.1) leaving $\mathcal{S}$ through $\mathcal{F}_0$ in finite time. We now give a formal definition of this problem. In order for Problem 3.1 to make sense, it is assumed that the dynamics (3.1) are extended to a small neighborhood of $\mathcal{S}$. However, as the control functions we investigate are normally trivially extendable to all of $\mathbb{R}^n$, this feature will not be recalled in the remainder of the thesis, except in Theorem 4.25 where the dynamics are extended in a non-obvious way.

Problem 3.1 (Reach Control Problem (RCP))**.** *Let $\mathcal{F}$ be a class of functions $u : \mathcal{S} \to \mathbb{R}^m$. Consider*

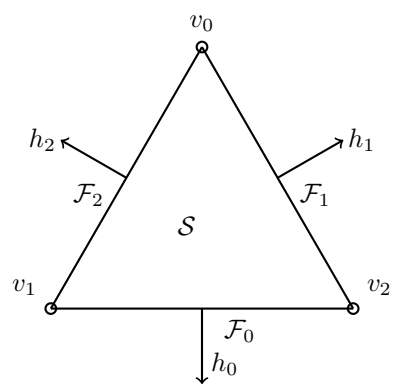

Figure 3.1: An illustration of basic notation relating to the simplex used in the thesis. The polytope $\mathcal{S} = \text{co}\{v_0, v_1, v_2\}$ is given by vertices $V_{\mathcal{S}} = \{v_0, v_1, v_2\}$ and facets $\mathcal{F}_0$, $\mathcal{F}_1$, and $\mathcal{F}_2$, with each facet indexed by the vertex it does not contain. h_i is the unit normal vector pointing out of $\mathcal{S}$.

system (3.1)*. Let* $\phi_u(\cdot, x_0)$ *be the trajectory generated by system* (3.1)*, with* $\phi_u(0, x_0) = x_0$*. Determine whether there exists a feedback control* $u \in \mathcal{F}$ *such that for every* $x_0 \in \mathcal{S}$*, there exist* $T \geq 0$ *and* $\varepsilon > 0$ *such that*

(i) $\phi_u(t, x_0) \in \mathcal{S}$ *for all* $t \in [0, T]$*.*

(ii) $\phi_u(T, x_0) \in \mathcal{F}_0$*.*

(iii) $\phi_u(T + \varepsilon, x_0) \notin \mathcal{S}$ *for all* $t \in (T, T + \varepsilon)$*.*

In line with the previous work, in the remainder of this chapter we use the shorthand notation $\mathcal{S} \xrightarrow{\mathcal{S}} \mathcal{F}_0$ to denote that Problem 3.1 is solved by some control.

As outlined in Section 1.3, a number of different classes $\mathfrak{F}$ have been investigated in previous work on the RCP, including affine feedback [47, 49, 50, 89, 121, 122, 125, 136], continuous state feedback [23, 125], continuous piecewise affine feedback [49, 50, 54, 56, 69], discontinuous piecewise affine feedback [24], multi-affine feedback [10, 55], and time-varying feedback [5]. With the partial exception of the material of Chapter 4, this thesis is focused on affine and continuous state feedbacks.

Remark 3.2. *Note that we assume that the trajectories* $\phi_u(\cdot, x_0)$ *defined in Problem 3.1 are unique. This is clearly true for affine, continuous piecewise affine and continuous state feedback. Moreover, it can be shown that it is also true for all classes of control laws currently investigated in the RCP; for more details on this, we invite the reader to consult Chapter 4.*

Conditions (i)-(iii) in Problem 3.1 ensure that closed-loop trajectories of system (3.1) leave $\mathcal{S}$ in finite time through the exit facet $\mathcal{F}_0$. We note that an analogous statement of Problem 3.1 can be given for

general polytopes, and this is the Reach Control Problem for polytopes. For the sake of completeness, such a statement is given as Problem 4.1 in Chapter 4. However, with the partial exception of Chapter 4 and Chapter 9, the work presented in this thesis concerns the RCP on a simplex, and for the sake of clarity, we seek to limit this exposition to the RCP on a simplex.

The overarching goal of reach control research has been to develop easily computable sufficient and necessary conditions for the solvability of the RCP, in various above mentioned classes of feedback.

3.2 Solvability of the RCP

A large part of the research progress that was attained on the RCP is ultimately based on two necessary conditions for solvability of the RCP. These conditions are fundamental in the work we are presenting in this thesis as well. We describe them below.

It is clear that a necessary condition for a control law u to solve Problem 3.1 is that it generates no equilibria in $\mathcal{S}$.

Proposition 3.3. *If a continuous state feedback $u : \mathcal{S} \to \mathbb{R}^m$ solves Problem 3.1, then*

$$Ax + Bu(x) + a \neq 0, \quad \textit{for all } x \in \mathcal{S}. \tag{3.2}$$

An additional necessary condition for solving the RCP by continuous state feedback is that the velocity vectors $Ax + Bu(x) + a$ must point into the simplex at points x in the set $\partial\mathcal{S}\backslash\mathcal{F}_0$ [50]. If this was not satisfied for some $x_0 \in \mathcal{S}$, the trajectory $\phi_u(t, x_0)$ would leave $\mathcal{S}$ through a facet other than $\mathcal{F}_0$. The formalization of this requirement are the *invariance conditions*. We first provide the geometric notions necessary to state these conditions.

Let $I = \{1, \ldots, n\}$. For $x \in \mathcal{S}$, let $I(x)$ be the smallest subset of I such that $x \in \mathrm{co}\{v_i \mid i \in I(x)\}$. Let $J(x) = I\backslash I(x)$. Define the closed, convex cone

$$\mathcal{C}(x) := \left\{ \; y \in \mathbb{R}^n \mid h_j \cdot y \leq 0, \; j \in J(x) \; \right\}, \tag{3.3}$$

with the convention that $\mathcal{C}(x) = \mathbb{R}^n$ if $J(x) = \emptyset$.

Remark 3.4. *Note that $J(x)$ is the set of all indices j such that $x \in \mathcal{F}_j$. Also, it can be easily shown that for any $x \in \mathcal{S} \setminus \mathcal{F}_0$, $\mathcal{C}(x)$ equals*

$$T_{\mathcal{S}}(x) = \{v \in \mathbb{R}^n \mid \liminf_{t \to 0+} \frac{\min_{y \in \mathcal{S}} \|x + tv - y\|}{t} = 0\},$$

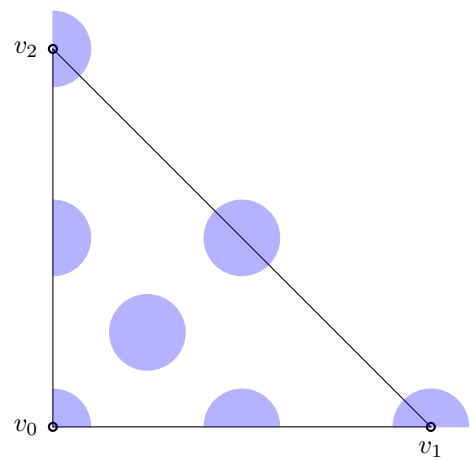

Figure 3.2: An illustration of cones $\mathcal{C}(x)$, depicted as blue cones at several points $x \in \mathcal{S}$.

the Bouligand tangent cone to $\mathcal{S}$ at x. For further details on the Bouligand tangent cone, see [31]. At points $x \in \mathcal{F}_0$, $\mathcal{C}(x)$ and $T_{\mathcal{S}}(x)$ are different since $\mathcal{C}(x)$ includes directions pointing out of $\mathcal{S}$ through $\mathcal{F}_0$.

Figure 3.2, modified from [103], illustrates the cones $\mathcal{C}(x)$ attached to corresponding $x \in \mathcal{S}$.

Finally, we formally state the invariance, or inward pointing, conditions.

Proposition 3.5 ([50])**.** *If a continuous state feedback $u : \mathcal{S} \to \mathbb{R}^m$ solves Problem 3.1, then*

$$Ax + Bu(x) + a \in \mathcal{C}(x), \quad \text{for all } x \in \mathcal{S}. \tag{3.4}$$

We note that the requirement of continuity of u is important in Proposition 3.5. If the right hand side of system (3.1) is discontinuous, the solution to Problem 3.1 is not defined in the usual sense. See Chapter 4 for further discussion on the solutions to Problem 3.1 in this setting.

It was shown in [47, 122] that, when combined, the above two necessary conditions for solvability of the RCP by continuous state feedback also form a sufficient condition for solvability of the RCP by affine feedback. This is the fundamental result of this section.

Theorem 3.6. *Let $u : \mathcal{S} \to \mathbb{R}^m$ be an affine function. Then, u solves Problem 3.1 if and only if conditions* (3.2) *and* (3.4) *hold.*

Bearing in mind the convexity of affine functions and of simplex $\mathcal{S}$, Theorem 3.6 can be alternatively posed in the following way:

Corollary 3.7. *Problem 3.1 is solvable by affine feedback if and only if there exist $u_0, \ldots, u_n \in \mathbb{R}^m$ such that:*

(i) $Av_i + Bu_i + a \in \mathcal{C}(v_i)$ *for all* $i \in \{0, \ldots, n\}$,

(ii) the affine function $u = Kx + g$ *given by* $u(v_i) = u_i$ *for all* $i \in \{0, \ldots, n\}$ *satisfies* $Ax + Bu(x) + a \neq 0$ *for all* $x \in \mathcal{S}$.

If conditions (i) and (ii) are satisfied, then affine feedback u *defined in (ii) solves Problem 3.1.*

Corollary 3.7 more vividly illustrates the problem with the necessary and sufficient conditions obtained in [47, 122]. While condition (i) is, by (3.3), merely a system of linear feasibility problems, condition (ii) does not lend itself to any obvious method of synthesis: for a given affine function u that satisfies (i), it is trivial to check whether it satisfies (ii), but it is difficult to find whether a function satisfying (i) and (ii) exists in general. This motivates the search for new sufficient and necessary conditions for the solvability of the RCP, which is the underlying motivation for Chapters 5–8. Before we attack this problem, however, let us discuss a fundamental technical issue with the definition of the RCP.

Chapter 4

Chattering in Reach Control

Having defined the RCP in the previous chapter, we now expand on a mid-level issue alluded to in the introductory chapter. In particular, we discuss how to ensure that a feedback control which solves the RCP on a single simplex of the state space also drives a system trajectory to automatically enter the desired adjacent simplex. Related to this question, this chapter also explores Zeno-like chattering phenomena and switching policies in reach control. With abridgements and modifications done in the interest of clarity and narrative of the thesis, the material of this chapter is largely equal to the material of [99].

4.1 Introduction and Related Literature

In the past period there has been a significant effort to formalize the mathematical foundations of switched and hybrid control systems, not limited to reach control. Due to the discontinuous nature of such systems, fundamental results guaranteeing the existence and uniqueness of solutions of classical ODEs with continuous vector fields no longer hold automatically. A new theory of existence and uniqueness of solutions of switched and hybrid systems has been formulated in, among other, [52, 63, 80].

An additional property specific to switched and hybrid systems is *Zeno behaviour*, in which a trajectory, even if guaranteed to exist and be unique, undergoes an infinite number of switches, i.e., discontinuous changes in the governing vector field, in a finite time interval. This property has been the subject of intense recent research, e.g., in [2, 28, 44, 60].

Additionally, a number of classical control notions such as controllability, observability, and Lyapunov stability do not apply to systems governed by discontinuous vector fields. There has been significant work to extend these concepts to switched and hybrid systems (see, e.g., [19, 39, 112]). For a comprehensive

treatment of discontinuous dynamical systems see [32, 76].

This chapter follows the above line of research on deepening the mathematical foundations of hybrid control system. While there is extensive material on reach control theory outlined in Section 1.3, this is the first effort to provide a formal and complete discussion of existence, uniqueness, and behaviour of solutions in reach control.

As outlined in Chapter 1, the intention is for the RCP to serve as a building block in reach control theory. To remind the reader, reach control theory rests upon partitioning the state space into simplices or polytopes to achieve some control objective. This chapter focuses on what happens when trajectories transition from one polytope to the next. In order to make the transitions between polytopes work, it is not only necessary for a trajectory to exit a simplex $\mathcal{S}$ with its last point in $\mathcal{S}$ lying on the desired exit facet $\mathcal{F}_0$, as stipulated by the RCP. We must ensure that this exit will simultaneously result in the trajectory entering the next simplex in the desired sequence. This chapter investigates a fundamental question in reach control theory which has not been addressed by previous work on general hybrid or switched systems: *what is the appropriate notion of leaving a simplex or a polytope through a facet?* In the sense of the results of this chapter, there is no major difference between the work on a simplex or on a polytope, so we largely focus on the polytopal setting.

The above question has never been explicitly examined. As mentioned in Section 1.3, in [49, 50] it was required that velocity vectors must point strictly outside the polytope $\mathcal{P}$ at points in its exit facet. This condition implies that a trajectory arriving at the exit facet will immediately enter the open half-space outside $\mathcal{P}$ and bounded by the exit facet. Sufficient conditions were given in [50] for a Lipschitz continuous feedback to solve this problem. The proof assumes strict inward or outward conditions on velocity vectors along the facets of the polytope. When these conditions are not strict, certain pathologies can arise, as this chapter shows, and arguments about whether trajectories lie in certain half-spaces with respect to facets are considerably more delicate. Lemma 3 of [122] regards trajectories exiting a polytope via a facet but without necessarily crossing into the outer open half-space. In Section 4.3.4 we provide a complete proof of a stronger version of this result.

This chapter achieves two objectives: first, it explores the relationship between the two notions for exiting a polytope. We consider the following question: *Is it possible for a trajectory to leave a polytope $\mathcal{P}$ but without crossing into an outer half-space?* When a trajectory exits $\mathcal{P}$ but does not cross into an outer half-space, we say it *chatters*. The second objective is to identify appropriate classes of feedback controls that do not allow chattering.

4.2 The RCP on a Polytope

We briefly introduce the polytopal version of the RCP set-up from Chapter 3. We note that the differences between the two set-ups are minimal, but we provide this description for the sake of completeness. Consider an n-dimensional convex polytope $\mathcal{P} = \mathrm{co}\{v_0, \ldots, v_p\} \subset \mathbb{R}^n$ with vertex set $\{v_0, \ldots, v_p\}$. Let $\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_r$ denote the facets of $\mathcal{P}$. The facet $\mathcal{F}_0$ is referred to as the *exit facet*, while $\mathcal{F}_1, \ldots, \mathcal{F}_r$ are called *restricted facets*. Let $J = \{1, \ldots, r\}$ and let h_i be the unit normal to each facet $\mathcal{F}_i$ pointing outside the polytope. We consider the affine control system 3.1 defined on $\mathcal{P}$, and give the polytopal version of Problem 3.1.

Problem 4.1 (Reach Control Problem (RCP))**.** *Consider system* (3.1) *defined on* $\mathcal{P}$*. Find a state feedback* $u : \mathcal{P} \to \mathbb{R}^n$ *such that for every* $x_0 \in \mathcal{P}$*, there exist* $T \geq 0$ *and* $\varepsilon > 0$ *such that*

(i) $\phi(t, x_0) \in \mathcal{P}$ *for all* $t \in [0, T]$*,*

(ii) $\phi(T, x_0) \in \mathcal{F}_0$*, and*

(iii) $\phi(t, x_0) \notin \mathcal{P}$ *for all* $t \in (T, T + \varepsilon)$*.*

Analogously to Problem 3.1, the statement of Problem 4.1 implicitly assumes that the trajectory $\phi(\cdot, x_0)$ exists and is unique. This is a known result in the case where u is continuous. In the case of discontinuous feedback u and open-loop controls, both discussed in Section 4.3.5, this is true in the sense of Carathéodory: see, e.g., Chapter I.5 of [51] for a longer discussion. In other words, there exists a unique (absolutely) continuous function $\phi(\cdot, x_0)$ such that

$$\phi(t, x_0) = x_0 + \int_0^t \big[A\phi(\tau, x_0) + Bu\big(\phi(\tau, x_0)\big) + a\big]\, d\tau, \quad \text{for all } t \geq 0.$$

All definitions from Chapter 3 transfer to the polytopal case of this chapter. Cones $\mathcal{C}(x)$ are particularly important. Analogously to what we noticed in Remark 3.4, we define $J(x) = \{j \in J \mid x \in \mathcal{F}_j\}$, and then define $\mathcal{C}(x)$ the same as in (3.3).

We now come to the central issue studied in this chapter: how trajectories exit $\mathcal{F}_0$. First, it can be seen that a notion of trajectories exiting a facet of $\mathcal{P}$ appears implicitly in the statement of the RCP. This notion is formalized as follows.

Definition 4.2. *We say* $\phi(\cdot, x_0)$ exits $\mathcal{P}$ through facet $\mathcal{F}_0$ *if there exist* $T \geq 0$ *and* $\varepsilon > 0$ *such that*

(i) $\phi(t, x_0) \in \mathcal{P}$ *for all* $t \in [0, T]$*;*

(ii) $\phi(T, x_0) \in \mathcal{F}_0$*;*

(iii) $\phi(t, x_0) \notin \mathcal{P}$ *for all* $t \in (T, T+\varepsilon)$.

We say $\phi(\cdot, x_0)$ exits $\mathcal{P}$ *if it exits* $\mathcal{P}$ *through some facet of* $\mathcal{P}$.

Despite the fact that the statement of the RCP in Problem 4.1 is clear on the meaning of exiting a facet, other notions are used in the literature, as discussed in Section 1.3 and Section 4.1. We now define a stronger notion of exiting a facet compared to Definition 4.2. To that end, let $\mathcal{H}_i$ be the closed half-space bounded by $\operatorname{aff}(\mathcal{F}_i)$ which contains $\mathcal{P}$; that is

$$\mathcal{H}_i = \{x \in \mathbb{R}^n \mid h_i \cdot (x - y) \leq 0\}, \tag{4.1}$$

where y is any point in $\mathcal{F}_i$.

Definition 4.3. *We say* $\phi(\cdot, x_0)$ crosses *facet* $\mathcal{F}_i$ *if there exist* $T \geq 0$ *and* $\varepsilon > 0$ *such that*

(i) $\phi(t, x_0) \in \mathcal{P}$ *for all* $t \in [0, T]$*;*

(ii) $\phi(T, x_0) \in \mathcal{F}_i$*;*

(iii) $\phi(t, x_0) \notin \mathcal{H}_i$ *for all* $t \in (T, T+\varepsilon)$.

Both notions of exiting the polytope — either *exiting through* a facet or by *crossing* a facet — have appeared in the literature, often with little distinction made between the two. Clearly, if a trajectory crosses a facet $\mathcal{F}_i$, then the trajectory exits $\mathcal{P}$ through $\mathcal{F}_i$. But the converse statement is not true. In fact, it is possible for a trajectory to exit $\mathcal{P}$, but not cross any of its facets. The following example illustrates the dichotomy.

Example 4.4. *Let* $\mathcal{P} = \operatorname{co}\{v_0, v_1, v_2, v_3\}$ *be a simplex with* $v_0 = (0,0,1)$*,* $v_1 = (1,0,0)$*,* $v_2 = (0,1,0)$*, and* $v_3 = (0,0,0)$*. Let the facets of* $\mathcal{P}$ *be* $\mathcal{F}_0, \ldots, \mathcal{F}_3$*, each facet indexed by the vertex it does not contain. Consider the nonlinear system* $\dot{x} = f(x)$*, where*

$$f(x) := \begin{pmatrix} \frac{e^{-1/x_3^2}(x_3\cos(1/x_3) - 2\sin(1/x_3) - 1.98)}{x_3^3} \\ \frac{e^{-1/x_3^2}(-2\cos(1/x_3) - x_3\sin(1/x_3) - 1.98)}{x_3^3} \\ 1 \end{pmatrix},$$

with

$$f(\cdot, \cdot, 0) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

It can be verified that $f \in C^\infty$. For the initial condition $x_0 = 0$, the solution is

$$\phi(t, x_0) = \begin{pmatrix} -e^{-1/t^2}(\sin\frac{1}{t} + 0.99) \\ -e^{-1/t^2}(\cos\frac{1}{t} + 0.99) \\ t \end{pmatrix}, \quad \text{for all } t > 0. \tag{4.2}$$

First, we will show that $\phi(\cdot, x_0)$ exits $\mathcal{P}$ at time $T = 0$ in the sense of Definition 4.2. Since $\mathcal{P} = \{x \in \mathbb{R}^3 \mid x_1, x_2, x_3 \geq 0, x_1 + x_2 + x_3 \leq 1\}$, it suffices to prove that for each $t > 0$, at least one of the coordinates of $\phi(t, x_0)$ is always negative. Suppose otherwise. Then, by (4.2), for some $t > 0$, $-e^{-1/t^2}(\cos(1/t) + 0.99) \geq 0$ and $-e^{-1/t^2}(\sin(1/t) + 0.99) \geq 0$. This implies $\cos(1/t), \sin(1/t) \leq -0.99$, which contradicts $\cos^2(1/t) + \sin^2(1/t) = 1$. Hence, $\phi(\cdot, x_0)$ exits $\mathcal{P}$ at time $T = 0$.

On the other hand, $\phi(\cdot, x_0)$ does not cross any facet $\mathcal{F}_i$ at $T = 0$ in the sense of Definition 4.3. Since $\phi(0, x_0) = 0 \notin \mathcal{F}_3$, $\phi(\cdot, x_0)$ clearly does not cross $\mathcal{F}_3$ at time $T = 0$. Observe that $\mathcal{H}_0 = \{x \in \mathbb{R}^3 \mid x_3 \geq 0\}$, and $\mathcal{H}_i = \{x \in \mathbb{R}^3 \mid x_i \geq 0\}$ for $i = 1, 2$. From (4.2) we note that $\phi(t, x_0) \in \mathcal{H}_0$ for all $t \geq 0$. Hence, $\phi(\cdot, x_0)$ does not cross $\mathcal{F}_0$ at $T = 0$.

Next suppose that $\phi(\cdot, x_0)$ crosses $\mathcal{F}_1$ at time $T = 0$. Then, by (iii) in Definition 4.3 and by (4.1), there exists $\varepsilon > 0$ such that $-e^{-1/t^2}(\cos(1/t) + 0.99) < 0$ for all $0 < t < \varepsilon$. However, this is impossible: take $t = 1/((2k+1)\pi)$ with $k \in \mathbb{N}$ to obtain $-e^{-1/t^2}(\cos(1/t)+0.99) = 0.01e^{-1/t^2} > 0$. Analogously, there does not exist $\varepsilon > 0$ such that $-e^{-1/t^2}(\sin(1/t)+0.99) < 0$ for all $0 < t < \varepsilon$, by taking $t = 1/((2k+3/2)\pi)$ for $k \in \mathbb{N}$. Hence, $\phi(\cdot, x_0)$ does not cross $\mathcal{F}_2$. We conclude that $\phi(\cdot, x_0)$ does not cross any facet of $\mathcal{P}$, yet, $\phi(\cdot, x_0)$ exits $\mathcal{P}$ at time $T = 0$.

Example 4.4 is clearly rather artificial: it was obtained by finding two smooth real functions which both oscillate between positive and negative values when approaching 0, but never attain a positive value at the same time. This results in the situation where the state trajectory constantly switches between $\mathcal{H}_1 \backslash \mathcal{H}_2$ and $\mathcal{H}_2 \backslash \mathcal{H}_1$. We will call such a pathological behaviour *chattering*.

Definition 4.5. *We say the trajectory $\phi(\cdot, x_0)$* chatters *if it exits $\mathcal{P}$, but does so without crossing any of its facets.*

As shown above, the trajectory in Example 4.4 satisfies the condition in Definition 4.5, i.e., this trajectory chatters. An illustration of a chattering trajectory in two dimensions was also already given in the left of Figure 1.2. The trajectory in that figure is a variation of the trajectory generated in Example 4.4. The set-up is two-dimensional, but that is because the control used to generate such a trajectory is time-varying, whereas in Example 4.4 the role of time has been replaced by x_3.

A question of practical interest in design, and a central question of this section, is to determine which classes of feedback controls used to solve the RCP guarantee that no chattering occurs. However, before proceeding, let us briefly note that chattering is indeed not the only undesirable behaviour that may occur when exiting a simplex. In the ideal scenario, we would want the trajectory to exit and cross the same set of facets. Definition 4.5, i.e., the notion of chattering, covers the pathological situation when the trajectory exits through a nonempty set of facets, but the set of crossed facets is empty. Other combinations of exiting and crossing facets are also possible. For instance, Figure 4.1 gives an illustration of a trajectory which exits a polytope through two facets, but only crosses one of them.

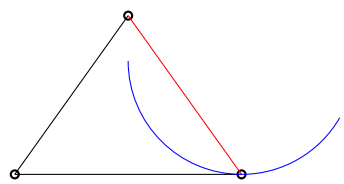

Figure 4.1: An example of non-chattering pathological behaviour in the RCP. The trajectory, marked in blue, exits simplex $\mathcal{S}$ through two facets, but crosses only one of them (marked in red).

While we will briefly discuss the type of behaviour from Figure 4.1 in the case of discontinuous feedback in Section 4.3.5, in this chapter we primarily focus on the issue of chattering as given in Definition 4.5 since it exhibits a more pathological behaviour. As mentioned, we are interested in the following problem:

Problem 4.6. *Find conditions on the vector field in* (3.1) *such that no trajectory of that system chatters.*

Example 4.4 shows that smoothness is not sufficient to satisfy Problem 4.6. In Section 4.3 we will show that analyticity of the vector field is such a sufficient condition. In turn, this finding answers Problem 4.6 for standard classes of feedback used in the RCP. In Section 4.3.2 we extend the results to continuous piecewise affine vector fields. We use this extension to provide the missing proof of Lemma 13 from [69] in Section 4.3.3. Additionally, using the result that affine feedback guarantees there is no chattering, we complete the proof of Lemma 3 from [122] in Section 4.3.4. Finally, in Section 4.3.5 we revisit the work of [24] on discontinuous piecewise affine feedback. We identify an issue with the control law proposed in that paper, propose an amended control law, and prove that chattering does not occur in the context of that control law.

4.3 Main Results

We begin our investigation of Problem 4.6 by identifying the key property of a vector field that closes the gap between Definition 4.2 and Definition 4.3 regarding how trajectories exit $\mathcal{P}$. Consider the nonlinear

system

$$\dot{x} = f(x) \tag{4.3}$$

where $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$.

Lemma 4.7. *Let $\mathcal{P}$ be an n-dimensional convex polytope. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be analytic at x_0, and let $\phi(\cdot, x_0)$ be the unique solution of (4.3) with $\phi(0, x_0) = x_0$. Then for every $h \in \mathbb{R}^n$, there exists $\varepsilon > 0$ such that either $h \cdot (\phi(t, x_0) - x_0) \leq 0$ for all $t \in (0, \varepsilon)$ or $h \cdot (\phi(t, x_0) - x_0) > 0$ for all $t \in (0, \varepsilon)$. Moreover, if $\phi(\cdot, x_0)$ exits $\mathcal{P}$ at time $T = 0$, then there exists a facet $\mathcal{F}_i$ of $\mathcal{P}$ such that $\phi(\cdot, x_0)$ crosses $\mathcal{F}_i$ at $T = 0$.*

Proof. By the Cauchy-Kowalevski theorem (see, e.g., [34, 40]), $\phi(t, x_0)$ is guaranteed to be analytic in t on some interval $(-\varepsilon_1, \varepsilon_1)$, where $\varepsilon_1 > 0$. Considering any $h \in \mathbb{R}^n$, $g(\cdot) := h \cdot (\phi(\cdot, x_0) - x_0)$ is a real analytic function of one variable. Thus, we can write

$$g(t) = \sum_{k=0}^{\infty} \frac{g^{(k)}(0)}{k!} t^k, \quad t \in (-\varepsilon_1, \varepsilon_1). \tag{4.4}$$

For the first claim, suppose it is not true that $h \cdot (\phi(t, x_0) - x_0) > 0$ on some interval $(0, \varepsilon)$. Also, suppose it is not true that $h \cdot (\phi(t, x_0) - x_0) \leq 0$ on some interval $(0, \varepsilon)$. Then there exists a sequence of times $(t_{n,1}^{(0)}) > 0$, $n \in \mathbb{N}$, with $t_{n,1}^{(0)} \to 0+$ such that $g(t_{n,1}^{(0)}) > 0$ for all $n \in \mathbb{N}$, and a sequence of times $(t_{n,2}^{(0)}) > 0$, $n \in \mathbb{N}$, with $t_{n,2}^{(0)} \to 0+$ such that $g(t_{n,2}^{(0)}) \leq 0$ for all $n \in \mathbb{N}$. Clearly, for every time $t_{n,1}^{(0)}$ there exists a time $0 < t_{n',2}^{(0)} < t_{n,1}^{(0)}$, and $g(t_{n',2}^{(0)}) \leq 0 < g(t_{n,1}^{(0)})$. By the Mean Value Theorem there exists a sequence of times $t_{n,1}^{(1)} \to 0+$ such that $g'(t_{n,1}^{(1)}) > 0$. Analogously, there exists a sequence of times $t_{n,2}^{(1)} \to 0+$ such that $g'(t_{n,2}^{(1)}) \leq 0$.

Proceeding in the same manner to the next derivative, we obtain that for all $k \geq 0$ there exist sequences $t_{n,1}^{(k)}, t_{n,2}^{(k)} \to 0+$ such that

$$g^{(k)}(t_{n,1}^{(k)}) > 0,\ g^{(k)}(t_{n,2}^{(k)}) \leq 0. \tag{4.5}$$

Since g is analytic around 0, all $g^{(k)}$ are continuous at 0. Hence, by letting $t_{n,1}^{(k)}, t_{n,2}^{(k)} \to 0+$, from (4.5) we have $g^{(k)}(0) = 0$. By (4.4), $g(t) = 0$ on some interval around 0. This contradicts our assumption that it is not true that $g(t) \leq 0$ on some interval $(0, \varepsilon)$. This proves our first claim.

For the second claim, suppose it is incorrect. That is, $\phi(\cdot, x_0)$ leaves $\mathcal{P}$ at time $T = 0$, but does not cross any facets $\mathcal{F}_i \subset \mathcal{P}$. Thus, for all facets $\mathcal{F}_i$ such that $\phi(0, x_0) = x_0 \in \mathcal{F}_i$, there is no interval $(0, \varepsilon)$ such that $h_i \cdot (\phi(t, x_0) - x_0) > 0$. By our first claim, this implies that there exists some $\varepsilon > 0$ such that for all $\mathcal{F}_i$ such that $x_0 \in \mathcal{F}_i$, $h_i \cdot (\phi(t, x_0) - x_0) \leq 0$ for all $t \in (0, \varepsilon)$. We also know that $x_0 \in \mathcal{P}$, so

for all $\mathcal{F}_i$ such that $x_0 \notin \mathcal{F}_i$, $h_i \cdot (\phi(t, x_0) - x_0) < 0$ for all $t \in (0, \varepsilon')$, for some sufficiently small $\varepsilon' > 0$. Combining the previous two statements, $\phi(t, x_0) \in \mathcal{P}$ for all $t \in (0, \min\{\varepsilon, \varepsilon'\})$. This contradicts that $\phi(\cdot, x_0)$ leaves $\mathcal{P}$ at time $T = 0$. □

4.3.1 Chattering Under Feedback

Lemma 4.7 provides the mathematical foundation to resolve the issue of chattering in previous work. We review the different classes of feedback control previously investigated in the literature and indicate how Lemma 4.7 applies to each class.

- **Affine feedback** was investigated in [47, 49, 50, 89, 121, 122, 125, 136], and is further explored in Chapter 6, Chapter 7, and Chapter 8 of this thesis. Because affine maps are analytic, Lemma 4.7 applies directly and shows that chattering does not occur in the RCP with affine feedback. In Section 4.3.4 we further discuss the exit behaviour of trajectories under affine feedback.

- **Continuous feedback** was investigated in [23, 125], and is also the setting of Chapter 5 of this thesis. As shown in Example 4.4, Lemma 4.7 cannot apply to continuous maps in general, or even to smooth maps. As partial recourse, we note that by the Weierstrass Approximation Theorem (see, e.g., [33]), every continuous function $f : C \to \mathbb{R}^n$ defined on a compact subset $C \subset \mathbb{R}^n$ can be approximated arbitrarily well by a map $(p_1, p_2, \ldots, p_n) : C \to \mathbb{R}^n$, where $p_1, \ldots, p_n$ are multivariate polynomials. Hence, every continuous system on a polytope can be perturbed slightly to obtain an analytical system, which is guaranteed to be non-chattering by Lemma 4.7.

- **Continuous piecewise affine feedback** was explored in [49, 50, 54, 56]. Unlike affine feedback, piecewise affine feedback results in the vector field f having a non-analytic structure. However, we will show that Lemma 4.7 is extendable to continuous piecewise affine vector fields. This is the topic of Section 4.3.2. This class of feedback arises in the *Output Reach Control Problem* (ORCP) investigated in [69], which we discuss in Section 4.3.3.

- **Discontinuous piecewise affine feedback** was explored in [24]. While the control law proposed in that paper suffers from minor inconsistencies discussed and resolved in Section 4.3.5, it can be shown that chattering does not occur in that setting as well. We provide a proof of that claim in Section 4.3.5.

- **Multi-affine feedback** was explored in [10, 55]. Multi-affine functions have the form

$$u(x_1, \ldots, x_n) = \sum_{i_1, \ldots, i_n \in \{0,1\}} c_i x_1^{i_1} \cdots x_n^{i_n},$$

so they are a special case of polynomials in $\mathbb{R}^n$. Hence, multi-affine feedback is analytic, and Lemma 4.7 proves that chattering does not occur in this case as well.

- **Time-varying feedback** was explored in [5]. Time-varying feedback does not fall into the setting explored in this chapter. However, Theorem 16 of [5] expresses the time-varying feedback proposed to solve the RCP in that paper as a multi-affine feedback on the extended state space $\mathcal{S} \times [0, 1]$, where $\mathcal{S}$ is the original state space. Thus, Lemma 4.7 guarantees that there will be no chattering in this extended system using the feedback given in [5].

4.3.2 Continuous Piecewise Affine Feedback

Because piecewise affine maps are not analytic at points where switches between two pieces occur, piecewise affine feedback generally results in the vector field f lacking an analytic structure. Hence, Lemma 4.7 does not apply directly. However, continuous piecewise affine maps still contain much more structure than general continuous maps. We will show that they admit a piecewise analytic structure, as defined in [128]. By invoking Theorem II of [128], we will prove that Lemma 4.7 can be extended to continuous piecewise affine maps. We recount the definition of a subanalytic set given in [128].

Definition 4.8. *The class of* subanalytic sets *is the smallest class $\mathfrak{C}$ of subsets E of finite-dimensional real analytic manifolds $\mathcal{M}$ such that:*

(i) $\mathfrak{C}$ contains all sets $\{x \in \mathbb{R}^n \mid g(x) = 0\}$ and $\{x \in \mathbb{R}^n \mid g(x) > 0\}$, where $g : \mathbb{R}^n \to \mathbb{R}$ is analytic.

(ii) $\mathfrak{C}$ is closed under complementation, locally finite unions and intersections.

(iii) $\mathfrak{C}$ is closed under an inverse image of an analytic map.

(iv) $\mathfrak{C}$ is closed under an image $f(E)$ of an analytic map $f : \mathcal{M} \to \mathcal{N}$ which is proper on $\overline{E}$, i.e., which preserves inverse images of compact sets in $\overline{E}$.

We also use the following definition of a piecewise affine function.

Definition 4.9. *Let L be a finite index set and let $\{\mathcal{A}_l \mid l \in L\}$ be a finite polyhedral partition of $\mathbb{R}^n$. That is, each $\mathcal{A}_l$ consists of finite unions and intersections of sets $\{x \in \mathbb{R}^n \mid g(x) = 0\}$ and $\{x \mid g(x) > 0\}$, where $g : \mathbb{R}^n \to \mathbb{R}$ are affine functions. We say that $f : \mathbb{R}^n \to \mathbb{R}^n$ is a* piecewise affine *(PWA) function if $f|_{\overline{\mathcal{A}_l}}$ is an affine function for all $l \in L$.*

Remark 4.10. *For the sake of simplicity we consider PWA functions $f : \mathbb{R}^n \to \mathbb{R}^n$, defined on all of $\mathbb{R}^n$, even if the system* (3.1) *is defined only on a polytope $\mathcal{P} \subset \mathbb{R}^n$. It was shown in [15] that any PWA*

function on $\mathcal{P}$ can be extended to a PWA function on $\mathbb{R}^n$, and if the original function was continuous PWA, the extension can be continuous PWA as well.

The setting explored by [128] is quite general. It deals with an extendably piecewise analytic vector field defined on a locally finite subanalytic partition of a real analytic manifold. We now give a version of Theorem II of [128] adapted to our needs. In our case, the sets $\mathcal{A}_j$ from Definition 4.9 are finite unions and intersections of sets in the form of case (i) of Definition 4.8. Thus, by property (ii) of Definition 4.8, they are subanalytic. The real analytic manifold that we deal with is $\mathbb{R}^n$ itself. Additional conditions (A1)-(A5) of Theorem II of [128] can be trivially verified to hold in the case of continuous PWA functions. Hence, we forgo the general statement of [128] in favour of a more wieldy version used in the setting of this chapter.

Theorem 4.11. *Assume $f : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous PWA function. Then for every compact set $K \subset \mathbb{R}^n$ and every $T > 0$ there exists a positive integer $N(K, T)$ such that if $x_0 \in K$ and $\phi(\cdot, x_0) : [0, T) \to K$ is any trajectory solving* (4.3) *with $\phi(0, x_0) = x_0$, then $\phi(\cdot, x_0)$ is a concatenation of at most $N(K, T)$ curves $\phi_1, \ldots, \phi_p$ such that $\phi_i \in A_j$ for some $j \in J$, for all $i \in \{1, \ldots, p\}$.*

Theorem 4.11 is proved in [128] in its entirety. We are now able to prove an extension of Lemma 4.7 to piecewise affine systems.

Lemma 4.12. *Let $\mathcal{P}$ be an n-dimensional convex polytope. Consider the system* (4.3) *and suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and PWA. Let $\phi(\cdot, x_0)$ be the unique solution of* (4.3) *with $\phi(0, x_0) = x_0$. Then for every $h \in \mathbb{R}^n$, there exists $\varepsilon > 0$ such that either $h \cdot (\phi(t, x_0) - x_0) \leq 0$ for all $t \in (0, \varepsilon)$ or $h \cdot (\phi(t, x_0) - x_0) > 0$ for all $t \in (0, \varepsilon)$. Moreover, if $\phi(\cdot, x_0)$ exits $\mathcal{P}$ at time $T = 0$, then there exists a facet $\mathcal{F}_i$ of $\mathcal{P}$ such that $\phi(\cdot, x_0)$ crosses $\mathcal{F}_i$ at $T = 0$.*

Proof. Let us first note that the existence of $\phi(\cdot, x_0)$ is guaranteed on some interval $(-\varepsilon', \varepsilon')$, where $\varepsilon' > 0$, by the Picard-Lindelöf theorem, because f is Lipschitz continuous on a neighbourhood of x_0 (see, e.g., [14]). Because $\phi(\cdot, x_0)$ is continuous, it is bounded on the interval $[0, \varepsilon'/2]$. Let K be any closed ball such that $\{\phi(t, x_0) \mid t \in [0, \varepsilon'/2]\} \subset K$.

Let $\mathbb{R}^n$ be partitioned into finitely many sets $\{A_j \mid j \in J\}$ as in Definition 4.9. By Theorem 4.11, trajectory $\phi(t, x_0)$, $t \in [0, \varepsilon'/2]$ is a concatenation of finitely many curves $\phi_1, \ldots, \phi_p$ such that $\phi_i \in A_j$ for some $j \in J$. Thus, there exist $0 < \delta < \varepsilon'/2$ and $j \in J$ such that $\phi(t, x_0) \in A_j$ for all $t \in (0, \delta)$. Hence, on the interval $[0, \delta)$, $\phi(\cdot, x_0)$ is governed by an affine system $\dot{x} = f|_{\overline{A}_j(x)}$. As an affine function is certainly analytic, the claim now holds by Lemma 4.7. $\square$

4.3.3 Output Reach Control Problem

The Output Reach Control Problem (ORCP) was introduced in [69]. The goal of the ORCP is to drive output trajectories starting in a given simplex $\mathcal{S}$ in the output space through a predetermined facet $\mathcal{F}_0$. The approach is to use viability theory to construct a polytope $\mathcal{P}$ in the state space such that if a state trajectory exits $\mathcal{P}$ through an exit facet $\mathcal{F}_0^{\mathcal{P}}$ of $\mathcal{P}$, then the output trajectory exits $\mathcal{S}$ through its exit facet $\mathcal{F}_0$. This problem is of particular relevance to our present investigation because a state trajectory exiting $\mathcal{P}$ through $\mathcal{F}_0^{\mathcal{P}}$ in the sense of Definition 4.2 does not imply the output trajectory exits $\mathcal{S}$. Instead it is necessary that state trajectories exit $\mathcal{P}$ in the sense of Definition 4.3. Resolving this dichotomy is the goal of Lemma 13 of [69]. A proof was not provided in that paper. This section provides a complete proof.

Problem 4.13 (Output Reach Control Problem (ORCP))**.** *Consider system* (3.1) *defined on* $\mathbb{R}^n$*. Let* $C \in \mathbb{R}^{n\times p}$*, and let* $y(\cdot, x_0) = C\phi(\cdot, x_0)$ *for all* $x_0 \in \mathbb{R}^n$*. Let* $\mathcal{S} \subset \mathbb{R}^p$ *be a simplex. Find a state feedback* $u(x)$ *such that for every* $x_0 \in \mathbb{R}^n$ *such that* $Cx_0 \in \mathcal{S}$*, there exist* $T \geq 0$ *and* $\varepsilon > 0$ *such that*

(i) $y(t, x_0) \in \mathcal{S}$ *for all* $t \in [0, T]$*,*

(ii) $y(T, x_0) \in \mathcal{F}_0$*, and*

(iii) $y(t, x_0) \notin \mathcal{S}$ *for all* $t \in (T, T + \varepsilon)$*.*

The main result of [69] is that solvability of the RCP on a particularly chosen bounded polytope $\mathcal{P} \subset \{x \in \mathbb{R}^n \mid Cx \in \mathcal{S}\}$, with some additional technical conditions, implies solvability of the ORCP on $\mathcal{S}$. As mentioned in Section 4.1, a key part of the proof is that trajectories exit a polytope through a restricted facet only if they also cross a non-restricted facet at the same time. This result is given in Lemma 13 of [69]. However, the proof is not provided.

It turns out that the issue of chattering is a salient one in proving Lemma 13. In other words, for the proof to go through, a non-chattering assumption should be made. This is noted in [69], but that work does not actually explicitly state whether a particular non-chattering condition is assumed, and does not discuss why chattering does not occur in the context of that paper. Here we clarify this inconsistency. While this is not explicitly stated in their paper, the viability approach used in [69] relies on continuous PWA feedback control, thus resulting in a continuous PWA dynamical system. Lemma 4.12 now proves that chattering indeed does not occur in that case.

The statement of the claim proved below is also slightly different than the statement of Lemma 13 of [69]. The original lemma states that a trajectory that exits $\mathcal{P}$ will cross a restricted facet only if it also crosses $\mathcal{F}_0$. Since we proved in Lemma 4.12 that a trajectory exiting $\mathcal{P}$ will necessarily cross at

least one facet, this statement is equivalent to saying that a trajectory exiting $\mathcal{P}$ will necessarily cross $\mathcal{F}_0$. This is the form in which we present the lemma. The proof is adapted from the standard proof of the Nagumo/Bony-Brezis theorem (see, e.g., [21]).

Lemma 4.14. *Consider an n-dimensional convex polytope $\mathcal{P}$ such that $0 \in \mathcal{F}_0 \cap \cdots \cap \mathcal{F}_k$ for some $k < r$. Consider the affine system* (3.1) *and let $u(x)$ be a continuous PWA feedback such that*

$$Ax + Bu(x) + a \in \mathcal{C}(x), \quad \text{for all } x \in \mathcal{F}_i,\ i = 1, \ldots, r. \tag{4.6}$$

Suppose $\phi(\cdot, 0)$ is the unique closed-loop trajectory of (3.1) *with $\phi(0,0) = 0$. If $\phi(\cdot, 0)$ exits $\mathcal{P}$ at time $T = 0$, then it does so by crossing $\mathcal{F}_0$.*

Remark 4.15. *Note that Lemma 4.14 does not say that trajectories do not cross $\mathcal{F}_1, \ldots, \mathcal{F}_k$ at* 0.

Proof. Suppose not; that is, $\phi(\cdot, 0)$ does not cross $\mathcal{F}_0$ at $T = 0$. Observe that (3.1) with a continuous piecewise affine feedback satisfies the requirements of Lemma 4.12. By Lemma 4.12, $\phi(\cdot, 0)$ crosses at least one facet $\mathcal{F}_i$. Without loss of generality assume that $\{h_0, \ldots, h_k\}$ is an orthonormal set and that for some $1 \le l \le k$, $\phi(\cdot, 0)$ crosses facets $\mathcal{F}_1, \ldots, \mathcal{F}_l$, and it does not cross $\mathcal{F}_0, \mathcal{F}_{l+1}, \ldots, \mathcal{F}_k$. By Lemma 4.12, there exists $\varepsilon > 0$ such that

$$\phi(t, 0) \in \mathcal{H}_0 \cap \left(\bigcap_{i=1}^{l} \mathcal{H}_i^c \right) \cap \left(\bigcap_{j=l+1}^{k} \mathcal{H}_j \right) =: \mathcal{G}, \quad t \in (0, \varepsilon).$$

Define $\mathcal{H} = \mathcal{H}_1 \cap \cdots \cap \mathcal{H}_l$. Note that $\mathcal{H}$ is closed as an intersection of closed sets. Hence we can define the point to set distance $d_{\mathcal{H}}(x) = \min_{z \in \mathcal{H}} \|x - z\|$.

Since $h_0, \ldots, h_k$ are orthonormal, every $x \in \mathbb{R}^n$ can be uniquely expressed as $x = \lambda_1(x)h_1 + \cdots + \lambda_l(x)h_l + \tilde{x}$, where $h_i \cdot \tilde{x} = 0$, $i = 1, \ldots, l$. Also, $h_i \cdot x = \lambda_i(x)$, $i = 1, \ldots, l$. Thus, if $z \in \mathcal{H}$, then $z = \lambda_1(z)h_1 + \ldots + \lambda_l(z)h_l + \tilde{z}$, with $\lambda_i(z) \le 0$, $i = 1, \ldots, l$. Now consider $x \in \mathcal{G}$, which means $\lambda_i(x) > 0$, $i = 1, \ldots, l$. Then $\|x - z\|^2 = \sum_{i=1}^{l} |\lambda_i(x) - \lambda_i(z)|^2 + \|\tilde{x} - \tilde{z}\|^2$, and the unique choice of $z \in \mathcal{H}$ that minimizes $\|x - z\|^2$ is given by

$$\lambda_i(z) = 0, \quad i = 1, \ldots, l \text{ and } \tilde{z} = \tilde{x}. \tag{4.7}$$

With this choice, for all $x \in \mathcal{G}$,

$$d_{\mathcal{H}}^2(x) = \lambda_1^2(x) + \cdots + \lambda_l^2(x) = (h_1 \cdot x)^2 + \cdots + (h_l \cdot x)^2. \tag{4.8}$$

Because $\phi(\cdot,0)$ is a solution of a continuous system (3.1), it is differentiable in t. Since $\phi(t,0)\in\mathcal{G}$ for all $t\in(0,\varepsilon)$, by assumption $d^2_{\mathcal{H}}(\phi(t,0))$ is differentiable on $(0,\varepsilon)$ by (4.8). Additionally, $d^2_{\mathcal{H}}(\phi(t,0))$ is continuous from the right at 0 by letting $x\to 0$ in (4.8) and noting that $d^2_{\mathcal{H}}(0)=0$.

Let $f(x)=Ax+Bu(x)+a$. We have from (3.1) and (4.8) that, for all $t\in(0,\varepsilon)$,

$$\frac{d}{dt}\left[d^2_{\mathcal{H}}(\phi(t,0))\right]=\sum_{i=1}^{l}2(h_i\cdot\phi(t,0))h_i\cdot f(\phi(t,0)). \tag{4.9}$$

By (4.7), we have $\phi(t,0)-z(t)=\sum_{i=1}^{l}\lambda_i(\phi(t,0))h_i=\sum_{i=1}^{l}(h_i\cdot\phi(t,0))h_i$, where $z(t)$ is the point in $\mathcal{H}$ closest to $\phi(t,0)$. Substituting into (4.9), we get

$$\frac{d}{dt}\left[d^2_{\mathcal{H}}(\phi(t,0))\right]=2(\phi(t,0)-z(t))\cdot f(\phi(t,0)). \tag{4.10}$$

Since $\phi(t,0)\in\mathcal{H}_0\cap\mathcal{H}_{l+1}\cap\cdots\cap\mathcal{H}_k$, from (4.7) we also have that $z(t)\in\mathcal{H}_0\cap\mathcal{H}_{l+1}\cap\cdots\cap\mathcal{H}_k$. Because $z(t)\in\mathcal{H}=\mathcal{H}_1\cap\ldots\cap\mathcal{H}_l$ by definition, $z(t)\in\mathcal{P}=\mathcal{H}_0\cap\cdots\cap\mathcal{H}_k$ for all $t\in(0,\varepsilon)$. In fact, since $\lambda_i(z(t))=h_i\cdot z(t)=0$ for all $i=1,\ldots,l$, and $0\in\mathcal{F}_1\cap\cdots\cap\mathcal{F}_l$,

$$z(t)\in\bigcap_{i=1}^{l}\mathcal{F}_i. \tag{4.11}$$

Let us rewrite (4.10) as

$$\begin{aligned}\frac{d}{dt}\left[d^2_{\mathcal{H}}(\phi(t,0))\right]&=2(\phi(t,0)-z(t))\cdot f(z(t))+\\&2(\phi(t,0)-z(t))\cdot\left[f(\phi(t,0))-f(z(t))\right].\end{aligned} \tag{4.12}$$

From (4.7) we know $\phi(t,0)-z(t)=\sum\limits_{i=1}^{l}\lambda_i(\phi(t,0))h_i$. Also, because $\phi(t,0)\in\mathcal{G}$ for all $t\in(0,\varepsilon)$,

$$\lambda_i(\phi(t,0))=h_i\cdot\phi(t,0)\geq 0\,,\quad i=1,\ldots,l\,. \tag{4.13}$$

Since $\mathcal{F}_1,\ldots,\mathcal{F}_l$ are restricted facets, from (4.6), (4.11) and (4.13) we get

$$2(\phi(t,0)-z(t))\cdot f(z(t))=2\sum_{i=1}^{l}\lambda_i(\phi(t,0))h_i\cdot f(z(t))\leq 0.$$

Thus, by (4.12)

$$\frac{d}{dt}\left[d^2_{\mathcal{H}}(\phi(t,0))\right]\leq 2(\phi(t,0)-z(t))\cdot(f(\phi(t,0))-f(z(t))). \tag{4.14}$$

We note that f is continuous and piecewise affine. Hence, it can easily be shown (see, e.g., [14]) that it is Lipschitz continuous on some neighbourhood of 0. Let $L > 0$ be the Lipschitz constant of f in that neighbourhood. We can reduce ε such that both $\phi(t,0)$ and $z(t)$ are in this neighbourhood for all $t \in (0,\varepsilon)$. Then from (4.14)

$$\frac{d}{dt}\left[d^2_{\mathcal{H}}(\phi(t,0))\right] \leq 2L\|\phi(t,0) - z(t)\|^2 = 2Ld^2_{\mathcal{H}}(\phi(t,0))\,.$$

Using this result, we find $\frac{d}{dt}\left[\mathrm{e}^{-2Lt}d^2_{\mathcal{H}}(\phi(t,0))\right] = -2L\mathrm{e}^{-2Lt}d^2_{\mathcal{H}}(\phi(t,0)) + \mathrm{e}^{-2Lt}\frac{d}{dt}\left[d^2_{\mathcal{H}}(\phi(t,0))\right] \leq 0$. Thus, $\mathrm{e}^{-2Lt}d^2_{\mathcal{H}}(\phi(t,0))$ is a non-increasing function on the interval $(0,\varepsilon)$. It is also nonnegative and continuous from the right at $t = 0$. Thus, since $\mathrm{e}^0 d^2_{\mathcal{H}}(\phi(0,0)) = 0$, we have $d^2_{\mathcal{H}}(\phi(t,0)) = 0$ for all $t \in [0,\varepsilon)$. This is in contradiction with our assumption that $\phi(t,0) \in \mathcal{G} \subset \mathcal{H}^c$. □

4.3.4 Affine Feedback

The results of Section 4.3.2 can be applied to the special case of affine feedback. We consider the situation investigated in [122] on the use of affine feedback to solve the RCP on simplices. The following result was stated in [122], with a partial proof provided in [121].

Lemma 4.16. *Consider the affine system* (3.1) *and consider an n-dimensional simplex $\mathcal{S}$ with facets $\mathcal{F}_0, \ldots, \mathcal{F}_n$. Suppose that $\mathcal{F}_0$ is the exit facet and $\mathcal{F}_1, \ldots, \mathcal{F}_n$ are restricted facets. Let $u(x) = Kx + g$ be an affine feedback such that the invariance conditions* (3.4) *hold. Then all trajectories originating in $\mathcal{S}$ that exit $\mathcal{S}$ do so through $\mathcal{F}_0$.*

Using Lemma 4.14 we are able to prove the following stronger result, showing that a trajectory exiting $\mathcal{S}$ does so not only by exiting through $\mathcal{F}_0$, but also by crossing this facet in the sense of Definition 4.3. This is given in the following lemma.

Lemma 4.17. *Consider the affine system* (3.1) *and consider an n-dimensional simplex $\mathcal{S}$ with facets $\mathcal{F}_0, \ldots, \mathcal{F}_n$. Suppose that $\mathcal{F}_0$ is the exit facet and $\mathcal{F}_1, \ldots, \mathcal{F}_n$ are restricted facets. Let $u(x) = Kx + g$ be an affine feedback such that* (3.4) *holds. Then all trajectories originating in $\mathcal{S}$ that exit $\mathcal{S}$ do so by crossing $\mathcal{F}_0$.*

Proof. Assume otherwise: a trajectory $\phi(\cdot, x_0)$ with $x_0 \in \mathcal{S}$ exits $\mathcal{S}$, but does so without crossing $\mathcal{F}_0$. Without loss of generality, we may assume that $\phi(\cdot, x_0)$ exits $\mathcal{S}$ at time $T = 0$, i.e., at point x_0. We distinguish between two cases: if $x_0 \in \mathcal{F}_0$, the conditions of Lemma 4.14 are satisfied, and it is impossible to exit $\mathcal{S}$ without crossing $\mathcal{F}_0$. If $x_0 \notin \mathcal{F}_0$, then all the facets that x_0 is contained in are restricted. Exactly

the same proof as in Lemma 4.14 works, just without any mention of the unrestricted facet $\mathcal{F}_0$: if we assume that $\phi(\cdot, x_0)$ exits $\mathcal{S}$ at time $T = 0$, we obtain a contradiction. □

Remark 4.18. *Lemma 4.17 also holds for continuous piecewise affine feedback controls, with the same proof.*

Analogously to Remark 4.15, Lemma 4.17 does not guarantee that a trajectory exiting $\mathcal{S}$ does not cross any restricted facets. The following example shows that even if the invariance conditions are solvable, and $f(x) \in \mathcal{C}(x)$ for all $x \in \mathcal{S}$, solutions may cross a restricted facet.

Example 4.19. *Let us consider $\mathcal{S} \subset \mathbb{R}^2$ as the two-dimensional simplex with vertices $(0,0)$, $(1,0)$ and $(0,1)$. Let $\mathcal{F}_0 = \{(x_1, 0) \mid 0 \leq x_1 \leq 1\}$, $\mathcal{F}_1 = \{(x_1, 1-x_1) \mid 0 \leq x_1 \leq 1\}$ and $\mathcal{F}_2 = \{(0, x_2) \mid 0 \leq x_2 \leq 1\}$. We consider the dynamics given by the system*

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -1 \end{pmatrix}. \tag{4.15}$$

We note that at $\mathcal{F}_0$ the vector field generated by this system points exactly down, through $\mathcal{F}_0$, while at $\mathcal{F}_1$ and $\mathcal{F}_2$ it points into the half-spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. Hence, $f(x) \in \mathcal{C}(x)$, with the definition of $\mathcal{C}(x)$ from (3.3). Thus, the invariance conditions are solvable, with $\mathcal{F}_0$ as the exit facet and $\mathcal{F}_1$ and $\mathcal{F}_2$ as restricted facets.

Now, let us consider the trajectory $\phi(\cdot, 0)$ generated by (4.15). We can easily calculate that

$$\phi(t, 0) = (-t^2/2, -t)$$

for all $t \geq 0$. Hence, this trajectory exits $\mathcal{S}$ by crossing $\mathcal{F}_0$, as guaranteed by Lemma 4.17, but it also crosses $\mathcal{F}_2$ at the same time.

4.3.5 Discontinuous Piecewise Affine Feedback

In this section we examine the question of chattering for the class of discontinuous PWA feedbacks defined on a possibly non-convex polytope formed by a so-called *chain of simplices* [56]. This class of feedbacks was also studied in [24] when the polytope is itself a simplex. We improve the results of [24] in several ways. First, we relax the requirement on open-loop controls that they satisfy the invariance conditions. Instead, we *prove* that there exists a set of open-loop controls satisfying the invariance conditions. Given this result, we then present a modified discrete supervisor rule based on the one in

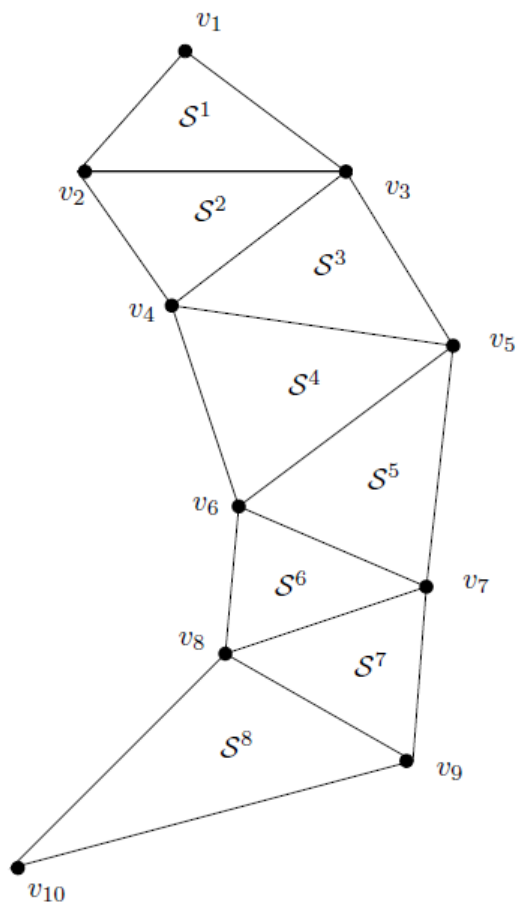

Figure 4.2: A chain of simplices.

[24]. This new rule closes a gap in [24] related to the behavior described in Example 4.19 above. More details are given below. Finally, we prove that the discontinuous affine feedback given here does not exhibit chattering. First we review the definition of a triangulation [71].

Definition 4.20. *A triangulation* $\mathbb{T} = \{\mathcal{S}^1, \ldots, \mathcal{S}^\chi\}$ *of a polytope* $\mathcal{P}$ *is a subdivision of* $\mathcal{P}$ *into full dimensional simplices* $\mathcal{S}^1, \cdots, \mathcal{S}^\chi$ *such that*

(i) $\mathcal{P} = \cup_{i=1}^{\chi} \mathcal{S}_i$,

(ii) for all $i,\ j \in \{1, \cdots, \chi\}$ *with* $i \neq j$, *the intersection* $\mathcal{S}^i \cap \mathcal{S}^j$ *is either empty or a common face of* $\mathcal{S}^i$ *and* $\mathcal{S}^j$.

Definition 4.21. *Let* $L := \{1, \ldots, \chi\}$. *Let* $\{\mathcal{S}^j \mid j \in L\}$ *be a collection of* n*-dimensional simplices. Define* $\mathcal{P} := \mathcal{S}^1 \cup \cdots \cup \mathcal{S}^\chi$. *We say that* $\mathcal{P}$ *is a* chain *if the following hold:*

(i) $\mathbb{T} = \{\mathcal{S}^1, \ldots, \mathcal{S}^\chi\}$ *is a triangulation of* $\mathcal{P}$ *such that* $\mathcal{S}^k \cap \mathcal{S}^{k-1}$ *is a facet of* $\mathcal{S}^k$ *and of* $\mathcal{S}^{k-1}$.

(ii) $\mathcal{F}_0^k := \mathcal{S}^k \cap \mathcal{S}^{k-1}$ *is designated to be the exit facet of* $\mathcal{S}^k$*, for each* $k = 2, \ldots, \chi$*.*

(iii) The exit facet of $\mathcal{P}$ *is designated to be* $\mathcal{F}_0 := \mathcal{F}_0^1$*, the exit facet of* $\mathcal{S}^1$*, and it is not a facet of any other simplex* $\mathcal{S}^k$*,* $k \neq 1$*.*

See Figure 4.2, adapted from [56], for an example of a chain. Since we allow discontinuous controls, a slightly stronger version of the RCP is needed in comparison to Problem 3.1 and Problem 4.1.

Problem 4.22 (Discontinuous Reach Control Problem (DRCP))**.** *Consider system (3.1) defined on a chain* $\mathcal{P}$*. Find a state feedback* $u(x)$ *such that:*

(i) For every $x_0 \in \mathcal{P}$*, there exist* $T \geq 0$ *and* $\varepsilon > 0$ *such that*

(i.1) $\phi(t, x_0) \in \mathcal{P}$ *for all* $t \in [0, T]$*,*

(i.2) $\phi(T, x_0) \in \mathcal{F}_0$*, and*

(i.3) $\phi(t, x_0) \notin \mathcal{P}$ *for all* $t \in (T, T + \varepsilon)$*.*

(ii) There exists $\gamma > 0$ *such that for every* $x \in \mathcal{P}$*,* $\quad \|Ax + Bu(x) + a\| > \gamma$*.*

Remark 4.23. *As already discussed in [24], the new condition (ii) appearing in Problem 4.22 compared to Problem 3.1 and Problem 4.1 is a robustness requirement. It circumvents a vanishingly small closed-loop vector field that could result in the appearance of an equilibrium in* $\mathcal{P}$ *if the system parameters are perturbed. When continuous feedbacks are used, condition (ii) is automatically satisfied, and Problem 4.22 reduces to Problem 4.1.*

We consider the following class of discontinuous PWA feedbacks for solving the DRCP on chains.

Definition 4.24. *Consider a chain* $\mathcal{P} = \mathcal{S}^1 \cup \cdots \mathcal{S}^\chi$*. Let* $u_k(x) := K_k x + g_k$*,* $k \in L$*, be a set of affine feedbacks where* $K_k \in \mathbb{R}^{m \times n}$ *and* $g_k \in \mathbb{R}^m$*. Define*

$$k(x) := \min\{k \in L \mid x \in \mathcal{S}^k\}\,. \tag{4.16}$$

We say $u(x)$ is a PWA feedback on $\mathcal{P}$ *if*

$$u(x) = u_{k(x)}(x)\,. \tag{4.17}$$

That is, $u(x)$ *is affine on the interior of each simplex, and at a point* $x \in \mathcal{P}$ *belonging to more than one simplex, the affine controller for the simplex with the smallest index is used.*

The requirement that the control law u uses the affine controller with the smallest index is designed to avoid the situation illustrated in Figure 4.1 from the beginning of the chapter. The next result shows that the class of discontinuous PWA affine feedbacks introduced above can indeed be used to solve the DRCP on a chain $\mathcal{P}$, assuming the constituent affine feedbacks solve a local RCP on each simplex.

Theorem 4.25. *Consider the system* (3.1) *defined on a chain* $\mathcal{P} = \mathcal{S}^1 \cup \cdots \mathcal{S}^\chi$. *Let* $u(x)$ *be a PWA feedback on* $\mathcal{P}$ *as in* (4.16)-(4.17). *Suppose that*

$$\mathcal{S}^k \xrightarrow{\mathcal{S}^k} \mathcal{F}_0^k\,, \qquad k \in L\,, \tag{4.18}$$

using $u = u_k(x)$ *in the sense of Problem 4.1. Then* $u(x)$ *solves Problem 4.22. In particular, for each* $x_0 \in \mathcal{P}$, *there exist* $T \geq 0$, $\varepsilon > 0$, $r \in \mathbb{N}$, *times* $0 < t_1 < \cdots < t_r = T$, *and a unique continuous solution* $\phi(\cdot, x_0) : [0, T+\varepsilon] \to \mathbb{R}^n$ *such that*

$$\frac{d}{dt}\phi(t, x_0) = A\phi(t, x_0) + Bu(\phi(t, x_0)) + a$$

for all $t \in (0, T+\varepsilon) \setminus \{t_1, \ldots, t_{r-1}\}$. *For each* $i = 1, \ldots, r-1$, $\phi(t_i, x_0) \in \mathcal{S}_j \cap \mathcal{S}_{j-1}$ *for some* $j \geq 1$. *Also, (i)-(ii) of Problem 4.22 are satisfied, and* $\phi(\cdot, x_0)$ *exits* $\mathcal{P}$ *by crossing the facet* $\mathcal{F}_0 = \mathcal{F}_0^1$.

Remark 4.26. *As discussed in Section 3.1, the system* (3.1) *must be defined on a small neighbourhood* $\mathcal{P} \subset \mathcal{N}$. *In this case, we assume that the control law on* $\mathcal{N} \backslash \mathcal{P}$ *is given by* $u(x) = u_1(x)$.

Proof. Let $k_0 := k(x_0)$ with k from (4.16). Assume that $k_0 > 1$ and $x_0 \in \mathcal{F}_0^{k_0}$. Then, by (ii) in Definition 4.21, $x_0 \in \mathcal{S}^{k_0-1}$, which is in contradiction to $k_0 = k(x_0)$. Thus, $x_0 \notin \mathcal{F}_0^{k_0}$ or $x_0 \in \mathcal{F}_0^1$. If the latter is true, part (i) of Problem 4.22 is trivially true by (4.18). Hence suppose $x_0 \notin \mathcal{F}_0^{k_0}$. By (4.18), there exists a $t_1 > 0$ and a unique trajectory $\phi_1(\cdot, x_0) \ : \ [0, t_1] \to \mathcal{S}^{k_0}$ with the following properties:

- (8.1) holds for all $t \in (0, t_1)$,
- $\phi_1(t_1, x_0) \in \mathcal{F}_0^{k_0}$,
- $\phi_1(t, x_0) \in \mathcal{S}^{k_0} \backslash \mathcal{F}_0^{k_0} \subset \mathcal{P}$ for all $t \in (0, t_1)$.

Consider now $x_1 := \phi_1(t_1, x_0)$. Since $x_1 \in \mathcal{F}_0^{k_0}$, either $k_0 = 1$ or $\min\{i \in L \mid x_1 \in \mathcal{S}^i\} \leq k_0 - 1$. If we define $k_1 := k(x_1)$, we have the same situation as from before: $x_1 \notin \mathcal{F}_0^{k_1}$ or $x_1 \in \mathcal{F}_0^1$. We proceed analogously and obtain a trajectory $\phi_2(\cdot, x_0) : [t_1, t_2) \to \mathcal{S}^{k_1}$. By iterating this procedure, we obtain a finite sequence of times $0 < t_1 < \cdots < t_r =: T$, indices $k_0 > k_1 > \cdots > k_r = 1$ and switching points $x_1, \ldots, x_r$, where $x_r \in \mathcal{F}_0^1$. By (4.18) there now exists a trajectory $\phi_{r+1}(\cdot, x_0)$ which leaves $\mathcal{S}_1$ through

$\mathcal{F}_0^1$ in the sense of Problem 4.1. We define ϕ as the concatenation of trajectories $\phi_1, \ldots, \phi_{r+1}$. ϕ clearly satisfies part (i) of Problem 4.22.

For part (ii) of Problem 4.22, define $f(x) = Ax+Bu(x)+a$ and $f_k(x) = Ax+Bu_k(x)+a$ for all $k \in L$. We note that all f_k are continuous functions, and by (4.18), contain no zeros on $\mathcal{S}^k$. Hence, for all $x \in \mathcal{S}^k$, it holds that $\|f_k(x)\| > \gamma_k$ for some $\gamma_k > 0$. Now, take any $x \in \mathcal{P}$ and define $k' := \min_{1\leq i\leq \chi}\{i \mid x_0 \in \mathcal{S}^i\}$. Then, $\|f(x)\| = \|f_{k'}(x)\| > \gamma_{k'}$. Thus, part (ii) of Problem 4.22 is satisfied with $\gamma := \min\{\gamma_1, \ldots, \gamma_\chi\}$.

Finally, we showed above that ϕ exits $\mathcal{P}$ through facet $\mathcal{F}_0$. However, in the time interval $[T, T+\varepsilon)$ when ϕ is exiting $\mathcal{F}_0$, ϕ is, by Remark 4.26, governed by the affine feedback control law $u_1(x)$. Hence, Lemma 4.7 applies directly, implying that ϕ will also cross the facet $\mathcal{F}_0$. □

Next we explore the extent to which PWA feedbacks on chains provide a complete solution to the DRCP; that is, if the DRCP is solvable by some reasonable class of open-loop controls, we want to show it is solvable by PWA feedback. We consider the special case, also investigated in [24], when $\mathcal{P}$ is itself a simplex denoted by $\mathcal{S}$.

We define the set of admissible open-loop controls for (3.1) as any measurable function $\mu : [0, \infty) \to \mathbb{R}^m$ that is bounded on compact time intervals. Solutions of (3.1) under an open-loop control $\mu(t)$ are in the sense of Carathéodory. Thus, as noted in Section 4.2, they exist and are unique. We reuse the notation $\phi_\mu(\cdot, x_0)$ to denote a trajectory of (3.1) starting at x_0 under an open-loop control $\mu(t)$.

Definition 4.27. *Consider system* (3.1) *defined on a chain* $\mathcal{S}$. *We say the* DRCP is solved by open-loop controls *if there exists* $\gamma > 0$ *such that for each* $x \in \mathcal{S}$, *there exist* $\varepsilon_x > 0$, $T_x \geq 0$, *and an open-loop control* $\mu_x(\cdot) : [0, T_x + \varepsilon_x] \to \mathbb{R}^m$ *such that*

(i) (i.1) $\phi_{\mu_x}(t, x) \in \mathcal{S}$ *for all* $t \in [0, T_x]$,

(i.2) $\phi_{\mu_x}(T_x, x) \in \mathcal{F}_0$, *and*

(i.2) $\phi_{\mu_x}(t, x) \notin \mathcal{S}$ *for all* $t \in (T_x, T_x + \varepsilon_x)$.

(ii) $\|A\phi_{\mu_x}(t, x) + B\mu_x(t) + a\| > \gamma$ *for all* $t \in [0, T_x]$.

In [24] an additional requirement on open-loop controls was that they satisfy the invariance conditions. Instead, in the next result we prove that there exists a set of open-loop controls that satisfy these conditions at $t = 0$.

Theorem 4.28. *If* $\mathcal{S} \xrightarrow{\mathcal{S}} \mathcal{F}_0$ *by open-loop controls in the sense of Definition 4.27, then* $\mathcal{S} \xrightarrow{\mathcal{S}} \mathcal{F}_0$ *by open-loop controls that also satisfy:*

(iii) $Ax + B\mu_x(0) + a \in \mathcal{C}(x)$ *for all* $x \in \mathcal{S} \setminus \mathcal{F}_0$.

Proof. Let $x \in \mathcal{S} \setminus \mathcal{F}_0$. By assumption there exists an open-loop control μ_x and a time $T_x > 0$ such that $\phi_{\mu_x}(t,x) \in \mathcal{S}$ for all $t \in [0, T_x]$. Since μ_x is an open-loop control, there exists $c \geq 0$ such that $\|\mu_x(t)\| \leq c$, for all $t \in [0, T_x]$. Define $\mathcal{Y}(z) := \{Az+Bw+a \mid w \in \mathbb{R}^m\}$ and $\mathcal{Y}_c(z) := \{Az+Bw+a \mid w \in \mathbb{R}^m, \|w\| \leq c\}$. Now take a sequence $\{t_i \mid t_i \in (0, T_x]\}$ with $t_i \to 0$. Note that

$$\frac{\|\phi_{\mu_x}(t_i,x) - x\|}{t_i} = \frac{1}{t_i}\int_0^{t_i} \big[A\phi_{\mu_x}(\tau,x) + B\mu_x(\tau) + a\big]d\tau. \tag{4.19}$$

Thus, since $\{y \in \mathcal{Y}_c(z) \mid z \in \mathcal{S}\}$ is bounded, there exists $M > 0$ such that $\|\phi_{\mu_x}(t_i,x) - x\| \leq Mt_i$. Therefore $\left\{\frac{\phi_{\mu_x}(t_i,x)-x}{t_i}\right\}$ is a bounded sequence, and there exists a convergent subsequence (with indices relabeled) such that $\lim_{i\to\infty} \frac{\phi_{\mu_x}(t_i,x)-x}{t_i} =: v$. Since $\phi_{\mu_x}(t_i,x) \in \mathcal{S}$, by the definition of the Bouligand tangent cone, $v \in T_{\mathcal{S}}(x)$. On the other hand, by taking the limit in (4.19), we get $v = Ax + B\lim_{i\to\infty}\mu_x(t_i) + a \in \mathcal{Y}(x)$. Note that $\lim_{i\to\infty}\mu_x(t_i)$ exists by passing to a subsequence, if necessary, because μ_x is bounded on compact intervals. Since $T_{\mathcal{S}}(x) = \mathcal{C}(x)$ for $x \in \mathcal{S} \setminus \mathcal{F}_0$, we conclude that $v \in \mathcal{Y}(x) \cap \mathcal{C}(x)$.

Now we construct a modified set of open loop controls $\tilde{\mu}_x$ as follows. If $x \in \mathcal{F}_0$, then let $\tilde{\mu}_x := \mu_x$. If $x \in \mathcal{S} \setminus \mathcal{F}_0$, then let $\tilde{\mu}_x(t) := \mu_x(t)$ for all $t \neq 0$. At $t = 0$ let $\tilde{\mu}_x(0) := v \in \mathcal{Y}(x) \cap \mathcal{C}(x)$, as above. Since each open-loop control μ_x was changed at no more than a single point, the trajectory generated by $\tilde{\mu}_x$ is unchanged, and we obtain the desired result. □

A special triangulation of the state space was studied in [24] which has proved to be useful both in theory and in applications. Under this triangulation, a *Subdivision Algorithm* was presented in [24] that partitions the original simplex $\mathcal{S}$ into $p+1$ simplices $\{\mathcal{S}^1, \ldots, \mathcal{S}^{p+1}\}$. By construction, the partition generates a chain $\mathcal{S} = \mathcal{S}^1 \cup \cdots \cup \mathcal{S}^{p+1}$. Our goal here is to apply our new result in Theorem 4.25 to solve the DRCP on this chain. As a byproduct we want to deduce there is no chattering. First, we introduce some assumptions from [24] in order to use existing results without extra duplication. It is sufficient for our purposes to say that these assumptions are primarily to set up the special triangulation. Further detailed explanations can be found in [24].

Assumption 4.29. *Simplex $\mathcal{S}$ and system* (3.1) *satisfy Assumption 12 of [24]. Additionally, the condition in Lemma 25 of [24] holds.*

Theorem 4.30. *Suppose Assumption 4.29 holds. If $\mathcal{S} \xrightarrow{\mathcal{S}} \mathcal{F}_0$ by open-loop controls in the sense of Definition 4.27, then Problem 4.22 is solvable by the discontinuous PWA feedback* (4.16)-(4.17)*. Moreover, all trajectories originating in $\mathcal{S}$ exit $\mathcal{S}$ by crossing $\mathcal{F}_0$.*

Proof. By Theorem 4.28 there exists a set of open loop controls $\{\mu_x \mid x \in \mathcal{S}\}$ such that (i)-(ii) of Definition 4.27 and (iii) of Theorem 4.28 hold. Then with a minor variation of the proof in [24],

Theorem 10 of [24] holds. Combining Theorem 9, Theorem 10 and Assumption 12 of [24], we can invoke Theorem 23 of [24] on the construction of the reach control indices $\{r_1, \ldots, r_p\}$. With Lemma 25 of [24], all the requirements to apply the Subdivision Algorithm of [24] are in place. This yields a chain $\{\mathcal{S}^1, \ldots, \mathcal{S}^{p+1}\}$. From the first part of the proof of Theorem 33 of [24], there exists a set of affine feedbacks $u_k(x) = K_k x + g_k$ such that (4.18) holds in the sense of Problem 4.1. Now we construct the discontinuous PWA feedback given in (4.16)-(4.17). By Theorem 4.25, the resulting feedback $u(x)$ solves condition (i) of Problem 4.22. A trivial compactness argument gives condition (ii) of Problem 4.22. Finally, by Theorem 4.25 again, all trajectories exit $\mathcal{S}$ by crossing $\mathcal{F}_0$. □

Comparing our new result Theorem 4.30 to the analogous result in [24], we have the following improvements. First, we have relaxed the requirement in [24] that the set of open-loop controls satisfy the invariance conditions on their interval of definition. Instead in Theorem 4.28 we prove that there exists a set of open loop controls that satisfy these conditions at the initial time only. Second, we show there is no chattering using our proposed discontinuous PWA controller; chattering was not discussed at all in [24]. Finally and most importantly, we close a gap in the result of [24]. Our discontinuous PWA feedback uses (4.16) to assign the affine feedback for the simplex with the smallest index at points that lie in more than one simplex of the chain. Instead the rule in [24] is to assign the affine feedback for the simplex with the largest index. This opens the possibility of the pathology shown in Example 4.19, where a trajectory may cross a restricted facet of a simplex; thus, never entering the next simplex of the chain and ultimately not leaving through the exit facet.

Chapter 5

A Topological Obstruction to Reach Control

After a segue into a technical issue of chattering in Chapter 4, in this chapter we return to exploring the solvability of the RCP. We discuss a strong necessary condition to the solvability of the RCP by continuous state feedback. As this problem can be transformed into a question in homotopy and obstruction theory, it is known as the problem of a topological obstruction. The results contained in this chapter have been compiled from a number of previous publications. The results for low-dimensional simplices are contained in [105]. The results concerning the case of $m = 2$ were the subject of [103], while the general results on the characterization of a topological obstruction are largely contained in [106]. The nerve theory argument contained in Section 5.5.4 is contained in [101]. We note that the work of Section 5.5.4 entirely supplants the result of Section 5.5.3. However, as the result of Section 5.5.3 was obtained using the basic machinery of Sperner's lemma, and Section 5.5.4 requires more advanced use of nerve theory, we include both results. Additionally, a small amount of preliminary work comes from [102].

5.1 Problem Definition

The previous two chapters helped rigorously define the RCP and examine the properties of trajectories in the RCP. However, they did not take us closer to the ultimate goal of reach control theory, which is to determine whether a control objective for a given system is achievable using a reach control approach, and if it is, to design a set of controllers achieving that objective.

The first step to that end is discussing the solvability of the RCP on a single simplex. In Chapter 3 we gave two trivial necessary conditions for the solvability of the RCP: Proposition 3.3 and Proposition 3.5. In Theorem 3.6 we noted that, in the case of affine feedback, these two conditions together form a necessary and sufficient condition for the solvability of the RCP [47, 122]. Affine feedback will be discussed in more detail in Chapter 6. While Theorem 3.6 is generally not true for continuous state feedback (see, e.g., [55]), it makes sense to combine these two conditions into a stronger necessary condition for the solvability of the RCP.

Let us consider the no-equilibria condition given in Proposition 3.3 further. If system (3.1) contains an equilibrium $x \in \mathcal{S}$ for some control function u, then it satisfies $Ax + Bu + a = 0$, i.e., $Ax + a$ equals $-Bu \in \mathrm{Im}(B) = \mathcal{B}$. Hence, all equilibria in the Reach Control Problem necessarily lie in the set

$$\mathcal{O}_{\mathcal{S}} = \{x \in \mathcal{S} \mid Ax + a \in \mathcal{B}\}. \tag{5.1}$$

From the foregoing discussion, to verify the no-equilibria condition of Proposition 3.3, it is sufficient to check that $Ax + Bu(x) + a \neq 0$ for all $x \in \mathcal{O}_{\mathcal{S}}$. The set $\mathcal{O}_{\mathcal{S}}$ is an intersection of an affine set

$$\mathcal{O} = \{x \in \mathbb{R}^n \mid Ax + a \in \mathcal{B}\} \tag{5.2}$$

and a simplex $\mathcal{S}$ and is hence a polytope [23, 122]. We note that

$$Ax + Bu(x) + a \in \mathcal{B}, \quad \text{for all } x \in \mathcal{O}_{\mathcal{S}} \tag{5.3}$$

by (5.1).

Going to the second necessary condition identified in Chapter 3, Proposition 3.5 states that (3.4) is a necessary condition for the solvability of the RCP with continuous state feedback. Now, let us define $f(x) = Ax + Bu(x) + a$. By combining (3.2), (3.4) and (5.3), a necessary condition for $u(x)$ to solve the RCP is that $f : \mathcal{O}_{\mathcal{S}} \to \mathcal{B} \setminus \{0\}$ is such that $f(x) \in \mathcal{C}(x)$ for all $x \in \mathcal{O}_{\mathcal{S}}$. Note that this is not a sufficient condition for solvability of the RCP: a counterexample is given in [125]. While in this chapter we are primarily interested in continuous state feedback, this problem has been posed in literature in two settings: one dealing with continuous state feedback, i.e., with the case where f is an continuous function, and another dealing with the affine case.

Problem 5.1 (Topological Obstruction)**.** *Let $\mathcal{S}$, $\mathcal{O}_{\mathcal{S}}$, and $\mathcal{B}$ be as above. Assume that $\mathcal{O}_{\mathcal{S}} \neq \emptyset$. Does*

there exist a continuous function $f : \mathcal{O}_{\mathcal{S}} \to \mathcal{B}\backslash\{0\}$ *such that*

$$f(x) \in \mathcal{C}(x) \quad \textit{for all } x \in \mathcal{O}_{\mathcal{S}}? \tag{5.4}$$

Problem 5.2 (Affine Obstruction)**.** *Let* $\mathcal{S}$*,* $\mathcal{O}_{\mathcal{S}}$*, and* $\mathcal{B}$ *be as above. Assume that* $\mathcal{O}_{\mathcal{S}} \neq \emptyset$*. Does there exist an affine function* $f : \mathcal{O}_{\mathcal{S}} \to \mathcal{B}\backslash\{0\}$ *such that* (5.4) *holds?*

Problem 5.1 was originally identified in [87]. However, prior to the results presented in this thesis, neither of the two problems above has been significantly illuminated. In this thesis, under some technical assumptions we provide an elegant sufficient and necessary conditions for the answers to both above questions to be correct. A reader who would like to skip to those answers immediately is directed to Theorem 5.49 for Problem 5.1 and Corollary 6.9 for Problem 5.2. Unfortunately, these results require introduction of quite a bit of previous notation.

While Problem 5.1 and Problem 5.2 are similar, the methods used to approach them are vastly different: Problem 5.2 can be attacked in a relatively straightforward manner using linear algebra, while Problem 5.1 gives rise to a deep topological problem. This chapter focuses on finding an elegant necessary and sufficient condition that gives an answer to Problem 5.1, while Problem 5.2 is considered in Chapter 6. However, some preliminary work applies to both problems equally.

5.2 Preliminaries

In this section we introduce a sufficient condition for solvability of Problems 5.1 and 5.2. Furthermore, we investigate the cases of $\dim \mathcal{O}_{\mathcal{S}} = 0$ and $\dim \mathcal{O}_{\mathcal{S}} = n$. All of the following results apply both to Problem 5.1 and Problem 5.2. Let us begin with a technical lemma which gives a comparison between cones of points in $\mathcal{O}_{\mathcal{S}}$.

Lemma 5.3 (Vertex Deletion)**.** *Let* $I(p) = \{0, i_1, i_2, \ldots, i_k\}$*, with* $k \geq 0$*. Furthermore, let* $I(q) = \{i_1, i_2, \ldots, i_k, i_{k+1}, \ldots, i_l\}$*, where* $l \geq k$*. We take all* i_j*'s to be different, and all different from* 0*. Then* $\mathcal{C}(p) \subset \mathcal{C}(q)$*.*

Proof. By the definition of $\mathcal{C}$ in (3.3),

$$\mathcal{C}(p) = \{y \in \mathbb{R}^n | h_j \cdot y \leq 0 \text{ for all } j \in \{1, \ldots, n\}\backslash\{0, i_1, i_2, \ldots, i_k\}\}.$$

On the other hand,

$$\mathcal{C}(q) = \{y \in \mathbb{R}^n | h_j \cdot y \leq 0 \text{ for all } j \in \{1, \ldots, n\} \backslash \{i_1, i_2, \ldots, i_k, i_{k+1}, \ldots, i_l\}\}.$$

Thus, as it is clear that the set of constraints in $\mathcal{C}(q)$ is a subset of the set of constraints in $\mathcal{C}(p)$, $\mathcal{C}(p) \subset \mathcal{C}(q)$. □

Remark 5.4. *From the proof, it is clear that it does not matter if $I(p)$ includes 0 or not. Analogously, it does not matter if q is in the convex hull of vertices that include v_0 or not.*

The above lemma can now be used to show that cones of points on the interior of a polytope in $\mathcal{O}_\mathcal{S}$ are less restrictive than cones of points at its boundary. This is given in Lemma 5.5.

Lemma 5.5. *Let $\mathcal{A} \subset \mathcal{O}_\mathcal{S}$ be a polytope, and let $x \in \text{int}(\mathcal{A})$. Also, let $y \in \partial\mathcal{A}$. Then, $\mathcal{C}(y) \subset \mathcal{C}(x)$.*

Proof. Consider the line going through points x and y. As $x \notin \partial\mathcal{A}$, by extending that line past x, we can determine a point $z \in \partial\mathcal{A}$ such that $x = \alpha y + \beta z$, where $\alpha, \beta > 0$, $\alpha + \beta = 1$. Let us assume that $y = \sum_{i=0}^n \alpha_i v_i$, $z = \sum_{i=0}^n \beta_i v_i$. Since both y and z are in $\mathcal{O}_\mathcal{S} \subset \mathcal{S}$, all α_i's and β_i's are nonnegative. Then, $x = \sum_{i=0}^n (\alpha\alpha_i + \beta\beta_i) v_i$. Now, for any i, if $i \in I(y)$, then $\alpha_i \neq 0$. We notice that, in that case, no matter what β, β_i and α are, $\alpha\alpha_i + \beta\beta_i > 0$. Thus, $i \in I(x)$. In other words, $I(y) \subset I(x)$ and thus, by Lemma 5.3, $\mathcal{C}(y) \subset \mathcal{C}(x)$. □

From now on, we will use the following notation:

$$\text{cone}(\mathcal{O}_\mathcal{S}) = \bigcap_{x \in \mathcal{O}_\mathcal{S}} \mathcal{C}(x). \tag{5.5}$$

We note that by Lemma 5.5

$$\text{cone}(\mathcal{O}_\mathcal{S}) = \bigcap_{i=1}^{r} \mathcal{C}(o_i),$$

where $o_1, \ldots, o_r$ are vertices of $\mathcal{O}_\mathcal{S}$.

The following result provides a sufficient condition for solving Problem 5.1 and Problem 5.2. Our discussion in Chapter 6 will show that this condition is not necessary in general. However, it holds a central position in treatment of a number of subcases when solving Problem 5.1 and Problem 5.2 for $n = 2, 3$.

Lemma 5.6. *If*

$$\mathcal{B} \cap \text{cone}(\mathcal{O}_\mathcal{S}) \neq \{0\}, \tag{5.6}$$

then the answer to Problem 5.1 and Problem 5.2 is affirmative.

Proof. Let $b \in \mathcal{B} \cap \operatorname{cone}(\mathcal{O}_\mathcal{S}) \backslash \{0\}$. We note that, by definition of $\operatorname{cone}(\mathcal{O}_\mathcal{S})$, b satisfies $b \in C(x)$ at every point $x \in \mathcal{O}_\mathcal{S}$. Thus, the function $f : \mathcal{O}_\mathcal{S} \to \mathcal{B} \backslash \{0\}$ defined by $f(x) = b$ for all $x \in \mathcal{O}_\mathcal{S}$ satisfies all the criteria of Problem 5.1 and of Problem 5.2. □

As a dual of sorts to Lemma 5.6, Lemma 5.7 gives an easy necessary condition for Problem 5.1 and Problem 5.2.

Lemma 5.7. *Assume that the function f from Problem 5.1 (Problem 5.2) exists. Then, for every $x \in \mathcal{O}_\mathcal{S}$, there exists $0 \neq b \in \mathcal{B} \cap \mathcal{C}(x)$.*

Proof. For any such x, take $b = f(x)$. By the conditions of Problems 5.1–5.2, $f(x) \in \mathcal{B} \backslash \{0\}$ and $f(x) \in \mathcal{C}(x)$. □

Finally, let us explore the dimension of $\mathcal{O}_\mathcal{S}$. Cases $\dim \mathcal{O}_\mathcal{S} = 0$ and $\dim \mathcal{O}_\mathcal{S} = n$, as well as the case of $v_0 \in \mathcal{O}_\mathcal{S}$, prove to be particularly easy to analyze. We do that as follows:

Lemma 5.8. *If $\dim \mathcal{O}_\mathcal{S} = 0$, the answer to Problem 5.1 (Problem 5.2) is affirmative if and only if* (5.6) *holds.*

Proof. We note that in this case, $\mathcal{O}_\mathcal{S}$ consists of a single point $x \in \mathcal{S}$. Thus, $\operatorname{cone}(\mathcal{O}_\mathcal{S}) = \mathcal{C}(x)$, sufficiency is proved by Lemma 5.6, and necessity is proved by Lemma 5.7. □

Lemma 5.9. *If $v_0 \in \mathcal{O}_\mathcal{S}$, then $\operatorname{cone}(\mathcal{O}_\mathcal{S}) = \mathcal{C}(v_0)$ and the answer to Problem 5.1 (Problem 5.2) is affirmative if and only if* (5.6) *holds.*

Proof. Sufficiency is proved by Lemma 5.6. Now, by the Vertex Deletion Lemma, $\mathcal{C}(v_0) \subset \mathcal{C}(x)$ for all $x \in \mathcal{O}_\mathcal{S}$. Thus,

$$\mathcal{C}(v_0) \supset \operatorname{cone}(\mathcal{O}_\mathcal{S}) = \bigcap_{x \in \mathcal{O}_\mathcal{S}} \mathcal{C}(x) \supset \operatorname{cone}(v_0).$$

So, $\operatorname{cone}(\mathcal{O}_\mathcal{S}) = \mathcal{C}(v_0)$, and necessity thus follows from Lemma 5.7. □

Corollary 5.10. *If $\dim \mathcal{O}_\mathcal{S} = n$, then $\operatorname{cone}(\mathcal{O}_\mathcal{S}) = \mathcal{C}(v_0)$ and the answer to Problems 5.1–5.2 is affirmative if and only if* (5.6) *holds.*

Proof. We note that $\dim \mathcal{O}_\mathcal{S} = n$ implies $\mathcal{O}_\mathcal{S} = \mathcal{S} \ni v_0$. The claim follows from Lemma 5.9. □

In the remainder of our treatment of Problem 5.1 and Problem 5.2, we assume that $1 \leq \dim \mathcal{O}_\mathcal{S} \leq n-1$. We analogously note that cases $\dim \mathcal{B} = 0$ and $\dim \mathcal{B} = n$ are trivial as well, and in the remainder of this text, we assume $1 \leq \dim \mathcal{B} \leq n-1$.

5.3 Results for Low-dimensional Simplices

This section presents the main results on solving Problem 5.1 in the setting of $n = 2$ and $n = 3$. We do this by showing that if $\dim(\mathcal{O}_{\mathcal{S}}) = n - 1$ for a general n, then it is possible to characterize the solution of Problem 5.1 in terms of a smaller polytope $\mathcal{O}'_{\mathcal{S}}$, and $\mathcal{O}'_{\mathcal{S}}$ will be amenable to a complete analysis of the problem for low values of n. This line of reasoning is founded on two assumptions: for $n = 2$ and $n = 3$ we can discuss the geometric structure of $\mathcal{O}_{\mathcal{S}}$ on a case-by-case basis. Generally, this is not easily done in higher dimensions. Additionally, in the case when $n = 2$ or $n = 3$, the dimension of $\mathcal{O}_{\mathcal{S}}$ is either equal to 0, which has been discussed in Section 5.2, 1, which has been previously solved, $n - 1$, which we will discuss in this section, or n, which has also been discussed in Section 5.2. For $n \geq 4$, there is a wider range of options for $\dim(\mathcal{O}_{\mathcal{S}})$, and such a simple analysis will not work.

The main result is presented in Theorem 5.13. The consequences of Theorem 5.13 to low-dimensional systems are presented in Theorem 5.19.

Let us assume $\mathcal{O}_{\mathcal{S}}$ is $(n-1)$-dimensional. According to [67, 88], this means $\mathcal{S}$ is cut by $\mathcal{O}$ into two parts: one part containing v_0 and $p \geq 0$ other vertices, and the other containing the other $n - p \geq 1$ vertices of $\mathcal{S}$. Without loss of generality we assume that $\{v_0, v_1, \ldots, v_p\}$ are on one side of $\mathcal{O}_{\mathcal{S}}$ and $\{v_{p+1}, \ldots, v_n\}$ are on the other side, where we assume vertices of $\mathcal{S}$ on $\mathcal{O}_{\mathcal{S}}$ are in the set $\{v_{p+1}, \ldots, v_n\}$. The vertices of $\mathcal{O}_{\mathcal{S}}$ lie on those edges of $\mathcal{S}$ connecting v_i's which are on different sides of $\mathcal{O}_{\mathcal{S}}$. For the sake of readability, we employ the notation o_{ij} to denote a vertex of $\mathcal{O}_{\mathcal{S}}$ with $I(o_{ij}) = \{i, j\}$, and denote the set of all vertices of $\mathcal{O}_{\mathcal{S}}$ by $V_{\mathcal{O}_{\mathcal{S}}}$. This notation will not be used outside this section.

If there are no vertices of $\mathcal{S}$ on $\mathcal{O}_{\mathcal{S}}$, then $\mathcal{O}_{\mathcal{S}}$ has $(p+1)(n-p)$ vertices [67], but if $\mathcal{O}_{\mathcal{S}}$ contains r vertices of $\mathcal{S}$, then $\mathcal{O}_{\mathcal{S}}$ has $(p+1)(n-p) - pr$ vertices. At this point we introduce a mild abuse of notation with the convention that if $v_j \in \mathcal{O}_{\mathcal{S}}$, then $o_{ij} = v_j$ for all $i = 0, \ldots, p$.

Let us introduce the following notation. Let

$$\{i_1, i_2, \ldots, i_k | j_1, j_2, \ldots, j_l\} = \mathrm{co}\{o_{i_\alpha j_\beta} : 1 \leq \alpha \leq k, 1 \leq \beta \leq l\}.$$

We observe that since $I(o_{ij}) = \{i, j\}$, if $x \in \{i_1, i_2, \ldots, i_k | j_1, j_2, \ldots, j_l\}$ then $I(x) \subset \{i_1, i_2, \ldots, i_k\} \cup \{j_1, j_2, \ldots, j_l\}$. Also observe that $\mathcal{O}_{\mathcal{S}} = \{0, \ldots, p | p+1, \ldots, n\}$.

In fact, it is noted in [67] that $\mathcal{O}_{\mathcal{S}}$ has the structure of a Cartesian product of two simplices: one of dimension p and another of dimension $n-p-1$. An obvious caveat is that if one of the vertices v_i lies on $\mathcal{O}_{\mathcal{S}}$, some points on those simplices might coincide, i.e., the edge between them is of length 0. However, in the interest of clearer exposition, we will still treat those points as separate and edges of length 0 as

legitimate 1-dimensional edges of $\mathcal{O}_{\mathcal{S}}$.

Lemma 5.11. *Let p be such that $\mathcal{O}_{\mathcal{S}} = \{0, \ldots, p|p+1, \ldots, n\}$. Let $\mathcal{A} = \{i_1, \ldots, i_k|j_1, \ldots, j_l\} \subset \mathcal{O}_{\mathcal{S}}$, and $\mathcal{A}' = \{i'_1, \ldots, i'_{k'}|j'_1, \ldots, j'_{l'}\} \subset \mathcal{O}_{\mathcal{S}}$. Let $L = \{i_1, \ldots, i_k\} \cap \{i'_1, \ldots, i'_{k'}\}$. Analogously, let $R = \{j_1, \ldots, j_l\} \cap \{j'_1, \ldots, j'_{l'}\}$. Then, $\mathcal{A} \cap \mathcal{A}' = \{L|R\}$.*

Proof. By definition every vertex of $\{L|R\}$ is a vertex of $\mathcal{A}$ and of $\mathcal{A}'$, so $\{L|R\} \subset \mathcal{A} \cap \mathcal{A}'$. Conversely, suppose $x \in \mathcal{A} \cap \mathcal{A}'$. Since $x \in \mathcal{A}$, $I(x) \subset \{i_1, \ldots, i_k, j_1, \ldots, j_l\}$ and since $x \in \mathcal{A}'$, $I(x) \subset \{i'_1, \ldots, i'_{k'}, j'_1, \ldots, j'_{l'}\}$. Hence, $I(x) \subset L \cup R$, where we use the fact that $\{i_1, \ldots, i_k\} \cap \{j'_1, \ldots, j'_{l'}\} = \emptyset$ and $\{i'_1, \ldots, i'_{k'}\} \cap \{j_1, \ldots, j_l\} = \emptyset$. It follows $x \in \{L|R\}$. □

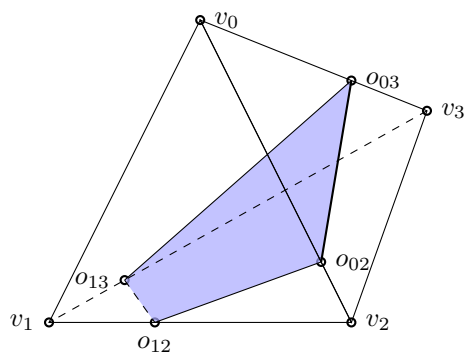

Figure 5.1: The set $\mathcal{O}_{\mathcal{S}}$ for Example 5.12. The edge from o_{02} to o_{03} forms $\mathcal{O}'_{\mathcal{S}}$.

Example 5.12. *Let us consider the case of $n = 3$, with $\mathcal{O}_{\mathcal{S}}$ a quadrilateral with vertices $o_{02} \in \mathrm{co}\{v_0, v_2\}$, $o_{03} \in \mathrm{co}\{v_0, v_3\}$, $o_{12} \in \mathrm{co}\{v_1, v_2\}$, and $o_{13} \in \mathrm{co}\{v_1, v_3\}$. By definition, $\mathcal{C}(o_{02}) \subset \mathcal{C}(o_{12})$ and $\mathcal{C}(o_{03}) \subset \mathcal{C}(o_{13})$. In fact, one can easily show that the cone of any point on the edge $\overline{o_{12}o_{02}}$ will be larger than $\mathcal{C}(o_{02})$, and analogously for the edge $\overline{o_{13}o_{03}}$. Thus, we reach the idea that if a continuous function satisfying Problem 5.1 exists on the convex set $\mathrm{co}\{o_{02}, o_{03}\}$ containing the most restrictive cones, then that function can easily be extended to the entire $\mathcal{O}_{\mathcal{S}}$.*

Theorem 5.13 will serve to prove the above claim. The procedure outlined in the proof of Theorem 5.13, adapted to this example, is as follows. If a function f satisfying Problem 5.1 can be defined on the edge $\overline{o_{02}, o_{03}}$, we can also define it on the edge $\overline{o_{12}o_{02}}$ by $f(x) = f(o_{02})$ and on the edge $\overline{o_{13}o_{03}}$ by $f(x) = f(o_{03})$. We note that such f is non-zero and satisfies the cone condition $f(x) \in \mathcal{C}(x)$ because $\mathcal{C}(o_{02}) \subset \mathcal{C}(o_{12})$ and $\mathcal{C}(o_{03}) \subset \mathcal{C}(o_{13})$. So far f has been defined on three edges of $\mathcal{O}_{\mathcal{S}}$. Then f can be defined on the remainder of $\mathcal{O}_{\mathcal{S}}$, which consists of its interior as well as the relative interior of the edge $\overline{o_{12}o_{13}}$, by using a retraction r — a continuous map from $\mathcal{O}_{\mathcal{S}}$ to the three edges of $\mathcal{O}_{\mathcal{S}}$ on which f is already defined, such that r is identity on those three edges. More formally, it can be shown, as in Theorem 5.13, that the function $f(x) = f(r(x))$ exists and solves Problem 5.1.

Theorem 5.13 (Dimension Reduction)**.** *Let* $\dim \mathcal{O}_{\mathcal{S}} = n-1$*,* $v_0 \notin \mathcal{O}_{\mathcal{S}}$*, and* $p > 0$*. Define* $V'_{\mathcal{O}_{\mathcal{S}}} = \{o \in V_{\mathcal{O}_{\mathcal{S}}} \mid o \in \mathrm{co}\{v_0, v_j\}$ *for some* $1 \leq j \leq n\}$*, and let* $\mathcal{O}'_{\mathcal{S}} = \mathrm{co}(V'_{\mathcal{O}_{\mathcal{S}}})$*. Then the answer to Problem 5.1 is affirmative if and only if it is affirmative for* $\mathcal{O}'_{\mathcal{S}}$*.*

Proof. Since $V'_{\mathcal{O}_{\mathcal{S}}} \subset V_{\mathcal{O}_{\mathcal{S}}}$ it follows that $\mathcal{O}'_{\mathcal{S}} \subset \mathcal{O}_{\mathcal{S}}$. Thus, if there exists $f : \mathcal{O}_{\mathcal{S}} \to \mathcal{B}\backslash\{0\}$ solving Problem 5.1 then $f|_{\mathcal{O}'_{\mathcal{S}}} : \mathcal{O}'_{\mathcal{S}} \to \mathcal{B}\backslash\{0\}$ also solves Problem 5.1. Conversely, suppose there exists $f' : \mathcal{O}'_{\mathcal{S}} \to \mathcal{B} \setminus \{0\}$ which solves Problem 5.1. From our notational convention, $\mathcal{O}_{\mathcal{S}} = \{0, 1, \ldots, p \mid p+1, \ldots, n\}$, $V'_{\mathcal{O}_{\mathcal{S}}} = \{o_{0(p+1)}, \ldots, o_{0n}\}$, and $\mathcal{O}'_{\mathcal{S}} = \{0 \mid p+1, \ldots, n\}$.

We now proceed with the main topological argument. Informally, we build a incomplete skeleton of $\mathcal{O}_{\mathcal{S}}$, starting with $\mathcal{O}'_{\mathcal{S}}$, and adding in each step additional edges and faces of $\mathcal{O}_{\mathcal{S}}$ until in the last step all of $\mathcal{O}_{\mathcal{S}}$ is added. We then use topological methods to show that Problem 5.1 for the set obtained in each step can be reduced to the same problem applied to the set from the previous step, thus going back from $\mathcal{O}_{\mathcal{S}}$ to $\mathcal{O}'_{\mathcal{S}}$.

We build an incomplete skeleton of $\mathcal{O}_{\mathcal{S}}$ as follows. Let $\mathcal{O}^1_{\mathcal{S}} = \mathcal{O}'_{\mathcal{S}}$, and for all $2 \leq k \leq n$, let

$$\mathcal{O}^k_{\mathcal{S}} = \mathcal{O}^{k-1}_{\mathcal{S}} \cup \bigcup_{\substack{0<i_1<\ldots<i_\alpha\leq p,\\ p<j_1<\ldots<j_\beta\leq n,\\ \alpha+\beta=k,\alpha,\beta\geq 1}} \{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}.$$

Note that $\mathcal{O}^k_{\mathcal{S}}$ is not the actual skeleton of $\mathcal{O}_{\mathcal{S}}$, i.e., $\mathcal{O}^k_{\mathcal{S}} \neq \partial^{(k)}\mathcal{O}_{\mathcal{S}}$. The set $\mathcal{O}^k_{\mathcal{S}}$ contains $\mathcal{O}'_{\mathcal{S}}$, and *some other* $k-1$-dimensional faces of $\mathcal{O}_{\mathcal{S}}$. On the other hand, $\partial^{(k)}\mathcal{O}_{\mathcal{S}}$, defined in Section 2.1, contains all k-dimensional faces of $\mathcal{O}_{\mathcal{S}}$.

Observe that each set $\mathcal{A}$ defined by $\mathcal{A} = \{0, i_1, \ldots, i_\alpha \mid j_1, \ldots, j_\beta\}$ is a closed, convex polytope of some dimension d, so by Theorem 2.22 it is homeomorphic to the closed ball $\mathbb{B}^d$, and its boundary is homeomorphic to the $(d-1)$-dimensional sphere $\mathbb{S}^{d-1}$. We claim that

$$\partial\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \setminus \mathcal{O}^{k-1}_{\mathcal{S}} = \mathrm{int}(\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}). \tag{5.7}$$

There are three points to the proof of the claim.

(i) First, we claim that $\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \in \partial\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$. To show this, we note that if $x \in \{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$, then

$$I(x) \subset \{i_1, \ldots, i_\alpha, j_1, \ldots, j_\beta\} \subset \{0, i_1, \ldots, i_\alpha, j_1, \ldots, j_\beta\},$$

so $x \in \{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$. Moreover, $\partial\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$ consists of points with $|I(x)| \leq$

$\alpha + \beta$.

(ii) $\operatorname{int}(\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}) \cap \mathcal{O}_{\mathcal{S}}^{k-1} = \emptyset$. Assume $x \in \operatorname{int}(\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\})$. Then $0 \notin I(x)$ and $I(x) = \alpha + \beta = k$. However, if $x \in \mathcal{O}_{\mathcal{S}}^{k-1}$, then either $0 \in I(x)$ or $|I(x)| \leq k - 1$.

(iii) $\partial\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \setminus \operatorname{int}(\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}) \subset \mathcal{O}_{\mathcal{S}}^{k-1}$. This follows because if $x \in \partial\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \setminus \operatorname{int}(\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\})$, then either $0 \in I(x)$ and $|I(x)| \leq k$ so $x \in \mathcal{O}_{\mathcal{S}}^{k-1}$; or $0 \notin I(x)$ and $|I(x)| \leq k - 1$, so again $x \in \mathcal{O}_{\mathcal{S}}^{k-1}$.

What we have shown so far is that $\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$ is homeomorphic to a closed ball, and $\partial\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \cap \mathcal{O}_{\mathcal{S}}^{k-1}$ is homeomorphic to its boundary sphere $\mathbb{S}^{d-1}$ with an open connected set $\operatorname{int}(\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\})$ of dimension $d - 1$ cut out of it. Now we show there exists a retraction from $\mathbb{B}^d$ to the punctured sphere $\mathbb{S}^{d-1} \setminus \mathcal{P}$, where $\mathcal{P}$ is homeomorphic to an open ball of dimension $d - 1$. Since $\mathcal{P}$ has dimension $d - 1$, $\mathbb{S}^{d-1}$ is split into two connected parts: $\mathcal{P}$ and $\mathbb{S}^{d-1} \setminus \mathcal{P}$. The retraction argument is standard in topology. We provide the proof since it is integral to our results.

First, let us note that $\mathbb{B}^d$ is homeomorphic to the upper half-ball $\mathbb{B}^{+d} = \{x \in \mathbb{B}^d : x_1 \geq 0\}$. The precise homeomorphism is not difficult to find, but one can simply imagine taking the ball and flattening its lower half. Now, our sphere $\mathbb{S}^{d-1}$ was mapped by this to the boundary of $\mathbb{B}^{+d}$. Furthermore, without loss of generality[1], we can assume that the closed part $\mathbb{S}^{d-1} \setminus \mathcal{P}$ makes up the bottom of the half-ball: $\{x \in \mathbb{B}^d : x_1 = 0\}$, while the open part $\{x \in \mathbb{B}^d : x_1 > 0\}$ corresponds to $\mathcal{P}$.

Let us define the function $r'_{\mathcal{A}} : \mathbb{B}^{+d} \to \{x \in \mathbb{B}^d : x_1 = 0\}$ by $r'_{\mathcal{A}}(x_1, x_2, \ldots, x_n) = (0, x_1, x_2, \ldots, x_n)$. Clearly, this is a valid retraction and thus, we have obtained a retraction from $\mathbb{B}^{+d}$ to $\{x \in \mathbb{B}^d : x_1 = 0\}$. Now, using the fact that $\mathbb{B}^{+d}$ is homeomorphic to $\mathbb{B}^d$, while the same homeomorphism takes $\{x \in \mathbb{B}^d : x_1 = 0\}$ to $\mathbb{S}^{d-1} \setminus \mathcal{P}$, we know there thus exists a retraction $r''_{\mathcal{A}} : \mathbb{B}^d \to \mathbb{S}^{d-1} \setminus \mathcal{P}$. Finally, reminding ourselves that there exists a homeomorphism between $\mathcal{A}$ and $\mathbb{B}^d$ which takes $\mathbb{S}^{d-1} \setminus \mathcal{P}$ to $\partial\mathcal{O}_{\mathcal{S}} \cap \mathcal{O}_{\mathcal{S}}^{k-1}$, by "pushing" $r''_{\mathcal{A}}$ through that homeomorphism, we obtain a retraction $r_{\mathcal{A}} : \{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \to \partial\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \cap \mathcal{O}_{\mathcal{S}}^{k-1}$.

Now we glue these retractions to each other. To do that, we need to know that for $\mathcal{A}$'s with constant $\alpha + \beta = k$, all the different retractions $r_{\mathcal{A}} : \{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \to \partial\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\} \cap \mathcal{O}_{\mathcal{S}}^{k-1}$ agree on the intersections of their domains. That is, if $\mathcal{A} \neq \mathcal{A}'$, then $r_{\mathcal{A}}|_{\mathcal{A} \cap \mathcal{A}'} \equiv r_{\mathcal{A}'}|_{\mathcal{A} \cap \mathcal{A}'}$. (The claim is obvious if $\mathcal{A} = \mathcal{A}'$.) Now, let $\mathcal{A} = \{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$ and $\mathcal{A}' = \{0, i'_1, \ldots, i'_{\alpha'} | j'_1, \ldots, j'_{\beta'}\}$. We noted in Lemma 5.11 that $\mathcal{A} \cap \mathcal{A}' = \{0, \{i_1, \ldots, i_\alpha\} \cap \{i'_1, \ldots, i'_{\alpha'}\} | \{j_1, \ldots, j_\beta\} \cap \{j'_1, \ldots, j'_{\beta'}\}\}$. Since we assumed that $\mathcal{A} \neq \mathcal{A}'$, there needs to be an element in $\{i_1, \ldots, i_\alpha, j_1, \ldots, j_\beta\}$ which is not an element

[1] Really, this is done through another homeomorphism: this time, imagine, before flattening the ball, choosing the part that needs to be flattened to be $\mathbb{S}^{d-1} \setminus \mathcal{P}$.

of the set $\{i'_1, \ldots, i'_{\alpha'}, j'_1, \ldots, j'_{\beta'}\}$ and vice versa (note that both of those sets have k elements, so one cannot be a subset of the other).

Thus, $\mathcal{A} \cap \mathcal{A}'$ will not contain more than $k-1$ non-zero vertices of $\mathcal{S}$ in its notation, and hence it will be in both $\mathcal{O}_{\mathcal{S}}^{k-1}$ (by the definition of $\mathcal{O}_{\mathcal{S}}^{k-1}$), and in $\partial\mathcal{A}$ (as none of its elements can be in the interior of $\mathcal{A}$: the expansion as a convex sum of every element in the interior needs to contain every vertex mentioned in the notation of $\mathcal{A}$). Analogously, $\mathcal{A} \cap \mathcal{A}' \in \partial\mathcal{A}' \cap \mathcal{O}_{\mathcal{S}}^{k-1}$, which is the image of the retraction $r_{\mathcal{A}'}$.

Hence, we know that $r_{\mathcal{A}'}|_{\mathcal{A}\cap\mathcal{A}'}$ is an identity map, and so is $r_{\mathcal{A}}|_{\mathcal{A}\cap\mathcal{A}'}$. Thus, these two retractions can indeed be glued together. By iterating this procedure for all $\mathcal{A}$, we obtain a glued retraction $r^k : \mathcal{O}_{\mathcal{S}}^k \to \mathcal{O}_{\mathcal{S}}^{k-1}$ which takes each k-dimensional edge in $\mathcal{O}_{\mathcal{S}}^k$ to its boundary. Let us note what this retraction does. For every point $x \in \mathcal{O}_{\mathcal{S}}^k$, if x is also in $\mathcal{O}_{\mathcal{S}}^{k-1}$, it will not do anything. Thus, $\mathcal{C}(r^k(x)) = \mathcal{C}(x)$.

If $x \notin \mathcal{O}_{\mathcal{S}}^{k-1}$, then either x is in the interior of some $\mathcal{A} = \{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$ that is being added to $\mathcal{O}_{\mathcal{S}}^k$, or it is in the interior of $\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$. Now, if x is in the interior of $\{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$, r^k maps it to a point in the boundary of $\mathcal{A}$. In that case, from Lemma 5.5 we note that $\mathcal{C}(r^k(x)) \subset \mathcal{C}(x)$.

If x is in the interior of $\{i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$, then $I(x) = \{i_1, \ldots, i_\alpha, j_1, \ldots, j_\beta\}$. On the other hand, $\mathcal{A} = \{0, i_1, \ldots, i_\alpha | j_1, \ldots, j_\beta\}$, so for any point $y \in \mathcal{A}$, $I(y) \subset \{i_1, \ldots, i_\alpha, j_1, \ldots, j_\beta\}$. Thus, $\mathcal{C}(y) \subset \mathcal{C}(x)$. Thus, specifically $\mathcal{C}(r^k(x)) \subset \mathcal{C}(x)$.

In all three cases, we deduce that $\mathcal{C}(r^k(x)) \subset \mathcal{C}(x)$. By composing $r = r^2 \circ r^3 \circ \cdots \circ r^n$, we obtain a retraction $r : \mathcal{O}_{\mathcal{S}} = \mathcal{O}_{\mathcal{S}}^n \to \mathcal{O}_{\mathcal{S}}^1 = \mathcal{O}'_{\mathcal{S}}$. Define $f(x) = f'(r(x))$. We obtained a nowhere vanishing function f on $\mathcal{O}_{\mathcal{S}}$ such that $f(x) = f'(r^2(r^3(\ldots(r^n(x))\ldots)))$. Thus, $f(x)$ is contained in $\mathcal{C}(r^2(r^3(\ldots(r^n(x))\ldots))) \subset \mathcal{C}(r^2(r^3(\ldots(r^{n-1}(x))\ldots))) \subset \ldots \subset \mathcal{C}(x)$. This function satisfies the conditions of Problem 5.1. $\square$

We now proceed to resolving Problem 5.1 for the case of $n = 2$ and $n = 3$. We have previously assumed that $\dim \mathcal{O}_{\mathcal{S}} \neq 0$ and $\dim \mathcal{O}_{\mathcal{S}} \neq n$, as these cases were analyzed in Lemma 5.8 and Corollary 5.10. The case of $n = 2$ is thus reduced to $\dim \mathcal{O}_{\mathcal{S}} = 1$. As we have also required $1 \leq \dim \mathcal{B} < n$, we conclude that $\dim \mathcal{B} = 1$ is the only interesting case. However, the case of $\dim \mathcal{B} = 1$ is resolved by Theorem 1 in [125].

This resolves the case of $n = 2$, as well as $\dim \mathcal{B} = 1$. The only cases that remain are when $n = 3$, $\dim \mathcal{B} = 2$, and $\dim \mathcal{O}_{\mathcal{S}}$ is either 1 or 2. We will see that, when $\dim \mathcal{O}_{\mathcal{S}} = 1$, an argument based on linear algebra applies. On the other hand, a purely topological argument applies when $\dim \mathcal{O}_{\mathcal{S}} = 2$.

First we examine why a sufficiently high dimension for $\mathcal{B}$ resolves Problem 5.1.

Lemma 5.14. *Suppose $\mathcal{O}_\mathcal{S} = \mathrm{co}\{o_1, \ldots, o_{\kappa+1}\}$ where the o_i's are the vertices of $\mathcal{O}_\mathcal{S}$. If there exists a linearly independent set $\{b_i \in \mathcal{B} \cap \mathcal{C}(o_i) \mid i = 1, \ldots, \kappa+1\}$, then the answer to Problem 5.1 is affirmative.*

Proof. Let $f : \mathcal{O}_\mathcal{S} \to \mathcal{B}$ be defined by $f(\sum_{i=1}^{\kappa+1} \alpha_i o_i) = \alpha_i b_i$, where $\sum \alpha_i = 1$ and $\alpha_i \geq 0$. Necessarily $f(x) \neq 0$ for $x \in \mathcal{O}_\mathcal{S}$ for otherwise the b_i's would be linearly dependent. Also, by a standard convexity argument $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$. □

The following is the key result in the case of $\dim \mathcal{O}_\mathcal{S} = 1$.

Lemma 5.15. *Let $n = 3$, $\dim \mathcal{B} = 2$, and let o_1 and o_2 be vertices of $\mathcal{O}_\mathcal{S}$. Then there exist linearly independent vectors $\{b_1, b_2 \mid b_i \in \mathcal{B} \cap \mathcal{C}(o_i)\}$. Moreover, if $\mathcal{O}_\mathcal{S} = \mathrm{co}\{o_1, o_2\}$, the answer to Problem 5.1 is affirmative.*

Proof. First we assume $o_1 \in \mathrm{int}(\mathcal{F}_i)$ for some $i \in \{0, 1, 2, 3\}$. By the definition of $\mathcal{C}(o_1)$, it is a closed half space or $\mathbb{R}^3$, so there exist linearly independent vectors $b_{11}, b_{12} \in \mathcal{B} \cap \mathcal{C}(o_1)$. We claim $\mathcal{B} \cap \mathcal{C}(o_2) \neq \{0\}$. If $o_2 \in \mathrm{int}(\mathcal{F}_i)$ for some $i \in \{0, 1, 2, 3\}$ then Lemma 5.14 proves the claim. Instead, assume without loss of generality that $o_2 \in \mathcal{F}_1 \cap \mathcal{F}_2$. Then $\mathcal{C}(o_2) = \{y \in \mathbb{R}^3 \mid h_1 \cdot y \leq 0, h_2 \cdot y \leq 0\}$. Let $\mathcal{B} = \mathrm{Ker}(F^T)$ for some $M \in \mathbb{R}^{3\times 1}$. Finding $0 \neq y \in \mathcal{B} \cap \mathcal{C}(o_2)$ is equivalent to solving

$$\begin{bmatrix} h_1^T \\ h_2^T \\ F^T \end{bmatrix} y = \begin{bmatrix} s_1 \\ s_2 \\ 0 \end{bmatrix} \tag{5.8}$$

for some $s_1, s_2 \in \mathbb{R}_0^-$ and $y \neq 0$. Because $\{h_1, h_2\}$ are linearly independent, $\mathrm{rank}(H) \geq 2$, where H is the matrix appearing on the left hand side of equation (5.8). If $\mathrm{rank}(H) = 3$, then let

$$[y_1 \;\; y_2] = H^{-1} \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}.$$

Since $(-1, 0, 0)$ and $(0, -1, 0)$ are linearly independent, y_1 and y_2 are linearly independent as well.

Next, assume $\mathrm{rank}(H) = 2$. In other words, $F = c_1 h_1 + c_2 h_2$ for some $c_1, c_2 \in \mathbb{R}$. Then, by taking $s_1 = s_2 = 0$, equation (5.8) reduces to $[h_1 \; h_2]^T y = 0$. By the rank-nullity theorem, there exists $y \neq 0$ satisfying this equation. Moreover, if without loss of generality $v_0 = 0$, then $y \in \mathcal{F}_1 \cap \mathcal{F}_2 = \mathrm{co}\{v_0, v_3\}$, and we can take $y = v_3$. We have shown there exist linearly independent $b_{11}, b_{12} \in \mathcal{C}(o_1)$ and there exists $0 \neq b_2 \in \mathcal{C}(o_2)$. We claim at least one of the pairs $\{b_{11}, b_2\}$ and $\{b_{12}, b_2\}$ is linearly independent.

For otherwise there exist $c_1, c_2 \in \mathbb{R}$ such that $b_2 = c_1 b_{11} = c_2 b_{12}$, implying b_{11} and b_{12} are linearly independent, a contradiction. We conclude there exists a linearly independent set $\{b_1, b_2 \mid b_i \in \mathcal{C}(o_i)\}$.

Next we assume neither o_1 nor o_2 lies in the interior of a facet. Without loss of generality suppose $o_1 \in \mathcal{F}_1 \cap \mathcal{F}_2$ and $o_2 \in \mathcal{F}_1 \cap \mathcal{F}_3$. If either $\mathcal{C}(o_1)$ or $\mathcal{C}(o_2)$ contains two linearly independent vectors, then by the previous argument, we are done. Otherwise, by the previous argument again $v_3 \in \mathcal{C}(o_1)$ and $v_2 \in \mathcal{C}(o_2)$. Since $\{v_2, v_3\}$ are linearly independent, we are done. Finally, if $\mathcal{O}_\mathcal{S} = \mathrm{co}\{o_1, o_2\}$, then by Lemma 5.14 the answer to Problem 5.1 is affirmative. □

Remark 5.16. *We note that the results of Lemma 5.14 and Lemma 5.15 also equally work for Problem 5.2.*

The remaining case to study is when $n = 3$, $\dim \mathcal{O}_\mathcal{S} = 2$, and $\dim \mathcal{B} = 2$. Assuming $v_0 \notin \mathcal{O}_\mathcal{S}$ (which is a trivial case discussed in Lemma 5.9), there are four topologically distinct cases for $\mathcal{O}_\mathcal{S}$, depending on the way $\mathcal{O}$ cuts $\mathcal{S}$. These are given in Figures 5.1 and 5.2. In Figure 5.1, $\mathcal{O}_\mathcal{S}$ is a quadrangle. In that case, $p = 1$; that is, there are two vertices of $\mathcal{S}$ on each side of $\mathcal{O}$. Then we can apply Theorem 5.13 to reduce $\mathcal{O}_\mathcal{S}$ to $\mathcal{O}'_\mathcal{S}$, and according to the construction in the proof, $\mathcal{O}'_\mathcal{S}$ has dimension 1, and we can apply Lemma 5.15. Similarly, in the cases given in middle and the rightmost configuration of Figure 5.2, we can apply Theorem 5.13 to reduce $\mathcal{O}_\mathcal{S}$ to $\mathcal{O}'_\mathcal{S}$ with $\dim \mathcal{O}'_\mathcal{S}$ being either 0 or 1, respectively. Finally, in the situation given by the leftmost configuration of Figure 5.2, we draw upon a proof method already utilized in [23], which is based on the KKM lemma. Here we employ a variant found in [16].

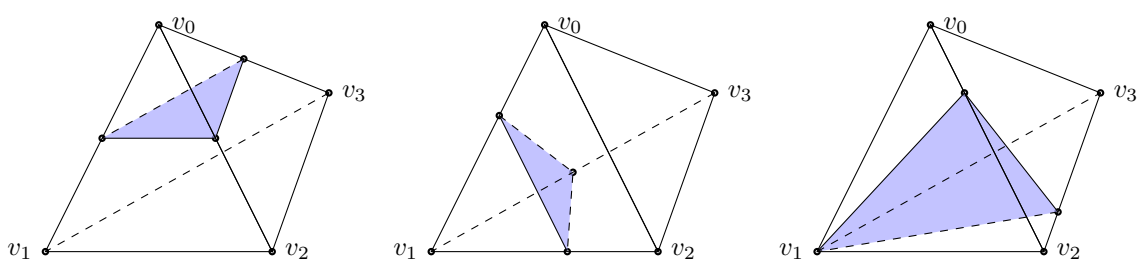

Figure 5.2: Three of the four possible configurations of set $\mathcal{O}_\mathcal{S}$ for $n = 3$ and $\dim \mathcal{O}_\mathcal{S} = 2$, with the fourth one given in Figure 5.1. The leftmost configuration is addressed by Theorem 5.18, while the other two can be reduced using Theorem 5.13.

Lemma 5.17. *Let $\mathcal{P} = \mathrm{co}\{w_1, \ldots, w_{n+1}\}$ be an n-dimensional simplex. Also, let $\{\mathcal{Q}_1, \ldots, \mathcal{Q}_{n+1}\}$ be a collection of sets covering $\mathcal{P}$ such that*

(P1) *Vertex $w_i \in \mathcal{Q}_i$ and $w_i \notin \mathcal{Q}_j$ for $j \neq i$.*

(P2) *If without loss of generality $x \in \mathrm{co}\{w_1, \ldots, w_l\}$ for some $1 \leq l \leq n+1$, then $x \in \mathcal{Q}_1 \cup \cdots \cup \mathcal{Q}_l$.*

Then $\bigcap_{1=1}^{n+1} \overline{\mathcal{Q}}_i \neq \emptyset$.

Theorem 5.18. *Let* $n = 3$ *and suppose* $\mathcal{O}_{\mathcal{S}} = \mathrm{co}\{o_1, o_2, o_3\}$ *with* $v_0 \notin \mathcal{O}_{\mathcal{S}}$ *and* $o_i \in (v_0, v_i]$, $i = 1, 2, 3$. *The answer to Problem 5.1 is affirmative if and only if*

$$\mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}}) \neq \{0\}.$$

Proof. Sufficiency is clear: if $0 \neq b \in \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$, a constant function $f(x) = b$ satisfies Problem 5.1. For necessity, suppose there exists $f : \mathcal{O}_{\mathcal{S}} \to \mathcal{B} \setminus \{0\}$ such that $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_{\mathcal{S}}$. By way of contradiction suppose $\mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}}) = \{0\}$. Since $o_i \in (v_0, v_i]$, $i = 1, 2, 3$, we have

$$\mathrm{cone}(\mathcal{O}_{\mathcal{S}}) = \{y \in \mathbb{R}^n \mid h_j \cdot y \leq 0\,, j = 1, 2, 3\}\,.$$

Define the sets

$$\mathcal{Q}_i := \{x \in \mathcal{O}_{\mathcal{S}} \mid h_i \cdot f(x) > 0\}\,, \quad i = 1, 2, 3\,. \tag{5.9}$$

Now we verify the conditions of Lemma 5.17.

Firstly, we claim that $\{\mathcal{Q}_i\}$ cover $\mathcal{O}_{\mathcal{S}}$. Suppose otherwise. Then there exists $x \in \mathcal{O}_{\mathcal{S}}$ such that $h_j \cdot f(x) \leq 0$, $j = 1, 2, 3$. Hence $f(x) \in \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$, so $f(x) = 0$, a contradiction to f being non-vanishing on $\mathcal{O}_{\mathcal{S}}$. Secondly, we verify property (P1). We claim that $o_i \in \mathcal{Q}_i$ for $i = 1, 2, 3$. For suppose not. Then $h_i \cdot f(x) \leq 0$. Additionally, because $f(o_i) \in \mathcal{C}(o_i)$, $h_j \cdot f(x) \leq 0$, $j \in \{1, 2, 3\} \setminus \{i\}$. We conclude $f(o_i) \in \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$, so $f(o_i) = 0$, a contradiction. Next we claim $o_i \notin \mathcal{Q}_j$, $j \neq i$. This is immediate since $f(o_i) \in \mathcal{C}(o_i)$ implies $h_j \cdot f(o_i) \leq 0$, $j \neq i$. Thirdly, we verify property (P2). Suppose without loss of generality (by reordering the indices $\{1, 2, 3\}$) $x \in \mathrm{co}\{o_1, \ldots, o_r\}$ for some $1 \leq r \leq 3$. We claim $x \in \mathcal{Q}_1 \cup \cdots \cup \mathcal{Q}_r$. For suppose not. Then $h_j \cdot f(x) \leq 0$, $j = 1, \ldots, r$. Also, it is easily verified that $\mathcal{C}(x) = \{y \in \mathbb{R}^n \mid h_j \cdot y \leq 0\,, j = r+1, \ldots, 3\}$. Thus, $h_j \cdot f(x) \leq 0$, $j = r+1, \ldots, 3$. Hence, $f(x) \in \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$, so $f(x) = 0$, a contradiction to f being non-vanishing on $\mathcal{O}_{\mathcal{S}}$.

We have verified (P1)-(P2) of Lemma 5.17. Applying the lemma, there exists $\overline{x} \in \bigcap_{i=1}^{3} \overline{\mathcal{Q}}_i$; that is, $h_j \cdot f(\overline{x}) \geq 0$, $j = 1, 2, 3$. We conclude that $-f(\overline{x}) \in \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$, so $f(\overline{x}) = 0$, a contradiction. □

The following result finally resolves Problem 5.1 in all cases of interest.

Theorem 5.19. *Let* $\mathcal{S}$, $\mathcal{B}$ *and* $\mathcal{O}_{\mathcal{S}}$ *be as above, and let* $n \in \{2, 3\}$. *If* $n = 3$, $\dim \mathcal{B} = 2$ *and* $\mathcal{O}_{\mathcal{S}}$ *does not satisfy the conditions of Theorem 5.18, then the answer to Problem 5.1 is affirmative. Otherwise, the answer to Problem 5.1 is affirmative if and only if* $\mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}}) \neq \{0\}$.

Proof. The discussion prior to Lemma 5.14, as well as Lemma 5.15 and Theorem 5.18, covered all the cases except for the one where $\dim \mathcal{B} = \dim \mathcal{O}_{\mathcal{S}} = 2$ and $\mathcal{O}_{\mathcal{S}}$ does not satisfy the conditions of Theorem 5.18. However, in that case, as described prior to Lemma 5.17, Theorem 5.13 reduces $\mathcal{O}_{\mathcal{S}}$ to $\mathcal{O}'_{\mathcal{S}}$ of dimension 0 or 1. Applying Lemma 5.15, we now obtain that the answer to Problem 5.1 is affirmative. □

5.4 Two-input Systems

In this section, we consider the case of a two-input system. Such a system is special for two reasons: first, we will reduce the problem of a topological obstruction to a problem of the existence of a particular map from $\mathbb{B}^2$ to $\mathbb{S}^1$. Since both the two-dimensional ball (i.e., disk) and one-dimensional sphere (i.e., circle) are commonly used and low-dimensional, we have a well-developed geometric intuition and formal theory for working with these objects. The second reason is deeper: as we will show in the remainder of this section, discussing the null-homotopy of particular maps from $\mathbb{S}^k \to \mathbb{S}^{m-1}$ is crucial to the solution of the problem of a topological obstruction. In the case when $m = 2$, all such maps with $k > 1$ are necessarily null-homotopic, which significantly simplifies our discussions. In the case when $m > 2$, Hopf showed in [61] that not all maps $\mathbb{S}^k \to \mathbb{S}^{m-1}$ are null-homotopic. Thus, the theory of this section cannot be easily directly extended to systems with more inputs.

Going back to the case when $m = 2$, we first transform the statement of Problem 5.1 so that explicit consideration of $\mathcal{B}$ can be avoided.

Since $\mathcal{B} \subset \mathbb{R}^n$ is an m-dimensional subspace, we can identify it with $\mathbb{R}^m$ through a linear transformation $Q \in \mathbb{R}^{n\times m}$ whose columns form an orthonormal basis of $\mathcal{B}$, such that $Q^TQ = I$ and $QQ^T|_{\mathcal{B}} \equiv id$. This interprets the function $f : \mathcal{O}_{\mathcal{S}} \to \mathcal{B}\backslash\{0\}$ as $\frac{Q^Tf}{\|Q^Tf\|} = \tilde{f} : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$. Next, the requirement $f(x) \in \mathcal{C}(x)$ says $h_j \cdot f(x) \leq 0$, $j \in J(x)$. Since $QQ^T|_{\mathcal{B}} \equiv id$, and $f(x) \in \mathcal{B}$ for $x \in \mathcal{O}_{\mathcal{S}}, h_j \cdot f(x) = h_j \cdot QQ^T f(x)$. Let $\tilde{h}_j := Q^T h_j \in \mathbb{R}^m$. Then $f(x) \in \mathcal{C}(x)$ is equivalent to $\tilde{h}_j \cdot \tilde{f}(x) \leq 0$, $j \in J(x)$. In what follows we remove the tilde's from the variables f and h_j. This is clearly an abuse of notation, and we apologize for the possible confusion, but we feel compelled to do it for the sake of readability. In Chapter 6, which is less involved, we do all the calculations without this abuse. We arrive at the main problem studied in this section.

Problem 5.20 (Topological Obstruction)**.** *Does there exist a continuous function $f : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$ satisfying*

$$f(x) \in \mathcal{C}(x)\,, \qquad x \in \mathcal{O}_{\mathcal{S}}\,. \tag{5.10}$$

From the foregoing discussion, problem 5.20 is equivalent to Problem 5.1.

To motivate the approach of this section, we give an example discussing the solvability of the RCP on a three-dimensional simplex with two inputs. Note that Problem 5.1 in this particular example can be answered using the results from Section 5.3. However, we use this example to illustrate the approach that we will be taking when determining the solvability of Problem 5.1 for general two-input systems.

5.4.1 Example

Let $m = 2$, and let $\mathcal{S}$ be the standard three-dimensional simplex with $v_0 = 0$ and $v_i = e_i$ for $i \in \{1, 2, 3\}$. Suppose that $\mathcal{O}_\mathcal{S} = \{x \in \mathcal{S} \mid Ax + a \in \mathcal{B}\}$ is given by a triangle with vertices $o_i = (v_0 + v_i)/2$, $i \in I = \{1, 2, 3\}$. In this case $J(o_i) = I \backslash \{i\}$ for all $i \in I$. Let

$$\mathcal{C}_j = \{y \in \mathbb{S}^{m-1} \mid h_j \cdot y \leq 0\} \tag{5.11}$$

and suppose these sets are as shown in Figure 5.3. Also, note that $\mathcal{C}(x) = \cap_{j \in J(x)} \mathcal{C}_j$.

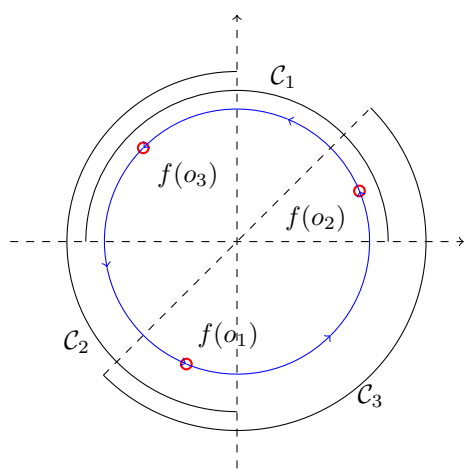

Figure 5.3: Positions of semicircles $\mathcal{C}_j$ from Section 5.4.1, with possible positions of $f(o_i)$, and "essential" appearance of function $f|_{\partial \mathcal{O}_\mathcal{S}}$.

Suppose that the answer to Problem 5.20 (i.e., Problem 5.1) is affirmative; that is, there exists a continuous function $f : \mathcal{O}_\mathcal{S} \to \mathbb{S}^1$ satisfying $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$. Then $f(o_1) \in \mathcal{C}_2 \cap \mathcal{C}_3$, $f(o_2) \in \mathcal{C}_1 \cap \mathcal{C}_3$, and $f(o_3) \in \mathcal{C}_1 \cap \mathcal{C}_2$. Additionally, for any other $x \in \overline{o_1 o_2}$, $J(x) = \{3\}$ and hence, $f(x) \in \mathcal{C}_3$. Analogously, $f(x) \in \mathcal{C}_1$ for all $x \in \overline{o_2 o_3} \backslash \{o_2, o_3\}$, and $f(x) \in \mathcal{C}_2$ for all $x \in \overline{o_1 o_3} \backslash \{o_1, o_3\}$.

By examining Figure 5.3, we observe that the above conditions imply that, as x travels along the edge of triangle $\mathcal{O}_\mathcal{S}$ from o_1 through o_2 and o_3 and back to o_1, $f(x)$ must make one complete counterclockwise encirclement of the origin. In other words, $f|_{\partial \mathcal{O}_\mathcal{S}}$ does not have degree 0 (the degree of a map to $\mathbb{S}^1$

is the number of times its image encircles the origin). This implies $f|_{\partial\mathcal{O}_\mathcal{S}}$, taken as a function from $\partial\mathcal{O}_\mathcal{S} \cong \mathbb{S}^1$ to $\mathbb{S}^1$, is not null-homotopic in $\mathbb{S}^1$, as null-homotopic functions in $\mathbb{S}^1$ are those of degree 0. By Theorem 2.34, there cannot exist a function $f : \mathcal{O}_\mathcal{S} \to \mathbb{S}^1$ satisfying Problem 5.20.

The main contribution of this section is to expand and formalize this idea of maps that encircle the origin to provide a characterization of a topological obstruction in the RCP.

5.4.2 Main Results

The main idea of our approach to Problem 5.20 is as follows. First, a technical lemma, Lemma 5.21, gives a useful property about the index sets $J(x)$. Then we examine the two dimensional polytopes in $\partial\mathcal{O}_\mathcal{S}$. We assume that on those polytopes there exists a continuous function which satisfies the requirements of Problem 5.20. Thanks to the conditions (3.4), this function can be shown, via Lemma 2.26 or Lemma 2.27, to be null-homotopic. In Lemma 5.22 an induction argument on the dimension of the boundary polytopes in $\mathcal{O}_\mathcal{S}$ shows that the proposed null-homotopic map defined on the two dimensional boundary polytopes of $\mathcal{O}_\mathcal{S}$ can be continuously extended to all of $\mathcal{O}_\mathcal{S}$ and still satisfy the conditions (3.4).

Once we have identified the existence of a continuous map satisfying (3.4) on the two dimensional boundary polytopes of $\mathcal{O}_\mathcal{S}$ as turnkey to the solution, we then turn, in Proposition 5.27, to identifying verifiable conditions for existence of such a map. The argument relies on homotopy properties of certain loops, expressed in Lemmas 5.25 and 5.26. Combining Lemma 5.22 and Proposition 5.27, we obtain the main result, which is Theorem 5.28.

Lemma 5.21. *Let* $\dim(\mathcal{O}_\mathcal{S}) = \kappa \geq 2$ *and* $\mathcal{O}_\mathcal{S} \cap \mathcal{F}_0 = \emptyset$. *Suppose* o_1, o_2 *are vertices of* $\mathcal{O}_\mathcal{S}$, *i.e., that* $\overline{o_1 o_2}$ *is a one-dimensional edge of* $\mathcal{O}_\mathcal{S}$. *Then,* $I(o_1) \cup I(o_2) \neq I$. *Equivalently,* $J(o_1) \cap J(o_2) \neq \emptyset$.

Proof. Suppose by way of contradiction that $I(o_1) \cup I(o_2) = I$. Let $x = \frac{1}{2}o_1 + \frac{1}{2}o_2$. Since $\dim(\mathcal{O}_\mathcal{S}) \geq 2$, $x \in \overline{o_1 o_2} \subset \partial\mathcal{O}_\mathcal{S}$. By Lemma 1 of [125], $\partial\mathcal{O}_\mathcal{S} \subset \partial\mathcal{S}$, so $x \in \partial\mathcal{S}$. Let $o_1 = \sum_{i\in I(o_1)} \alpha_i v_i$ and $o_2 = \sum_{i\in I(o_2)} \beta_i v_i$ for some $\alpha_i > 0$ and $\beta_i > 0$. Then $x = \frac{1}{2}\sum_{i\in I(o_1)} \alpha_i v_i + \frac{1}{2}\sum_{i\in I(o_2)} \beta_i v_i = \sum_{k\in I} \lambda_k v_k$, where $\lambda_k = \alpha_k/2$ if $k \in I(o_1) \setminus I(o_2)$, $\lambda_k = \beta_k/2$ if $k \in I(o_2) \setminus I(o_1)$, and $\lambda_k = (\alpha_k + \beta_k)/2$ if $k \in I(o_1) \cap I(o_2)$. Note that hence $\lambda_k > 0$, $k \in I$. Thus, $x \in \mathcal{S}^\circ \cup \mathcal{F}_0$ but $\mathcal{O}_\mathcal{S} \cap \mathcal{F}_0 = \emptyset$, so $x \in \mathcal{S}^\circ$, a contradiction. □

As defined in Section 2.1, let for each k, $0 \leq k \leq \kappa$, $\partial^{(k)}\mathcal{O}_\mathcal{S}$ be the *k-skeleton* of $\mathcal{O}_\mathcal{S}$, i.e., the union of all k-dimensional edges of $\mathcal{O}_\mathcal{S}$. Note that, while the underlying method is similar, this is not the same skeleton as used in Section 5.3.

The following result allows us to reduce the problem of finding the topological obstruction on $\mathcal{O}_{\mathcal{S}}$ with an arbitrary dimension to determining whether there exists an obstruction on the skeleton of two-dimensional edges of $\mathcal{O}_{\mathcal{S}}$. The argument rests on the fact that every continuous map $g : \mathbb{S}^k \to \mathbb{S}^1$, $k \geq 2$, is null-homotopic. As mentioned, Hopf showed in [61] that such a claim is not generally true for maps $g : \mathbb{S}^k \to \mathbb{S}^r$ with $r \geq 2$.

Lemma 5.22. *Let* $\dim(\mathcal{O}_{\mathcal{S}}) = \kappa \geq 2$ *and* $m = 2$. *If there exists* $\partial^{(2)} f : \partial^{(2)}\mathcal{O}_{\mathcal{S}} \to \mathbb{S}^1$ *satisfying* $\partial^{(2)} f(x) \in \mathcal{C}(x)$ *for* $x \in \partial^{(2)}\mathcal{O}_{\mathcal{S}}$, *then* $\partial^{(2)} f$ *can be extended to* $f : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^1$ *satisfying* $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_{\mathcal{S}}$.

Proof. The proof is by induction on the dimension k. The base case when $k = 2$ is the assumption of the lemma statement. Next, if $2 \leq k \leq \kappa - 1$ and there exists $\partial^{(k)} f : \partial^{(k)}\mathcal{O}_{\mathcal{S}} \to \mathbb{S}^1$ satisfying (5.10), then we will show that there exists $\partial^{(k+1)} f : \partial^{(k+1)}\mathcal{O}_{\mathcal{S}} \to \mathbb{S}^1$ satisfying (5.10) and such that $\partial^{(k+1)} f|_{\partial^{(k)}\mathcal{O}_{\mathcal{S}}} \equiv \partial^{(k)} f$. Finally, the result follows by noting that $\partial^{(\kappa)}\mathcal{O}_{\mathcal{S}} = \mathcal{O}_{\mathcal{S}}$.

Let $2 \leq k \leq \kappa - 1$ and suppose there exists $\partial^{(k)} f : \partial^{(k)}\mathcal{O}_{\mathcal{S}} \to \mathbb{S}^1$ satisfying (5.10). Since $\mathcal{O}_{\mathcal{S}}$ is a convex polytope, $\partial^{(k+1)}\mathcal{O}_{\mathcal{S}}$ is comprised of $(k+1)$-dimensional convex polytopes, and those polytopes intersect in polytopes of lower dimension. Take any $(k+1)$-dimensional polytope $\mathcal{A}$ in $\partial^{(k+1)}\mathcal{O}_{\mathcal{S}}$. Let $\partial\mathcal{A} \subset \partial^{(k)}\mathcal{O}_{\mathcal{S}}$ be the relative boundary of $\mathcal{A}$, consisting of k-dimensional polytopes in $\partial^{(k)}\mathcal{O}_{\mathcal{S}}$. By Theorem 2.22, $\mathcal{A}$ is homeomorphic to $\mathbb{B}^{k+1}$, and the same homeomorphism sends $\partial\mathcal{A}$ to $\partial\mathbb{B}^{k+1} = \mathbb{S}^k$. Via this homeomorphism $\partial^{(k)} f|_{\partial\mathcal{A}} : \partial\mathcal{A} \to \mathbb{S}^1$ can be understood as $\partial^{(k)} f|_{\partial\mathcal{A}} : \mathbb{S}^k \to \mathbb{S}^1$. Let $\mathcal{Y} = \mathrm{Im}(\partial^{(k)} f|_{\partial\mathcal{A}})$.

We consider two cases. First, suppose $\mathcal{Y} \neq \mathbb{S}^1$. Since $\partial^{(k)} f|_{\partial\mathcal{A}}$ is a continuous map and $\partial\mathcal{A}$ is connected, $\mathcal{Y}$ must be a circular arc. Hence, it is trivially homeomorphic to $[0, 1]$. The interval $[0, 1]$ and hence $\mathcal{Y}$ are contractible. By Lemma 2.26 we obtain that $\partial^{(k)} f|_{\partial\mathcal{A}} : \mathbb{S}^{k+1} \to \mathcal{Y}$ is null-homotopic. By Theorem 2.34, $\partial^{(k)} f|_{\partial\mathcal{A}}$ can be extended to $f_{\mathcal{A}} : \mathcal{A} \to \mathcal{Y}$. Second, suppose $\mathcal{Y} = \mathbb{S}^1$. By Lemma 2.27, $\partial^{(k)} f|_{\partial\mathcal{A}} : \mathbb{S}^k \to \mathbb{S}^1$ is null-homotopic. By Theorem 2.34, $\partial^{(k)} f|_{\partial\mathcal{A}}$ can be extended to $f_{\mathcal{A}} : \mathcal{A} \to \mathcal{Y}$. (It is worth noting that we cannot use Lemma 2.27 for the first case because we are specifically interested in extensions that map into $\mathrm{Im}(\partial^{(k)} f|_{\partial\mathcal{A}})$ in order for the argument below to go through.)

Next we show that $f_{\mathcal{A}}$ satisfies (5.10). Consider any $x \in \mathcal{A}$. If $x \in \partial\mathcal{A}$, then $f_{\mathcal{A}}(x) = \partial^{(k)} f(x)$, and by assumption $\partial^{(k)} f$ satisfies (5.10). Instead suppose $x \in \mathcal{A}^\circ$. In the two cases above we have constructed an extension that maps to $\mathcal{Y} = \mathrm{Im}(\partial^{(k)} f|_{\partial\mathcal{A}})$. Thus, there exists $y \in \partial\mathcal{A}$ such that $f_{\mathcal{A}}(x) = \partial^{(k)} f(y)$. By assumption, $\partial^{(k)} f(y) \in \mathcal{C}(y)$. On the other hand, $\mathcal{C}(y) \subset \mathcal{C}(x)$. To see that, consider the line going through y and x. Since $\mathcal{A}$ is a convex polytope, there exists $z \in \partial\mathcal{A}$ such that $x \in \overline{yz}$. Then x is a convex combination of y and z. Thus, $I(y) \cup I(z) \subset I(x)$. In particular, $J(x) \subset J(y)$. Using the definition of

$\mathcal{C}(x)$, we get $f_{\mathcal{A}}(x) \in \mathcal{C}(y) \subset \mathcal{C}(x)$.

Finally, we want to show that different $f_{\mathcal{A}}$ can be "glued" together to obtain $\partial^{(k+1)} f$. Precisely, if $\mathcal{A}$ and $\mathcal{A}'$ are different $(k+1)$-dimensional faces of $\mathcal{O}_{\mathcal{S}}$ and $x \in \mathcal{A} \cap \mathcal{A}'$, then $f_{\mathcal{A}}(x) = f_{\mathcal{A}'}(x)$. This follows because the $(k+1)$-dimensional faces of $\mathcal{O}_{\mathcal{S}}$ intersect on k-dimensional faces. Thus, if $x \in \mathcal{A} \cap \mathcal{A}'$, then $x \in \partial^{(k)}\mathcal{O}_{\mathcal{S}}$, and $f_{\mathcal{A}}(x) = \partial^{(k)} f(x) = f_{\mathcal{A}'}(x)$. The result is that $\partial^{(k+1)} f(x) := f_{\mathcal{A}}(x)$ if $x \in \mathcal{A}$ defines a continuous map on $\partial^{(k+1)}\mathcal{O}_{\mathcal{S}} = \cup \mathcal{A}$. □

We now make the following assumptions:

Assumption 5.23.

(i) $h_i \neq 0$ *for all* $i \in I$.

(ii) $h_i \neq \lambda h_j$ *for all* $i, j \in I$ *and all* $\lambda < 0$.

(iii) $\mathcal{O}_{\mathcal{S}} \cap \mathcal{F}_0 = \emptyset$.

The requirements of Assumption 5.23 can be removed, which we will do in Theorem 5.29 but they significantly contribute to the elegance of the stated results. So, unless noted otherwise, we are now assuming that Assumption 5.23 is true.

Let $\mathcal{A}$ be a two-dimensional polytope in $\partial^{(2)}\mathcal{O}_{\mathcal{S}}$ such that $\mathcal{A} = \mathrm{co}\{o_1, \ldots, o_r\}$, where we are assuming without loss of generality that o_i's are ordered counterclockwise. In other words, the edges of $\mathcal{A}$ are $\overline{o_1 o_2}$, $\overline{o_2 o_3}$, $\ldots$, $\overline{o_r o_1}$. Consider a continuous map $F : \mathcal{A} \to \mathbb{S}^1$ such that $F(x) \in \mathcal{C}(x)$, $x \in \mathcal{A}$. Define $f_i := F|_{\overline{o_i o_{i+1}}}$, taking $o_{r+1} \equiv o_1$. Also let $\tilde{f}_i$ be the path traversing the shorter circular arc between $F(o_i)$ and $F(o_{i+1})$. Since reparametrizations do not change the homotopy properties of paths, without loss of generality we assume that f_i and $\tilde{f}_i$ traverse the arc in $\mathbb{S}^1$ with uniform speed. Let $b_1, \ldots, b_r \in \mathbb{S}^1$ be any vectors such that $b_i \in \mathcal{C}(o_i)$, $i = 1, \ldots, r$. Such b_i's exist, as otherwise some $\mathcal{C}(o_i)$ would be empty and the problem would clearly not be solvable (see Proposition 5.27 and Theorem 5.28). Define g_i to be the shorter arc in $\mathbb{S}^1$ between b_i and b_{i+1}. Let m_i be the path from $F(o_i)$ to b_i through the shorter circular arc in $\mathbb{S}^1$. An illustrative example of the paths defined above is given in Figure 5.4.

Remark 5.24. *In the remainder of this text, it will often be necessary to deal with the "shorter arc" between two points* $b_i \in \mathcal{C}(o_i)$ *and* $b_{i+1} \in \mathcal{C}(o_{i+1})$. *This presents an issue if both arcs between those points are of equal length* π. *Let those arcs be labeled* $\mathcal{K}_1$ *and* $\mathcal{K}_2$. *We will show that in that case, exactly one of the arcs* $\mathcal{K}_i$ *satisfies the following: for each point* y *on the arc,* $h_j \cdot y \leq 0$ *for all* $j \in J(o_i) \cap J(o_{i+1})$.

Let us prove that. We note that by Assumption 5.23 and Lemma 5.21, there exists $k \in J(o_i) \cap J(o_{i+1})$. *Now,* $b_i \in \mathcal{C}(o_i)$ *implies that* $h_k \cdot b_i \leq 0$, *and* $h_k \cdot b_{i+1} \leq 0$. *Since both arcs between* b_i *and* b_{i+1} *are of*

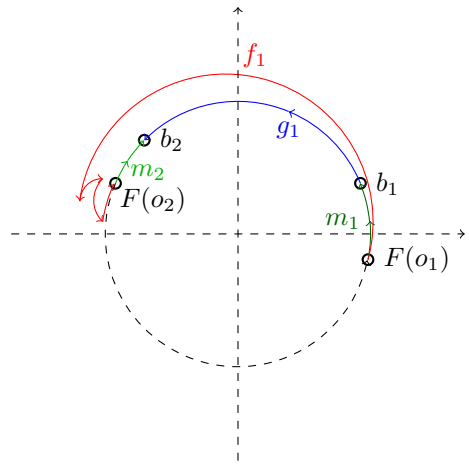

Figure 5.4: An illustration of the different paths defined below. Paths m_1 and m_2 are the shortest paths from $F(o_1)$ to b_1 and $F(o_2)$ to b_2, respectively. Path g_1 is the shortest path from b_1 to b_2, and f_1 is a path (not necessarily the shortest) from $F(o_1)$ to $F(o_2)$. Path f_1 is drawn away from the circle for visibility. Path $\tilde{f}_1$ is not pictured, but is the shortest path from $F(o_1)$ to $F(o_2)$. In this case, $\tilde{f}_1$ equals $m_1 g_1 \overline{m}_2$.

length π, this means that $b_i = -b_{i+1}$. So, $h_k \cdot b_i = h_k \cdot b_{i+1} = 0$. Thus, b_i and b_{i+1} are on the edges of the semicircle $\{y \in \mathbb{S}^1 | h_k \cdot y \leq 0\}$. This semicircle is hence exactly an arc between b_i and b_{i+1}. Hence, for each $k \in J(o_i) \cap J(o_{i+1})$, $\{y \in \mathbb{S}^1 | h_k \cdot y \leq 0\} = \mathcal{K}_1$ or $\{y \in \mathcal{S}^1 | h_k \cdot y \leq 0\} = \mathcal{K}_2$.

We note that $h_k \cdot b_i = h_k \cdot b_{i+1} = 0$ also implies that, for every $k \in J(o_i) \cap J(o_{i+1})$, h_k is perpendicular to b_i. As we are working with vectors in $\mathbb{R}^2$, there is only one perpendicular line to b_i. Hence, all h_k's are scalar multiples of each other. The second part of Assumption 5.23 implies that they are indeed positive multiples of each other. In other words, all the semicircles $\{y \in \mathbb{S}^1 | h_k \cdot y \leq 0\}$ for $k \in J(o_i) \cap J(o_{i+1})$ are the same, and equal exactly one of the arcs $\mathcal{K}_1$ or $\mathcal{K}_2$.

The next two lemmas will provide a connection between f_i's, $\tilde{f}_i$'s and g_i's. This will lead to an easily checkable characterization of Problem 5.20 in terms of null-homotopic loops.

Lemma 5.25. *The paths f_i and $\tilde{f}_i$ are path-homotopic.*

Proof. Let $\overline{\tilde{f}_i}$ be the reverse path of $\tilde{f}_i$. We will show that $f_i \overline{\tilde{f}_i}$ is null-path-homotopic. First, f_i and $\tilde{f}_i$ have the same start and end points, so $f_i \overline{\tilde{f}_i}$ is a loop in $\mathbb{S}^1$. For any $x \in \overline{o_i o_{i+1}}$, $I(x) \subset I(o_i) \cup I(o_{i+1})$ so $J(x) \supset J(o_i) \cap J(o_{i+1})$. By assumption, $f_i(x) \in \mathcal{C}(x)$, $x \in \overline{o_i o_{i+1}}$, so $h_j \cdot f_i(x) \leq 0$, $x \in \overline{o_i o_{i+1}}$, $j \in J(x) \supset J(o_i) \cap J(o_{i+1})$. Now consider $\tilde{f}_i(x)$, $x \in \overline{o_i o_{i+1}}$. Since it is the shorter arc from $F(o_i)$ to $F(o_{i+1})$, $\tilde{f}_i(x)$ is a positive multiple of a convex combination of $F(o_i)$ and $F(o_{i+1})$. Since $F(o_i) \in \mathcal{C}(o_i)$ and $F(o_{i+1}) \in \mathcal{C}(o_{i+1})$, we get $h_j \cdot \tilde{f}_i(x) \leq 0$, $x \in \overline{o_i o_{i+1}}$, $j \in J(x) \supset J(o_i) \cap J(o_{i+1})$. By Lemma 5.21,

there exists $k \in J(o_i) \cap J(o_{i+1})$. We conclude $f_i\overline{\tilde{f}_i} \subset \{y \in \mathbb{S}^1 \mid h_k \cdot y \leq 0\}$. This implies $f_i\overline{\tilde{f}_i}$ is not surjective. Hence, $f_i\overline{\tilde{f}_i}$ covers an arc, which is homeomorphic to $[0,1]$. So, we can take the range of $f_i\overline{\tilde{f}_i}$ to be $[0,1]$, and $f_i\overline{\tilde{f}_i}(0) = f_i\overline{\tilde{f}_i}(1) = 0$. Now, define $H : [0,1] \times [0,1] \to [0,1]$ by $H(x,t) = (f_i\overline{\tilde{f}_i})(xt)$. This is a homotopy of paths between a constant map and the loop $f_i\overline{\tilde{f}_i}$. Thus, $f_i\overline{\tilde{f}_i}$ is null-path-homotopic. By Lemma 2.29, $f_i \simeq \tilde{f}_i$, as desired. □

Lemma 5.26. *The paths $\tilde{f}_i m_{i+1}\overline{g}_i$ and m_i are path-homotopic.*

Proof. Both $\tilde{f}_i m_{i+1}\overline{g}_i$ and m_i are paths from $F(o_i)$ to b_{i+1}. Thus, we will prove that the loop $\tilde{f}_i m_{i+1}\overline{g}_i\overline{m}_i$ is null-path-homotopic. We showed in the proof above that there exists $k \in J(o_i) \cap J(o_{i+1})$ such that $h_k \cdot y \leq 0$ for all y in $\tilde{f}_i$. Also, $h_k \cdot y \leq 0$ for all y in g_i since $b_i \in \mathcal{C}(o_i)$, $b_{i+1} \in \mathcal{C}(o_{i+1})$, and any y in g_i is a positive scalar multiple of a convex combination of b_i and b_{i+1}. Next consider m_i. Since $b_i, F(o_i) \in \mathcal{C}(o_i)$, we have $h_k \cdot b_i \leq 0$, $h_k \cdot F(o_i) \leq 0$. Then, since every y in m_i is a positive multiple of a convex combination of b_i and $F(o_i)$, we get $h_k \cdot y \leq 0$ for all y in m_i. By an analogous argument we find $h_k \cdot y \leq 0$ for all y in m_{i+1}. We conclude that $h_k \cdot y \leq 0$ for all y in $\tilde{f}_i m_{i+1}\overline{g}_i\overline{m}_i$. This implies $\tilde{f}_i m_{i+1}\overline{g}_i\overline{m}_i$ is not surjective so by the same procedure as at the end of the proof of Lemma 5.25, it is null-path-homotopic. By Lemma 2.29, $\tilde{f}_i m_{i+1}\overline{g}_i \simeq m_i$, as desired. □

We now present our main technical tool for characterizing the topological obstruction. We note that indices are taken modulo r, i.e., $o_{r+1} \equiv o_1$. Recall that, for any $b_i, b_{i+1} \in \mathbb{S}^1$, g_i is defined as the shorter arc in $\mathbb{S}^1$ between b_i and b_{i+1}.

Proposition 5.27 solves Problem 5.20 in the case of a two-dimensional polytope $\mathcal{A}$ by exploring the homotopy classes of maps defined on its boundary ∂A. This will serve as the foundation of our final result given in Theorem 5.28.

Proposition 5.27. *Let Assumption 5.23 hold, $m = 2$, and $\kappa \geq 2$. Let $\mathcal{A} = \mathrm{co}\{o_1, \ldots, o_r\}$ be a two-dimensional polytope in $\partial^{(2)}\mathcal{O}_{\mathcal{S}}$. Then there exists $F : \mathcal{A} \to \mathbb{S}^1$ such that $F(x) \in \mathcal{C}(x)$, $x \in \mathcal{A}$, if and only if*

(i) There exists $\{b_1, \ldots, b_r\}$ with $b_i \in \mathbb{S}^1 \cap \mathcal{C}(o_i)$, $i = 1, \ldots, r$.

(ii) For any selection $\{b_1, \ldots, b_r \mid b_i \in \mathbb{S}^1 \cap \mathcal{C}(o_i)\}$, the map $g : \partial A \to \mathbb{S}^1$ defined by $g(o_i) = b_i$, $i = 1, \ldots, r$ and $g|_{\overline{o_i o_{i+1}}} := g_i$ is null-homotopic.

Proof. ($\Longleftarrow$) Suppose (i)-(ii) are satisfied. Let g be as in (ii). We claim g satisfies $g(x) \in \mathcal{C}(x)$, $x \in \partial\mathcal{A}$. First, $g(o_i) = b_i \in \mathcal{C}(o_i)$, $g(o_{i+1}) = b_{i+1} \in \mathcal{C}(o_{i+1})$ by (ii). Second, let $x \in \partial\mathcal{A}\backslash\{o_1, \ldots, o_r\}$. Then x is in the relative interior of $\overline{o_i o_{i+1}}$ for some i. We know $I(x) = I(o_i) \cup I(o_{i+1})$, so $J(x) = J(o_i) \cap J(o_{i+1})$.

Since g_i is on the shorter arc between b_i and b_{i+1}, $g(x)$ is a positive multiple of a convex combination of b_i and b_{i+1}. Since $h_j \cdot b_i \leq 0$, $j \in J(o_i)$, $h_j \cdot b_{i+1} \leq 0$, $j \in J(o_{i+1})$, then $h_j \cdot g(x) \leq 0$, $j \in J(o_i) \cap J(o_{i+1})$, so $g(x) \in \mathcal{C}(x)$.

By assumption, g is null-homotopic. Now, if $\mathrm{Im}(g) = \mathbb{S}^1$, g can be extended to $F : \mathcal{A} \to \mathrm{Im}(g)$ by Theorem 2.34. If g is not surjective, its image is a circular arc. As an arc is contractible, by Lemma 2.26, $g : \mathcal{A} \to \mathrm{Im}(g)$ is null-homotopic. Hence, by Theorem 2.34, it can again be extended to $F : \mathcal{A} \to \mathrm{Im}(g)$.

We claim $F(x) \in \mathcal{C}(x)$, $x \in \mathcal{A}$. For $x \in \partial\mathcal{A}$, we have $F(x) = g(x) \in \mathcal{C}(x)$. For $x \in \mathcal{A}^\circ$, $F(x) = g(z)$ for some $z \in \partial\mathcal{A}$. As in Lemma 5.22, $\mathcal{C}(z) \subset \mathcal{C}(x)$. Thus, $F(x) = g(z) \in \mathcal{C}(z) \subset \mathcal{C}(x)$.

($\Longrightarrow$) For the converse direction, suppose there exists $F : \mathcal{A} \to \mathbb{S}^1$ such that $F(x) \in \mathcal{C}(x)$, $x \in \mathcal{A}$. Then (i) is immediately satisfied by taking $b_i := F(o_i) \in \mathbb{S}^1 \cap \mathcal{C}(o_i)$, $i = 1, \ldots, r$. To prove (ii), let $g : \partial\mathcal{A} \to \mathbb{S}^1$ be any map such that $g(o_i) = b_i \in \mathcal{C}(o_i)$, $i = 1, \ldots, r$ and $g|_{\overline{o_i o_{i+1}}} = g_i$, the shorter arc in $\mathbb{S}^1$ between b_i and b_{i+1}. We will show that g is null-homotopic. First, we claim that $F|_{\partial\mathcal{A}}$ and g are homotopic. Since $F|_{\partial\mathcal{A}}$ extends to F on $\mathcal{A}$, by Theorem 2.34 $F|_{\partial\mathcal{A}}$ is null-homotopic. Then if $F|_{\partial\mathcal{A}} \simeq g$, g is also null-homotopic. Therefore, we must only show $F|_{\partial\mathcal{A}} \simeq g$.

To that end, recall that $f_i := F|_{\overline{o_i o_{i+1}}}$ and $\tilde{f}_i$ denotes the shorter arc in $\mathbb{S}^1$ between $F(o_i)$ and $F(o_{i+1})$. Define $\tilde{f} : \partial\mathcal{A} \to \mathbb{S}^1$ by $\tilde{f}|_{\overline{o_i o_{i+1}}} = \tilde{f}_i$, $i = 1, \ldots, r$, with indices again taken modulo r. Also define $g : \partial\mathcal{A} \to \mathbb{S}^1$ to be the concatenation of g_i, $i = 1, \ldots, r$. In our loop notation, $F|_{\partial\mathcal{A}} = f_1 \cdots f_r$, $g = g_1 \cdots g_r$, and $\tilde{f} = \tilde{f}_1 \cdots \tilde{f}_r$.

By Lemma 5.25, $f_i \simeq \tilde{f}_i$ so by iterating on Proposition 7.10 of [72], $F|_{\partial\mathcal{A}} = f_1 \cdots f_r \simeq \tilde{f}_1 \cdots \tilde{f}_r = \tilde{f}$. Now consider $\tilde{f}_1 \cdots \tilde{f}_r m_1 \overline{g}_r \cdots \overline{g}_1 \overline{m}_1$. By Lemma 5.26, $\tilde{f}_r m_1 \overline{g}_r \simeq m_r$, so $\tilde{f}_1 \cdots \tilde{f}_r m_1 \overline{g}_r \cdots \overline{g}_1 \overline{m}_1 \simeq \tilde{f}_1 \cdots \tilde{f}_{r-1} m_r \overline{g}_{r-1} \cdots \overline{g}_1 \overline{m}_1$. Iterating on this argument, we get $\tilde{f} m_1 \overline{g m}_1 \simeq \tilde{f}_1 m_2 \overline{g}_1 \overline{m}_1$. Again by Lemma 5.26, $\tilde{f}_1 m_2 \overline{g}_1 \overline{m}_1 \simeq m_1 \overline{m}_1$, and by Lemma 2.29, that loop is null-path-homotopic. Thus, $\tilde{f} m_1 \overline{g m}_1$ is null-path-homotopic. Now, $m_1 g \overline{m}_1$ is a loop from $F(o_1)$ through $b_1, \ldots, b_r$ back to $F(o_1)$. Equivalently, it can be expressed as $g \overline{m}_1 m_1$, a loop starting and ending at b_1. By Lemma 2.29, $g \overline{m}_1 m_1 \overline{g} = g \overline{m}_1 \overline{g \overline{m}_1}$ is null-path-homotopic. Again applying Lemma 2.29, this implies $g \overline{m}_1 m_1 \simeq g$. We conclude $\tilde{f} \simeq m_1 g \overline{m}_1 \simeq g$. We already showed $F|_{\partial\mathcal{A}} \simeq \tilde{f}$. We conclude $F|_{\partial\mathcal{A}} \simeq g$, as desired. $\square$

Using Lemma 5.22 and Proposition 5.27, we are now ready to prove our main result.

Theorem 5.28. *Let Assumption 5.23 hold and let* $\dim(\mathcal{B}) = 2$. *There exists a continuous function* $F : \mathcal{O}_\mathcal{S} \to \mathcal{B} \backslash \{0\}$ *such that* $F(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$, *if and only if:*

(i) For all vertices o_i *of* $\mathcal{O}_\mathcal{S}$, *there exists* $b_i \in \mathcal{B} \cap \mathcal{C}(o_i)$, $b_i \neq 0$.

(ii) For every polytope $\mathcal{A} = \mathrm{co}\{o_1, \ldots, o_r\} \subset \partial^{(2)}\mathcal{O}_\mathcal{S}$ *and for any selection* $\{b_1, \ldots, b_r \mid b_i \in \mathcal{B} \cap$

$\mathcal{C}(o_i), b_i \neq 0\}$, the map $g : \partial\mathcal{A} \to \mathcal{B}\backslash\{0\} \simeq \mathbb{S}^1$ defined by $g(o_i) = b_i$, $i = 1, \ldots, r$ and $g|_{\overline{o_i o_{i+1}}} := g_i$ is null-homotopic.

Proof. First we consider the case that $\dim(\mathcal{O}_\mathcal{S}) = \kappa \geq 2$. Suppose (i)-(ii) hold. By Proposition 5.27, for every 2-dimensional $\mathcal{A}$ there exists a function $F_\mathcal{A} : \mathcal{A} \to \mathbb{S}^1$ satisfying (5.10). We also know that all functions $F_\mathcal{A}$ agree on the 1-dimensional edges of $\mathcal{O}_\mathcal{S}$, as $F_\mathcal{A}|_{\partial A} \equiv g$, and the definition of g on some edge $\overline{o_i o_j}$ of $\mathcal{O}_\mathcal{S}$ only depends on b_i and b_j, i.e., does not depend on $\mathcal{A}$. Hence, all functions $F_\mathcal{A}$ can be "glued" together into a continuous function $F : \partial^{(2)}\mathcal{O}_\mathcal{S} \to \mathbb{S}^1 \simeq \mathcal{B}\backslash\{0\}$ satisfying (5.10). By Lemma 5.22, F can be extended to a function $F : \mathcal{O}_\mathcal{S} \to \mathcal{B}\backslash\{0\}$ satisfying $F(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$.

In the converse direction, suppose $\kappa \geq 2$ and there exists a function $F : \mathcal{O}_\mathcal{S} \to \mathcal{B}\backslash\{0\}$ such that $F(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$. Then (i) is automatically satisfied by taking $b_i = F(o_i)$. For (ii), we note that the function $F|_\mathcal{A}$ satisfies the Proposition 5.27. Hence, (i)-(ii) in Proposition 5.27 hold, and by (ii), g is null-homotopic.

The only case remaining is $\kappa = 1$. In that case, (ii) is vacuous, i.e., there are no two-dimensional polytopes $\mathcal{A}$. Now assume that (i) holds. Let $\mathcal{O}_\mathcal{S} = \overline{o_1 o_2}$. We define $F : \mathcal{O}_\mathcal{S} \to \mathbb{S}^1 \simeq \mathcal{B}\backslash\{0\}$ as the shorter arc connecting b_1 and b_2. If the length of arc between b_1 and b_2 is exactly π, by Assumption 5.23 and our discussion in Remark 5.24, we know that at least one of those two arcs lies in $\{y \in \mathbb{S}^1 | \tilde{h}_j \cdot y \leq 0\}$, for all $j \in J(o_1) \cap J(o_2)$. We choose that arc as the "shorter". (The set $J(o_1) \cap J(o_2)$ may be empty, in which case both of those arcs satisfy our conditions.) By the same discussion as in the proof of Proposition 5.27, we know that F so defined satisfies (5.10) on all of $\mathcal{O}_\mathcal{S}$, and hence solves Problem 5.20. On the other hand, if we assume that there exists a function $F : \mathcal{O}_\mathcal{S} \to \mathcal{B}\backslash\{0\}$ satisfying $F(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$, (i) is automatically satisfied by taking $b_i = F(o_i)$. □

At the end of this section, let us devote some time to discussing the use of Assumption 5.23. We note that it was used to mainly to invoke Lemma 5.21, which showed that of the two arcs between b_i and b_{i+1}, the path going through the shorter arc will be the one satisfying the cone conditions (5.10) on $\overline{o_i o_{i+1}}$.

By removing this assumption we face two possible issues.

First, it is now possible that $I(o_i) \cup I(o_{i+1}) = I$. In this case, moving from b_i to b_{i+1} by going through either of the two arcs between those vectors will still satisfy the cone conditions (5.10) on $\overline{o_i o_{i+1}}$. Hence, there are two plausible possibilities of how to proceed from b_i to b_{i+1}, and both need to be checked.

The second part of Assumption 5.23 also serves to ensure that, if both arcs between b_i and b_{i+1} are the same length, at least one of them will satisfy the cone conditions (5.10) on $\overline{o_i o_{i+1}}$. This now needs to be manually checked, and if not true, the only way to travel from b_i to b_{i+1} is if indeed $b_i = b_{i+1}$.

From the above discussion, by removing Assumption 5.23, it is clear that we can modify Theorem 5.28 into the following result:

Theorem 5.29. *Let* $\dim(\mathcal{B}) = 2$. *Then, there exists a function* $F : \mathcal{O}_{\mathcal{S}} \to \mathcal{B}\backslash\{0\}$ *satisfying the conditions of Problem 5.20 if and only if all of the following conditions are satisfied:*

(i) There exists a selection $\{b_1, \ldots, b_r \mid b_i \in \mathbb{S}^1 \cap \mathcal{C}(o_i)\}$ *satisfying the following requirement:*

(ia) For every one-dimensional edge $\overline{o_i o_j}$ *of* $\mathcal{O}_{\mathcal{S}}$, *if there exist* $j_1, j_2 \in J(o_i) \cap J(o_j)$, $\lambda < 0$ *such that* $h_{j_1} = -\lambda h_{j_2} \neq 0$, *then* $b_i = b_j$, *and both* b_i *and* b_j *equal one of the two vectors perpendicular to* h_{j_1}. *In that case, we will say that* o_i *and* o_j *were* invoked in (ia), *and that* b_i *and* b_j *are* constrained.

(ii) There exists a choice of constrained b_i*'s such that, for any selection of non-constrained* b_i*'s satisfying (i), the following holds:*

(iia) For every one-dimensional edge $\overline{o_i o_j}$ *of* $\mathcal{O}_{\mathcal{S}}$, *if* $J(o_i) \cap J(o_j) \neq \emptyset$, *let* $\widehat{a_{ij}}$ *be the shorter arc between* b_i *and* b_j. *If* $J(o_i) \cap J(o_j) = \emptyset$, *let* $\widehat{a_{ij1}}$ *and* $\widehat{a_{ij2}}$ *be the arcs between* b_i *and* b_j. *Then, there exists a choice of such arcs (i.e.* $\widehat{a_{ij}} \in \{\widehat{a_{ij1}}, \widehat{a_{ij2}}\}$*) such that the following holds: For every two-dimensional edge* A *of* $\mathcal{O}_{\mathcal{S}}$ *with vertices* $o_{i_1}, o_{i_2}, \ldots, o_{i_r}$ *in that order, the map* $g : \partial A \to \mathcal{B}\backslash\{0\} \simeq \mathbb{S}^1$ *given by* $g|_{\overline{o_{i_k} o_{i_{k+1}}}} = \widehat{a_{i_k i_{k+1}}}$ *is null-homotopic.*

This provides the complete characterization of a topological obstruction for $\dim(\mathcal{B}) = 2$.

5.5 General Solution

We now turn to the solution in the general case, without any restrictions on m or n.

5.5.1 Assumptions and Background Results

In the remainder of this section, we use the definition of $\mathcal{C}_j$ from (5.11). We now introduce our main assumptions.

Assumption 5.30.

(A1) The pair (A, B) *is controllable.*

(A2) $2 \leq m \leq n - 1$.

(A3) For any non-empty index set $I' \subset I$, *if* $\mathcal{Y} := \bigcap_{j \in I'} \mathcal{C}_j \neq \emptyset$, *then* $\mathcal{Y} \cong \mathbb{B}^\rho$ *for some* $\rho \in \{0, \ldots, m-1\}$.

(A4) $\mathcal{O}_{\mathcal{S}} = \mathrm{co}\{o_1, \ldots, o_{\kappa+1}\}$, *a* κ*-dimensional simplex with vertices* $o_1, \ldots, o_{\kappa+1}$.

(A5) $v_0 \notin \mathcal{O}_{\mathcal{S}}$.

(A6) $\mathcal{O}_{\mathcal{S}} \cap \mathrm{int}(\mathcal{S}) \neq \emptyset$.

(A7) $\mathcal{C}(o_i) \neq \emptyset$, $i = 1, \ldots, \kappa + 1$.

Assumption (A1) implies that $\dim(\mathcal{O}) = m$; see Lemma 5.31(i). Regarding Assumption (A2), we do not consider the case $m = 1$ because it was resolved by Theorem 1 in [125]. Also we do not consider the case $m = n$, since then there is a trivial solution to the problem (see Lemma 5.10). Assumption (A3) is a non-degeneracy assumption which is discussed in greater detail below. Assumptions (A4)-(A6) regard the interaction between the simplex $\mathcal{S}$ and the set $\mathcal{O}$. While these assumptions are seemingly restrictive, the interactions between $\mathcal{O}$ and $\mathcal{S}$ arise from the choice of triangulation of the original state space. Thus, these interactions are therefore under the designer's discretion. The same set of assumptions has been used in previous work, and is one of the several standard assumptions on the triangulation that are used in the RCP. For a longer discussion, we invite the reader to see [125]. Assumption (A7) incorporates a necessary condition for solvability of the RCP: if there exists $f : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$ satisfying $f(x) \in \mathcal{C}(x)$ for all $x \in \mathcal{O}_{\mathcal{S}}$, then $f(o_i) \in \mathcal{C}(o_i)$. We enumerate the relevant known results under Assumption 5.30; see [125].

Lemma 5.31.

(i) If (A, B) *is controllable and* $\mathcal{O}_{\mathcal{S}} \cap \mathrm{int}(\mathcal{S}) \neq \emptyset$, *then* $\dim(\mathcal{O}_{\mathcal{S}}) = m$.

(ii) If $\dim(\mathcal{O}_{\mathcal{S}}) \geq 1$, *then* $\partial\mathcal{O}_{\mathcal{S}} \subset \partial\mathcal{S}$.

Now we discuss (A3). Assumption (A3) ensures that non-empty intersections of the $\mathcal{C}_j$'s form sets homeomorphic to disks. This guarantees that the $\mathcal{C}_j$'s form "nice" closed covers of the spaces we are observing; we note that a similar condition was imposed in Section 5.4.2. Developing the theory of topological obstructions in the RCP without this assumption is possible, but it results in a number of degenerate cases. In particular, (A3) ensures that each $\mathcal{C}_j$ must be contractible, so each $\mathcal{C}_j \neq \mathbb{S}^{m-1}$, which, in turn, implies each $h_j \neq 0$, $j \in I$. Since the cones $\mathcal{C}_j$ are projections of $\mathcal{B} \cap \mathcal{C}(x)$, $x \in \mathcal{F}_j$, onto $\mathcal{B}$ in the original set-up, this non-degeneracy assumption imposes constraints on the interaction between $\mathcal{B}$ and $\mathcal{S}$. It is possible to test whether (A3) is satisfied by a numerical procedure. First we have the following characterization of (A3).

Lemma 5.32. *Let $I' \subset I$ be a non-empty index set. Define $\mathcal{Y} = \bigcap_{j \in I'} \mathcal{C}_j$ and suppose $\mathcal{Y} \neq \emptyset$. Additionally, assume that there exists $x \in \mathcal{Y}$ such that $h_j \cdot x < 0$ for some $j \in I'$. Then, $\mathcal{Y} \cong \mathbb{B}^\rho$ for some $\rho \in \{0, \ldots, m-1\}$.*

Proof. Clearly $-x \notin \mathcal{Y}$. Thus, we can use a stereographic projection centered at $-x$ to homeomorphically project $\mathcal{Y}$ on $\mathbb{R}^{m-1}$. By Lemma 5.35, this projection morphs $\mathcal{Y}$ into an intersection of closed balls and half-spaces, at least one of which, corresponding to $\mathcal{C}_j$, is a ball. Hence, the projection of $\mathcal{Y}$ on $\mathbb{R}^{m-1}$ is: closed, as an finite intersection of closed sets; bounded, as a subset of the ball corresponding to $\mathcal{C}_j$; convex, as an intersection of convex sets. By Theorem 2.23, $\mathcal{Y}$ is homeomorphic to a ball of some dimension $0 \leq \rho \leq m-1$. □

Suppose that $I' = \{j_1, \ldots, j_p\}$ in the previous lemma. Then the statement that there exists $x \in \mathcal{Y}$ such that $h_j \cdot x < 0$ for some $j \in I'$ is equivalent to the following: there exists an $x \in \mathbb{R}^m$ which is a solution to

$$\begin{bmatrix} h_{j_1}^T \\ h_{j_2}^T \\ \vdots \\ h_{j_p}^T \\ h_{j_1}^T + \ldots + h_{j_p}^T \end{bmatrix} x \leq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix}. \tag{5.12}$$

This is a standard linear feasibility problem. We also have a converse statement to Lemma 5.32.

Lemma 5.33. *Suppose (A3) holds and let $I' \subset I$ be a non-empty index set. Define $\mathcal{Y} := \bigcap_{j \in I'} \mathcal{C}_j$ and suppose $\mathcal{Y} \neq \emptyset$. Then there exists $x \in \mathcal{Y}$ such that $h_j \cdot x < 0$ for all $j \in I'$.*

Proof. We prove the statement by induction on $|I'|$, the cardinality of I'. Consider $I' \subset I$ such that $|I'| = 1$ and $\mathcal{Y} := \cap_{j \in I'} \mathcal{C}_j \neq \emptyset$. Assumption (A3) implies $h_j \neq 0$, so there clearly exists $x \in \mathcal{Y}$ such that $h_j \cdot x < 0$, $j \in I'$. Now supppose the statement holds for all index sets with cardinality less than $k+1$. Consider any index set $I' \subset I$ such that $|I'| = k+1$ and $\mathcal{Y} := \cap_{j \in I'} \mathcal{C}_j \neq \emptyset$.

First we claim there exist $x \in \mathcal{Y}$ and $j' \in I'$ such that $h_{j'} \cdot x < 0$. For if not, then for all $x \in \mathcal{Y}$ and $j' \in I'$, $h_{j'} \cdot x = 0$. This implies $\mathcal{Y} \subset \{z \in \mathbb{S}^{m-1} \mid h_j \cdot z = 0, j \in I'\}$. But by definition of $\mathcal{Y}$, $\{z \in \mathbb{S}^{m-1} \mid h_j \cdot z = 0, j \in I'\} \subset \mathcal{Y}$, and thus $\mathcal{Y} = \{z \in \mathbb{S}^{m-1} \mid h_j \cdot z = 0, j \in I'\}$, the intersection of $(m-2)$-dimensional spheres. Hence, it is itself a sphere, which contradicts (A3).

Thus, there exist $x \in \mathcal{Y}$ and $j' \in I'$ such that $h_{j'} \cdot x < 0$. By the induction step there exists x' such that $h_j \cdot x' < 0$ for all $j \in I' \backslash \{j'\}$ (since $|I' \setminus \{j'\}| = k < k+1$). We observe that $x \neq -\lambda x'$ for any $\lambda > 0$ because $x \in \mathcal{Y}$ implies $h_j \cdot x \leq 0$ for all $j \in I'$. But if $x = -\lambda x'$, then $h_j \cdot (-\lambda x') = h_j \cdot x > 0$,

$j \in I' \setminus \{j'\}$, a contradiction. Now consider $\tilde{x} = (x + \lambda x')/\|x + \lambda x'\|$. Since $h_j \cdot x \leq 0$ and $h_j \cdot x' < 0$ for $j \in I' \setminus \{j'\}$, we have $h_j \cdot \tilde{x} = \frac{1}{\|x+\lambda x'\|}\big(h_j \cdot x + h_j \cdot x'\big) < 0$ for $j \in I' \setminus \{j'\}$. Also since $h_{j'} \cdot x < 0$, we can choose λ sufficiently small such that $h_{j'} \cdot \tilde{x} < 0$. Thus, $\tilde{x} \in \mathcal{Y}$ satisfies the statement. □

In light of (A4), we define $I_{\mathcal{O}_\mathcal{S}} = \{1, \ldots, \kappa + 1\}$, and we denote the facets of the simplex $\mathcal{O}_\mathcal{S}$ as $\mathcal{F}_k^\mathcal{O}$, $k = 1, \ldots, \kappa + 1$, where $\mathcal{F}_k^\mathcal{O}$ is the facet not containing the vertex o_k. Consider any $\mathcal{F}_j^\mathcal{O}$. We observe that $I(x) = \cup_{i \in \{1,\ldots,j-1,j+1,\ldots,\kappa+1\}} I(o_i)$ for all $x \in \mathrm{int}(\mathcal{F}_j^\mathcal{O})$, so $\mathcal{C}(x)$ are the same for every $x \in \mathrm{int}(\mathcal{F}_j^\mathcal{O})$. Therefore we can define

$$\mathcal{H}_j := \mathcal{C}(x)\,, \quad x \in \mathrm{int}(\mathcal{F}_j^\mathcal{O}), \quad j \in I_{\mathcal{O}_\mathcal{S}}\,. \tag{5.13}$$

Notice that $\mathcal{H}_j$ is a closed subset in $\mathbb{S}^{m-1}$. Also define $\mathcal{H}_j^c := \mathbb{S}^{m-1} \setminus \mathcal{H}_j$, $j \in I_{\mathcal{O}_\mathcal{S}}$, which is an open subset in $\mathbb{S}^{m-1}$. We have the following fact about the sets $\mathcal{H}_j$.

Lemma 5.34. *Suppose (A3)-(A6) hold. At most one $\mathcal{H}_j$, $j \in I_{\mathcal{O}_\mathcal{S}}$, satisfies $\mathcal{H}_j = \mathbb{S}^{m-1}$.*

Proof. Suppose not. First we take any $j \in I_{\mathcal{O}_\mathcal{S}}$ such that $\mathcal{H}_j = \mathbb{S}^{m-1}$ and $x \in \mathrm{int}(\mathcal{F}_j^\mathcal{O})$. By definition $\mathcal{H}_j = \bigcap_{i \in I \setminus I(x)} \mathcal{C}_i$. In order that $\mathcal{H}_j = \mathbb{S}^{m-1}$, either $\mathcal{C}_i = \mathbb{S}^{m-1}$ for all $i \in I \setminus I(x)$ or $I \setminus I(x) = \emptyset$. First, $\mathcal{C}_i = \mathbb{S}^{m-1}$ contradicts (A3). Second, suppose $I(x) = \{0, 1, \ldots, n\}$. This is impossible by Lemma 5.31(ii), since $x \in \mathcal{F}_j^\mathcal{O} \subset \partial\mathcal{O}_\mathcal{S} \subset \partial\mathcal{S}$. Third, suppose $I(x) = I$. Then $x \in \mathcal{F}_0$. Since $I \setminus I(z)$ is the same for all $z \in \mathrm{int}(\mathcal{F}_j^\mathcal{O})$, we deduce $\mathrm{int}(\mathcal{F}_j^\mathcal{O}) \subset \mathcal{F}_0$. Since $\mathcal{F}_j^\mathcal{O}$ and $\mathcal{F}_0$ are both polytopes, $\mathcal{F}_j^\mathcal{O} \subset \mathcal{F}_0$. Since $\mathcal{O}_\mathcal{S}$ is a simplex, $o_i \in \mathcal{F}_0$ for all $i \neq j$. Now we repeat this argument for $j' \neq j$ such that $\mathcal{H}_{j'} = \mathbb{S}^{m-1}$. Then $\mathcal{F}_{j'}^\mathcal{O} \subset \mathcal{F}_0$. Thus, $\mathcal{O}_\mathcal{S}$ is a simplex with two facets in $\mathcal{F}_0$, which implies $\mathcal{O}_\mathcal{S} \subset \mathcal{F}_0$. This contradicts (A6). □

Lemma 5.35 ([73])**.** *Suppose (A3) holds. Let $y \in \mathbb{S}^{m-1}$. Suppose $y \in \mathcal{C}_j$. The stereographic projection of $\mathcal{C}_j \setminus \{-y\}$ centered at $-y$ equals:*

(i) a closed half-space in $\mathbb{R}^{m-1}$ if $h_j \cdot y = 0$,

(ii) an $m-1$-dimensional closed ball in $\mathbb{R}^{m-1}$ if $h_j \cdot y < 0$.

Lemma 5.36. *Let $I' \subset I$ be a non-empty index set. Define $\mathcal{Y} := \bigcap_{j \in I'} \mathcal{C}_j$ and suppose $\mathcal{Y} \neq \emptyset$. Then $\mathcal{Y} \cong \mathbb{B}^\rho$ for some $\rho \in \{0, \ldots, m-1\}$, or $\mathcal{Y} \cong \mathbb{S}^{\rho-1}$ for some $\rho \in \{1, \ldots, m\}$.*

Proof. We consider two cases. First, suppose there exist $y \in \mathcal{Y}$ and $j_0 \in I'$ such that $h_{j_0} \cdot y < 0$. Clearly $-y \notin \mathcal{Y}$. Thus, we can use a stereographic projection centered at $-y$ to homeomorphically project $\mathcal{Y}$ on $\mathbb{R}^{m-1}$. By Lemma 5.35, this projection morphs $\mathcal{Y}$ into an intersection of closed balls and half-spaces, at least one of which, corresponding to $\mathcal{C}_{j_0}$, is a ball. Hence, the projection of $\mathcal{Y}$ on $\mathbb{R}^{m-1}$ is: closed,

as an finite intersection of closed sets; bounded, as a subset of the ball corresponding to $\mathcal{C}_{j_0}$; convex, as an intersection of convex sets. By Theorem 2.23, $\mathcal{Y}$ is homeomorphic to a ball of some dimension $0 \leq \rho \leq m-1$.

Second, suppose that for all $y \in \mathcal{Y}$ and $j \in I'$, $h_j \cdot y = 0$. Then $\mathcal{Y} \subset \mathbb{S}^{m-1}$ is given by solutions of the linear equation $H^T y = 0$, where H is a matrix with columns equal to h_j, $j \in I'$. This means $\mathcal{Y}$ is an intersection of $\mathbb{S}^{m-1}$ and a subspace of $\mathbb{R}^m$. By a change of coordinates we may assume without loss of generality that $\mathcal{Y} = \mathbb{S}^{m-1} \cap \{x \in \mathbb{R}^m \mid x_{\rho+1} = \ldots = x_m = 0\}$. Thus, $\mathcal{Y} = \mathbb{S}^{\rho-1} \subset \mathbb{S}^{m-1}$ with $\rho \in \{1, \ldots, m\}$. □

Finally we define

$$\mathcal{H} := \bigcup_{\mathcal{H}_j \neq \mathbb{S}^{m-1}} \mathcal{H}_j \,, \qquad \mathcal{H}^\star := \bigcup_{\mathcal{H}_j \neq \mathbb{S}^{m-1}} \operatorname{int}(\mathcal{H}_j) \,. \tag{5.14}$$

Lemma 5.37. *Suppose (A3) and (A4) hold. $\mathcal{H}$ is locally contractible.*

Proof. Since local contractibility is a local property, we need only study a neighborhood of any point in $\mathcal{H}$. To that end, let $x \in \mathcal{H}$ and suppose without loss of generality $x \in \mathcal{H}_1 \cap \cdots \cap \mathcal{H}_r$, and $x \notin \mathcal{H}_{r+1}, \ldots, \mathcal{H}_{\kappa+1}$. Since all $\mathcal{H}_j$'s are closed in $\mathbb{S}^{m-1}$ there is a neighborhood $\mathcal{W}$ of x such that $\mathcal{W} \cap \mathcal{H} = \mathcal{W} \cap \bigcup_{j=1}^r \mathcal{H}_j$. Now consider $-x$. It is certainly outside some neighborhood of x. We will shrink $\mathcal{W}$ so that $-x \notin \mathcal{W}$. We will prove that $\mathcal{T} = \left(\bigcup_{j=1}^r \mathcal{H}_j\right) \backslash \{-x\}$ is locally contractible, from which it follows $\mathcal{W} \cap \mathcal{H}$ is locally contractible.

We use a stereographic projection centered at $-x$ of $\mathbb{S}^{m-1} \backslash \{-x\}$ into $\mathbb{R}^{m-1}$. By Lemma 5.35, this projection homeomorphically maps $\mathcal{C}_i \backslash \{-x\}$ to either a closed half-space in $\mathbb{R}^{m-1}$ (if $h_i \cdot x = 0$), or to a closed ball in $\mathbb{R}^{m-1}$, if $h_i \cdot x < 0$. Since $\mathcal{H}_j = \mathcal{C}(x)$ for $x \in \operatorname{int}(\mathcal{F}_j^{\mathcal{O}})$, $j \in I_{\mathcal{O}_\mathcal{S}}$, each $\mathcal{H}_j \backslash \{-x\}$ is the intersection of sets $\mathcal{C}_i \backslash \{-x\}$, so $\mathcal{T}$ is the union of intersections of sets $\mathcal{C}_i \backslash \{-x\}$. By Lemma 5.35, each $\mathcal{C}_i \backslash \{-x\}$ is mapped by the same homeomorphism into a convex set: either a half-space or a closed ball. Thus, each $\mathcal{H}_j \backslash \{-x\}$ is mapped into a convex set. Finally, $\mathcal{T}$ is homeomorphically deformed into a finite union of convex sets. By Lemma 2.30, it is locally contractible. □

Lemma 5.38. *Suppose (A3) and (A4) hold. Also suppose $\mathcal{H} \neq \mathbb{S}^{m-1}$ and $\bigcap_{\mathcal{H}_j \neq \mathbb{S}^{m-1}} \mathcal{H}_j \neq \emptyset$. Then $\mathcal{H}$ is contractible.*

Proof. Let $\mathcal{Y} := \cap_{\mathcal{H}_j \neq \mathbb{S}^{m-1}} \mathcal{H}_j$. Since each $\mathcal{H}_j$ is itself an intersection of $\mathcal{C}_j$'s, $\mathcal{Y}$ satisfies Lemma 5.33. Let $I' \subset I$ be the index set of $\mathcal{C}_j$'s whose intersection forms $\mathcal{Y}$. By Lemma 5.33 there exists $x \in \mathcal{Y} \subset \mathcal{H}$ such that $h_k \cdot x < 0$ for all $k \in I'$. Since $h_j \cdot (-x) > 0$ for all $k \in I'$, we know $-x \notin \mathcal{H}_j$ for any $\mathcal{H}_j \subset \mathcal{H}$. Thus, $-x \notin \mathcal{H}$. Consider geodesics on $\mathbb{S}^{m-1}$ coming out of x. Because the antipodal point $-x$ is not in $\mathcal{H}$, there exists a unique geodesic $f_{x'}$ between x and any point $x' \in \mathcal{H}_j$ for any $\mathcal{H}_j \subset \mathcal{H}$. Since each $\mathcal{H}_j$

is Robinson-convex (see [119, 35]), the entire path of geodesic $f_{x'}$ lies inside some $\mathcal{H}_j \subset \mathcal{H}$, as both x and x' are in $\mathcal{H}_j$. Thus, $\mathcal{H}$ is a star-shaped set with respect to geodesics on a sphere. By a repetition of the standard proof for star-shaped sets in Euclidean spaces, $\mathcal{H}$ is contractible. □

5.5.2 Extension Problem

In this section we begin our study of Problem 4.22 by investigating when it is possible to extend a continuous boundary map ∂f from $\partial\mathcal{O}_\mathcal{S}$, the boundary of $\mathcal{O}_\mathcal{S}$, to its interior. Our main tool will be the Extension Theorem 2.34.

The main idea is as follows. By (A7) we can construct a vertex map $f(o_i) \in \mathcal{C}(o_i)$, $i \in I_{\mathcal{O}_\mathcal{S}}$. We want to continuously extend this map to all of $\mathcal{O}_\mathcal{S}$ such that $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$. We observe that if $\mathcal{C}(x) = \mathbb{S}^{m-1}$ for some $x \in \mathcal{O}_\mathcal{S}$, then the constraint $f(x) \in \mathcal{C}(x)$ is trivially satisfied. For example, if $x \in \text{int}(\mathcal{O}_\mathcal{S})$ then by Lemma 5.31(ii), $x \in \text{int}(\mathcal{S})$, and $\mathcal{C}(x) = \mathbb{S}^{m-1}$. Thus, the only relevant constraints on $f(x)$ arise on points $x \in \partial\mathcal{O}_\mathcal{S}$, and moreover, points in $\partial\mathcal{O}_\mathcal{S}$ where $\mathcal{C}(x) \neq \mathbb{S}^{m-1}$. We can think of $\mathcal{H}$ as capturing the co-domain of a boundary map $\partial f : \partial\mathcal{O}_\mathcal{S} \to \mathcal{H}$ where the constraints are non-vacuous. If this co-domain does not cover all of $\mathbb{S}^{m-1}$, that is $\mathcal{H} \neq \mathbb{S}^{m-1}$, then the boundary map is not surjective and Lemma 2.25 and Theorem 2.34 allow us to argue that the boundary map can be extended to all of $\mathcal{O}_\mathcal{S}$. A key step in this construction is to extend a map $f : \partial\mathcal{P} \to \mathcal{H}$ from the boundary of a face $\mathcal{P}$ of $\mathcal{O}_\mathcal{S}$ to its interior. We do that in the following lemma by showing that the codomain of the map on $\mathcal{P}$, namely $\mathcal{Y} := \mathcal{C}(x) \cap \mathcal{H}$, $x \in \text{int}(\mathcal{P})$, is AR. The main tool to establish that $\mathcal{Y}$ is an algebraic retract is Theorem 2.33. This requires showing that $\mathcal{H}$ is compact, contractible, and locally contractible, which was done in Lemmas 5.37 and 5.38.

Lemma 5.39. *Suppose Assumption 5.30 holds, and suppose $\mathcal{H} \neq \mathbb{S}^{m-1}$. Let $\mathcal{P}$ be any k-dimensional face of $\mathcal{O}_\mathcal{S}$ with $0 \leq k \leq \kappa - 1$. Let $\mathcal{Y} := \mathcal{C}(x) \cap \mathcal{H}$ for any $x \in \text{int}(\mathcal{P})$. Then, $\mathcal{Y}$ is AR.*

Proof. We consider two cases. First suppose there exists a facet $\mathcal{F}_j^\mathcal{O}$ of $\mathcal{O}_\mathcal{S}$ with $\mathcal{P} \subset \mathcal{F}_j^\mathcal{O}$ such that $\mathcal{H}_j \neq \mathbb{S}^{m-1}$. By Lemma 5.5, $\mathcal{C}(x) \subset \mathcal{H}_j \subset \mathcal{H}$. Since $\mathcal{Y} := \mathcal{C}(x) \cap \mathcal{H}$, we get $\mathcal{Y} = \mathcal{C}(x)$ for any $x \in \text{int}(\mathcal{P})$. Further, since $\mathcal{C}(x) = \cap_{j \in I \setminus I(x)} \mathcal{C}_j$, by (A3), $\mathcal{Y} = \mathcal{C}(x)$ is homeomorphic either to a ball or a sphere, so, it is locally contractible. Also, by definition $\mathcal{Y}$ is compact. In sum, by Theorem 2.33, $\mathcal{Y}$ is AR.

Second, suppose $\mathcal{H}_j = \mathbb{S}^{m-1}$ for all facets $\mathcal{F}_j^\mathcal{O}$ of $\mathcal{O}_\mathcal{S}$ such that $\mathcal{P} \subset \mathcal{F}_j^\mathcal{O}$. By Lemma 5.34, there is at most one facet $\mathcal{F}_j^\mathcal{O}$ such that $\mathcal{H}_j = \mathbb{S}^{m-1}$. Thus, $\mathcal{P} = \mathcal{F}_j^\mathcal{O}$ with $\mathcal{H}_j = \mathbb{S}^{m-1}$, and $\mathcal{Y} = \mathcal{H}_j \cap \mathcal{H} = \mathbb{S}^{m-1} \cap \mathcal{H} = \mathcal{H}$. Recall that o_j is the vertex of $\mathcal{O}_\mathcal{S}$ not contained in $\mathcal{P}$. By Lemma 5.5 and (5.13), $\mathcal{C}(o_j) \subset \mathcal{H}_i$ for all $i = 1, \ldots, j-1, j+1, \ldots, \kappa+1$. Then by (A7), $\bigcap_{\mathcal{H}_i \neq \mathbb{S}^{m-1}} \mathcal{H}_i \neq \emptyset$. By Lemma 5.38, $\mathcal{H}$ is contractible; by Lemma 5.37, $\mathcal{H}$ is locally contractible; and by definition, $\mathcal{H}$ is compact. In sum,

by Theorem 2.33, $\mathcal{Y} = \mathcal{H}$ is AR. $\square$

We now give the main result on extending vertex maps. Here is the roadmap for our proof strategy. It is similar in flavour to the strategy of Section 5.4. We must show that if $\mathcal{H} \neq \mathbb{S}^{m-1}$, then there exists a continuous map $f : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$ satisfying the cone conditions (5.10). We employ an induction argument on the dimension of the faces of $\mathcal{O}_{\mathcal{S}}$. For the base step, we define a vertex map ∂f on the vertices $\{o_1, \ldots, o_{\kappa+1}\}$ of $\mathcal{O}_{\mathcal{S}}$ such that $\partial f(o_i) \in \mathcal{C}(o_i)$. To inductively build up ∂f on the faces of $\mathcal{O}_{\mathcal{S}}$, we assume a map with the required properties is defined on the k-th skeleton of $\mathcal{O}_{\mathcal{S}}$ consisting of all k-dimensional faces. This map is extended to the $(k+1)$-th skeleton by invoking Lemma 5.39 on each $(k+1)$-dimensional face $\mathcal{P}$. Specifically, the set $\mathcal{Y} := \mathcal{C}(x)$, $x \in \operatorname{int}(\mathcal{P})$, which is effectively the codomain of the map, is an algebraic retract. We continue with this procedure until we reach $\partial f : \partial\mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$ satisfying (5.10). Based on the assumption that the union of cones $\mathcal{C}(x)$, $x \in \partial\mathcal{O}_{\mathcal{S}}$, does not cover $\mathbb{S}^{m-1}$, the map ∂f is not surjective. Then, as mentioned above, Lemma 2.25 and Theorem 2.34 are used to conclude that ∂f can be continuously extended to the interior of $\mathcal{O}_{\mathcal{S}}$. Fortunately, the interior of $\mathcal{O}_{\mathcal{S}}$ lies in the interior of $\mathcal{S}$, where $\mathcal{C}(x) = \mathbb{S}^{m-1}$ and the conditions $f(x) \in \mathcal{C}(x)$ are trivially satisfied.

Theorem 5.40. *Suppose Assumption 5.30 holds. Suppose $\mathcal{H} \neq \mathbb{S}^{m-1}$. Then there exists a continuous map $f : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$ such that $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_{\mathcal{S}}$.*

Proof. The plan of the proof is, first, to construct a boundary map $\partial f : \partial\mathcal{O}_{\mathcal{S}} \to \mathcal{H}$ such that $\partial f(x) \in \mathcal{C}(x)$ for all $x \in \partial\mathcal{O}_{\mathcal{S}}$. Second, this map is extended to all of $\mathcal{O}_{\mathcal{S}}$. By (A7), $\mathcal{C}(o_i) \neq \emptyset$, $i \in I_{\mathcal{O}_{\mathcal{S}}}$, so take any $b_i \in \mathcal{C}(o_i)$ and defined $\partial f(o_i) := b_i$, $i \in I_{\mathcal{O}_{\mathcal{S}}}$. Each o_i lies in κ facets, where $\kappa = m \geq 2$, by (A2) and Lemma 5.31(i). By Lemma 5.34, at most one $j \in I_{\mathcal{O}_{\mathcal{S}}}$ satisfies $\mathcal{H}_j = \mathbb{S}^{m-1}$. Thus, $o_i \in \mathcal{F}_j^{\mathcal{O}}$ for some $j \in I_{\mathcal{O}_{\mathcal{S}}}$ such that $\mathcal{H}_j \neq \mathbb{S}^{m-1}$. By Lemma 5.5, $\mathcal{C}(o_i) \subset \mathcal{H}_j \subset \mathcal{H}$.

We have defined ∂f on $\partial^{(0)}\mathcal{O}_{\mathcal{S}}$. We now define it on $\partial\mathcal{O}_{\mathcal{S}} = \partial^{(m-1)}\mathcal{O}_{\mathcal{S}}$ by inductively building up the map on $\partial^{(k)}\mathcal{O}_{\mathcal{S}}$. Assume that ∂f is defined on $\partial^{(k)}\mathcal{O}_{\mathcal{S}}$, and that it satisfies $\partial f(x) \in \mathcal{C}(x) \cap \mathcal{H}$ for all $x \in \partial^{(k)}\mathcal{O}_{\mathcal{S}}$. Take any $(k+1)$-dimensional face $\mathcal{P}$ of $\mathcal{O}_{\mathcal{S}}$, where $k+1 < \kappa$. It suffices to show that ∂f can be extended on $\mathcal{P}$ such that $\partial f(x) \in \mathcal{C}(x) \cap \mathcal{H}$ for $x \in \operatorname{int}(\mathcal{P})$. Define

$$\mathcal{Y} := \mathcal{C}(x) \cap \mathcal{H} \tag{5.15}$$

for some $x \in \operatorname{int}(\mathcal{P})$. By Lemma 5.5 and the definition of $\partial f(x)$, $\partial f(x) \in \mathcal{Y}$ for all $x \in \partial\mathcal{P}$. Thus, $\partial f|_{\partial\mathcal{P}} : \partial\mathcal{P} \to \mathcal{Y}$. By Lemma 5.39, $\mathcal{Y}$ is AR. Hence, the map $\partial f|_{\partial\mathcal{P}}$ can be extended to a map $\partial f : \mathcal{P} \to \mathcal{Y}$. Such a map satisfies $\partial f(x) \in \mathcal{Y} = \mathcal{C}(x) \cap \mathcal{H}$ for $x \in \operatorname{int}(\mathcal{P})$, so we are hence done.

Now we have constructed $\partial f : \partial\mathcal{O}_{\mathcal{S}} \to \mathcal{H}$ such that $\partial f(x) \in \mathcal{C}(x)$, $x \in \partial\mathcal{O}_{\mathcal{S}}$. Next we interpret

∂f as a map into $\mathbb{S}^{m-1}$; that is, $\partial f : \partial\mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$. Since $\mathcal{H} \neq \mathbb{S}^{m-1}$, ∂f is not surjective. By Lemma 2.25, $\partial f : \partial\mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$ is null-homotopic. By Lemma 5.31 and Theorem 2.22, $\partial\mathcal{O}_{\mathcal{S}} \cong \mathbb{S}^{m-1}$. By Theorem 2.34, ∂f extends to a map $f : \mathcal{O}_{\mathcal{S}} \cong \mathbb{B}^{m-1} \to \mathbb{S}^{m-1}$. Finally, since by construction $\partial f(x) \in \mathcal{C}(x)$ for all $x \in \partial\mathcal{O}_{\mathcal{S}}$, we have $f(x) = \partial f(x) \in \mathcal{C}(x)$ for $x \in \partial\mathcal{O}_{\mathcal{S}}$. For $x \in \text{int}(\mathcal{O}_{\mathcal{S}})$, by (A7) and Lemma 5.31, $x \in \text{int}(\mathcal{S})$. Thus, $\mathcal{C}(x) = \mathbb{S}^{m-1}$, and it immediately follows that $f(x) \in \mathcal{C}(x)$, $x \in \text{int}(\mathcal{O}_{\mathcal{S}})$. □

We note that the sufficient condition $\mathcal{H} \neq \mathbb{S}^{m-1}$, while elegant, is difficult to numerically verify. The set $\mathcal{H}$ is a union of finitely many sets $\mathcal{C}(x) = \{y \in \mathbb{S}^{m-1} \mid h_j \cdot y \leq 0, j \in I\backslash I(x)\}$. By defining $\mathcal{C}'(x) = \{\lambda y \mid \lambda \geq 0, y \in \mathcal{C}(x)\}$, the problem of verifying $\mathcal{H} \neq \mathbb{S}^{m-1}$ is equivalent to the problem of verifying whether a union of polyhedral cones $\mathcal{C}'(x) \subset \mathbb{R}^m$ covers the entire Euclidean space. The same problem also appears in a solution of Problem 5.2 in Chapter 6.

5.5.3 An Obstruction — Sperner's Lemma Argument

In this section we study when it is not possible to extend a continuous boundary map ∂f from $\partial\mathcal{O}_{\mathcal{S}}$, the boundary of $\mathcal{O}_{\mathcal{S}}$, to its interior. The condition that forces the obstruction is that $\mathcal{H}^\star$ covers $\mathbb{S}^{m-1}$. We note that this condition is almost the exact opposite of the sufficient condition for solvability in the previous section, namely $\mathcal{H} \neq \mathbb{S}^{m-1}$. The only gap arises when $\mathcal{H}^\star \neq \mathbb{S}^{m-1}$, but $\mathcal{H}$, the closure of $\mathcal{H}^\star$, equals $\mathbb{S}^{m-1}$. Thus, we arrive at necessary and sufficient conditions, up to a small gap due to a degenerate case. This gap will be removed in Section 5.5.4, but at a price of using significantly deeper mathematical results.

As in Section 5.3, our main tool for the proof is Sperner's Lemma. The essential argument here follows by contradiction: suppose a continuous map $f : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$ exists such that $f(x) \in \mathcal{C}(x)$. Each vertex o_i of $\mathcal{O}_{\mathcal{S}}$ is assigned its own distinct "color" corresponding to whether $f(o_i) \notin \mathcal{H}_i$. Next, successively finer triangulations are constructed on $\mathcal{O}_{\mathcal{S}}$ following the rules of a proper labelling of colors to satisfy Sperner's lemma, as defined below. By Sperner's lemma, for each such triangulation, there exists a simplex of the triangulation whose vertex set includes all colors. Taking the limit as the diameter of the triangulations goes to zero, we deduce there exists a point $\overline{x}$ where $f(\overline{x}) \notin \mathcal{H}_i$, $i \in I_{\mathcal{O}_{\mathcal{S}}}$. This will contradict that $\mathcal{H}^\star$ covers $\mathbb{S}^{m-1}$.

We now present the setup of Sperner's lemma (see, e.g., [16]). Let $\mathbb{T}$ be a triangulation of an n-dimensional simplex $\mathcal{S}$, in the sense of Definition 4.20. That is, $\mathbb{T}$ is a subdivision of $\mathcal{S}$ into n-dimensional simplices such that any two simplices of $\mathbb{T}$ intersect in a common face or not at all. A *proper labeling* of the vertices of $\mathbb{T}$ is as follows:

(L1) Each vertex of the original simplex $\mathcal{S}$ has its own distinct label.

(L2) Vertices of $\mathbb{T}$ on a face of $\mathcal{S}$ are labeled using only the labels of the vertices forming the face.

Given a properly labeled triangulation of $\mathcal{S}$, we say a simplex in $\mathbb{T}$ is *distinguished* if its vertices have all $n+1$ labels.

Lemma 5.41 (Sperner's lemma)**.** *Every properly labeled triangulation of $\mathcal{S}$ has an odd number of distinguished simplices.*

Before applying Sperner's lemma to our problem, we need the following result which underlies our choice of a proper labelling.

Lemma 5.42. *Suppose Assumption 5.30 holds. Also suppose $\mathcal{H} = \mathbb{S}^{m-1}$. Then $\bigcap_{\mathcal{H}_j \neq \mathbb{S}^{m-1}} \mathcal{H}_j = \emptyset$.*

Proof. Suppose by way of contradiction that $\mathcal{Y} := \bigcap_{\mathcal{H}_j \neq \mathbb{S}^{m-1}} \mathcal{H}_j \neq \emptyset$. Since each $\mathcal{H}_j$ is itself an intersection of $\mathcal{C}_k$'s, we can apply Lemma 5.33. That is, there exists $y \in \mathcal{Y} \subset \mathcal{H}$ such that $h_k \cdot y < 0$ for all $\mathcal{C}_k$'s contributing to $\mathcal{Y}$. Then $h_k \cdot (-y) > 0$ for all such $\mathcal{C}_k$'s so $-y \notin \mathcal{H}_j$ when $\mathcal{H}_j \neq \mathbb{S}^{m-1}$. This implies $-y \notin \mathcal{H}$, which contradicts that $\mathcal{H} = \mathbb{S}^{m-1}$. □

Theorem 5.43. *Suppose Assumption 5.30 holds. Also suppose $\mathcal{H}^\star = \mathbb{S}^{m-1}$. There does not exist a continuous map $f : \mathcal{O}_\mathcal{S} \to \mathbb{S}^{m-1}$ such that $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$.*

Proof. First we note that if $\mathcal{H}^\star = \mathbb{S}^{m-1}$, then $\mathcal{H} = \mathbb{S}^{m-1}$ by (5.14). Suppose by way of contradiction that there exists a continuous map $f : \mathcal{O}_\mathcal{S} \to \mathbb{S}^{m-1} = \mathcal{H}$ such that $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$. Suppose that, after possibly reordering indices, $\mathcal{H}_1 = \mathbb{S}^{m-1}$. By our index convention, $o_1 \in \mathcal{F}_j^\mathcal{O}$, $j \in I_{\mathcal{O}_\mathcal{S}} \setminus \{1\}$, so by Lemma 5.5, $\mathcal{C}(o_1) \subset \mathcal{H}_j$, $j \in I_{\mathcal{O}_\mathcal{S}} \setminus \{1\}$. Then since $f(o_1) \in \mathcal{C}(o_1)$, we have that $f(o_1) \in \mathcal{H}_j$, $j \in I_{\mathcal{O}_\mathcal{S}} \setminus \{1\}$. Then $f(o_1) \in \cap_{\mathcal{H}_j \neq \mathbb{S}^{m-1}} \mathcal{H}_j$, which contradicts Lemma 5.42. We deduce that $\mathcal{H}_j \neq \mathbb{S}^{m-1}$, $j \in I_{\mathcal{O}_\mathcal{S}}$. This implies $\mathcal{H} = \bigcup_{j \in I_{\mathcal{O}_\mathcal{S}}} \mathcal{H}_j$ and $\mathcal{H}^\star = \bigcup_{j \in I_{\mathcal{O}_\mathcal{S}}} \mathrm{int}(\mathcal{H}_j)$. Now we apply Sperner's lemma. The first step is to obtain a proper labeling of $\mathcal{O}_\mathcal{S}$. We define the sets

$$\mathcal{Q}_i := f^{-1}(\mathcal{H}_i^c)\,, \qquad i \in I_{\mathcal{O}_\mathcal{S}}\,.$$

Since $f : \mathcal{O}_\mathcal{S} \to \mathbb{S}^{m-1}$ is continuous and $\mathcal{H}_i^c$ is open, each $\mathcal{Q}_i$ is an open subset of $\mathcal{O}_\mathcal{S}$. Because $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$, we observe that

$$f(o_i) \in \mathcal{H}_j\,, \quad i \in I_{\mathcal{O}_\mathcal{S}}\,, j \in I_{\mathcal{O}_\mathcal{S}} \setminus \{i\}\,. \tag{5.16}$$

From (5.16) it is immediate that $o_i \notin \mathcal{Q}_j$ when $i \neq j$. Also, $o_i \in \mathcal{Q}_i$, for otherwise, invoking (5.16), we would have $f(o_i) \in \cap_{j \in I_{\mathcal{O}_\mathcal{S}}} \mathcal{H}_j$, which contradicts Lemma 5.42. Thus, inclusion in a set $\mathcal{Q}_i$ provides a distinct label for the vertices o_i of $\mathcal{O}_\mathcal{S}$. This satisfies (L1) of a proper labeling of $\mathcal{O}_\mathcal{S}$. Next, let $\mathbb{T}$ be any triangulation of $\mathcal{O}_\mathcal{S}$ and consider a vertex v of $\mathbb{T}$ which is not a vertex of $\mathcal{O}_\mathcal{S}$ and lies in $\partial\mathcal{O}_\mathcal{S}$. Without loss of generality, let $v \in \text{co}\{o_1, \ldots, o_l\}$ for some $2 \leq l \leq \kappa$. Then it must be that $v \in \mathcal{Q}_k$ for some $1 \leq k \leq l$. For suppose not, that is $f(v) \in \mathcal{H}_j$, $j = 1, \ldots, l$. Since $v \in \text{co}\{o_1, \ldots, o_l\}$, $v \in \mathcal{F}^{\mathcal{O}}_{l+1} \cap \cdots \cap \mathcal{F}^{\mathcal{O}}_{\kappa+1}$, which means $f(v) \in \mathcal{H}_{l+1} \cap \ldots \cap \mathcal{H}_{\kappa+1}$. In sum, $f(v) \in \cap_{j \in I_{\mathcal{O}_\mathcal{S}}} \mathcal{H}_j$, which contradicts Lemma 5.42. Clearly this labeling of v satisfies the second condition (L2) for a proper labeling. Finally, for vertices v of $\mathbb{T}$ in the interior of $\mathcal{O}_\mathcal{S}$, any label $\mathcal{Q}_i$ such that $v \in \mathcal{Q}_i$ can be used (at least one such exists because otherwise $f(v) \in \cap_{j \in I_{\mathcal{O}_\mathcal{S}}} \mathcal{H}_j$, leading to the same contradiction).

Now for each $\alpha > 0$, $\alpha \in \mathbb{N}$, define a triangulation $\mathbb{T}^\alpha$ of $\mathcal{O}_\mathcal{S}$ such that each simplex of $\mathbb{T}^\alpha$ has diameter $\frac{1}{\alpha}$. Apply Sperner's lemma for each $\mathbb{T}^\alpha$ to obtain a distinguished simplex $\text{co}\{v^\alpha_1, \ldots, v^\alpha_{\kappa+1}\}$ and its barycenter x^α. The set $\{x^\alpha\}$ defines a bounded sequence in $\mathcal{O}_\mathcal{S}$ which has a convergent subsequence, which by abuse of notation is again denoted $\{x^\alpha\}$. We have $\lim_{\alpha \to \infty} x^\alpha = \overline{x} \in \mathcal{O}_\mathcal{S}$, since $\mathcal{O}_\mathcal{S}$ is closed. Also, by construction $v^\alpha_i \to \overline{x}$, $i \in I_{\mathcal{O}_\mathcal{S}}$. By Sperner's lemma we know that $v^\alpha_i \in \mathcal{Q}_i$, $i \in I_{\mathcal{O}_\mathcal{S}}$, so by continuity of $f(x)$ this implies $\overline{x} \in \overline{\mathcal{Q}}_i$, $i \in I_{\mathcal{O}_\mathcal{S}}$. That is, $\cap_{i \in I_{\mathcal{O}_\mathcal{S}}} \overline{\mathcal{Q}}_i \neq \emptyset$.

We conclude there exists a point $\overline{x} \in \mathcal{O}_\mathcal{S}$ such that $\overline{x} \in \cap_{i \in I_{\mathcal{O}_\mathcal{S}}} \overline{\mathcal{Q}}_i$. We claim that if $\overline{x} \in \overline{\mathcal{Q}}_i$, then $f(\overline{x}) \notin \text{int}(\mathcal{H}_i)$. First, if $\overline{x} \in \mathcal{Q}_i$, then by definition, $f(\overline{x}) \in \mathcal{H}^c_i$ so $f(\overline{x}) \notin \text{int}(\mathcal{H}_i)$. Second, if $\overline{x} \in \partial\mathcal{Q}_i$, one can show by continuity of f that $f(\overline{x}) \in \partial\mathcal{H}_i$, so again $f(\overline{x}) \notin \text{int}(\mathcal{H}_i)$. We conclude $f(\overline{x}) \notin \text{int}(\mathcal{H}_i)$, $i \in I_{\mathcal{O}_\mathcal{S}}$. This contradicts that $\mathcal{H}^\star = \mathbb{S}^{m-1}$. □

5.5.4 An Obstruction — Nerve Theory

As mentioned, Theorem 5.43 leaves a small gap in the sufficient condition for a topological obstruction: if $\mathcal{H}^* \subsetneq \mathcal{H} = \mathbb{S}^{m-1}$, neither Theorem 5.43 nor Theorem 5.40 will be able to state whether there exists a topological obstruction for given problem data. In this section we will fix that gap by showing that the claim of Theorem 5.43 in fact holds even if $\mathcal{H} = \mathbb{S}^{m-1}$.

Let us illustrate the technique that we will be using by returning to the example in Section 5.4.1. Since $o_i = (v_0 + v_i)/2$ for all $i \in \{1, 2, 3\}$, we note that for x in the interior of $\overline{o_j o_k}$, we have $I(x) = \{0, j, k\}$. Hence, by (5.11) and (5.13), $\mathcal{H}_i = \mathcal{C}_i$ for all $i \in \{1, 2, 3\}$. By looking at the Figure 5.3, we notice that we can split $\mathbb{S}^1$ into six (non-disjoint) arcs: $\mathcal{H}_1$, $\mathcal{H}_1 \cap \mathcal{H}_2$, $\mathcal{H}_2$, $\mathcal{H}_2 \cap \mathcal{H}_3$, $\mathcal{H}_3$, and $\mathcal{H}_3 \cap \mathcal{H}_1$. This is shown in Figure 5.5.

Let us contract $\mathcal{H}_1 \cap \mathcal{H}_2$, $\mathcal{H}_2 \cap \mathcal{H}_3$ and $\mathcal{H}_1 \cap \mathcal{H}_3$ from Figure 5.5 into single points, and "straighten

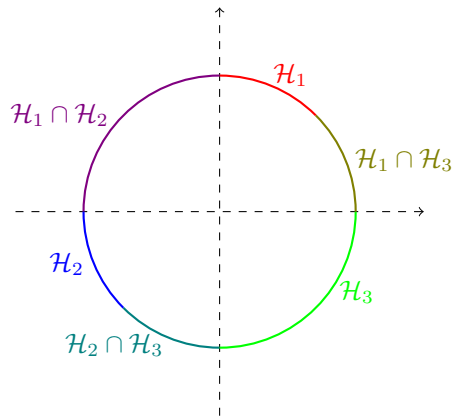

Figure 5.5: Different areas of $\mathbb{S}^1$ from Section 5.4.1. $\mathcal{H}_1$ is denoted by red, $\mathcal{H}_2$ by blue, and $\mathcal{H}_3$ by green, while their intersections are denoted by a mixture of the corresponding colours.

out" $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$. We are being imprecise here: this procedure will be formally justified in Theorem 5.48. We obtain a simplex as on the right side of Figure 5.6.

Now, consider a function $\partial f : \partial\mathcal{O}_{\mathcal{S}} \to \mathbb{S}^1$ which satisfies $f(x) \in \mathcal{C}(x)$ for all $x \in \partial\mathcal{O}_{\mathcal{S}}$. Hence, by (3.3) and (5.13), ∂f satisfies $\partial f(o_1) \in \mathcal{H}_2 \cap \mathcal{H}_3$, $\partial f(o_2) \in \mathcal{H}_1 \cap \mathcal{H}_3$ and $\partial f(o_3) \in \mathcal{H}_1 \cap \mathcal{H}_2$. Also, ∂f satisfies $\partial f(\mathcal{F}_i^{\mathcal{O}}) \subset \mathcal{H}_i$ for all $i \in \{1, 2, 3\}$. We obtain that ∂f is a map from one simplex ($\mathcal{O}_{\mathcal{S}}$) to the boundary of another (modified $\mathbb{S}^1$) which preserves faces. This is illustrated in Figure 5.6.

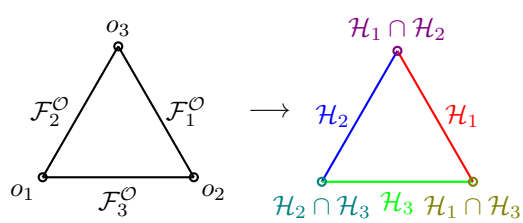

Figure 5.6: An illustration of a nerve-theoretic approach to a topological obstruction. On the left side is $\mathcal{O}_{\mathcal{S}}$, while the right side contains a $\mathbb{S}^1$ modified from Figure 5.5 into a simplex, as outlined in the text.

However, by Corollary 2.36 such a map cannot be extended into a map on the entire $\mathcal{O}_{\mathcal{S}}$.

The above discussion gives an outline of the method that we use to prove that $\mathcal{H} = \mathbb{S}^{m-1}$ implies a negative answer to Problem 5.20. In the above example, we manually transformed the six arcs of $\mathbb{S}^1$ from Figure 5.5 into vertices and edges of a simplex. After this transformation, we obtained a map from one simplex to another which preserves faces. In this example, we did not give a precise definition of this transformation. To formalize this process and generalize it to higher dimensions we use a tool that we introduced beforehand, and that automatically transforms sets $\mathcal{H}_{i_1} \cap \cdots \mathcal{H}_{i_k}$ into faces of a simplex. This tool is nerve theory. We now proceed to a presentation of the formal results.

In order to solve Problem 5.20, we will reduce it to the following question:

Problem 5.44. *Let $\mathcal{F}_1^{\mathcal{O}}, \ldots, \mathcal{F}_{\kappa+1}^{\mathcal{O}}$ be the facets of $\mathcal{O}_{\mathcal{S}}$. Does there exist a continuous map $f : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1}$ which satisfies*

$$f(\mathcal{F}_j^{\mathcal{O}}) \subset \mathcal{H}_j, j = 1, \ldots, \kappa + 1?$$

Lemma 5.45. *If f satisfies the conditions of Problem 5.20, then it satisfies the conditions of Problem 5.44.*

Proof. Assume f solves Problem 5.20. Let $j \in \{1, \ldots, \kappa+1\}$ and let $x \in \operatorname{int}(\mathcal{F}_j^{\mathcal{O}})$. We have $f(x) \in \mathcal{C}(x)$. By (5.13), then $f(x) \in \mathcal{H}_j$. Thus, $f\left(\operatorname{int}(\mathcal{F}_j^{\mathcal{O}})\right) \subset \mathcal{H}_j$. As f is continuous and $\mathcal{H}_j$ is closed, we get $f(\mathcal{F}_j^{\mathcal{O}}) \subset \mathcal{H}_j$. □

Now, we present a technical lemma that establishes a structure-preserving homeomorphism between a k-dimensional simplex Δ^k and a simplex whose vertices are the barycentric centers of the facets of Δ^k. The proof, while long, is rather straightforward.

Lemma 5.46. *Let Δ^k be a simplex with vertices $1, 2, \ldots, k+1$. Let $\mathcal{A}_1, \ldots, \mathcal{A}_{k+1}$ be the facets of Δ^k, and let their barycentric centers be $t_1, \ldots, t_{k+1}$, respectively. Then, $\Delta' = \operatorname{co}\{t_1, \ldots, t_{k+1}\}$ is a simplex, and there exists a homeomorphism $h : \partial\Delta^k \to \partial\Delta'$ such that $h\left(\operatorname{bst}(j)\right) = \operatorname{co}\{t_i \mid i = 1, \ldots, k+1, i \neq j\}$ for all $j \in \{1, \ldots, k+1\}$.*

Proof. We introduce the following notation: $C(j_1, \ldots, j_r)$ is the barycentric center of $\operatorname{co}\{j_1, \ldots, j_r\}$. In a special case, $t_i = C(1, 2, \ldots, i-1, i+1, \ldots, k+1)$. It is trivial to computationally show that t_i's are affinely independent. Hence, Δ' is a simplex. Let us denote its facets by $\mathcal{A}_i'$.

Let $J = \{1, 2, \ldots, k+1\}$. We define $h : \partial\Delta^k \to \partial\Delta'$ to be piecewise affine on each simplex of the barycentric subdivision of $\mathcal{N}$. In other words,

$$h(C(j_1, \ldots, j_r)) = C(t_j \mid j \in J \backslash \{j_1, \ldots, j_r\}), \tag{5.17}$$

with an affine extension on each simplex $\operatorname{co}\{C(j_1), C(j_1, j_2), \ldots, C(j_1, \ldots, j_k)\}$. h maps each simplex in the barycentric subdivision of $\partial\Delta^k$ to a simplex in the barycentric subdivision of $\partial\Delta'$. h is clearly well-defined, bijective, continuous and piecewise affine, and has a continuous and piecewise affine inverse.

Let us observe $h(\operatorname{bst}(j))$. If $x \in \operatorname{bst}(j)$, then x is in some simplex of the barycentric subdivision of $\partial\Delta^k$ which contains j:

$$x = \alpha_1 C(j) + \alpha_2 C(j, j_2) + \ldots + \alpha_k C(j, j_2, \ldots, j_k).$$

By (5.17),

$$h(x) = \sum_{r=2}^{k} \alpha_r C(t_i \mid i \in J \backslash \{j, j_2, \ldots, j_r\}) + \alpha_1 C(t_i \mid i \in J \backslash \{j\}).$$

None of the sets $J \backslash \{j, j_2, \ldots, j_r\}$ contain j. Hence, $h(x) \in \mathcal{A}'_j$. Thus,

$$h(\mathrm{bst}(j)) \subset \mathcal{A}'_j. \tag{5.18}$$

Now, if $x \in \mathcal{A}'_j$, then

$$x = \alpha_1 C(t_{j_1}) + \alpha_2 C(t_{j_1}, t_{j_2}) + \ldots + \alpha_k C(t_{j_1}, \ldots, t_{j_k}),$$

where none of j_i's equal j. By (5.17),

$$h^{-1}(x) = \sum_{r=1}^{k} \alpha_{k+1-r} C(i \mid i \neq j_1, \ldots, j_r).$$

Since $j \neq j_1, \ldots, j_k$, $h^{-1}(x) \in \mathrm{bst}(j)$, i.e.,

$$h^{-1}(\mathcal{A}'_j) \subset \mathrm{bst}(j). \tag{5.19}$$

Combining (5.18) and (5.19) we get

$$h(\mathrm{bst}(j)) = \mathcal{A}'_j.$$

□

An illustration of Lemma 5.46 on a two-dimensional simplex is provided in Figure 5.7.

In the proof of our main result, we will also need the following technical statement. It builds on the homotopy equivalence of space $\mathcal{X}$ and its cover $\mathcal{N}$ from Theorem 2.37 by showing that each element $\mathcal{X}_j \subset \mathcal{X}$ corresponds to a barycentric star of a vertex of $\mathcal{N}$.

Lemma 5.47 (Proof of Theorem 3.3, [95])**.** *Let $\mathcal{N}$ be the nerve of a closed finite regular cover $\{\mathcal{X}_j \mid j \in J\}$ with respect to $\mathcal{X}$. Let us denote vertices of $\mathcal{N}$ by $1, \ldots, j$. Then there exist continuous functions $f' : \mathcal{X} \to \mathcal{N}$ and $g' : \mathcal{N} \to \mathcal{X}$ such that*

$$f'(\mathcal{X}_j) \subset \mathrm{bst}(j)$$

and

$$g'(\mathrm{bst}(j)) \subset \mathcal{X}_j.$$

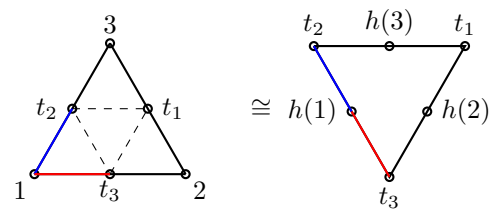

Figure 5.7: An illustration of the proof of Lemma 5.46. On the left side is Δ^2. Simplex $\mathrm{co}\{t_1, t_2, t_3\}$ is denoted by dashed lines on the left, and is enlarged on the right. Homeomorphism h essentially "straightens out" barycentric stars $\mathrm{bst}(j)$ and maps them into the facets of the simplex on the right. In particular, simplex $\mathrm{co}\{1, t_2\} = \mathrm{co}\{C(1), C(1,3)\}$ is marked in blue, and is mapped by h into simplex $\mathrm{co}\{C(t_2, t_3), C(t_2)\} = \mathrm{co}\{h(1), t_2\}$ on the right. Analogously, $\mathrm{co}\{1, t_3\} = \mathrm{co}\{C(1), C(1,2)\}$ is marked in red, and is mapped into $\mathrm{co}\{h(1), t_3\}$.

Finally, we reach the main result of this section.

Theorem 5.48. *Suppose Assumption 7.8 holds, and let $\mathcal{H} = \mathbb{S}^{m-1}$. Then, the answer to Problem 5.20 is negative.*

Proof. Assume otherwise. Let $f : \mathcal{O}_{\mathcal{S}} \to \mathbb{S}^{m-1} = \mathcal{H}$ be a solution to Problem 5.20. By Lemma 5.45, f is then a solution to Problem 5.44.

By Lemma 5.31 (i), $\kappa = \dim(\mathcal{O}_{\mathcal{S}}) = m$, i.e., the vertices of $\mathcal{O}_{\mathcal{S}}$ are $o_1, \ldots, o_{m+1}$. By the conditions of Problem 5.44,

$$f(o_i) \in \mathcal{H}_j, \quad i, j = 1, \ldots, m+1, \quad i \neq j. \tag{5.20}$$

Let $\mathcal{N}$ be the nerve of $\{\mathcal{H}_j \mid \mathcal{H}_j \neq \mathbb{S}^{m-1}\}$ with respect to $\mathcal{H}$. We note that the cover $\{\mathcal{H}_j \mid \mathcal{H}_j \neq \mathbb{S}^{m-1}\}$ is closed and finite. Additionally, by (A3) all $\mathcal{H}_j$ which are not whole spheres are homeomorphic to closed balls. It can easily be verified that a closed ball is contractible and locally contractible, so the cover $\{\mathcal{H}_j \mid \mathcal{H}_j \neq \mathbb{S}^{m-1}\}$ is also regular by Theorem 2.33. Hence, invoking Theorem 2.37,

$$\mathcal{N} \simeq \mathcal{H} = \mathbb{S}^{m-1}. \tag{5.21}$$

We note that $\mathcal{N}$ has at most $m+1$ vertices. Let us show that it has exactly $m+1$ vertices, i.e., that $\mathcal{H}_j \neq \mathbb{S}^{m-1}$ for all $j \in \{1, \ldots, m+1\}$. By Lemma 5.34, there is at most one $\mathcal{H}_j$ such that $\mathcal{H}_j = \mathbb{S}^{m-1}$. Assume that such an $\mathcal{H}_j$ indeed exists, and without loss of generality, let $\mathcal{H}_{m+1} = \mathbb{S}^{m-1}$. By (5.20), $f(o_{m+1}) \in \mathcal{H}_j$ for all $j = 1, \ldots, m$, and also $f(o_{m+1}) \in \mathbb{S}^{m-1} = \mathcal{H}_{m+1}$. Thus, all $\mathcal{H}_j$'s intersect. Hence, by (2.3), $\mathcal{N}$ is a full m-dimensional simplex Δ^m. Since $\Delta^m \cong \mathbb{B}^m \not\simeq \mathbb{S}^{m-1}$, this is in contradiction with (5.21).

Thus, let the vertices of $\mathcal{N}$ be denoted by $1, 2, \ldots, m+1$. Now, by (2.3) and (5.20), simplicial complex $\mathcal{N}$ includes all faces of an m-dimensional simplex Δ^m. As $\Delta^m \cong \mathbb{B}^m \not\cong \mathbb{S}^{m-1}$, $\mathcal{N} \neq \Delta^m$. Thus,

$$\mathcal{N} = \partial\Delta^m. \tag{5.22}$$

By Lemma 5.47, there exists a continuous function $f' : \mathcal{H} \to \mathcal{N}$ such that $f'(\mathcal{H}_j) \subset \mathrm{bst}(j)$. By the conditions of Problem 5.44, continuous map $f' \circ f : \mathcal{O}_{\mathcal{S}} \to \mathcal{N}$ satisfies

$$(f' \circ f)(\mathcal{F}_j^{\mathcal{O}}) \subset f'(\mathcal{H}_j) \subset \mathrm{bst}(j), \quad j = 1, \ldots, m+1. \tag{5.23}$$

Let $\mathcal{N}'$ be the simplex generated by barycentric centers of facets of $\mathcal{N}$, and let $\mathcal{A}'_1, \ldots, \mathcal{A}'_m$ be the facets of $\mathcal{N}'$. By (5.22) and Lemma 5.46, there exists a homeomorphism $h : \mathcal{N} \to \mathcal{N}'$ such that

$$h(\mathrm{bst}(j)) = \mathcal{A}'_j. \tag{5.24}$$

Now, observe a continuous map $\overline{f} = h \circ f' \circ f : \mathcal{O}_{\mathcal{S}} \to \mathcal{N}'$. By (5.23) and (5.24),

$$\overline{f}(\mathcal{F}_j^{\mathcal{O}}) \subset h(\mathrm{bst}(j)) = \mathcal{A}'_j, \quad j = 1, \ldots, m+1. \tag{5.25}$$

We already established that $\mathcal{O}_{\mathcal{S}} \cong \Delta^m$. By Lemma 5.46 and (5.22), $\mathcal{N}' \cong \mathcal{N} = \partial\Delta^m$. Thus, $\overline{f} : \mathcal{O}_{\mathcal{S}} \to \mathcal{N}'$ can be viewed as a continuous function $\overline{f} : \Delta^m \to \partial\Delta^m$ which by (5.25) satisfies

$$\overline{f}(\mathcal{P}_j) \subset \mathcal{P}_j \tag{5.26}$$

for all facets $\mathcal{P}_j \subset \Delta^m$. By taking intersections of different $\mathcal{P}_j$'s, it is clear from (5.26) that $\overline{f} : \Delta^m \to \partial\Delta^m$ not only preserves facets of Δ^m, but also all faces. However, by Corollary 2.36, such a map cannot exist. □

We finally obtain the following result, which completely characterizes the solvability of the problem of a topological obstruction in the RCP.

Theorem 5.49. *Suppose Assumption 7.8 holds. The answer to Problem 5.20 is affirmative if and only if $\mathcal{H} \neq \mathbb{S}^{m-1}$.*

Proof. By Theorem 5.40 that, if $\mathcal{H} \neq \mathbb{S}^{m-1}$, the answer to Problem 5.20 is affirmative. In the other direction, if the answer to Problem 5.20 is affirmative, condition $\mathcal{H} \neq \mathbb{S}^{m-1}$ holds by Theorem 5.48. □

5.6 Discussion

In this section we compare the results contained in different sections of this chapter, along with the results from previous literature. First, this chapter only regards the case when $m > 1$, but we can easily recover the result for $m = 1$ found in Theorem 1 of [125]. When $m = 1$, the problem of a topological obstruction becomes a problem of finding a continuous function $f : \mathcal{O}_\mathcal{S} \to \mathbb{S}^0$ satisfying $f(x) \in \mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$. As $\mathbb{S}^0$ consists of only two points, this means that such a continuous f can only be a constant function. Thus, a necessary and sufficient condition for finding the required f is that all $\mathcal{C}(x)$, $x \in \mathcal{O}_\mathcal{S}$, contain the same point. Using the definition of $\text{cone}(\mathcal{O}_\mathcal{S})$ from (5.5), this is equivalent to saying $\mathcal{B} \cap \text{cone}(\mathcal{O}_\mathcal{S}) \neq \{0\}$, which is precisely the necessary and sufficient condition given in [125].

Let us first concentrate on the general results from Section 5.5. The case opposite to Assumption 5.30, i.e., the situation where the state space has been triangulated so that any intersection of $\mathcal{O}$ with $\mathcal{S}$ is along a face of $\mathcal{S}$ not containing v_0 has been extensively studied; see [5, 23, 24]. The key difference with the case of Assumption 5.30 is that when $\mathcal{O}_\mathcal{S}$ is a face of $\mathcal{S}$, then $\text{int}(\mathcal{O}_\mathcal{S}) \subset \partial\mathcal{S}$, so there are constraints on f in the relative interior of $\mathcal{O}_\mathcal{S}$. In the case Assumption 5.30 holds, $\text{int}(\mathcal{O}_\mathcal{S}) \subset \text{int}(\mathcal{S})$, so f has no constraints in the relative interior of $\mathcal{O}_\mathcal{S}$. For the sake of the present argument, suppose that $m = \kappa < n - 1$ (the general case can also be handled). Following [23], suppose that $\mathcal{O}_\mathcal{S} = \text{co}\{v_1, \ldots, v_{m+1}\}$. Then $\mathcal{O}_\mathcal{S} \subset \mathcal{F}_j$, $j = m+2, \ldots, n$, and for all $x \in \mathcal{O}_\mathcal{S}$, $\mathcal{C}(x) \subset \mathcal{C}_j$, $j = m+2, \ldots, n$. Hence, $\mathcal{H}_k \subset \mathcal{C}_j$, $k \in I_{\mathcal{O}_\mathcal{S}}$, $j = m+2, \ldots, n$, so $\mathcal{H} \subset \mathcal{C}_j$, $j = m+2, \ldots, n$. Now if we assume $\mathcal{H} = \mathbb{S}^{m-1}$ yet $\mathcal{H} \subset \mathcal{C}_j$, $j = m+2, \ldots, n$, it means that $\mathcal{B}$ is parallel to $\mathcal{F}_j$, $j = m+2, \ldots, n$. This is the essence of Proposition 8.2 and Remark 8.2 of [23] saying that with $v_0 = 0$, $\mathcal{B} \subset \text{sp}\{v_1, \ldots, v_{m+1}\}$. It is shown in Theorem 7.3 of [23] that when $\mathcal{B} \cap \text{cone}(\mathcal{O}_\mathcal{S}) = \{0\}$ then there exists a closed-loop equilibrium in $\mathcal{O}_\mathcal{S}$ using any continuous state feedback satisfying the invariance conditions. We recover this finding here via the following result that relates the statement that $\mathcal{H} = \mathbb{S}^{m-1}$ to the condition on $\mathcal{B} \cap \mathcal{C}(\mathcal{O}_\mathcal{S})$.

Lemma 5.50. *Suppose $v_0 = 0$, $\mathcal{B} = \text{sp}\{v_1, \ldots, v_{m+1}\}$, and $\mathcal{O}_\mathcal{S} = \text{co}\{v_1, \ldots, v_{m+1}\}$. If $\mathcal{B} \cap \text{cone}(\mathcal{O}_\mathcal{S}) = \{0\}$, then $\mathcal{H} = \mathbb{S}^{m-1}$.*

Proof. Using the definition of the matrix M from Section 5.4, we note that $\mathcal{B} = \text{sp}\{v_1, \ldots, v_{m+1}\}$ implies that, with $\tilde{h}_i = M^T h_i$ and dropping the tilde's, we have

$$h_{m+2}, \ldots, h_n = 0 \quad h_1, \ldots, h_{m+1} \neq 0\,. \tag{5.27}$$

By our index convention, the i-th facet of $\mathcal{O}_\mathcal{S}$ is $\mathcal{F}_i^\mathcal{O} = \{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{m+1}\}$. Thus, for $x \in \text{int}(\mathcal{F}_i^\mathcal{O})$, $I(x) = \{1, 2, ..., i-1, i+1, ..., m+1\}$. By (3.3), $\mathcal{C}(x) = \{y \in \mathbb{S}^{m-1} \mid h_j \cdot y \leq 0, j = i, m+2, \ldots, n\}$.

By (5.27), $\mathcal{C}(x) = \mathcal{C}_i$, so by definition $\mathcal{H}_i = \mathcal{C}_i$, which is a half-sphere. After the transformation through M, the assumption $\mathcal{B} \cap \text{cone}(\mathcal{O}_\mathcal{S}) = \{0\}$ becomes

$$\cap_{i=1}^{m+1} \mathcal{C}_i = \emptyset. \tag{5.28}$$

Suppose by way of contradiction that $\mathcal{H} \neq \mathbb{S}^{m-1}$. Consider the complements $\mathcal{H}_i^c = \mathbb{S}^{m-1} \setminus \mathcal{H}_i$. Since $\mathcal{H} \neq \mathbb{S}^{m-1}$, and $\mathcal{H}_i = \mathcal{C}_i$, there exists $x \in \mathbb{S}^{m-1}$ such that

$$h_i \cdot x > 0\,, i = 1, \ldots, m+1\,. \tag{5.29}$$

Now consider $-x \in \mathbb{S}^{m-1}$. By (5.29), $-x \in \mathcal{C}_i$ for all $i = 1, \ldots, m+1$, a contradiction with (5.28). □

This deals with the triangulations which do not satisfy Assumption 5.30. Let us briefly discuss the relationship between the sections of this chapter. Generally speaking, the results contained in Section 5.5 are significantly deeper than the results contained in Section 5.3 and Section 5.4.

In Section 5.4, the dimension of $\mathcal{B}$ is limited to $m = 2$. However, there are no strong assumptions on the geometric structure and dimension of $\mathcal{O}_\mathcal{S}$. The basic strategies used in Section 5.4 and Section 5.5 are similar. In both sections, the approach is to iteratively build up a map on the skeleton of $\mathcal{O}_\mathcal{S}$ to a sphere. In the case of $m = 2$, all maps $f : \mathbb{S}^k \to \mathbb{S}^1$, $k > 1$, are automatically null-homotopic. As mentioned is not the case for maps $f : \mathbb{S}^k \to \mathbb{S}^{m-1}$, $m > 2$ (see [61]). Thus, extending boundary maps presents a key issue largely dealt with in 5.5. This was done using the theory of absolute retracts. The main result of 5.4 is that a solution to the problem of a topological obstruction in the case of $m = 2$ exists if and only if any valid boundary function is null-homotopic. Section 5.5 gives a solvability condition in terms of cones that relate $\mathcal{B}$ to $\mathcal{O}_\mathcal{S}$. These conditions arise more directly from the problem data. Additionally, we can partly recover result for $m = 2$ as follows. Assuming that $\mathcal{H}_j \neq \mathbb{S}^{m-1}$ for all $j = 1, \ldots, m$, every boundary function satisfying the invariance conditions will have its image in $\mathcal{H} = \cup_j \mathcal{H}_j$. Hence, if $\mathcal{H} \neq \mathbb{S}^{m-1}$, which we proved to be a sufficient and necessary condition for the existence of the solution to the problem of a topological obstruction, then any valid boundary function will be non-surjective. Thus, it will be null-homotopic by Lemma 2.25.

The work of Section 5.3 deals with the case of $n = 2$ and $n = 3$, and the central case of interest is when $\dim \mathcal{O}_\mathcal{S} = m = 2$. Unlike Section 5.5, those results do not require the assumption that $\mathcal{O}_\mathcal{S} \cap \mathcal{F}_0 = \emptyset$. In the context of Section 5.5, this would mean that $\mathcal{H}_i \subset \mathbb{S}^1$ only exist for $i = 1, 2, 3$, and it is possible that some of those $\mathcal{H}_i$'s in fact equal the entire sphere, thus not satisfying the assumptions present in that section. However, the results of Section 5.3 are severely limited by the fact that the method of

dealing with Problem 5.1 in that setting is largely case-by-case, and does not allow for an immediate generalization for higher dimensions.

An additional paper [25], the results of which are not included in this chapter, uses a purely algebraic strategy to solve the problem of a topological obstruction in a case where the position of $\mathcal{O}_{\mathcal{S}}$ with respect to $\mathcal{S}$ is very constrained, i.e., $\mathcal{O}_{\mathcal{S}}$ is positioned symmetrically within $\mathcal{S}$ in some sense. There also exist additional constraints on the position of $\mathcal{B}$ with respect to $\mathcal{S}$. Paper [25] uses the KKM lemma, also used in this chapter, to again show that, under the assumptions of that paper, (5.6) is a sufficient and necessary condition for the existence of a solution to the problem of a topological obstruction. However, the methods used in that paper are not topological. Instead, they depend on the position of $\mathcal{O}_{\mathcal{S}}$.

Chapter 6

An Affine Obstruction to Reach Control

The previous chapter explored an obstruction to the solvability of the RCP when using continuous state feedback. This chapter explores an analogous problem when using affine feedback. This problem was already defined in the previous chapter, as Problem 5.2. However, as mentioned previously, apart from some preliminary results, the methods used to solve the problem in the general case are significantly different from the ones in Chapter 5.

The work of this chapter is a modified compilation of two previous papers. The work on low-dimensional simplices was presented in [102], while the remainder of the chapter, including the general solution, is the subject of [89]. Unlike in Chapter 5, the computationally efficient solution for two-input systems in this case follows directly from the (computationally difficult) general solution of the problem, and hence it appears after the general solution. At the end of this chapter we also present some computationally efficient necessary, but insufficient, conditions for the solvability of Problem 5.2.

6.1 Low-dimensional simplices

In parallel to the work done on Problem 5.1 in Chapter 5, we first explore a method to solve Problem 5.2 in the case of $n = 2, 3$. Our investigation is made much easier by the fact that a number of results on Problem 5.1 in lower dimensions also equally hold for Problem 5.2.

Analogously to the approach to Problem 5.1 in Section 5.3, our approach is to investigate Problem 5.2 on a case-by-case basis, employing methods from linear algebra. We note that the results from

Lemma 5.8 and Corollary 5.10 settle the case $\dim \mathcal{O}_{\mathcal{S}} \in \{0, n\}$. Since $0 \leq \dim \mathcal{O}_{\mathcal{S}} \leq n$, when $n = 2, 3$ this reduces the problem to $\dim \mathcal{O}_{\mathcal{S}} = 1$ and $\dim \mathcal{O}_{\mathcal{S}} = 2$, with the latter only relevant when $n = 3$. In both of these, the sufficient condition (5.6) from Lemma 5.6 will again make an appearance, as it will be shown that, depending on the case, Problem 5.2 is either always solvable, or (5.6) is a necessary condition.

We note that in the case when $n = 2, 3$, the only relevant cases for the dimension of $\mathcal{B}$ are $\dim \mathcal{B} = 1$, $\dim \mathcal{B} = 2$, as $\dim \mathcal{B} \leq n$, and $\dim \mathcal{B} = n$ is solved by Lemma 5.6.

6.1.1 Single-input Systems

In this case, the matter is clear, and solved by the argument in [125]. If $\dim \mathcal{B} = 1$, by the Intermediate Value Theorem, the vectors $f(x)$ assigned at the segment between any two points in $\mathcal{O}_{\mathcal{S}}$ need to be positive scalar multiples of each other. Thus, in order to satisfy invariance conditions (3.4) on $\mathcal{O}_{\mathcal{S}}$, there needs to exist $b \in \mathcal{C}(x)$ for all $x \in \mathcal{O}_{\mathcal{S}}$. By Lemma 5.6, the answer to Problem 5.2 is affirmative if and only if (5.6) holds.

6.1.2 Two-input Systems

We note that by Remark 5.16, Lemma 5.15 solves Problem 5.2 in the case of $\dim \mathcal{O}_{\mathcal{S}} = 1$ and $\dim \mathcal{B} = 2$. Let us now assume $\dim \mathcal{O}_{\mathcal{S}} = 2$. We additionally assume that $v_0 \notin \mathcal{O}_{\mathcal{S}}$, for that case has been settled by Lemma 5.9. We also assume that $\mathcal{B} \cap \operatorname{cone}(\mathcal{O}_{\mathcal{S}}) = \{0\}$, i.e., that (5.6) does not hold. Otherwise, we are done by Lemma 5.6.

Let us observe the structure of $\mathcal{O}_{\mathcal{S}}$. As given by the formula for product of simplices in [67], $\mathcal{O}_{\mathcal{S}}$ can either be a product of a 2-simplex and a 0-simplex, i.e., a triangle, or a product of two 1-simplices, i.e., a quadrilateral. We also must allow for $\mathcal{O}_{\mathcal{S}}$ passing through one of the vertices of $\mathcal{S}$, resulting in a triangle (essentially, a degenerate quadrilateral).

$\mathcal{O}_{\mathcal{S}}$ is a triangle

First, let us assume that $\mathcal{O}_{\mathcal{S}}$ satisfies $o_i \in (v_0, v_i]$ for all $i \in \{1, 2, 3\}$. Then, Theorem 5.18 provides a solution to both Problem 5.1 and Problem 5.2.

From now on, we can assume that $\mathcal{O}_{\mathcal{S}}$ is not a triangle satisfying the conditions of Theorem 5.18. Thus, if $\mathcal{O}_{\mathcal{S}}$ is a triangle, by the discussion of simplicial products in [67], it can either pass through one of the vertices of $\mathcal{S}$, or have all its vertices on the edges of $\mathcal{S}$ which connect a single vertex in $\mathcal{F}_0$, say v_1, to the others.

In the latter case, say those vertices are $o_1 \in \text{co}\{v_1, v_2\}$, $o_2 \in \text{co}\{v_0, v_1\}$, $o_3 \in \text{co}\{v_1, v_3\}$. We assumed above that $\mathcal{O}_\mathcal{S}$ does not pass through any of vertices v_i. Thus, by Lemma 5.3, we note that $\mathcal{C}(o_2) \supset \mathcal{C}(o_1) = \mathcal{C}(o_3)$. Hence, for any point $x \in \mathcal{O}_\mathcal{S}$, $\mathcal{C}(x) \supset \mathcal{C}(o_1)$. Thus, $\text{cone}(\mathcal{O}_\mathcal{S}) = \mathcal{C}(o_1)$ and hence by Lemma 5.6 and Lemma 5.7, the answer to Problem 5.2 is affirmative if and only if

$$\mathcal{B} \cap \mathcal{C}(o_1) = \mathcal{B} \cap \text{cone}(\mathcal{O}_\mathcal{S}) \neq \{0\}.$$

In the case where $\mathcal{O}_\mathcal{S}$ passes through one of the vertices of $\mathcal{S}$, say without loss of generality that $o_1 = v_1$, $o_2 \in \text{co}\{v_0, v_2\}$ and $o_3 \in \text{co}\{v_2, v_3\}$ (where neither o_2 nor o_3 coincide with any v_i's). Now, as $\dim \mathcal{B} = 2$, we know by Lemma 5.15 that there exist linearly independent vectors $b_1 \in \mathcal{B} \cap \mathcal{C}(o_1)$ and $b_2 \in \mathcal{B} \cap \mathcal{C}(o_2)$. Define $f : \mathcal{O}_\mathcal{S} \to \mathcal{B}$ by $f(x) = Fx$, where $Fo_1 = b_1$, $Fo_2 = b_2$ and $Fo_3 = b_2$. We first note that o_1, o_2 and o_3 are linearly independent; this can be trivially computationally verified. Thus, the above assignment can be accomplished. Next, we note that $\mathcal{C}(o_2) \subset \mathcal{C}(o_3)$ by Lemma 5.3. Thus, the assignments on the vertices of $\mathcal{O}_\mathcal{S}$ satisfy the cone condition.

Now, let us write any $x \in \mathcal{O}_\mathcal{S}$ as $x = \alpha_1 o_1 + \alpha_2 o_2 + \alpha_3 o_3$, where $\alpha_1 + \alpha_2 + \alpha_3 = 1$ and α_i's are nonnegative. We note that if $\alpha_i \neq 0$, $I(x) \supset I(o_i)$. Thus, by Lemma 5.3, $\mathcal{C}(x) \supset \mathcal{C}(o_i)$. As $\mathcal{C}(x)$ is clearly convex from (3.3), then

$$\mathcal{C}(x) \supset \text{co}\{\mathcal{C}(o_i) : \alpha_i \neq 0\}. \tag{6.1}$$

On the other hand, $Fx = \alpha_1 Fo_1 + \alpha_2 Fo_2 + \alpha_3 Fo_3$. Hence, as we have proved that for every i, $Fo_i \in \mathcal{C}(o_i)$, Fx is a convex combination of vectors from $\mathcal{C}(o_i)$, where $\alpha_i \neq 0$. From (6.1), we conclude that $f(x) = Fx$ satisfies the invariance conditions (5.4).

We observe that the above proof works for any affine function: if it satisfies the cone criteria on the vertices, it will satisfy those criteria on the rest of $\mathcal{O}_\mathcal{S}$ as well.

Finally, we note the following: with the above notation,

$$f(x) = Fx = \alpha_1 Fo_1 + \alpha_2 Fo_2 + \alpha_3 Fo_3 = \alpha_1 b_1 + (\alpha_2 + \alpha_3) b_2.$$

As b_1 and b_2 are linearly independent and α_1 and $\alpha_2 + \alpha_3$ can not both be zero, $f(x) \neq 0$ for any $x \in \mathcal{O}_\mathcal{S}$. We have thus given a constructive solution for Problem 5.2 in this case.

$\mathcal{O}_\mathcal{S}$ is a quadrilateral

Say without loss of generality that $o_1 \in \text{co}\{v_0, v_2\}$, $o_2 \in \text{co}\{v_0, v_3\}$, $o_3 \in \text{co}\{v_1, v_2\}$ and $o_4 \in \text{co}\{v_1, v_3\}$ (where none of o_i's actually coincide with any v_j's). This is the set-up in Figure 5.1. Now, we again

know by Lemma 5.15 that there exist linearly independent vectors $b_1 \in \mathcal{B} \cap \mathcal{C}(o_1)$ and $b_2 \in \mathcal{B} \cap \mathcal{C}(o_2)$.

From the definition of a cone in (3.3), we know that $h_3 \cdot b_1 \leq 0$ and $h_2 \cdot b_2 \leq 0$. We distinguish between two cases: in the first one, without loss of generality, $h_2 \cdot b_2 < 0$.

Since $\{o_1, o_2, o_3\}$ is a linearly independent set, which can be trivially verified, we know that there exist unique coefficients α_i such that

$$o_4 = \sum_{i=1}^{3} \alpha_i o_i.$$

Furthermore, $\alpha_2 > 0$, as $o_4 = \lambda_1 v_1 + \lambda_3 v_3$, with $\lambda_3 > 0$, and $v_3 \notin \mathrm{co}\{o_1, o_3\}$. Now, let us define $f : \mathcal{O}_{\mathcal{S}} \to \mathcal{B}$ by $f(x) = Fx$, where $Fo_1 = \varepsilon b_1$, $Fo_2 = b_2$, $Fo_3 = \varepsilon b_1$ (we note that $\mathcal{C}(o_3) \supset \mathcal{C}(o_1)$ by Lemma 5.3), and

$$\varepsilon = \frac{-\alpha_2(h_2 \cdot b_2)}{|2(\alpha_1 + \alpha_3)(h_2 \cdot b_1)|}.$$

(If $(\alpha_1 + \alpha_3)(h_2 \cdot b_1) = 0$, let $\varepsilon = 1$.) As $\{o_1, o_2, o_3\}$ is a linearly independent set in $\mathbb{R}^3$, F is well-defined.

Now, we note that, since $o_4 = \alpha_1 o_1 + \alpha_2 o_2 + \alpha_3 o_3$, $Fo_4 = \varepsilon(\alpha_1 + \alpha_3) b_1 + \alpha_2 b_2$ and thus

$$h_2 \cdot f(o_4) = \pm \frac{\alpha_2}{2}(h_2 \cdot b_2) + \alpha_2(h_2 \cdot b_2) \leq \frac{\alpha_2}{2}(h_2 \cdot b_2) < 0.$$

(If $(\alpha_1 + \alpha_3)(h_2 \cdot b_1) = 0$, then $f(o_4) \cdot h_2 = \alpha_2(h_2 \cdot b_2) < 0$.) Thus, $Fo_4 \in \mathcal{C}(o_4)$. Hence, $f(x) = Fx$ satisfies the invariance conditions at all four vertices of $\mathcal{O}_{\mathcal{S}}$. By the proof same as in the case of the triangle, since f is affine, it hence satisfies (5.4).

Finally, we note that for any $x \in \mathcal{O}_{\mathcal{S}}$, $f(x) = Fx = F(\kappa_1 o_1 + \kappa_2 o_2 + \kappa_3 o_3 + \kappa_4 o_4)$, where $\kappa_1 + \kappa_2 + \kappa_3 + \kappa_4 = 1$, and all κ_i's are nonnegative. Thus, $f(x) = (\kappa_1 + \kappa_3 + \kappa_4\alpha_1 + \kappa_4\alpha_3)\varepsilon b_1 + (\kappa_2 + \kappa_4\alpha_2) b_2$. Since b_1 and b_2 are linearly independent, $f(x) = 0$ is thus equivalent to $\kappa_1 + \kappa_3 + \kappa_4\alpha_1 + \kappa_4\alpha_3 = 0$ and $\kappa_2 + \kappa_4\alpha_2 = 0$. Since $\alpha_2 > 0$, and κ_i's are nonnegative, the latter equation implies $\kappa_2 = \kappa_4 = 0$. Hence, the first equation implies $\kappa_1 + \kappa_3 = 0$, which implies $\kappa_1 = \kappa_3 = 0$. As $\kappa_1 + \kappa_2 + \kappa_3 + \kappa_4 = 1$, this is clearly impossible. Thus, f is nowhere zero. We are done, having defined a function satisfying Problem 5.2 on $\mathcal{O}_{\mathcal{S}}$.

Now, let us assume that $h_3 \cdot b_1 = h_2 \cdot b_2 = 0$. If we remind ourselves that $o_1 \in \mathrm{co}\{v_0, v_2\}$ and $o_2 \in \mathrm{co}\{v_0, v_3\}$, we also know (from the definition of cones at o_1 and o_2) that $h_1 \cdot b_1, h_1 \cdot b_2 \leq 0$. Finally, we can assume that $h_2 \cdot b_1, h_3 \cdot b_2 > 0$. Otherwise, we would have that b_1 or b_2 is in $\mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$, which was solved by Lemma 5.6.

Now, let us first assume that $h_1 \cdot b_1 < 0$ or $h_1 \cdot b_2 < 0$. Without loss of generality we choose the first

option. In that case, let $b_1' = b_1 - cb_2$, where

$$c = \frac{h_1 \cdot b_1}{2(h_1 \cdot b_2)} > 0.$$

(If $h_1 \cdot b_2 = 0$, take $c = 1$ instead.) Now, we note that $h_1 \cdot b_1' = h_1 \cdot b_1 - ch_1 \cdot b_2 = \frac{1}{2} h_1 \cdot b_1 < 0$. (If $h_1 \cdot b_2 = 0$, $h_1 \cdot b_1' = h_1 \cdot b_1 < 0$.) Also, $h_3 \cdot b_1' = h_3 \cdot b_1 - ch_3 \cdot b_2 = -ch_3 \cdot b_2 < 0$. Thus, $b_1' \in \mathcal{B} \cap \mathcal{C}(o_1)$. Furthermore, b_1' and b_2 are still linearly independent and we already established $h_3 \cdot b_1' < 0$. Now, since we have that $h_3 \cdot b_1'$ is strictly negative, we can go a few paragraphs back, just using b_1' and b_2 instead of b_1 and b_2 in order to find a constructive answer to Problem 5.2.

Finally, we have $h_1 \cdot b_1 = 0$ and $h_1 \cdot b_2 = 0$. However, we also know from before that $h_3 \cdot b_1 = h_2 \cdot b_2 = 0$, and that $h_2 \cdot b_1, h_3 \cdot b_2 > 0$. Now, let $b_1' = b_1 - b_2$. Then, $h_1 \cdot b_1' = 0$ and $h_3 \cdot b_1' = h_3 \cdot b_1 - h_3 \cdot b_2 = -h_3 \cdot b_2 < 0$. Thus, again, $b_1' \in \mathcal{B} \cap \mathcal{C}(o_1)$, b_1' and b_2 are linearly independent and $b_1' \cdot h_3 < 0$. Again, as before, we can go a few paragraphs back and obtain a solution to Problem 5.2.

We note that this long and drawn out affair proved the following:

Theorem 6.1. *Let $n = 3$, $\dim \mathcal{O}_{\mathcal{S}} = 2$, $\dim \mathcal{B} = 2$ and $v_0 \notin \mathcal{O}_{\mathcal{S}}$. Assume that $\mathcal{O}_{\mathcal{S}}$ does not satisfy the conditions of Theorem 5.18. Then, the answer to Problem 5.2 is affirmative.*

By combining Theorem 5.18, Theorem 5.19 and Theorem 6.1, we obtain the following:

Theorem 6.2. *Let $\mathcal{S}$, $\mathcal{B}$ and $\mathcal{O}_{\mathcal{S}}$ be as above, and let $n \in \{2, 3\}$. Then, the answer to Problem 5.2 is affirmative if and only if the answer to Problem 5.1 is affirmative. The answers to these problems are affirmative if and only if at least one of the following holds:*

(i) $n = 3$, $\dim \mathcal{B} = 2$ *and $\mathcal{O}_{\mathcal{S}}$ does not satisfy the conditions of Theorem 5.18,*

(ii) condition (5.6) *holds.*

Thus, we've answered the last remaining case.

6.2 General Solution

As in Section 5.3, the low-dimensional solution of Section 6.1 provides some geometric intuition into the obstruction problem, but does not offer a clear path to a solution in higher dimensions. In this section, we build a more substantial linear algebraic machinery to solve Problem 5.2.

We start with a technical result which resembles the *flow condition* introduced in [122].

Since the map f in Problem 5.2 is required to be affine, it is sufficient to check the invariance condition $f(x) \in \mathcal{C}(x)$ just at the vertices of $\mathcal{O}_{\mathcal{S}}$. The following lemma formulates the non-vanishing condition on f in Problem 5.2 as $\kappa + 1$ inequalities, where $\mathcal{O}_{\mathcal{S}}$ has $\kappa + 1$ vertices.

Lemma 6.3. *Let $V_{\mathcal{O}_{\mathcal{S}}} = \{o_1, \ldots, o_{\kappa+1}\}$ be the set of vertices of $\mathcal{O}_{\mathcal{S}}$. Assume that there exists a map $f' : V_{\mathcal{O}_{\mathcal{S}}} \to \mathcal{B}$ that is extendible on $\mathcal{O}_{\mathcal{S}} \neq \emptyset$ to an affine map $f : \mathcal{O}_{\mathcal{S}} \to \mathcal{B}$. The following statements are equivalent:*

(i) The affine map f satisfies the conditions of Problem 5.2.

(ii) The vertex map f' satisfies

$$\begin{aligned} f'(o_i) &\in \mathcal{C}(o_i), && i \in \{1, \ldots, \kappa+1\}, \\ (\exists\, \xi \in \mathbb{R}^n) \quad \xi \cdot f'(o_i) &< 0, && i \in \{1, \ldots, \kappa+1\}. \end{aligned}$$

Proof. (i) $\Longrightarrow$ (ii) The arguments in this part are similar to the proof of Theorem 6 in [122]. First, from Problem 5.2 we have $f(x) \in \mathcal{C}(x)$ for all $x \in \mathcal{O}_{\mathcal{S}}$, which trivially implies $f'(o_i) \in \mathcal{C}(o_i)$, $i \in \{1, \ldots, \kappa+1\}$. Next, since $\mathcal{O}_{\mathcal{S}}$ is compact and convex it follows that the image of $\mathcal{O}_{\mathcal{S}}$ under the affine map f, denoted by $C_1 := f(\mathcal{O}_{\mathcal{S}})$, is also compact and convex and does not contain the origin, by condition (i) of the Lemma. Thus letting $C_2 = \{0\}$ and using the Separating Hyperplane Theorem [120] there exists a hyperplane that separates C_1 and C_2 strongly. In other words, there exists $\varepsilon > 0$ and some $\xi \in \mathbb{R}^n$ such that for all $x \in \mathcal{O}_{\mathcal{S}}$, $\xi \cdot f(x) \leq -\varepsilon$, or $\xi \cdot f(x) < 0$. Therefore, $\xi \cdot f'(o_i) < 0$, $i \in \{1, \ldots, \kappa+1\}$.

(ii) $\Longrightarrow$ (i) First, note that any point $x \in \mathcal{O}_{\mathcal{S}}$ can be written as $x = \sum_{i=1}^{\kappa+1} \alpha_i o_i \in \mathcal{O}_{\mathcal{S}}$, where $\alpha_i \geq 0$ and $\sum_{i=1}^{\kappa+1} \alpha_i = 1$. Define the non-empty set $I^+ := \{i \mid 1 \leq i \leq \kappa+1,\ \alpha_i > 0\}$. Therefore, $x \in \mathrm{co}\{o_i \mid i \in I^+\}$, which yields $x \in \mathrm{co}\{v_j \mid j \in I(o_i),\ i \in I^+\}$. Hence, $I(o_i) \subset I(x)$ for any $i \in I^+$, which yields $\bigcup_{i \in I^+} I(o_i) \subset I(x)$. Therefore, $I \backslash \bigcup_{i \in I^+} I(o_i) \supset I \backslash I(x)$, or equivalently $\bigcap_{i \in I^+} (I \backslash I(o_i)) \supset I \backslash I(x)$. Since f is the affine extension of f' on $\mathcal{O}_{\mathcal{S}}$, we can write $f(x) = \sum_{i=1}^{\kappa+1} \alpha_i f'(o_i) = \sum_{i \in I^+} \alpha_i f'(o_i)$. Assume $f'(o_i) \in \mathcal{C}(o_i)$, $i \in \{1, \ldots, \kappa+1\}$. For any $j \in I \setminus I(x) \subset \bigcap_{i \in I^+} (I \setminus I(o_i))$, equation (3.3) yields $h_j \cdot f(x) = \sum_{i \in I^+} \alpha_i h_j \cdot f'(o_i) \leq 0$. Therefore, $f(x) \in \mathcal{C}(x)$ for all $x \in \mathcal{O}_{\mathcal{S}}$. Next, assume there exists $\xi \in \mathbb{R}^n$ such that $\xi \cdot f'(o_i) < 0$, $i \in \{1, \ldots, \kappa+1\}$. For any $x \in \mathcal{O}_{\mathcal{S}}$, we can write $\xi \cdot f(x) = \sum_{i \in I^+} \alpha_i \xi \cdot f'(o_i) < 0$. Therefore, $f(x) \neq 0$ for all $x \in \mathcal{O}_{\mathcal{S}}$, i.e., the affine map f is non-vanishing. $\square$

It is worth mentioning that the second condition above is equivalent to the statement that the vector field points outside the set $\mathcal{O}_{\mathcal{S}}$. In light of Lemma 6.3, we restate Problem 5.2 as follows:

Problem 6.4 (Affine Obstruction)**.** *Assume that $\mathcal{O}_\mathcal{S} \neq \emptyset$ is a simplex. Find a vertex map $f : V_{\mathcal{O}_\mathcal{S}} \to \mathcal{B}$ such that $f(o_i) \in \mathcal{C}(o_i)$, $i \in \{1, \ldots, \kappa+1\}$, and that there exists a vector $\xi \in \mathbb{R}^n$ satisfying $\xi \cdot f(o_i) < 0$, $i \in \{1, \ldots, \kappa+1\}$.*

We note that the assumption that $\mathcal{O}_\mathcal{S}$ is a simplex is both natural and necessary. If it did not hold, we might not be able to independently assign the values of an affine function at vertices of $\mathcal{C}(o_i)$. This presents a different kind of obstruction than the one we want to concentrate on. Additionally, if $\mathcal{O}_\mathcal{S}$ is not a simplex, Lemma 6.3 might not hold. Hence, in the remainder of the chapter, $\mathcal{O}_\mathcal{S}$ is a κ-dimensional simplex with the vertex set $V_{\mathcal{O}_\mathcal{S}} = \{v_1, \ldots, v_{\kappa+1}\}$.

Based on Lemma 6.3, if Problem 6.4 is infeasible then there exists an obstruction to solvability of the RCP using affine feedback.

As in Chapter 5, let the columns of $Q \in \mathbb{R}^{n\times m}$ form an orthonormal basis for $\mathcal{B}$, such that $Q^TQ = I_{m\times m}$ and $QQ^Tb = b$ for all $b \in \mathcal{B}$. Such a Q is produced by the *QR factorization* of matrix B [45]. Our objective is to find a vertex map f such that $f(o_i) \in \mathcal{B} \cap \mathcal{C}(o_i)$, $i \in \{1, \ldots, \kappa+1\}$. To this end, define the convex cone $\mathcal{C}^i$, $i \in \{1, \ldots, \kappa+1\}$, as

$$\mathcal{C}^i := \{w \in \mathbb{R}^m \mid (Q^Th_j) \cdot w \leq 0,\ j \in I \setminus I(o_i)\} = \bigcap_{j\in I\setminus I(o_i)} \mathcal{C}_j, \tag{6.2}$$

where $\mathcal{C}_j$ is defined analogously to (5.11), i.e., by

$$\mathcal{C}_j := \{w \in \mathbb{R}^m \mid (Q^Th_j) \cdot w \leq 0\}, \qquad j \in I. \tag{6.3}$$

Cones $\mathcal{C}^i$ are analogues of the cones $\mathcal{C}(o_i)$ in the original problem set-up. In fact, as $\mathcal{B}$ is mapped by into $\mathbb{R}^m$ by Q^T, the same mapping transforms $\mathcal{C}(o_i) \cap \mathcal{B}$ into $\mathcal{C}^i$. As in Chapter 5, cones $\mathcal{C}_j$ identify each of the possible constraints impacting $\mathcal{C}^i$.

Lemma 6.5. *Let $y \in \mathcal{B}$. Then $Q^Ty \in \mathcal{C}^i$ if and only if $y \in \mathcal{B} \cap \mathcal{C}(o_i)$.*

Proof. ($\Longrightarrow$) Since $y \in \mathcal{B}$ we can write $QQ^Ty = y$. Assume $Q^Ty \in \mathcal{C}^i$. For $j \in I \setminus I(o_i)$, (6.2) yields $(Q^Th_j) \cdot (Q^Ty) \leq 0 \Longleftrightarrow h_j^TQQ^Ty \leq 0 \Longleftrightarrow h_j \cdot y \leq 0$. Considering (3.3), $y \in \mathcal{B} \cap \mathcal{C}(o_i)$. ($\Longleftarrow$) The result follows from arguments similar to the first part of the proof. □

Remark 6.6. *We note that there are three different notions of essentially the same cone: in the original definition from Chapter 3, $\mathcal{C}(o_i)$ is a cone in $\mathbb{R}^n$. This differs from $\mathcal{C}(o_i)$ produced in Chapter 5 from $\mathcal{B} \cap \mathcal{C}(o_i)$ by abuse of notation, which is a subset of $\mathbb{S}^{m-1}$. Finally, $\mathcal{C}^i$ in this chapter is a cone in $\mathbb{R}^m$, again analogous to $\mathcal{B} \cap \mathcal{C}(o_i)$, as shown by Lemma 6.5. Nonetheless, while these three cones are slightly*

different to account for different approaches to the solvability of the RCP, they are obtained from each other by simple transformations, and ultimately serve the same purpose: to constrain the possible values of $f(o_i)$.

Same as in Chapter 5, if $Q^T h_j \neq 0$ then $\mathcal{C}_j$ represents a closed half space in $\mathbb{R}^m$. If $Q^T h_j = 0$ then $\mathcal{C}_j = \mathbb{R}^m$. Since $\mathcal{C}^i$ is the intersection of a finite number of half spaces it is a convex polyhedral cone, and it can be written as the convex hull of a finite set of rays. For each $i \in \{1, \ldots, \kappa+1\}$, let the unit vectors corresponding to these rays be $v_{i,r}$, $r \in \{1, \ldots, r_i\}$. Therefore, any vector $w \in \mathcal{C}^i$ can be written as the conic combination of $v_{i,r}$, $r \in \{1, \ldots, r_i\}$, i.e., $w = \sum_{r=1}^{r_i} \lambda_r v_{i,r}$, for some $\lambda_r \geq 0$. Next, for $i \in \{1, \ldots, \kappa+1\}$, define

$$\mathcal{N}_i := \bigcup_{r=1}^{r_i} \{h \in \mathbb{R}^m \mid h \cdot v_{i,r} < 0\}. \tag{6.4}$$

Each $\mathcal{N}_i$ represents the set of all vectors $h \in \mathbb{R}^m$ for which there exists a vector $w \in \mathcal{C}^i$ such that $h \cdot w < 0$. In other words, each $\mathcal{N}_i$ contains all the vectors h for which there exists an element w which is contained in $\mathcal{C}^i$ and in the halfspace *opposite* from h. This definition may be inelegant; however, we will soon show that it can be naturally interpreted in relation to the standard notion of dual cones. Note that $\mathcal{N}_i$ is a (not necessarily convex) *blunt cone* (i.e., a cone that does not contain 0) because $h \in \mathcal{N}_i$ implies $\lambda h \in \mathcal{N}_i$ for any $\lambda > 0$, and $0 \notin \mathcal{N}_i$. Theorem 6.7 formulates a necessary and sufficient condition for the obstruction described by Problem 6.4 in terms of the intersection of $\mathcal{N}_i$, $i \in \{1, \ldots, \kappa+1\}$.

Theorem 6.7. *Let $\mathcal{O}_{\mathcal{S}} \neq \emptyset$ be a simplex. There exists a solution to Problem 6.4, i.e., to Problem 5.2, if and only if*

$$\bigcap_{i=1}^{\kappa+1} \mathcal{N}_i \neq \emptyset. \tag{6.5}$$

Proof. ($\Longrightarrow$) Suppose Problem 6.4 is solvable, i.e., there exist $f(o_i) \in \mathcal{B} \cap \mathcal{C}(o_i)$, $i \in \{1, \ldots, \kappa+1\}$, and $\xi \in \mathbb{R}^n$ such that $\xi \cdot f(o_i) < 0$ for all $i \in \{1, \ldots, \kappa+1\}$. By Lemma 6.5, $Q^T f(o_i) \in \mathcal{C}^i$ and we can write $Q^T f(o_i) = \sum_{r=1}^{r_i} \lambda_r v_{i,r}$, for some $\lambda_r \geq 0$. Since $f(o_i) \in \mathcal{B}$ we have $QQ^T f(o_i) = f(o_i)$. Therefore, for $i \in \{1, \ldots, \kappa+1\}$, $\xi \cdot f(o_i) = \xi^T QQ^T f(o_i) = (Q^T \xi) \cdot (\sum_{r=1}^{r_i} \lambda_r v_{i,r}) < 0$. Since $\lambda_r \geq 0$ for all $r \in \{1, \ldots, r_i\}$, there exists r^* such that $(Q^T \xi) \cdot v_{i,r^*} < 0$. Therefore, $Q^T \xi \in \mathcal{N}_i$, for all $i \in \{1, \ldots, \kappa+1\}$, and (6.5) is satisfied.

($\Longleftarrow$) Suppose (6.5) holds. Since $\mathcal{N}_i$ does not contain 0, there exists a non-zero vector $h^\circ \in \mathbb{R}^m$ such that for each $i \in \{1, \ldots, \kappa+1\}$ there exists a vector v_{i,r_i^*} such that $h^\circ \cdot v_{i,r_i^*} < 0$. Let $f(o_i) := Qv_{i,r_i^*}$, $i \in \{1, \ldots, \kappa+1\}$. Clearly, $f(o_i) \in \mathcal{B} = \mathrm{Im}(Q)$. Since $Q^T f(o_i) = Q^T Q v_{i,r_i^*} = v_{i,r_i^*} \in \mathcal{C}^i$, by Lemma 6.5, $f(o_i) \in \mathcal{B} \cap \mathcal{C}(o_i)$. Next, let $\xi = Qh^\circ$. Then for $i \in \{1, \ldots, \kappa+1\}$, $\xi \cdot f(o_i) = (Qh^\circ) \cdot (Qv_{i,r_i^*}) = h^{\circ T} Q^T Q v_{i,r_i^*} = h^\circ \cdot v_{i,r_i^*} < 0$. $\square$

Theorem 6.7 provides a necessary and sufficient condition for solvability of Problem 6.4 for multi-input systems. However, since the cones $\mathcal{N}_i$ are not necessarily convex, condition (6.5) in Theorem 6.7 leads to a non-convex feasibility problem. Nonetheless, assuming that (6.5) is satisfied and a vector $h^\circ \in \bigcap_{i=1}^{\kappa+1} \mathcal{N}_i$ is known, Problem 6.4 can be formulated as a computationally efficient feasibility program in terms of linear inequalities, as shown in the next corollary.

Corollary 6.8. *Suppose there exists a vector $h^\circ \in \bigcap_{i=1}^{\kappa+1} \mathcal{N}_i$. The vertex map $f(o_i) = Qw_i$, $i \in \{1, \ldots, \kappa+1\}$, is a solution to Problem 6.4, where w_i always exists and is a solution of the following feasibility program:*

$$\begin{aligned} &\text{find} && w_i \in \mathcal{C}^i \\ &\text{subject to} && h^\circ \cdot w_i < 0. \end{aligned} \tag{6.6}$$

Proof. Suppose $h^\circ \in \bigcap_{i=1}^{\kappa+1} \mathcal{N}_i$. Considering (6.4), for each $i \in \{1, \ldots, \kappa+1\}$, there exists $r_i^* \in \{1, \ldots, r_i\}$ such that $h^\circ \cdot v_{i,r_i^*} < 0$. Therefore, optimization program (6.6) is always feasible since clearly $v_{i,r_i^*} \in \mathcal{C}^i$. Since $Q^T Q w_i = w_i \in \mathcal{C}^i$, by Lemma 6.5, $f(o_i) = Qw_i \in \mathcal{B} \cap \mathcal{C}(o_i)$. For $i \in \{1, \ldots, \kappa+1\}$, observe that $h^\circ \cdot w_i = h^{\circ T} Q^T Q w_i = (Qh^\circ) \cdot (Qw_i) < 0$. Define $\xi := Qh^\circ \in \mathbb{R}^n$. Therefore, $\xi \cdot f(o_i) < 0$, which completes the proof. □

The necessary and sufficient condition in Theorem 6.7 is based on the cones $\mathcal{N}_i$, which are not necessarily convex. Instead, it is appealing to reformulate the necessary and sufficient condition (6.5) in terms of the more standard notion of the dual cones, which are convex. This new dual formulation simplifies the presentation of the results in the remaining sections of the chapter. To this end, note that the *dual cone* of a cone $C \in \mathbb{R}^n$ is defined as ([18])

$$C^* := \{y \in \mathbb{R}^n \mid y \cdot c \geq 0, \ \forall\, c \in C\}. \tag{6.7}$$

Since every vector in $\mathcal{C}^i$ is a conic combination of the vectors $v_{i,r}$, $r \in \{1, \ldots, r_i\}$, the dual cone of $\mathcal{C}^i$, $i \in \{1, \ldots, \kappa+1\}$, is written as

$$\mathcal{C}^{i*} = \{y \in \mathbb{R}^m \mid y \cdot v_{i,r} \geq 0, \ \forall\, r = 1, \ldots, r_i\}. \tag{6.8}$$

Corollary 6.9. *There exists a solution to Problem 6.4 if and only if*

$$\bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*} \neq \mathbb{R}^m. \tag{6.9}$$

Proof. Considering (6.4) and (6.8), it is easy to see that $\mathcal{N}_i = \mathbb{R}^m \setminus \mathcal{C}^{i*}$. Therefore,

$$\bigcap_{i=1}^{\kappa+1} \mathcal{N}_i = \bigcap_{i=1}^{\kappa+1} (\mathbb{R}^m \setminus \mathcal{C}^{i*}) = \mathbb{R}^m \setminus \bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*}. \tag{6.10}$$

Thus, conditions (6.5) and (6.9) are equivalent. □

Corollary 6.9 presents necessary and sufficient conditions for the obstruction identified in Problem 6.4 as a compact and plausible cone condition. According to [38], however, determining if the union of a set of polyhedral cones covers $\mathbb{R}^m$ is an NP-complete problem. Therefore, Corollary 6.9 (and analogously Theorem 6.7) cannot be efficiently solved using optimization software. We note that a similar problem arose in Section 5.5. In the rest of the chapter, we present computationally efficient conditions for solvability of Problem 6.4. In particular, (i) for two-input systems, necessary and sufficient conditions are presented in Section 6.3 in terms of easily verifiable convexity relations, and (ii) for general systems, necessary conditions are presented in Section 6.4 as feasibility programs in terms of linear inequalities. The results of Sections 6.3 and 6.4 can be easily programmed and solved using available optimization software.

6.3 Two-input Systems

This section is focused on two-input systems, i.e., systems for which $\mathcal{B}$ is a 2-dimensional subspace of $\mathbb{R}^n$. This allows us to present a graphical representation of Corollary 6.9, and propose computationally efficient necessary and sufficient conditions for solving Problem 6.4 for two-input systems. To this end, consider (6.2) where $\mathcal{C}^i$ is defined as the intersection of a set of closed convex cones $\mathcal{C}_j$. According to Theorem 2 in [124], for $i \in \{1, \ldots, \kappa+1\}$ we can write

$$\mathcal{C}^{i*} = \mathrm{co}\Big\{ \bigcup_{j \in I \setminus I(o_i)} \mathcal{C}_j^* \Big\} = \mathrm{co}\{\mathcal{C}_j^* \mid j \in I \setminus I(o_i)\}, \tag{6.11}$$

where $\mathcal{C}_j^*$ is the dual cone of $\mathcal{C}_j$. The set $\mathcal{C}_j^*$ is either a ray or the singleton $\{0\}$, as proved in the following lemma.

Lemma 6.10. *If $Q^T h_j \neq 0$ then $\mathcal{C}_j^*$ is a ray given by $\mathcal{C}_j^* := \{-\alpha Q^T h_j \mid \alpha \geq 0\}$. Otherwise, $\mathcal{C}_j^* = \{0\}$.*

Proof. Considering (6.3) and (6.7), for all $j \in I$, we can write

$$\mathcal{C}_j^* = \{y \mid y \cdot w \geq 0 \text{ for all } w \in \mathcal{C}_j\} = \{y \mid y \cdot w \geq 0 \text{ for all } w \text{ such that } -(Q^T h_j) \cdot w \geq 0\} \tag{6.12}$$

If $Q^T h_j = 0$ then $\mathcal{C}_j^* = \{y \mid y \cdot w \geq 0,\ \forall\, w \in \mathbb{R}^m\} = \{0\}$. Next, assume $Q^T h_j \neq 0$. Any vector αy^*, where $y^* = -Q^T h_j$, lies in $\mathcal{C}_j^*$ if and only if $\alpha \geq 0$. We claim that any vector y that is not collinear with y^* does not lie in $\mathcal{C}_j^*$. Without loss of generality, let $y = \beta y^* + \beta_\perp y_\perp^*$, where $y^* \cdot y_\perp^* = 0$, $y_\perp^* \neq 0$, $\beta, \beta_\perp \in \mathbb{R}$, and $\beta_\perp \neq 0$. Equation (6.3) yields $y_\perp^* \in \mathcal{C}_j$ and $-y_\perp^* \in \mathcal{C}_j$. However, $y \cdot y_\perp^* = \beta_\perp |y_\perp^*|^2$ and $y \cdot (-y_\perp^*) = -\beta_\perp |y_\perp^*|^2$. Therefore, either $y \cdot y_\perp^*$ or $y \cdot (-y_\perp^*)$ is less than zero. Hence, by (6.12), $y \notin \mathcal{C}_j^*$. $\square$

Without loss of generality, by reordering indices, assume $Q^T h_j \neq 0$, $j \in \{1, \ldots, n'\}$, where $0 \leq n' \leq n$. Now assume the system has two inputs, i.e., assume $m = 2$. Without loss of generality, by reordering indices $j \in \{2, \ldots, n'\}$, assume the rays $\mathcal{C}_j^*$, $j \in \{1, \ldots, n'\}$, are arranged in clockwise order (see Figure 6.2 for an example). Let $\mathcal{C}_0^* := \mathcal{C}_{n'}^*$ and define $-\mathcal{C}_j^* := \{y \mid -y \in \mathcal{C}_j^*\}$, $j \in \{0, \ldots, n'\}$. The following lemma presents three special cases where the obstruction described by Problem 6.4 does not exist.

Lemma 6.11. *Suppose the affine system has two inputs. Problem 6.4 is solvable if one of the following conditions is satisfied:*

(i) $n' \leq 2$,

(ii) *there exists* $j' \in \{0, \ldots, n'-1\}$ *such that* $\mathcal{C}_{j'+1}^* = -\mathcal{C}_{j'}^*$,

(iii) *there exists* $j' \in \{0, \ldots, n'-1\}$ *such that* $\mathcal{C}_j' \subset \mathrm{co}\{\mathcal{C}_{j'}^*, \mathcal{C}_{j'+1}^*\}$, $\forall j \in \{1, \ldots, n'\}$.

Proof. *(i)* Assume $n' \leq 2$. Considering (6.11), we have $\mathcal{C}^{i*} \subset \mathrm{co}\{\{0\} \cup \big(\bigcup_{1 \leq j \leq n'} \mathcal{C}_j^*\big)\}$, $i \in \{1, \ldots, \kappa+1\}$. Hence, $\bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*} \subset \mathrm{co}\{\{0\} \cup \big(\bigcup_{1 \leq j \leq n'} \mathcal{C}_j^*\big)\}$. Considering Figure 6.1, the set $\mathrm{co}\{\{0\} \cup \big(\bigcup_{1 \leq j \leq n'} \mathcal{C}_j^*\big)\}$ is equal to the singleton $\{0\}$ if $n' = 0$, the ray $\mathcal{C}_1^*$ if $n' = 1$, or the cone $\mathrm{co}\{\mathcal{C}_1^*, \mathcal{C}_2^*\} \neq \mathbb{R}^m$ if $n' = 2$. Therefore, there always exists a vector $h^\circ \notin \mathrm{co}\{\{0\} \cup \big(\bigcup_{1 \leq j \leq n'} \mathcal{C}_j^*\big)\}$ and, by Corollary 6.9, Problem 6.4 is solvable.

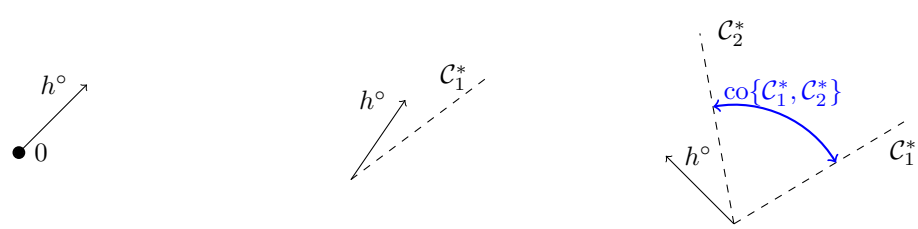

Figure 6.1: An illustration describing the proof of part (i) of Lemma 6.11. The set $\mathrm{co}\{\{0\} \cup \big(\bigcup_{1 \leq j \leq n'} \mathcal{C}_j^*\big)\}$ is equal to (a) the singleton $\{0\}$ (if $n' = 0$), (b) the ray $\mathcal{C}_1^*$ (if $n' = 1$), or (c) the cone $\mathrm{co}\{\mathcal{C}_1^*, \mathcal{C}_2^*\}$ (if $n' = 2$).

(ii) Assume there exists $j' \in \{0, \ldots, n'-1\}$ such that $\mathcal{C}_{j'+1}^* = -\mathcal{C}_{j'}^*$. Since the rays are arranged in clockwise order, all rays $\mathcal{C}_j^*$, $j \in \{1, \ldots, n'\}$, lie in the same side of the line $\mathcal{C}_{j'}^* \cup \mathcal{C}_{j'+1}^*$, in a closed

half-plane $\mathcal{P}$ (see Figure 6.2). Considering (6.11), the cones $\mathcal{C}^{i*}$, $i \in \{1, \ldots, \kappa+1\}$, are subsets of $\mathcal{P}$. Hence by Corollary 6.9, Problem 6.4 is solvable.

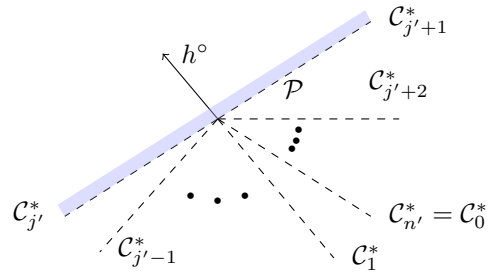

Figure 6.2: An illustration corresponding to the proof of part *(ii)* of Lemma 6.11.

(iii) Assume there exists $j' \in \{0, \ldots, n'-1\}$ such that $\mathcal{C}^*_j \subset \mathrm{co}\{\mathcal{C}^*_{j'}, \mathcal{C}^*_{j'+1}\}$ for all $j \in \{1, \ldots, n'\}$. Hence, considering Lemma 6.10, each vector $-Q^T h_j$, $j \in \{1, ..., n'\}$, can be written as a conic combination of the two vectors $-Q^T h_{j'}$ and $-Q^T h_{j'+1}$. Recall that the indices were reordered without loss of generality such that $Q^T h_j = 0$, $j \in \{n'+1, ..., n\}$. Furthermore, according to (6.11) and Lemma 6.10, any vector in $\mathcal{C}^{i*}$ can be written as a conic combination of the vectors $-Q^T h_j$, $j \in I \setminus I(o_i)$. Therefore, any vector in $\mathcal{C}^{i*}$ can be written as a conic combination of the two vectors $-Q^T h_{j'}$ and $-Q^T h_{j'+1}$, which yields $\mathcal{C}^{i*} \subset \mathrm{co}\{\mathcal{C}^*_{j'}, \mathcal{C}^*_{j'+1}\}$. Pick a vector $h^\circ \notin \mathrm{co}\{\mathcal{C}^*_{j'}, \mathcal{C}^*_{j'+1}\}$. Clearly, $h^\circ \notin \mathcal{C}^{i*}$, $i \in \{1, \ldots, \kappa+1\}$ (see Figure 6.3). Therefore, by Corollary 6.9, Problem 6.4 is solvable. □

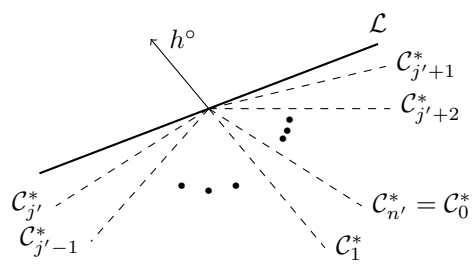

Figure 6.3: An illustration corresponding to the proof of part *(iii)* of Lemma 6.11 and the second part of Theorem 6.14.

The results of the following Lemma are used to address the case where none of the conditions in Lemma 6.11 is satisfied.

Lemma 6.12. *The following statements hold for two-input systems.*

(i) *The dual cones $\mathcal{C}^{i*}$, $i \in \{1, \ldots, \kappa+1\}$, are convex and their boundaries are among $\mathcal{C}^*_j$, $j \in I$.*

(ii) *If $\cup_{j=0}^{n'-1} \mathrm{co}\{\mathcal{C}^*_j, \mathcal{C}^*_{j+1}\} = \mathbb{R}^2$, then $\bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*} = \mathbb{R}^2$ if and only if each $\mathrm{co}\{\mathcal{C}^*_j, \mathcal{C}^*_{j+1}\}$, $j \in \{0, \ldots, n'-1\}$, is contained in some $\mathcal{C}^{i*}$.*

Proof. *(i)* This can be verified by (6.11) and Theorem 2 in [41].

(ii) In one direction, assume each $\text{co}\{\mathcal{C}_j^*, \mathcal{C}_{j+1}^*\}$, $j \in \{0, \ldots, n'-1\}$, is contained in some $\mathcal{C}^{i*}$. Then, using the assumption in part (ii) of the lemma, $\mathbb{R}^2 = \cup_{j=0}^{n'-1}\text{co}\{\mathcal{C}_j^*, \mathcal{C}_{j+1}^*\} \subset \bigcup_{i=1}^{\kappa+1}\mathcal{C}^{i*}$, i.e. $\bigcup_{i=1}^{\kappa+1}\mathcal{C}^{i*} = \mathbb{R}^2$.

In the other direction, let $\bigcup_{i=1}^{\kappa+1}\mathcal{C}^{i*} = \mathbb{R}^2$. Now assume there exists $j^* \in \{0, \ldots, n'-1\}$ such that $\text{co}\{\mathcal{C}_{j^*}^*, \mathcal{C}_{j^*+1}^*\} \not\subset \mathcal{C}^{i*}$, $i \in \{1, \ldots, \kappa+1\}$. Since the boundaries of $\mathcal{C}^{i*}$ are among $\mathcal{C}_j^*$ (part (i) of the lemma) and considering the fact that $\mathcal{C}_{j^*}^*$ and $\mathcal{C}_{j^*+1}^*$ are in consecutive (clockwise) order, we can conclude that the intersection of $\text{co}\{\mathcal{C}_{j^*}^*, \mathcal{C}_{j^*+1}^*\}$ with $\mathcal{C}^{i*}$, $i \in \{1, \ldots, \kappa+1\}$, is either $\{0\}$ or any of the rays $\mathcal{C}_{j^*}^*$ or $\mathcal{C}_{j^*+1}^*$. Hence, there exists a vector $h^\circ \in \text{co}\{\mathcal{C}_{j^*}^*, \mathcal{C}_{j^*+1}^*\}$ such that $h^\circ \notin \mathcal{C}^{i*}$, $i \in \{1, \ldots, \kappa+1\}$, which is in contradiction with $\bigcup_{i=1}^{\kappa+1}\mathcal{C}^{i*} = \mathbb{R}^2$. □

Based on Lemma 6.12, if the union of convex hulls $\cup_{j=0}^{n'-1}\text{co}\{\mathcal{C}_j^*, \mathcal{C}_{j+1}^*\}$ covers the whole space, then $\bigcup_{i=1}^{\kappa+1}\mathcal{C}^{i*} = \mathbb{R}^2$ if and only if each convex hull $\text{co}\{\mathcal{C}_j^*, \mathcal{C}_{j+1}^*\}$ is contained in some $\mathcal{C}^{i*}$. Furthermore, according to Corollary 6.9, Problem 6.4 is infeasible if and only if the union of the cones $\mathcal{C}^{i*}$ covers $\mathbb{R}^2$. In other words, if $\cup_{j=0}^{n'-1}\text{co}\{\mathcal{C}_j^*, \mathcal{C}_{j+1}^*\} = \mathbb{R}^2$, Problem 6.4 is infeasible if and only if each pair of rays $\mathcal{C}_j^*$ and $\mathcal{C}_{j+1}^*$ is contained in some $\mathcal{C}^{i*}$. Thus, we need to know which cones $\mathcal{C}^{i*}$, $i \in \{1, \ldots, \kappa+1\}$ each ray $\mathcal{C}_j^*$, $j \in \{1, \ldots, n'\}$, is contained in. To this end, for $j \in \{1, \ldots, n'\}$, define

$$J_j := \{i \in \{1, \ldots, \kappa+1\} \mid j \in \ I \setminus I(o_i)\}. \tag{6.13}$$

Considering (6.11), if $i \in J_j$ then $\mathcal{C}_j^* \subset \mathcal{C}^{i*}$. However, there may exist j^* and i^* such that $\mathcal{C}_{j^*}^* \subset \mathcal{C}^{i^**}$ but $i^* \notin J_{j^*}$. For example, assume $I = \{1, 2, 3, 4\}$ and $I(o_1) = \{0, 4\}$. Let the rays $\mathcal{C}_2^*$, $\mathcal{C}_3^*$, and $\mathcal{C}_4^*$ be as shown in Figure 6.4. Here, $\mathcal{C}_4^* \subset \mathcal{C}^{1*}$, but $1 \notin J_4$. Therefore, we should update the sets J_j by adding new indices i^* that satisfy $\mathcal{C}_j^* \subset \mathcal{C}^{i^**}$. Note that if $\mathcal{C}^{i*}$ contains the rays $\mathcal{C}_{j_1}^*$ and $\mathcal{C}_{j_2}^*$, then by convexity $\mathcal{C}^{i*}$ contains any ray $\mathcal{C}_{j_3}^* \subset \text{co}\{\mathcal{C}_{j_1}^*, \mathcal{C}_{j_2}^*\}$, where $j_1, j_2, j_3 \in \{1, \ldots, n'\}$. Based on this fact, Algorithm 6.13 updates the sets J_j and stores them in new sets $\widetilde{J}_j$. The members of each set $\widetilde{J}_j$ are indices $i \in \{1, \ldots, \kappa+1\}$ such that $\mathcal{C}_j^* \subset \mathcal{C}^{i*}$. If $\widetilde{J}_j \cap \widetilde{J}_{j+1} \neq \emptyset$ for all $j \in \{0, \ldots, n'-1\}$, then there exists some i^* such that $\mathcal{C}_j^*, \mathcal{C}_{j+1}^* \subset \mathcal{C}^{i^**}$ and thus the union of the cones $\mathcal{C}^{i*}$ covers $\mathbb{R}^2$. The following theorem presents computationally efficient necessary and sufficient conditions for checking the obstruction associated with Problem 6.4 for two-input systems.

Theorem 6.14. *Suppose the affine system has two inputs and none of the conditions in Lemma 6.11 is satisfied. Let the index sets $\widetilde{J}_j$ for all $j \in \{0, \ldots, n'\}$ be computed as in Algorithm 6.13. Problem 6.4 is solvable if and only if there exists $j^\circ \in \{0, \ldots, n'-1\}$ such that $\widetilde{J}_{j^\circ} \cap \widetilde{J}_{j^\circ+1} = \emptyset$.*

Proof. ($\Longleftarrow$) Suppose there exists $j^\circ \in \{0, \ldots, n'-1\}$ such that $\widetilde{J}_{j^\circ} \cap \widetilde{J}_{j^\circ+1} = \emptyset$. We study the following

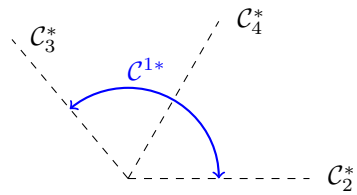

Figure 6.4: An example of rays $\mathcal{C}_2^*$, $\mathcal{C}_3^*$, and $\mathcal{C}_4^*$. In this example $\mathcal{C}_4^* \subset \mathcal{C}^{1*}$, but $1 \notin J_4$, because in this hypothetical scenario $I(o_1) = \{0, 4\}$ and $I = \{1, 2, 3, 4\}$.

Algorithm 6.13.

$\widetilde{J}_j := J_j$, $j \in \{1, \ldots, n'\}$
for $j_1, j_2, j_3 \in \{1, \ldots, n'\}$
  if $\mathcal{C}_{j_3}^* \subset \mathrm{co}\{\mathcal{C}_{j_1}^*, \mathcal{C}_{j_2}^*\}$
    $\widetilde{J}_{j_3} := \widetilde{J}_{j_3} \cup (J_{j_1} \cap J_{j_2})$
  end if
end for
$\widetilde{J}_0 := \widetilde{J}_{n'}$

two cases separately: *(i)* $\mathcal{C}_{j^\circ}^* = \mathcal{C}_{j^\circ+1}^*$ and *(ii)* $\mathcal{C}_{j^\circ}^* \neq \mathcal{C}_{j^\circ+1}^*$.

(i) Assume $\widetilde{J}_{j^\circ} \cap \widetilde{J}_{j^\circ+1} = \emptyset$ and $\mathcal{C}_{j^\circ}^* = \mathcal{C}_{j^\circ+1}^*$. We claim

$$J_{j'} = \emptyset, \text{ for all } j' \in \{1, \ldots, n'\} \text{ such that } \mathcal{C}_{j^\circ}^* = \mathcal{C}_{j^\circ+1}^* = \mathcal{C}_{j'}^*. \tag{6.14}$$

Suppose not. Let $J_{j'} \neq \emptyset$. Since $\mathcal{C}_{j^\circ}^* = \mathcal{C}_{j^\circ+1}^* = \mathrm{co}\{\mathcal{C}_{j'}^*, \mathcal{C}_{j'}^*\}$, Algorithm 6.13 guarantees that $J_{j'}$ is added to $\widetilde{J}_{j^\circ}$ and $\widetilde{J}_{j^\circ+1}$. This yields $\widetilde{J}_{j^\circ} \cap \widetilde{J}_{j^\circ+1} \neq \emptyset$, a contradiction. Therefore, (6.14) holds. By (6.13), observe that (6.14) implies that there does not exist

$$i \in \{1, \ldots, \kappa+1\} \text{ such that } j' \in I \setminus I(o_i), \mathcal{C}_{j^\circ}^* = \mathcal{C}_{j^\circ+1}^* = \mathcal{C}_{j'}^*. \tag{6.15}$$

Next, for $l_1, l_2 \in \{1, \ldots, n'\}$, we claim

$$\mathcal{C}_{j^\circ}^* = \mathcal{C}_{j^\circ+1}^* \subset \mathrm{co}\{\mathcal{C}_{l_1}^*, \mathcal{C}_{l_2}^*\} \implies J_{l_1} \cap J_{l_2} = \emptyset. \tag{6.16}$$

Suppose not. Then Algorithm 6.13 guarantees that $J_{l_1} \cap J_{l_2} \neq \emptyset$ is added to $\widetilde{J}_{j^\circ}$ and $\widetilde{J}_{j^\circ+1}$. This yields $\widetilde{J}_{j^\circ} \cap \widetilde{J}_{j^\circ+1} \neq \emptyset$, a contradiction. Therefore, (6.16) holds. By (6.13), observe that

$$J_{l_1} \cap J_{l_2} = \emptyset \implies \text{ there is no } i \in \{1, \ldots, \kappa+1\} \text{ such that } l_1, l_2 \in I \setminus I(o_i). \tag{6.17}$$

The sets $\mathcal{C}^{i*}$, $i \in \{1, \dots, \kappa+1\}$, are m'-dimensional cones, where $0 \leq m' \leq 2$. By (6.11), a 1-dimensional cone $\mathcal{C}^{i*}$ is the union of some linearly dependent rays $\mathcal{C}^*_j$, $j \in I \setminus I(o_i)$. Since by (6.15), no index $j' \in \{1, \dots, n'\}$ such that $\mathcal{C}^*_{j'} = \mathcal{C}^*_{j^\circ} = \mathcal{C}^*_{j^\circ+1}$ appears in a set $I \setminus I(o_i)$, $i \in \{1, \dots, \kappa+1\}$, the ray $\mathcal{C}^*_{j^\circ} = \mathcal{C}^*_{j^\circ+1}$ cannot be a subset of a 1-dimensional cone. Furthermore, a 2-dimensional cone $\mathcal{C}^{i*}$ is the convex hull of the union of rays $\mathcal{C}^*_{j_1}$ and $\mathcal{C}^*_{j_2}$, $j_1, j_2 \in I \setminus I(o_i)$. Combining (6.16) and (6.17), however, if $\mathcal{C}^*_{j^\circ} = \mathcal{C}^*_{j^\circ+1}$ lies in the convex hull of two rays $\mathcal{C}^*_{l_1}$ and $\mathcal{C}^*_{l_2}$, then there is no index $i \in \{1, \dots, \kappa+1\}$ such that $l_1, l_2 \in I \setminus I(o_i)$. Therefore, the ray $\mathcal{C}^*_{j^\circ} = \mathcal{C}^*_{j^\circ+1}$ cannot be a subset of a 2-dimensional cone either. Hence, $\mathcal{C}^*_{j^\circ} = \mathcal{C}^*_{j^\circ+1} \not\subset \bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*}$ and, according to Corollary 6.9, Problem 6.4 is solvable.

(ii) Assume $\widetilde{J}_{j^\circ} \cap \widetilde{J}_{j^\circ+1} = \emptyset$ and $\mathcal{C}^*_{j^\circ} \neq \mathcal{C}^*_{j^\circ+1}$. Pick any vector h° in the relative interior of $\mathrm{co}\{\mathcal{C}^*_{j^\circ}, \mathcal{C}^*_{j^\circ+1}\}$ (note that since condition *(ii)* in Lemma 6.11 is not satisfied, the rays $\mathcal{C}^*_{j^\circ}$ and $\mathcal{C}^*_{j^\circ+1}$ are linearly independent and the relative interior is non-empty). We prove by contradiction that $h^\circ \notin \bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*}$. Assume $h^\circ \in \mathcal{C}^{i^\circ *}$ for some $i^\circ \in \{1, \dots, \kappa+1\}$. Since $h^\circ \in \mathcal{C}^{i^\circ *}$ and $h^\circ \notin \mathcal{C}^*_j$, $j \in \{1, \dots, n'\}$, by (6.11), $\mathcal{C}^{i^\circ *}$ is a 2-dimensional cone. Using (6.11), there exist $l_1, l_2 \in \{1, \dots, n'\} \setminus I(o_{i^\circ})$ such that $h^\circ \in \mathrm{co}\{\mathcal{C}^*_{l_1}, \mathcal{C}^*_{l_2}\} \subset \mathcal{C}^{i^\circ *}$, where $\mathcal{C}^*_{l_1}$ and $\mathcal{C}^*_{l_2}$ are two linearly independent rays. Notice that $i^\circ \in J_{l_1} \cap J_{l_2}$. Since $h^\circ \in \mathrm{co}\{\mathcal{C}^*_{j^\circ}, \mathcal{C}^*_{j^\circ+1}\} \cap \mathrm{co}\{\mathcal{C}^*_{l_1}, \mathcal{C}^*_{l_2}\} \neq \emptyset$ (see Figure 6.5) and $\mathcal{C}^*_{j^\circ}$ and $\mathcal{C}^*_{j^\circ+1}$ are in consecutive (clockwise) order, we have $\mathrm{co}\{\mathcal{C}^*_{j^\circ}, \mathcal{C}^*_{j^\circ+1}\} \subset \mathrm{co}\{\mathcal{C}^*_{l_1}, \mathcal{C}^*_{l_2}\}$. Therefore, Algorithm 6.13 guarantees that i° is added to the sets $\widetilde{J}_{j^\circ}$ and $\widetilde{J}_{j^\circ+1}$. This contradicts $\widetilde{J}_{j^\circ} \cap \widetilde{J}_{j^\circ+1} = \emptyset$. Hence, $h^\circ \notin \bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*}$ and, according to Corollary 6.9, Problem 6.4 is solvable.

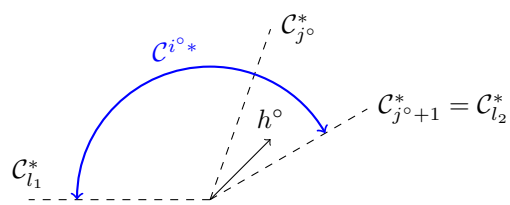

Figure 6.5: An illustration corresponding to the proof of the first part of Theorem 6.14.

($\Longrightarrow$) First, we claim that $\bigcup_{j=0}^{n'-1} \mathrm{co}\{\mathcal{C}^*_j, \mathcal{C}^*_{j+1}\} = \mathbb{R}^m$. Suppose not. Then there exists a ray $h^\circ \not\subset \mathrm{co}\{\mathcal{C}^*_j, \mathcal{C}^*_{j+1}\}$, $j \in \{0, \dots, n'-1\}$. Let $\mathcal{C}^*_{j^\circ}$ and $\mathcal{C}^*_{j^\circ+1}$ be the rays that have the smallest angular distance to h° on each side (see Figure 6.3). Since h° and the relative interior of $\mathrm{co}\{\mathcal{C}^*_{j^\circ}, \mathcal{C}^*_{j^\circ+1}\}$ have no points in common, there exists a line $\mathcal{L}$ that separates them. Since the rays are arranged in clockwise order, all the rays $\mathcal{C}^*_j$, $j \in \{1, \dots, n'\}$, lie in the same side of the line $\mathcal{L}$, i.e., all the rays lie in an open half-plane. Therefore, $\mathcal{C}^*_j \subset \mathrm{co}\{\mathcal{C}^*_{j^\circ}, \mathcal{C}^*_{j^\circ+1}\}$, $\forall j \in \{1, \dots, n'\}$, and condition *(iii)* in Lemma 6.11 is satisfied, which is a contradiction.

Now, suppose Problem 6.4 is solvable. According to Corollary 6.9, $\bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*} \neq \mathbb{R}^m$. Suppose by

way of contradiction $\widetilde{J}_j \cap \widetilde{J}_{j+1} \neq \emptyset$ for all $j \in \{0, \ldots, n'-1\}$. Therefore, for each $j \in \{0, \ldots, n'-1\}$, there exists $i^\circ \in \{1, \ldots, \kappa+1\}$ such that $j, j+1 \in I \setminus I(o_{i^\circ})$. By (6.11), $\mathrm{co}\{\mathcal{C}_j^*, \mathcal{C}_{j+1}^*\} \subset \mathcal{C}^{i*}$. Therefore, $\bigcup_{j=0}^{n'-1} \mathrm{co}\{\mathcal{C}_j^*, \mathcal{C}_{j+1}^*\} \subset \bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*}$, which is a contradiction because $\bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*} \neq \mathbb{R}^m$ and we showed in the previous argument that $\bigcup_{j=0}^{n'-1} \mathrm{co}\{\mathcal{C}_j^*, \mathcal{C}_{j+1}^*\} = \mathbb{R}^m$. □

Remark 6.15. (Computational complexity) *It was assumed that the rays $\mathcal{C}_j^*$, $j \in \{1, \ldots, n'\}$, are arranged in clockwise order. The time complexity of efficient sorting algorithms is known to be in $O(n \log n)$ [68]. Lemma 6.11 requires verification of three conditions: condition (i) is trivially checked in constant time. Conditions (ii) and (iii) require at most $O(n)$ and $O(n^2)$ operations, each of which can be formulated as scalar linear equalities or inequalities. Algorithm 6.13 contains, at worst, $O(n^3)$ repetitions of a set union and intersection operation. Thus, the total complexity of Algorithm 6.13 may vary depending on the data structure used in the implementation, but it can certainly be bounded by $O(n^4)$. Theorem 6.14 has the same computational complexity.*

Next, we use the results of this section to solve Problem 6.4 for an affine system with two inputs. Example 6.16, which is four-dimensional as by Section 6.1 lower-dimensional examples will not be particularly illustrative, is important for two reasons. First, it shows that the necessary condition (5.6) proposed in [125] for single-input systems, i.e., $\mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}}) \neq \{0\}$, where $\mathrm{cone}(\mathcal{O}_{\mathcal{S}})$ is defined in (5.5), is not a necessary condition for solvability of Problem 6.4 for multi-input systems. Second, as mentioned earlier, since $\mathcal{O}_{\mathcal{S}}$ is not a face of $\mathcal{S}$, previous results in the literature cannot address the solvability of the RCP in Example 6.16, without resorting to some triangulation of $\mathcal{S}$.

Example 6.16. *Consider a standard orthogonal simplex $\mathcal{S} \subset \mathbb{R}^4$, i.e., a simplex $\mathcal{S} = \mathrm{co}\{v_0, \ldots, v_4\} \subset \mathbb{R}^4$, where $v_0^T = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$, and for $j \in I$, $v_j = e_j$, the jth Euclidean basis vector. We note that $h_j := -e_j$, $j \in I$. Let the parameters of the affine system* (3.1) *be*

$$A = \begin{bmatrix} -8 & -8 & 2 & -6 \\ 4 & 16 & 0 & 4 \\ 4 & 10 & 0 & 4 \\ -4 & 2 & 0 & 0 \end{bmatrix}, \; B = \begin{bmatrix} -3 & -1 \\ 3 & 2 \\ 2 & 1 \\ 0 & -1 \end{bmatrix}, \; a = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Based on $\mathcal{O} = \{x \in \mathbb{R}^4 \mid Ax + a \in \mathcal{B}\}$, it is easy to verify that

$$\mathcal{O} = \left\{ x \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} -8 & 16 & 4 & 0 \\ 0 & -6 & -2 & -2 \end{bmatrix} x = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\},$$

and $\mathcal{O}_{\mathcal{S}} = \mathrm{co}\{o_1, o_2, o_3\}$, *where* $o_1 = (0, 0, 0, \frac{1}{2})$, $o_2 = (\frac{1}{4}, 0, \frac{1}{2}, 0)$, *and* $o_3 = (\frac{1}{3}, \frac{1}{6}, 0, 0)$. *Note that* $I(o_1) = \{0, 4\}$, $I(o_2) = \{0, 1, 3\}$, *and* $I(o_3) = \{0, 1, 2\}$. *The QR factorization of* B *is computed as*

$$Q = \begin{bmatrix} 0.64 & -0.64 & -0.43 & 0 \\ -0.41 & -0.41 & 0 & 0.82 \end{bmatrix}^T, \quad R = \begin{bmatrix} -4.69 & -2.35 \\ 0 & -1.22 \end{bmatrix}.$$

The cones $\mathcal{C}^i$, $i \in \{1, 2, 3\}$, *are shown on the left of Figure 6.6 and it is easy to see that their intersection is the singleton* $\{0\}$.

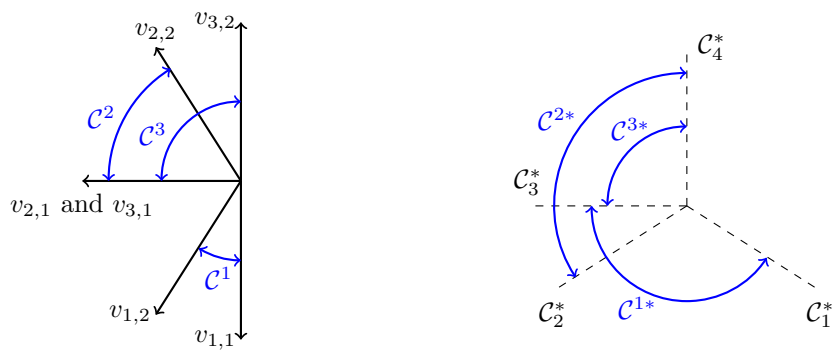

Figure 6.6: Illustrations corresponding to Example 6.16. The rays are shown with dashed lines, the vectors are shown with solid arrows, and the cones are illustrate by blue arcs. The left figure shows cones $\mathcal{C}^i$, $i \in \{1, 2, 3\}$, and the right figure shows the corresponding dual cones $\mathcal{C}^{i*}$, $i \in \{1, 2, 3\}$. $v_{i,1}, v_{i,2}$ are the boundary rays for cones $\mathcal{C}^i$.

Using Lemma 6.5, we conclude that $\mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}}) = \{0\}$. *Next, we use Theorem 6.14 to solve Problem 6.4. This will show that the necessary condition proposed in [125] for single-input systems, i.e.,* (5.6), *is no longer necessary for two-input systems. Note that* $Q^T h_j \neq 0$ *for all* $j \in I$ *and the rays* $\mathcal{C}_j^*$, $j \in I$, *are already in clockwise order as illustrated on the right of Figure 6.6. It is easy to see that* $J_1 = \{1\}$, $J_2 = \{1, 2\}$, $J_3 = \{1, 3\}$, *and* $J_4 = \{2, 3\}$. *Next, Algorithm 6.13 yields* $\widetilde{J}_1 = \{1\}$, $\widetilde{J}_2 = \{1, 2\}$, $\widetilde{J}_3 = \{1, 2, 3\}$, *and* $\widetilde{J}_0 = \widetilde{J}_4 = \{2, 3\}$. *Since* $\widetilde{J}_1 \cap \widetilde{J}_4 = \emptyset$, *by Theorem 6.14, Problem 6.4 is solvable. This is also in accordance with Corollary 6.9, since* $\bigcup_{i=1}^3 \mathcal{C}^{i*} \neq \mathbb{R}^2$ *as shown in the second part of Figure 6.6. Next, we use Corollary 6.8 to compute a vertex map* $f : V_{\mathcal{O}_{\mathcal{S}}} \to \mathcal{B}$ *that solves Problem 6.4. Consider the vector* $h^\circ = [1.5 \quad 0.42]^T \in \mathrm{co}\{\mathcal{C}_1^*, \mathcal{C}_4^*\}$. *Clearly,* $h^\circ \notin \bigcup_{i=1}^3 \mathcal{C}^{i*}$ *and, by* (6.10), $h^\circ \in \bigcap_{i=1}^3 \mathcal{N}_i$. *Solving* (6.6), *we find the following feasible vectors:* $w_1 = [-33.6 \quad -120.46]^T \in \mathcal{C}^1$, $w_2 = [-67.71 \quad 35.1]^T \in \mathcal{C}^2$, *and* $w_3 = [-90.59 \quad 47.59]^T \in \mathcal{C}^3$. *The corresponding closed-loop vertex map* $f(o_i) = Qw_i$, $i \in \{1, 2, 3\}$, *is then computed as* $f(o_1) = (27.69, 70.67, 14.33, -98.36)$, $f(o_2) = (-57.64, 28.98, 28.87, 28.66)$, *and* $f(o_3) = (-77.37, 38.51, 38.63, 38.86)$. *We can use the vertex map* f *to*

find an affine feedback $u(x) = Kx + g$ *such that the closed-loop vector field is non-vanishing on* $\mathcal{O}_{\mathcal{S}}$ *and satisfies* $Ax + Bu(x) + a \in \mathcal{C}(x),\ x \in \mathcal{O}_{\mathcal{S}}$. *To this end, the vertex map* f *is extendible on* $\mathcal{O}_{\mathcal{S}}$ *to the affine map* $F = K_F x + g_F$, *where*

$$K_F = \begin{bmatrix} -233.52 & 2.8 & 1.48 & 55.37 \\ 112.11 & 0.86 & -0.09 & 139.34 \\ 115.21 & 1.36 & 0.14 & 28.66 \\ 113.41 & 6.34 & 0.61 & -196.71 \end{bmatrix}, \quad g_F = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Note that $K_F = A + BK = A + QRK$ *and* $g_F = Bg + a = QRg + a$. *Hence, the affine feedback controller gains* $K = R^{-1}Q^T(K_F - A)$ *and* $g = R^{-1}Q^T(g_F - a)$ *are computed as*

$$K = \begin{bmatrix} 114.31 & -2.15 & 0.38 & -86.03 \\ -117.41 & -4.34 & -0.61 & 196.71 \end{bmatrix}, \quad g = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

It is easy to verify that the closed-loop affine system defined over simplex S *satisfies the conditions in Corollary 9 in [122] with* $\xi = -[5.37 \quad 5.83 \quad 0.87 \quad 5.27]^T$. *Therefore, the RCP is solvable even though* $\mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}}) = \{0\}$.

We note that, while it might have been preferable to have a lower-dimensional example in lieu of Example 6.16, the results of Section 6.1 indicate that the first truly illustrative examples can only appear when $n \geq 4$.

6.4 Computationally Efficient Necessary Conditions for Existence of a Non-vanishing Affine Extension on $\mathcal{O}_{\mathcal{S}}$

In this section, we present necessary conditions for solvability of Problem 6.4 as a feasibility program in terms of linear inequalities. Feasibility problems subject to linear inequalities can easily be programmed and solved using available optimization software. In contrast to Section 6.3, the necessary conditions are applicable to all systems regardless of the number of inputs.

Let $\mathcal{I}$ be an $(m+1)$-subset of $\{1, \ldots, \kappa+1\}$, i.e. $\mathcal{I}$ is a subset of $\{1, \ldots, \kappa+1\}$ with $m+1$ elements. For each $i \in \mathcal{I}$, let $E_{\mathcal{I}}(o_i) \subset I(o_i)$ be the set of non-zero *exclusive members* of $I(o_i)$ in the set $\bigcup_{i \in \mathcal{I}} I(o_i)$, defined as $E_{\mathcal{I}}(o_i) := \{l \in I(o_i) \mid l \notin I(o_j),\ \forall\, j \in \mathcal{I},\ j \neq i\} \backslash \{0\}$. Next, define $S_{\mathcal{I}}$ to be the set of non-zero

shared vertices in $\bigcup_{i\in\mathcal{I}} I(o_i)$ given by

$$S_{\mathcal{I}} := \Big(\bigcup_{i\in\mathcal{I}} I(o_i)\Big) \setminus \Big(\bigcup_{i\in\mathcal{I}} E_{\mathcal{I}}(o_i) \cup \{0\}\Big). \tag{6.18}$$

The following lemma is used in the proof of the main result of this section.

Lemma 6.17. *Each ray $\mathcal{C}_j^*$, $j \in I \setminus S_{\mathcal{I}}$, is contained in all, except (potentially) one, cones $\mathcal{C}^{i*}$, $i \in \mathcal{I}$.*

Proof. Equation (6.18) yields $I \setminus S_{\mathcal{I}} = \big(I \setminus \bigcup_{i\in\mathcal{I}} I(o_i)\big) \cup \big(\bigcup_{i\in\mathcal{I}} E_{\mathcal{I}}(o_i)\big)$. First, assume $j \in I \setminus \bigcup_{i'\in\mathcal{I}} I(o_{i'})$. Since $I \setminus \bigcup_{i'\in\mathcal{I}} I(o_{i'}) \subset I \setminus I(o_i)$, $i \in \mathcal{I}$, (6.11) yields $\mathcal{C}_j^* \subset \mathcal{C}^{i*}$, $i \in \mathcal{I}$. Therefore, each ray $\mathcal{C}_j^*$, $j \in I \setminus \bigcup_{i'\in\mathcal{I}} I(o_{i'})$, is contained in all $m+1$ cones $\mathcal{C}^{i*}$, $i \in \mathcal{I}$. Second, assume $j \in E_{\mathcal{I}}(o_{i'})$, $i' \in \mathcal{I}$. Since

$$E_{\mathcal{I}}(o_{i'}) \subset I \setminus I(o_i),\ i' \in \mathcal{I},\ i \in \mathcal{I} \setminus \{i'\},$$

(6.11) yields $\mathcal{C}_j^* \subset \mathcal{C}^{i*}$, $i \in \mathcal{I} \setminus \{i'\}$. Therefore, each ray $\mathcal{C}_j^*$, $j \in E_{\mathcal{I}}(o_{i'})$, $i' \in \mathcal{I}$, is contained in all cones $\mathcal{C}^{i*}$, $i \in \mathcal{I}$, except potentially $\mathcal{C}^{i'*}$. This completes the proof. □

Given the set $\mathcal{I}$, let the cone $M_{\mathcal{I}}$ be defined as

$$M_{\mathcal{I}} := \{y \in \mathbb{R}^n \mid h_j \cdot y \le 0,\ j \in I \setminus S_{\mathcal{I}}\}. \tag{6.19}$$

The next result shows that if $\mathcal{B}$ is not of sufficiently high dimension, and there are no non-zero vectors lying in $\mathcal{B} \cap M_{\mathcal{I}}$, then the obstruction described by Problem 6.4 exists.

Theorem 6.18. *Assume $1 \le m < \kappa + 1$ and let $\mathcal{I}$ be an $(m+1)$-subset of $\{1, \ldots, \kappa+1\}$. If Problem 6.4 is solvable, then $\mathcal{B} \cap M_{\mathcal{I}} \neq \{0\}$.*

Proof. By way of contradiction, assume Problem 6.4 is solvable, but $\mathcal{B} \cap M_{\mathcal{I}} = \{0\}$. Define the cone $\mathcal{M}_{\mathcal{I}}$ as

$$\mathcal{M}_{\mathcal{I}} := \{w \in \mathbb{R}^m \mid (Q^T h_j) \cdot w \le 0,\ j \in I \setminus S_{\mathcal{I}}\} = \bigcap_{j\in I\setminus S_{\mathcal{I}}} \mathcal{C}_j,$$

where $\mathcal{C}_j$ is defined in (6.3). Similar to Lemma 6.5, it is easy to show that if $y \in \mathcal{B}$, then $Q^T y \in \mathcal{M}_{\mathcal{I}}$ if and only if $y \in \mathcal{B} \cap M_{\mathcal{I}}$. By assumption, $\mathcal{B} \cap M_{\mathcal{I}} = \{0\}$. Therefore, $\mathcal{M}_{\mathcal{I}} = \{0\}$, and $\mathcal{M}_{\mathcal{I}}^* = \mathbb{R}^m$, where $\mathcal{M}_{\mathcal{I}}^*$ is the dual cone of $\mathcal{M}_{\mathcal{I}}$. Using Theorem 2 in [124], we can write

$$\mathcal{M}_{\mathcal{I}}^* = \mathrm{co}\Big\{\bigcup_{j\in I\setminus S_{\mathcal{I}}} \mathcal{C}_j^*\Big\} = \mathrm{co}\{\mathcal{C}_j^* \mid j \in I \setminus S_{\mathcal{I}}\}.$$

Based on Theorem 3.2 in [9], the m-dimensional cone $\mathcal{M}_{\mathcal{I}}^*$ can be triangulated into *simplicial cones*, where each simplicial cone is defined as the convex hull of precisely m linearly independent rays $\mathcal{C}_j^*$, $j \in I \setminus S_{\mathcal{I}}$. Therefore,

$$\mathcal{M}_{\mathcal{I}}^* = \bigcup_{\mathcal{J}_m} \mathrm{co}\Big\{ \bigcup_{j \in \mathcal{J}_m} \mathcal{C}_j^* \Big\}, \tag{6.20}$$

where $\mathcal{J}_m$ is an m-subset[1] of $I \setminus S_{\mathcal{I}}$. Evidently, not all m-subsets of $I \setminus S_{\mathcal{I}}$ might consist of linearly independent rays, however, their union is equal to $\mathcal{M}_{\mathcal{I}}^*$. Consider the rays $\mathcal{C}_j^*$, $j \in \mathcal{J}_m \subset I \setminus S_{\mathcal{I}}$. Now, let us note $|\mathcal{I}| = m+1 \geq 2$. Hence, from Lemma 6.17, there exists at least one $i^\circ \in \mathcal{I}$ such that $\mathcal{C}_j^* \subset \mathcal{C}^{i^\circ *}$, $j \in \mathcal{J}_m$. Since dual cones are convex, we also have $\mathrm{co}\{\bigcup_{j \in \mathcal{J}_m} \mathcal{C}_j^*\} \subset \mathcal{C}^{i^* *}$. Now, (6.20) yields

$$\mathcal{M}_{\mathcal{I}}^* = \bigcup_{\mathcal{J}_m} \mathrm{co}\Big\{ \bigcup_{j \in \mathcal{J}_m} \mathcal{C}_j^* \Big\} \subset \bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*}.$$

Since $\mathcal{M}_{\mathcal{I}}^* = \mathbb{R}^m$, we have $\bigcup_{i=1}^{\kappa+1} \mathcal{C}^{i*} = \mathbb{R}^m$. Therefore, according to Corollary 6.9, Problem 6.4 is unsolvable, which is a contradiction. □

Based on Theorem 6.18, if Problem 6.4 is solvable, the cone condition $\mathcal{B} \cap M_{\mathcal{I}} \neq \{0\}$ must hold for any $(m+1)$-subset of $\{1, \ldots, \kappa+1\}$. The number of $(m+1)$-subsets of $\kappa+1$ elements is given by the binomial coefficient. Therefore, Theorem 6.18 presents $\binom{\kappa+1}{m+1}$ sets of necessary conditions for solvability of Problem 6.4. These conditions rely on determining an intersection of up to n half-spaces in $\mathbb{R}^m$, which was shown in [26] to be equivalent to the multi-dimensional convex hull problem. This problem can be solved in time polynomial in n. One of the algorithms was given in [126], with the worst case complexity of $O(n^{\lfloor m/2 \rfloor + 1})$.

We finish this section with an interesting special case where the familiar condition (5.6) emerges from Theorem 6.18 as both a necessary and sufficient condition.

Assumption 6.19. *Any shared vertex index $j \in I$ is shared by all $I(o_i)$, $i = 1, \ldots, \kappa+1$.*

Corollary 6.20. *Suppose Assumption 6.19 holds and $1 \leq m < \kappa+1$. Problem 6.4 is solvable if and only if* (5.6) *holds.*

[1] Since the convex hull of rays $\mathcal{C}_j^*$, $j \in I \setminus S_{\mathcal{I}}$, covers the whole space $\mathbb{R}^m$, the cardinality of the set $I \setminus S_{\mathcal{I}}$ is greater than or equal to m.

Proof. In one direction, from (5.5) we can write

$$\begin{aligned}\operatorname{cone}(\mathcal{O}_{\mathcal{S}}) &= \bigcap_{i=1}^{\kappa+1} \mathcal{C}(o_i) = \bigcap_{i=1}^{\kappa+1} \{y \mid h_j \cdot y \leq 0,\ j \in I \setminus I(o_i)\} \\ &= \{y \mid h_j \cdot y \leq 0,\ j \in [I \setminus I(o_1)] \cup \cdots \cup [I \setminus I(o_{\kappa+1})]\} \\ &= \big\{y \mid h_j \cdot y \leq 0,\ j \in I \setminus \big[I(o_1) \cap \cdots \cap I(o_{\kappa+1})\big]\big\}.\end{aligned} \tag{6.21}$$

Suppose Assumption 6.19 holds. Then for any set $\mathcal{I}$ defined as an $(m+1)$-subset of $\{1, \ldots, \kappa+1\}$, the set of non-zero shared vertices $S_{\mathcal{I}}$ can be written as $S_{\mathcal{I}} = I(o_1) \cap \cdots \cap I(o_{\kappa+1}) \setminus \{0\}$. Thus, $I \setminus \big[I(o_1) \cap \cdots \cap I(o_{\kappa+1})\big] = I \setminus S_{\mathcal{I}}$. Therefore, (6.19) and (6.21) yield $\operatorname{cone}(\mathcal{O}_{\mathcal{S}}) = M_{\mathcal{I}}$. The result follows from Theorem 6.18.

In the other direction, Problem 6.4 is trivially solvable by setting $f(o_i) = y$ for any $0 \neq y \in \mathcal{B} \cap \operatorname{cone}(\mathcal{O}_{\mathcal{S}})$. $\square$

Chapter 7

A Graph-theoretic Approach to the RCP

Going back to previous work, Theorem 3.6 established necessary and sufficient conditions for the solvability of the RCP using affine feedback. However, as discussed previously, these conditions are of limited use. Thus, the research direction presented in Chapter 5 and Chapter 6 is to modify them into a set of necessary conditions which can be utilized more easily in order to determine the solvability of the RCP on a particular system. While we obtained substantial results, methodology of Chapter 5 and Chapter 6 is tailor-made for the RCP. In this chapter we seek to reinterpret necessary and sufficient conditions of Theorem 3.6 in terms of standard positive systems theory and reconcile the RCP with previous work done in that area. In that case, the solvability of the RCP can be stated in terms of existence of nonnegative solutions of linear equations with Z-matrices. These conditions also have an appealing graph-theoretical interpretation. The results of this chapter are taken from [100].

7.1 The RCP on a Standard Orthogonal Simplex

Going back to the set-up of the RCP in Chapter 3, for the purposes of computational work it would clearly be advantageous to assume that the vertices of $\mathcal{S}$ lie on the coordinate axes. Indeed, by observing (3.1), it is clear that an affine coordinate transformation can be applied in order to transform $\mathcal{S}$ into a so-called standard orthogonal simplex. However, the full methodology for such a transformation was never presented, and for the sake of completeness, we provide it here.

Definition 7.1. *A simplex $\mathcal{S} \subset \mathbb{R}^n$ with vertices $v_0, \ldots, v_n$ is a* standard orthogonal simplex *if $v_0 = 0$*

and $v_i = e_i$ for all $i \in \{1, \ldots, n\}$. It has the following properties:

- $\mathcal{S} = \{x \in \mathbb{R}^n \mid x \geq 0, \mathbb{1}_n^T x \leq 1\}$,
- $h_0 = \mathbb{1}_n$, *and* $h_i = -e_i$ *for all* $i \in \{1, \ldots, n\}$,
- $\mathcal{F}_0 = \{x \in \mathbb{R}^n \mid x \geq 0, \mathbb{1}_n^T x = 1\}$, *and* $\mathcal{F}_i = \{x \in \mathbb{R}^n \mid x \geq 0, \mathbb{1}_n^T x \leq 1, [x]_i = 0\}$ *for all* $i \in \{1, \ldots, n\}$.

Let $\mathfrak{F}$ be a class of functions defined on $\mathbb{R}^n$. $\mathfrak{F}$ is *affine-invariant* if $f \in \mathfrak{F} \Rightarrow f \circ p \in \mathfrak{F}$ for all invertible affine functions $p : \mathbb{R}^n \to \mathbb{R}^n$.

Remark 7.2. *Clearly, affine functions, continuous functions, and piecewise affine functions are all affine-invariant.*

Lemma 7.3. *Let $\mathcal{S} \subset \mathbb{R}^n$ be a simplex, and let $\Delta \subset \mathbb{R}^n$ be the standard orthogonal simplex. Let $\mathfrak{F}$ be an affine-invariant class of feedback controls $u : \mathbb{R}^n \to \mathbb{R}^m$. Let $p : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible affine map $p(x) = Kx + g$ with $K(e_i) = v_i$ for all $i \in \{0, \ldots, n\}$, and $g = v_0$. Then, RCP is solvable for $(A, B, a, \mathcal{S})$ by $\mathfrak{F}$ if and only if it is solvable for $(K^{-1}AK, K^{-1}B, K^{-1}a + K^{-1}Ag, \Delta)$.*

Proof. We note that p maps each vertex of Δ into a corresponding vertex of $\mathcal{S}$. It also maps Δ into $\mathcal{S}$ and the exit facet $\mathcal{F}_0^\Delta$ of Δ, into the exit facet $\mathcal{F}_0^\mathcal{S}$ of $\mathcal{S}$.

Let u be any feedback in $\mathfrak{F}$, and let $\phi_u(\cdot, x_0)$ be the trajectory generated by system (3.1), with the initial condition $\phi_u(0, x_0) = x_0$. Consider the feedback $u' := u \circ p$. We first note that $u' \in \mathfrak{F}$ by the definition of affine-invariance. Now, consider the system

$$\dot{y} = K^{-1}AKy + K^{-1}Bu'(y) + K^{-1}a + K^{-1}Ag. \tag{7.1}$$

Let $\phi'_{u'}(\cdot, y_0)$ be the trajectory generated by (7.1), with $\phi'_{u'}(0, y_0) = y_0$. We claim that

$$p \circ \phi'_{u'}(\cdot, y_0) = \phi_u(\cdot, p(y_0)). \tag{7.2}$$

This is easily shown: $d(p \circ \phi'_{u'}(t, y_0))/dt = Kd(\phi'_{u'}(t, y_0))/dt = AK\phi'_{u'}(t, y_0) + Bu'(\phi'_{u'}(t, y_0)) + a + Ag = A(p \circ \phi'_{u'}(t, y_0)) + Bu(p \circ \phi'_{u'}(t, y_0)) + a$. Thus, $p \circ \phi'_{u'}(\cdot, y_0)$ satisfies (3.1), and $p \circ \phi'_{u'}(0, y_0) = p(y_0)$. Hence, $p \circ \phi_{u'}(\cdot, y_0) = \phi_u(\cdot, p(y_0))$.

Now, assume that the RCP is solvable for $(A, B, a, \mathcal{S})$ by $\mathfrak{F}$. Let $u \in \mathfrak{F}$ be the feedback that solves this RCP. We claim that $u' = u \circ p$ solves the RCP for $(K^{-1}AK, K^{-1}B, K^{-1}a + K^{-1}Ag, \Delta)$. Let $y_0 \in \Delta$. We verify conditions (i)-(iii) from the definition of the RCP:

(i) Since p maps Δ to $\mathcal{S}$, $\phi'_{u'}(t, y_0) \in \Delta$ for all $t \in [0, T]$ is equivalent to $p \circ \phi'_{u'}(t, y_0) \in \mathcal{S}$ for all $t \in [0, T]$. By (7.2), this is equivalent to $\phi_u(t, p(y_0)) \in \mathcal{S}$ for all $t \in [0, T]$, with $p(y_0) \in \mathcal{S}$. This is the condition (i) applied to the RCP on $(A, B, a, \mathcal{S})$, which holds by assumption.

(ii) Since p maps $\mathcal{F}_0^\Delta$ to $\mathcal{F}_0^\mathcal{S}$, $\phi'_{u'}(T, y_0) \in \mathcal{F}_0^\Delta$ if and only if $p \circ \phi'_{u'}(T, y_0) \in \mathcal{F}_0^\mathcal{S}$. This holds by (7.2) and condition (ii) applied to the RCP on $(A, B, a, \mathcal{S})$.

(iii) Since p maps $\mathbb{R}^n \backslash \Delta$ to $\mathbb{R}^n \backslash \mathcal{S}$, $\phi'_{u'}(t, y_0) \notin \Delta$ is equivalent to $\phi_u(t, p(y_0)) \notin \mathcal{S}$, by the same discussion as in (i). Hence, (iii) follows from the condition (iii) in the RCP applied to $(A, B, a, \mathcal{S})$.

Hence, feedback control $u' \in \mathfrak{F}$ solves the RCP for $(K^{-1}AK, K^{-1}B, K^{-1}a + K^{-1}Ag, \Delta)$. That is, the RCP is solvable for $(K^{-1}AK, K^{-1}B, K^{-1}a + K^{-1}Ag, \Delta)$ by $\mathfrak{F}$. The other direction is entirely analogous. □

Lemma 7.3 ensures that we can always assume without loss of generality that $\mathcal{S}$ is the standard orthogonal simplex, and discuss the solvability of the RCP in that setting. Hence, in the remainder of this chapter it is assumed that $\mathcal{S}$ is the standard orthogonal simplex. This is a first step towards positioning the RCP in terms of theory of positive systems. The same assumption will be made in Chapter 8 as well.

We define

$$M := -(A + a\mathbb{1}_n^T). \tag{7.3}$$

The following lemma is the key in all the work presented further in this chapter.

Lemma 7.4. *Let $\mathcal{S}$ be the standard orthogonal simplex. Then, $Av_i + a \in \mathcal{C}(v_i)$, $i \in \{0, \ldots, n\}$, if and only if the following two conditions hold:*

(i) M is a Z-matrix,

(ii) $a \geq 0$.

Proof. We note that, since $\mathcal{S}$ is the standard orthogonal simplex,

$$h_j \cdot (Av_i + a) = -([A]_{ji} + [a]_j) = [M]_{ji}$$

for all $i, j \in \{1, \ldots, n\}$, $i \neq j$. Hence, for any $i \in \{1, \ldots, n\}$, the invariance condition $Av_i + a \in \mathcal{C}(v_i)$ holds if and only if $[M]_{ji} \leq 0$ for all $j \neq i$. By going through all $i \in \{1, \ldots, n\}$, we obtain the definition of a Z-matrix.

For (ii), we notice that the invariance condition $Av_0 + a \in \mathcal{C}(v_0)$ holds if and only if $h_j \cdot (Av_0 + a) = -e_j \cdot a = -[a]_j \leq 0$ for all $j \in I$. That is, if and only if $a \geq 0$. □

We note that conditions (i) and (ii) of Lemma 7.4 are standard conditions for positive invariance of an affine dynamical system $\dot{x} = Ax + a$ [1]. Analogously, the invariance conditions (3.4) can just be interpreted as conditions for positive invariance of an affine control system (3.1).

7.2 Main Results

Similarly to $\mathcal{O}_\mathcal{S}$ and $\mathcal{O}$ defined in (5.1) and (5.2), respectively, we define the set of *open-loop equilibria* of system (3.1) by

$$\mathcal{E}_\mathcal{S} = \{x \in \mathcal{S} \mid Ax + a = 0\} \tag{7.4}$$

and

$$\mathcal{E} = \{x \in \mathbb{R}^n \mid Ax + a = 0\}. \tag{7.5}$$

Solvability of the RCP by affine feedback is intimately connected with the geometry of the set $\mathcal{E}_\mathcal{S}$. Suppose that the RCP is solvable by an affine feedback $u(x) = Kx + g$. By Proposition 3.3, $\{x \in \mathcal{S} \mid (A + BK)x + (a + Bg) = 0\} = \emptyset$. However, this set is exactly $\mathcal{E}_\mathcal{S}$ for a new affine system

$$\dot{x} = \tilde{A}x + \tilde{a}, \tag{7.6}$$

with $\tilde{A} = A + BK$, $\tilde{a} = a + Bg$.

In order to examine the set $\mathcal{E}_\mathcal{S}$, we partition it into disjoint sets $\mathcal{E}_\mathcal{S} \cap \mathcal{F}_0$ and $\mathcal{E}_\mathcal{S} \backslash \mathcal{F}_0$. This has two benefits: first, we will show that it is possible to easily obtain necessary and sufficient graph-theoretical conditions for $\mathcal{E}_\mathcal{S} \cap \mathcal{F}_0 = \emptyset$ and $\mathcal{E}_\mathcal{S} \backslash \mathcal{F}_0 = \emptyset$. Secondly, this partition naturally follows from the previous work on the RCP. As briefly mentioned in Chapter 5, because the choice of a triangulation of the state space into simplices is left to the designer, the designer can choose this triangulation so that the set of potential equilibria $\mathcal{O}_\mathcal{S}$ from (5.1) lies in the exit facet $\mathcal{F}_0$, or that it does not intersect $\mathcal{F}_0$. Since $\mathcal{E}_\mathcal{S} \subset \mathcal{O}_\mathcal{S}$, choosing $\mathcal{O}_\mathcal{S} \subset \mathcal{F}_0$ automatically ensures that $\mathcal{E}_\mathcal{S} \backslash \mathcal{F}_0 \subset \mathcal{O}_\mathcal{S} \backslash \mathcal{F}_0 = \emptyset$. On the other hand, choosing $\mathcal{O}_\mathcal{S} \subset \mathcal{S} \backslash \mathcal{F}_0$ would ensure that $\mathcal{E}_\mathcal{S} \cap \mathcal{F}_0 = \emptyset$. We discuss the results obtained by these two triangulations in Section 7.3.

We now examine the sets $\mathcal{E}_\mathcal{S} \cap \mathcal{F}_0$ and $\mathcal{E}_\mathcal{S} \backslash \mathcal{F}_0$.

Proposition 7.5. *Let $\mathcal{S}$ be the standard orthogonal simplex. Let M be as in* (7.3)*. There exists $x \in \mathcal{E}_\mathcal{S} \cap \mathcal{F}_0$ if and only if the linear equation $My = 0$ has a solution $y > 0$.*

Proof. Assume that there exists $x \in \mathcal{E}_{\mathcal{S}} \cap \mathcal{F}_0$. Then, by (7.4), $Ax + a = 0$. Also, by Definition 7.1, $x > 0$ and $\mathbb{1}_n^T x = 1$. Thus, $Ax + a(\mathbb{1}_n^T x) = 0$, i.e., $Mx = 0$.

In the other direction, assume that $My = 0$ has a solution $y > 0$. Define $x = y/(\mathbb{1}_n^T y)$. We note that $x > 0$, and $\mathbb{1}_n^T x = \mathbb{1}_n^T y/(\mathbb{1}_n^T y) = 1$. Hence, by Definition 7.1, $x \in \mathcal{F}_0$. Additionally, $Mx = 0$, i.e., $(A + a\mathbb{1}_n^T)x = Ax + a = 0$. Thus, $x \in \mathcal{E}_{\mathcal{S}}$. □

Proposition 7.6. *Let $\mathcal{S}$ be the standard orthogonal simplex. Then, there exists $x \in \mathcal{E}_{\mathcal{S}} \backslash \mathcal{F}_0$ if and only if the linear equation $My = a$ has a solution $y \geq 0$.*

Proof. Assume that there exists $x \in \mathcal{E}_{\mathcal{S}} \backslash \mathcal{F}_0$. Then, by (7.4), $Ax + a = 0$. Hence, $Ax + a\mathbb{1}_n^T x = a(\mathbb{1}_n^T x - 1)$, i.e., $Mx = a(1 - \mathbb{1}_n^T x)$. Also, by Definition 7.1, $x \geq 0$ and $\mathbb{1}_n^T x < 1$. Thus, if we define $y = x/(1 - \mathbb{1}_n^T x)$, we obtain that $y \geq 0$ is a solution to $My = a$.

In the other direction, assume that $y \geq 0$ is a solution to $My = a$. Define $x = y/(1 + \mathbb{1}_n^T y)$. We note that $x \geq 0$ and

$$\mathbb{1}_n^T x = (\mathbb{1}_n^T y)/(1 + \mathbb{1}_n^T y) < 1. \tag{7.7}$$

Hence, $x \in \mathcal{S} \backslash \mathcal{F}_0$. Additionally, $Mx = My/(1 + \mathbb{1}_n^T y) = a/(1 + \mathbb{1}_n^T y)$. Thus, $Ax + a\mathbb{1}_n^T x = -a/(1 + \mathbb{1}_n^T y)$, i.e., $Ax + a = a(1 - \mathbb{1}_n^T x - 1/(1 + \mathbb{1}_n^T y))$. From (7.7) it follows that $Ax + a = 0$, i.e., $x \in \mathcal{E}$. We are done. □

At this point, we can provide a general characterization of solvability of the RCP for the case of affine feedback.

Theorem 7.7. *Let $\mathcal{S}$ be the standard orthogonal simplex. Then, RCP is solvable by affine feedback if and only if there exist $K \in \mathbb{R}^{m \times n}, g \in \mathbb{R}^m$ such that the following conditions are satisfied:*

(i) $-((A + BK) + (a + Bg)\mathbb{1}_n^T)$ *is a Z-matrix,*

(ii) $a + Bg > 0$,

(iii) neither of the following equations:

(iii.a) $-((A + BK) + (a + Bg)\mathbb{1}_n^T)y = 0$,

(iii.b) $-((A + BK) + (a + Bg)\mathbb{1}_n^T)y = a + Bg$

admits a solution $y > 0$.

If the above conditions are satisfied for some K, g, then $u = Kx + g$ solves the RCP.

Proof. Assume that the RCP is solvable by affine feedback. Let $u = Kx + g$ solve the RCP. Then, (i) holds by Lemma 7.4 and the discussion at the beginning of this section. Also by Lemma 7.4, $a + Bg \geq 0$. However, $a + Bg \neq 0$, because $a + Bg = 0$ would imply that $v_0 = 0$ is an equilibrium of (3.1), which is prohibited by Proposition 3.3. Hence, (ii) holds as well. Finally, if the RCP is solvable by affine feedback, then by Proposition 3.3 it neither has an equilibrium in $\mathcal{F}_0$ nor in $\mathcal{S}\backslash\mathcal{F}_0$. Hence, by Proposition 7.5, equation (iii.a) admits no solutions $y > 0$, and by Proposition 7.6, (iii.b) admits no solutions $y \geq 0$.

In the other direction, assume that (i), (ii) and (iii) hold. Then, from (i), (ii) and Lemma 7.4, we obtain that the invariance conditions (3.4) hold for $u = Kx+g$. Additionally, from (iii.a) and Proposition 7.5, we obtain that the set of equilibria of (3.1) contained in $\mathcal{F}_0$ is empty. We note that (ii) and (iii.b) imply that the equation in (iii.b) also admits no solutions $y \geq 0$. Thus, by Proposition 7.5, system (3.1) has no equilibria in $\mathcal{S}\backslash\mathcal{F}_0$. Hence, as noted in Theorem 3.6, the RCP is solvable by affine feedback u. □

Finally, let us dig into the graph-theoretic conditions for solvability of the RCP implied by Proposition 7.5 and Proposition 7.6. Let us permute the rows and columns of matrix M from (7.3) so that M becomes a block lower triangular matrix in a Frobenius normal form with blocks M_{ij}, $i, j \in \{1, \ldots, p\}$. Hence, from now on we assume that $M_{ij} = 0$ for $i < j$, and each block M_{ii} is either irreducible or a 1×1 block with $M_{ii} = 0$. The components of a are permuted according to the same permutation, with blocks $a_1, \ldots, a_p$. This process simply corresponds to the relabeling of vertices $v_1, \ldots, v_n$ of $\mathcal{S}$.

In order to connect the results of Proposition 7.5 and Proposition 7.6 with results on graph theory from Section 2.9, we need to ensure that M is a Z-matrix. As we saw previously, this property follows naturally from invariance conditions. If conditions (3.4) are solvable, then one can always pre-apply a control $u(x) = K'x + g'$ such that (3.4) holds, and then define $\tilde{A} = A + BK'$ and $\tilde{a} = a + Bg'$. Then, $\tilde{A}v_i + \tilde{a} \in \mathcal{C}(v_i)$ for all $i \in \{0, \ldots, n\}$. If conditions (3.4) are not solvable, then the RCP will not be solvable at all, as discussed in Section 3.2. Hence, by abusing notation and removing the tilde's from $\tilde{A}$ and $\tilde{a}$, we can without loss of generality make the following assumption:

Assumption 7.8. *Let $Av_i + a \in \mathcal{C}(v_i)$ for all $i \in \{0, \ldots, n\}$.*

With this assumption, by Lemma 7.4, M is a Z-matrix.

The following theorem is an immediate consequence of previous work, Proposition 2.40 and Proposition 2.41:

Theorem 7.9. *Assume that $\mathcal{S}$ is the standard orthogonal simplex, and suppose Assumption 7.8 holds. Then, $\mathcal{E}_{\mathcal{S}} = \emptyset$ if and only if both following statements hold:*

(i) $S \subset \text{above}(T)$,

(ii) $\text{supp}(a) \cap \text{above}(T) \neq \emptyset$.

Proof. We note that, by Lemma 7.4, M is a Z-matrix and $a \geq 0$. Also, $\mathcal{E}_{\mathcal{S}} = \emptyset$ if and only if $\mathcal{E}_{\mathcal{S}} \cap \mathcal{F}_0 = \emptyset$ and $\mathcal{E}_{\mathcal{S}} \backslash \mathcal{F}_0 = \emptyset$.

First consider $\mathcal{E}_{\mathcal{S}} \cap \mathcal{F}_0$. By Proposition 7.5, $\mathcal{E}_{\mathcal{S}} \cap \mathcal{F}_0 = \emptyset$ if and only if there does not exist $y > 0$ such that $My = 0$. By Proposition 2.40, this is equivalent to $S \backslash \text{above}(T) = \emptyset$, i.e., $S \subset \text{above}(T)$. This gives (i).

Next consider $\mathcal{E}_{\mathcal{S}} \backslash \mathcal{F}_0$. By Proposition 7.6, $\mathcal{E}_{\mathcal{S}} \backslash \mathcal{F}_0 = \emptyset$ if and only if there does not exist $y \geq 0$ such that $My = a$. By Proposition 2.41, this is equivalent to $\text{supp}(a) \cap \text{above}(S \cup T) \neq \emptyset$. However, using (i), if $S \subset \text{above}(T)$, then $\text{above}(S \cup T) = \text{above}(S) \cup \text{above}(T) = \text{above}(T)$. This gives (ii). □

Theorem 7.9 can also be rephrased in terms of solvability of the RCP, as follows:

Corollary 7.10. *Assume that $\mathcal{S}$ is the standard orthogonal simplex. Then, the RCP is solvable by affine feedback if and only if there exist $K \in \mathbb{R}^{m \times n}$ and $g \in \mathbb{R}^n$ such that:*

(i) $\tilde{A} = -(A + BK) - (a + Bg)\mathbb{1}_n^T$ *is a Z-matrix,*

(ii) $\tilde{a} = a + Bg \geq 0$,

(iii) the conditions of Theorem 7.9 are satisfied, with $\mathcal{E}_{\mathcal{S}}$ being defined with respect to $\tilde{A}$ and $\tilde{a}$.

Proof. Analogously to the discussion at the beginning of Section 7.2, let $\tilde{A} = A + BK$ and $\tilde{a} = a + Bg$. Then, the RCP is solvable for system (3.1) by affine feedback $u = Kx + g$ if and only if it is solvable for (7.6). By Lemma 7.4, the invariance conditions (3.4) for that system are equivalent to (i) and (ii). Lack of equilibria in system (7.6) is characterized by the conditions of Theorem 7.9, i.e., by (iii). Thus, system (7.6) will satisfy the invariance conditions and will not contain any equilibria in $\mathcal{S}$ if and only if conditions (i)-(iii) hold. By Theorem 3.6, this is equivalent to the solvability of the RCP for system (7.6). □

7.3 Necessary and Sufficient Conditions for Special Triangulations

In the remainder of this chapter, we give some results on solvability of the RCP using affine feedback for two particular cases of the geometric structure of $\mathcal{O}_{\mathcal{S}}$. As previously mentioned, since the choice of triangulation in reach control theory is left to the designer, the triangulation can always be performed in such a way that $\mathcal{O}_{\mathcal{S}}$ is at the "bottom" of the simplex, i.e., $\mathcal{O}_{\mathcal{S}} \subset \mathcal{F}_0$, or that $\mathcal{O}_{\mathcal{S}}$ does not touch $\mathcal{F}_0$. The former case was investigated in, e.g., [55, 99], while the latter was explored in [22, 24].

7.3.1 First Triangulation

In this section, we assume $\mathcal{O}_{\mathcal{S}} \cap \mathcal{F}_0 = \emptyset$. Here we give yet another set of necessary and sufficient conditions for the set $\mathcal{E}_{\mathcal{S}}$ to be empty.

The following result is a consequence of Proposition 2.41 and Proposition 7.6.

Proposition 7.11. *Assume that $\mathcal{S}$ is the standard orthogonal simplex with $\mathcal{O}_{\mathcal{S}} \cap \mathcal{F}_0 = \emptyset$. Additionally, suppose that Assumption 1 holds and that $a \gg 0$. Then, $\mathcal{E}_{\mathcal{S}} \neq \emptyset$ if and only if M is a nonsingular M-matrix.*

Proof. Since $a \gg 0$, by Proposition 2.41 and Proposition 7.6 we know that $\mathcal{E}_{\mathcal{S}} \neq \emptyset$ if and only if above$(S \cup T) = \emptyset$. By (2.5), this is equivalent to $S \cup T = \emptyset$. Now, first assume that $S \cup T = \emptyset$. Consider any $i \in \{1, \ldots, p\}$. Since $i \notin S \cup T$, matrix M_{ii} satisfies $l(M_{ii}) > 0$ by (2.6). We recount that M is a block triangular matrix with blocks M_{ii}, $i \in \{1, \ldots, p\}$, on the diagonal. Thus, the eigenvalues of M are the union of eigenvalues of all M_{ii}. Hence, $l(M) > 0$, i.e., M is a non-singular M-matrix.

In the other direction, if M is a non-singular M-matrix, $l(M) > 0$. Thus, since M is a block triangular matrix, $l(M_{ii}) > 0$ for all $i \in \{1, \ldots, p\}$. Thus, by (2.6), $S \cup T = \emptyset$, which means that $\mathcal{E}_{\mathcal{S}} \neq \emptyset$. □

Remark 7.12. *Proposition 7.11 can also be proved more directly from Proposition 7.6, by invoking properties (I_{28}) and (N_{39}) of non-singular M-matrices from [11]. However, we chose to present a proof which uses Proposition 2.41 as it once again illustrates the graph-theoretic nature of M-matrices.*

7.3.2 Second Triangulation

In this section, we assume $\mathcal{O}_{\mathcal{S}} \subset \mathcal{F}_0$.

The primary contribution of this section is an easily checkable strong sufficient condition for solvability of the RCP. Let M be given by (7.3). Furthermore, let $M + \mathcal{B} \subset \mathbb{R}^{n \times n}$ denote the set of matrices $M + X$ where all the columns of X are in $\mathcal{B}$. Equivalently,

$$M + \mathcal{B} = \{M + BK \mid K \in \mathbb{R}^{m \times n}\}. \tag{7.8}$$

Theorem 7.13. *Assume that $\mathcal{S}$ is the standard orthogonal simplex, $a > 0$, and $\mathcal{O}_{\mathcal{S}} \subset \mathcal{F}_0$. Let $\mathcal{D} = \mathcal{Z}_n \cap (M + \mathcal{B})$. Suppose that there exists a non-singular matrix $D \in \mathcal{D}$. Then, the RCP is solvable by affine feedback.*

Proof. Let $D \in \mathcal{D}$. Then, $D = M - BK$ for some matrix $K \in \mathbb{R}^{m \times n}$. We observe system (3.1), and claim that the closed-loop feedback $u(x) = Kx + g$ with $g = 0$ solves the RCP. We will do that by

showing that the conditions of Theorem 7.7 are satisfied.

First, we note that $-((A+BK)+(a+Bg)\mathbb{1}_n^T) = -(A+a\mathbb{1}_n^T) - BK = M - BK = D \in \mathcal{Z}_n$. Thus, condition (i) of Theorem 7.7 is satisfied. Since $a+Bg = a > 0$, condition (ii) is satisfied as well. Since D is non-singular, condition (iii.a) is satisfied. Finally, let us consider condition (iii.b). By Proposition 7.6 and the discussion in the proof of Theorem 7.7, this condition is equivalent to (3.1) not containing an equilibrium in $\mathcal{S}\backslash\mathcal{F}_0$. However, all possible equilibria of (3.1) are contained in $\mathcal{O}_\mathcal{S}$ [125], and $\mathcal{O}_\mathcal{S} \subset \mathcal{F}_0$. Hence, condition (iii.b) is automatically verified. Thus, $u(x)$ indeed solves the RCP. □

We note that the set $\mathcal{D}$ from Theorem 7.13 is either empty or a polyhedron, as $\mathcal{Z}_n$ is defined by linear inequalities, and $M + \mathcal{B} \subset \mathbb{R}^{n\times n}$ is an affine space. Hence, it is easy to compute $\mathcal{D}$. Additionally, verifying that there exists a non-singular matrix on a polyhedron is equivalent to verifying the existence of a non-singular matrix in an affine space generated by that polyhedron. This is shown in the following corollary:

Corollary 7.14. *Let $\mathcal{D} \neq \emptyset$. Then, there exists a non-singular matrix $D \in \mathcal{D}$ if and only if there exists a non-singular matrix $D' \in \mathrm{aff}(\mathcal{D})$.*

Proof. One direction is obvious. In the other direction, let $D' \in \mathrm{aff}(\mathcal{D})$ be non-singular, and assume $\det(D) = 0$ for all $D \in \mathcal{D}$. Take any $D_0 \in \mathrm{int}(\mathcal{D})$. Then, $\mathrm{aff}(\mathcal{D}) = \{D_0+\alpha_1 D_1+\ldots+\alpha_k D_k \mid \alpha_1,\ldots,\alpha_k \in \mathbb{R}\}$, where $D_1,\ldots,D_k$ are some matrices, not necessarily in $\mathcal{D}$. Now, $p(\alpha_1,\ldots,\alpha_k) = \det(D_0 + \alpha_1 D_1 + \ldots + \alpha_k D_k)$ is a multivariate polynomial in $\mathbb{R}[\alpha_1,\ldots,\alpha_k]$, and because $\det(D') \neq 0$, it is not always 0. However, $p(\alpha_1,\ldots,\alpha_k)$ is always zero for some small ball $\mathbb{B}^k(0,\epsilon)$, because $D_0 \in \mathrm{int}(\mathcal{D})$, so $D_0 + \alpha_1 D_1 + \ldots + \alpha_k D_k \subset \mathcal{D}$ for small α_i. Thus, the Taylor expansion of p around $(0,0,\ldots,0)$ is zero. Since the Taylor expansion of p is p itself, all coefficients of p are 0. Hence, $\det(D') = 0$, which is a contradiction. □

By combining Theorem 7.13 and Corollary 7.14 we obtain the following:

Corollary 7.15. *Assume that $\mathcal{S}$ is the standard orthogonal simplex, $a > 0$, and $\mathcal{O}_\mathcal{S} \subset \mathcal{F}_0$. Let $D_0 \in \mathcal{D}$, and let $D_1,\ldots,D_k \in \mathbb{R}^{n\times n}$ be the basis elements of the vector space $\{D' - D_0 \mid D' \in \mathrm{aff}(\mathcal{D})\}$. Then, if $\det(D_0 + \alpha_1 D_1 + \ldots + \alpha_k D_k)$ is not a zero polynomial in $\alpha_1,\ldots,\alpha_k$, the RCP is solvable.*

Proof. Assume that $\det(D_0+\alpha_1 D_1+\ldots+\alpha_k D_k)$ is not a zero polynomial. Then, there exist $\alpha'_1,\ldots,\alpha'_k \in \mathbb{R}^n$ such that $\det(D_0 + \alpha'_1 D_1 + \ldots + \alpha'_k D_k) \neq 0$. Let $D' = D_0 + \alpha'_1 D_1 + \ldots + \alpha'_k D_k \in \mathrm{aff}(\mathcal{D})$. Then, $\det(D') \neq 0$. By Corollary 7.14, there exists $D \in \mathcal{D}$ such that $\det(D) \neq 0$. By Theorem 7.13, the RCP is solvable. □

Corollary 7.15 shows that the sufficient condition for the solvability of the RCP in Theorem 7.13 can be verified by just checking whether all coefficients of a certain easily computable polynomial are 0. Thus, it gives an easily computable sufficient condition for solvability of the RCP. It is a strong condition: assume that the RCP is solvable by affine feedback. Let $u(x) = Kx + g$ be the affine feedback that solves the RCP. Then, by Theorem 7.7,

$$-((A + BK) + (a + Bg)\mathbb{1}_n^T) = M - B(K + g\mathbb{1}_n^T) \in \mathcal{Z}_n.$$

Additionally, $M - B(K + g\mathbb{1}_n^T) \in M + \mathcal{B}$ by (7.8). Hence, $\mathcal{D} = \mathcal{Z}_n \cap (M + \mathcal{B}) \neq \emptyset$. Then, either the condition of Theorem 7.13 is satisfied, or the entire polyhedron $\mathcal{D}$ consists solely of singular matrices.

Remark 7.16. *This chapter introduces new mathematical tools for formulating and solving the RCP. It exposes a previously unknown correspondence between reach control theory and the theory of positive systems. Nevertheless, some mathematical constructs that are ubiquitous for positive systems have also been previously applied in the RCP under the conditions of Assumption 7.8. Particularly, Z-matrices and M-matrices were used in [5, 23, 24] to develop the reach control indices, and syntheses of time-varying affine and piecewise affine feedbacks. The reach control indices are intrinsically tied to certain nonsingular M-matrices, but their development is rather involved and still seems quite remote from the present work.*

While it is outside the scope of presented work to delve into the machinery of the reach control indices, as a final remark, we point out that the methods introduced in this chapter shed new light on the indices, thus lending further evidence that our results here provide a useful new perspective on reach control theory. It can be shown that, supposing that Assumption 7.8 holds, and under the additional assumptions used by the theory of reach control indices, the Frobenius form of M is given in block-form by

$$\left[\begin{array}{c|cccc} M_0 & 0 & 0 & \cdots & 0 \\ \hline * & M_{11} & 0 & \cdots & 0 \\ * & 0 & M_{22} & \cdots & 0 \\ * & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & 0 & 0 & \cdots & M_{pp} \end{array}\right].$$

In the above matrix, M_0 is an arbitrary lower triangular matrix in the Frobenius normal form. It can be shown that each M_{kk}, $k = 1, \ldots, p$ are $r_k \times r_k$ matrices, where $\{r_1, \ldots, r_p\}$ are the reach control indices. Moreover, each M_{kk}, $k = 1, \ldots, p$ is a singular, irreducible M-matrix.

Chapter 8

Geometric Structure of Open-Loop Equilibria

In this chapter, we take yet another approach to the solvability of the RCP. Inspired by [125], we note that one can interpret the RCP as the application of feedback control to push closed-loop equilibria off the simplex $\mathcal{S}$. In an effort to characterize when the system is sufficiently actuated to do so, notions such as reach control indices and reach controllability were formulated in [24] and [125], respectively. In particular, reach controllability was designed to describe how an *infinitesimal* control actuation serves to push the closed-loop equilibria out of $\mathcal{S}$. In that sense, it acts as an analogue to the standard notion of local controllability, which uses infinitesimal control actuation to move the state within the neighbourhood of the starting point. This chapter builds on the work of [125]. We explore the geometric structure of open-loop equilibria by a series of algebraic manipulations based on the invariance conditions in the RCP. We also use the same approach to improve previous results on reach controllability. We note that, with minor modifications, this chapter is taken from [104].

8.1 Assumptions and Preliminaries

In this chapter, we make the following assumptions. They are motivated by the work in [125].

Assumption 8.1.

(A1) $\mathcal{E}_{\mathcal{S}} = \mathrm{co}\{\varepsilon_1, \cdots, \varepsilon_{\kappa_0+1}\}$, *a* κ_0*-dimensional simplex with* $0 \leq \kappa_0 \leq \kappa$,

(A2) $\mathcal{O} \cap \mathrm{int}(\mathcal{S}) \neq \emptyset$,

(A3) $\mathcal{O} \cap \mathcal{F}_0 = \emptyset$,

(N1) $Av_i + a \in \mathcal{C}(v_i)\,, \quad i = 0, \ldots, n$,

(N2) $\mathcal{B} \cap \mathrm{cone}(\mathcal{O}_\mathcal{S}) \neq \{0\}$.

Assumptions (A2) and (A3) allow $\mathcal{O}$ to intersect the interior of simplices. Arguments motivating this choice of triangulation are found in [125], but we remove the restriction of [125] that $\mathcal{O}_\mathcal{S}$ is itself a simplex. We use a different numbering for assumptions (N1) and (N2) as these arise from necessary conditions. Proposition 3.5 shows that a necessary condition for solvability of the RCP by an affine feedback $u = Kx + g$ is that the invariance conditions (3.4) hold. (N1) is the same as Assumption 7.8 in Chapter 7. As mentioned in previous chapters, assumption (N2) was shown in [125] to be a necessary condition for solvability of the RCP by continuous state feedback in the case of single-input systems. Accordingly, we will only be making use of this assumption when discussing single-input systems.

The mathematical machinery laid out in Lemmas 8.2-8.6 to derive the main arguments on characterization of equilibria is based on manipulating index sets $I(o_k)$ in order to book-keep the constraints arising from (N1). This machinery enables us to relate the combinatorial properties of index sets to geometric properties of polytopes and cones.

Lemma 8.2. *Let* $I(o_1) \cap \cdots \cap I(o_{\kappa+1}) = \{0, 1, 2, \ldots, l\}$. *Then,* $\mathrm{cone}(\mathcal{O}_\mathcal{S}) = \{y \mid h_j \cdot y \leq 0$ *for all* $j = l+1, \ldots, n\}$.

Proof. Follows trivially by definitions (3.3) and (5.5). □

Lemma 8.3. *Let* $\mathcal{S} = \mathrm{co}\{v_0, \ldots, v_n\}$ *be an n-dimensional simplex, and let* $\mathcal{P} = \mathrm{co}\{w_1, \ldots, w_k\} \subset \mathcal{S}$. *Let* $q \leq n$ *be such that*

$$\bigcup_{i=1}^{k} I(w_i) = \{0, \ldots, q\}.$$

Define $\mathcal{S}' = \mathrm{co}\{v_0, \ldots, v_q\}$. *Then, there exists* $x \in \mathcal{P} \cap \mathrm{int}(\mathcal{S}')$.

Proof. We note that $x = (w_1 + \ldots + w_k)/n$ satisfies the conditions of the problem. □

We now give a proof of three fundamental claims which we use as building blocks for our main results. One of the key properties is the following generalization of Lemma 2 of [125].

Lemma 8.4. *Let* $\mathcal{S} = \mathrm{co}\{v_0, \ldots, v_n\}$ *be an n-dimensional simplex. Let* $\mathcal{S}'$ *be a q-dimensional face of* $\mathcal{S}$, *with* $q \leq n$. *Let* $\mathcal{P} = \mathrm{co}\{w_1, \ldots, w_{p+1}\}$ *be a p-dimensional simplex. Suppose* $\mathcal{P} \subset \mathcal{S}'$, $\mathcal{P} \cap \mathrm{int}(\mathcal{S}') \neq \emptyset$ *and* $\partial\mathcal{P} \subset \partial\mathcal{S}'$. *Then each index set* $I(w_k)$, $k \in \{1, \ldots, p+1\}$, *has an exclusive member. That is, there exists* $i_k \in I(w_k)$ *such that* $i_k \notin I(w_j)$ *for all* $j \in \{1, \ldots, p+1\} \setminus \{k\}$.

Proof. Without loss of generality, assume the vertices of $\mathcal{S}'$ are $v_0, \ldots, v_q$. By the assumption that $\mathcal{P} \cap \text{int}(\mathcal{S}') \neq \emptyset$, we have $\cup_{i=1}^{p+1} I(w_i) = \{0, \ldots, q\}$. If $p = 0$ we are done. Instead suppose without loss of generality that $I(w_1) \subset \cup_{j=2}^{p+1} I(w_j)$. Thus, $\cup_{j=2}^{p+1} I(w_j) = \{0, \ldots, q\}$. Define $\mathcal{P}' = \text{co}\{w_2, \ldots, w_{p+1}\}$. Since $\mathcal{P}$ is a simplex, $\mathcal{P}'$ is a $(p-1)$-dimensional facet of $\mathcal{P}$ so $\mathcal{P}' \subset \partial\mathcal{P}$. However, $\cup_{j=2}^{p+1} I(w_j) = \{0, \ldots, q\}$ implies $\mathcal{P}' \cap \text{int}(\mathcal{S}') \neq \emptyset$. This contradicts that $\partial\mathcal{P} \subset \partial\mathcal{S}'$. □

The following lemma relates the dimension of $\mathcal{E}_\mathcal{S}$ with the dimension of the affine space generated by the vertices v_i present in $I(\varepsilon_k)$.

Lemma 8.5. *Let $\mathcal{E}_\mathcal{S} = \text{co}\{\varepsilon_1, \cdots, \varepsilon_{\kappa_0+1}\}$ be a simplex with $\dim(\mathcal{E}_\mathcal{S}) = \kappa_0$. Let $q \leq n$ be such that $\cup_{k=1}^{\kappa_0} I(\varepsilon_k) = \{0, 1, \ldots, q\}$. Then, $\dim(\mathcal{E} \cap \text{aff}\{v_0, \ldots, v_q\}) = \kappa_0$.*

Proof. Let $\mathcal{S}' = \text{co}\{v_0, \ldots, v_q\}$, and define $\mathcal{E}'_\mathcal{S} = \mathcal{E} \cap \mathcal{S}'$. Since $\mathcal{E}_\mathcal{S} \subset \mathcal{S}'$ and $\mathcal{E}_\mathcal{S} \subset \mathcal{E}$, we note $\mathcal{E}_\mathcal{S} \subset \mathcal{E}'_\mathcal{S}$. On the other hand, $\mathcal{E}'_\mathcal{S} = \mathcal{E} \cap \mathcal{S}' \subset \mathcal{E} \cap \mathcal{S} = \mathcal{E}_\mathcal{S}$. Hence, $\mathcal{E}'_\mathcal{S} = \mathcal{E}_\mathcal{S}$ and thus, $\dim \mathcal{E}'_\mathcal{S} = \dim \mathcal{E}_\mathcal{S} = \kappa_0$.

Now, since $\mathcal{S}' \subset \text{aff}(\mathcal{S}')$, $\mathcal{E}'_\mathcal{S} = \mathcal{E} \cap \mathcal{S}' = \mathcal{E} \cap \mathcal{S}' \cap \text{aff}(\mathcal{S}')$. By Lemma 8.3, there exists $x \in \mathcal{E}_\mathcal{S} = \mathcal{E}'_\mathcal{S}$ such that $x \in \text{int}(\mathcal{S}')$. By observing the dimension of $\mathcal{E} \cap \text{aff}(\mathcal{S}')$ locally around x, we note that $\dim(\mathcal{E} \cap \text{aff}(\mathcal{S}')) = \dim(\mathcal{E} \cap \text{aff}(\mathcal{S}') \cap \mathcal{S}') = \dim(\mathcal{E}'_\mathcal{S}) = \dim(\mathcal{E}_\mathcal{S}) = \kappa_0$. □

Finally, the following lemma will present the key rank argument appearing in Theorem 8.7 and Theorem 8.13. This is a generalization of Proposition 1 from [125]. The assumptions of Proposition 1 from [125] have been relaxed, and the scope of the result has been significantly extended. In particular, it now covers a larger class of simplices on $\mathcal{S}$, instead of solely $\mathcal{O}_\mathcal{S}$ itself.

Lemma 8.6. *Let $\mathcal{P} = \text{co}\{w_1, \ldots, w_{p+1}\} \subset \mathcal{S}$ be a simplex in $\mathcal{S}$ with vertices $w_1, \ldots, w_{p+1}$. We assume the following:*

(I1) $\cup_{i=1}^{p+1} I(w_i) = \{0, 1, \ldots, n\}$.

(I2) There exists $r_0 > 0$ such that $\{0, 1, \ldots, r_0\}$ is the set of all indices that appear in more than one index set $I(w_i)$.

(I3) Each $I(w_i)$, $i = 1, \ldots, p+1$, has at least one non-zero exclusive member.

Further we assume

(N1) $Av_i + a \in \mathcal{C}(v_i)$ *for all* $i = 0, \ldots, n$.

(E1) $h_j \cdot (Av_i + a) = 0$ *for all* $i = 0, \ldots, r_0$, $j = r_0 + 1, \ldots, n$.

(E2) $h_j \cdot (Aw_k + a) = 0$ *for all* $k = 1, \ldots, p+1$, $j = r_0 + 1, \ldots, n$, $j \notin I(w_k)$.

(E3) $h_j \cdot (Aw_k + a) = 0$ *for all* $k = 1, \ldots, q+1$, $q \leq p$, $j = r_0 + 1, \ldots, n$, j *is an exclusive member of* $I(w_k)$.

Then, $\operatorname{rank}(A) < n - q$.

Proof. Without loss of generality, the vertices $v_0, \ldots, v_n$ can be ordered according to the non-zero exclusive members of $I(w_k)$. That is, the indices are ordered as $\{0, 1, \ldots, r_0, r_0+1, \ldots, r_1, \ldots, r_p+1, \ldots, r_{p+1}\}$, where $r_{p+1} = n$. Here, $\{0, 1, \ldots, r_0\}$ appear in more than one index set $I(w_k)$. Indices $\{r_{k-1}+1, \ldots, r_k\}$ only appear in $I(w_k)$. We assume that $\mathcal{S}$ is the standard orthogonal simplex: $v_0 = 0$, $v_i = e_i$, $h_i = -e_i$ for $i = 1, \ldots, n$. This is done without loss of generality by Section 7.1.

Now we examine the consequences of (N1), (E1)-(E2) on the forms of A and a. First consider (E1). Setting $i = 0$ we have

$$h_j \cdot (Av_0 + a) = h_j \cdot a = (-e_j) \cdot a = a_j = 0,\ j = r_0 + 1, \ldots, n. \tag{8.1}$$

Then again from (E1):

$$h_j \cdot Av_i = [A]_{ji} = 0,\ i = 1, \ldots, r_0,\ j = r_0 + 1, \ldots, n. \tag{8.2}$$

Next we examine (N1) and (E2). First we use (8.1) to simplify (E2):

$$h_j \cdot Aw_k = 0, \quad k = 1, \ldots, p+1, \quad j = r_0 + 1, \ldots, n,\ j \notin I(w_k). \tag{8.3}$$

Let $w_k = \sum_{i \in I(w_k)} \alpha_i^{w_k} v_i$, with $\alpha_i^{w_k} > 0$, $\sum_{i \in I(w_k)} \alpha_i^{w_k} = 1$. Then for each $k = 1, \ldots, p+1$, (8.3) becomes

$$\sum_{i \in I(w_k)} \alpha_i^{w_k} h_j \cdot Av_i = 0,\ j = r_0 + 1, \ldots, n,\ j \notin I(w_k). \tag{8.4}$$

By (N1) and (8.1):

$$h_j \cdot (Av_i + a) = h_j \cdot Av_i \leq 0,\ j = r_0 + 1, \ldots, n,\ i \neq j. \tag{8.5}$$

Combining (8.4) and (8.5) and using $\alpha_i^{w_k} > 0$, we get

$$h_j \cdot Av_i = [A]_{ji} = 0, \quad k = 1, \ldots, p+1, \quad i \in I(w_k),\ j \in \{r_0 + 1, \ldots, n\} \backslash I(w_k). \tag{8.6}$$

Consider $j = r_{k-1} + 1, \ldots, r_k$, the exclusive indices of $I(w_k)$. By exclusivity, $r_{k-1} + 1, \ldots, r_k \in$

$\{r_0+1,\ldots,n\}\backslash I(w_l)$ for all $l=1,\ldots,p+1$, $l\neq k$. Also,

$$\bigcup_{l\neq k} I(w_l) = \{0,\ldots,n\}\backslash\{r_{k-1}+1,\ldots,r_k\}.$$

Applying these observations to (8.6) we get

$$[A]_{ji}=0,\quad k=1,\ldots,p+1,\quad i\in\{0,\ldots,n\}\backslash\{r_{k-1}+1,\ldots,r_k\},\quad j\in\{r_{k-1}+1,\ldots,r_k\}. \tag{8.7}$$

Putting together the information in (8.1), (8.2) and (8.7), the forms of A and a are:

$$A=\begin{bmatrix} A_{00} & A_{01} & \cdots & A_{0,p+1} \\ & A_{11} & & \\ & & \ddots & \\ & & & A_{p+1,p+1} \end{bmatrix},\ a=\begin{bmatrix} a_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{8.8}$$

These forms are obtained as follows:

- The block of zero elements below $a_0\in\mathbb{R}^{r_0}$ in a is due to (8.1).
- The block of zero elements below $A_{00}\in\mathbb{R}^{r_0\times r_0}$ in A is due to (8.2).
- The off-diagonal zero blocks in the same rows as $A_{11},\ldots,A_{p+1,p+1}$ are due to (8.7).
- The first r_0 rows of A correspond to the values of j in (8.2), (8.7) for which we have no constraints.
- There are $p+1$ blocks of rows corresponding to sets of indices $\{r_{k-1}+1,\ldots,r_k\}$, $k=1,\ldots,p+1$. Thus, $A_{kk}\in\mathbb{R}^{(r_k-r_{k-1})\times(r_k-r_{k-1})}$.

Finally, we consider (E3). First, partition

$$w_k=(w_k^0,w_k^1,\ldots,w_k^{p+1})$$

according to $\{1,\ldots,r_0,r_0+1,\ldots,r_1,\ldots,r_p+1,\ldots,r_{p+1}\}$. Combining (8.8) and (E3):

$$\begin{bmatrix} A_{11} & & \\ & \ddots & \\ & & A_{q+1,q+1} \end{bmatrix}\begin{bmatrix} w_k^1 \\ \vdots \\ w_k^{q+1} \end{bmatrix}=0,\ k=1,\ldots,q+1. \tag{8.9}$$

Since $\{r_{k-1}+1,\ldots,r_k\}$ is exactly the set of exclusive members of $I(w_k)$, we know $w_k^k \neq 0$. Thus, A_{kk} is singular for all $k=1,\ldots,q+1$. Hence, from (8.9), $\operatorname{rank}(A) < n-q$. □

8.2 Equilibrium Set

We now proceed to the main contributions of this chapter. The central result of this section is that $\dim(\mathcal{E}_\mathcal{S}) = 0$. A characterization of $\mathcal{E}_\mathcal{S}$ was previously explored in [55], where it was also shown that $\mathcal{E}_\mathcal{S}$ is a point. However, [55] dealt only with single-input systems, whereas we deal with multi-input systems. It was shown in [125] that the equilibria in single-input systems can lie only on the boundary of $\mathcal{S}$. As the set of possible equilibria in controllable single-input systems is a line segment, the result of [55] was not unreasonable to expect. Our multi-input result, in contrast, is more surprising, and represents a significant improvement. Moreover, the fact that $\mathcal{E}_\mathcal{S}$ is a point was used in [55] to apply multi-affine feedback to solve the RCP for single-input systems in the case when affine feedback fails. Our result may provide an avenue to apply multi-affine feedback in the case of multiple inputs. This may ultimately serve to answer the fundamental question of an appropriate class of feedback for solvability of the RCP. We remark that there are also other differences from the result of [55]. In particular, we will not be assuming that $\mathcal{O}_\mathcal{S}$ is a simplex and we do not require assumptions (A2) and (N2).

Theorem 8.7. *Consider the system* (3.1) *defined on a simplex* $\mathcal{S}$*. Suppose assumptions (A1), (A3) and (N1) hold. If* $\mathcal{E}_\mathcal{S} \neq \emptyset$*, then* $\dim(\mathcal{E}_\mathcal{S}) = 0$*.*

Proof. Suppose $\dim(\mathcal{E}_\mathcal{S}) = \kappa_0 > 0$. Without loss of generality, let $q \leq n$ be such that $\cup_{k=1}^{\kappa_0+1} I(\varepsilon_k) = \{0,1,\ldots,q\}$. Let $\mathcal{S}' = \operatorname{co}\{v_0,\ldots,v_p\}$.

Let $\mathcal{O}'_\mathcal{S} = \mathcal{O} \cap \mathcal{S}'$ and $\mathcal{E}'_\mathcal{S} = \mathcal{E} \cap \mathcal{S}'$. By a variant of Lemma 1 in [125], $\partial\mathcal{E}_\mathcal{S} \subset \partial\mathcal{S}'$.

By Lemma 8.3, $\mathcal{E}_\mathcal{S} \cap \operatorname{int}(\mathcal{S}') \neq \emptyset$, so Lemma 8.4 applies with $\mathcal{P} = \mathcal{E}_\mathcal{S}$. By (A3), 0 is not an exclusive member of any $I(\varepsilon_k)$, $k=1,\ldots,\kappa_0+1$. Thus, by Lemma 8.4, the vertices of $\mathcal{S}$ can be ordered according to non-zero exclusive members of $I(\varepsilon_k)$. That is, the indices are ordered as $\{0,1,\ldots,r_0,r_0+1,\ldots,r_1,\ldots,r_{\kappa_0}+1,\ldots,r_{\kappa_0+1},q+1,\ldots,n\}$, with $r_0 < r_1 < \ldots < r_{\kappa_0+1} = q$. Here $\{0,1,\ldots,r_0\}$ are the indices appearing in more than one index set $I(\varepsilon_k)$, $k=1,\ldots,\kappa_0+1$. Indices $\{r_{k-1}+1,\ldots,r_k\}$ only appear in $I(\varepsilon_k)$.

Consider any vertex $\varepsilon_k \in \mathcal{E}_\mathcal{S}$, $k=1,\ldots,\kappa_0+1$. We have $A\varepsilon_k + a = 0$, and thus $h_j \cdot (A\varepsilon_k + a) = 0$, $j \in I$. Let $\varepsilon_k = \sum_{i\in I(\varepsilon_k)} \alpha_i^{\varepsilon_k} v_i$ with $\alpha_i^{\varepsilon_k} > 0$ and $\sum_{i\in I(\varepsilon_k)} \alpha_i^{\varepsilon_k} = 1$. Then

$$\sum_{i\in I(\varepsilon_k)} \alpha_i^{\varepsilon_k} h_j \cdot (Av_i + a) = 0\,, \quad j \in I\,. \tag{8.10}$$

By (N1),

$$h_j \cdot (Av_i + a) \leq 0, \quad i \in I(\varepsilon_k), \quad j \in I \setminus I(\varepsilon_k)\,. \tag{8.11}$$

Combining (8.10), (8.11), and the fact that $\alpha_i^{\varepsilon_k} > 0$ for $i \in I(\varepsilon_k)$ we get

$$h_j \cdot (Av_i + a) = 0, \quad k = 1, \ldots, \kappa_0 + 1, \quad i \in I(\varepsilon_k),\, j \in I \setminus I(\varepsilon_k). \tag{8.12}$$

Consider $j = r_{k-1} + 1, \ldots, r_k$. By exclusivity, $r_{k-1} + 1, \ldots, r_k \in I \setminus I(\varepsilon_l)$ for all $l = 1, \ldots, \kappa_0 + 1$, $l \neq k$. Also, $\cup_{l \neq k} I(\varepsilon_l) = \{0, \ldots, q\} \setminus \{r_{k-1} + 1, \ldots, r_k\}$. Applying these observations to (8.12) we get

$$\begin{aligned} h_j \cdot (Av_i + a) = 0, \quad & k = 1, \ldots, \kappa_0 + 1\,, \\ & i \in \{0, \ldots, q\} \setminus \{r_{k-1} + 1, \ldots, r_k\}\,, \\ & j \in \{r_{k-1} + 1, \ldots, r_k\}\,. \end{aligned} \tag{8.13}$$

Let us now invoke Lemma 8.6 for the simplex $\mathcal{E}_{\mathcal{S}}$. We note that Lemma 8.6 requires assumption (I1). This does not necessarily apply directly. However, we are interested only in solutions $x \in \mathcal{E} \cap \mathrm{aff}(\mathcal{S}')$. Thus, instead of looking at the whole simplex $\mathcal{S}$, we will be observing only $\mathcal{S}' = \mathrm{co}\{v_0, \ldots, v_q\}$, which was chosen exactly in a way that (I1) applies on it. (I2) and (I3) are satisfied.

(N1) is satisfied by the assumptions of the theorem, and it still holds for the reduced system on $\mathcal{S}'$. (E1) is satisfied by (8.13), (E2) is satisfied since $A\varepsilon_k + a = 0$ for all $k = 1, \ldots, \kappa_0 + 1$, and (E3) is also satisfied for that reason, with q from Lemma 8.6 equalling κ_0 in this theorem.

Hence, by Lemma 8.6 $\mathrm{rank}(\tilde{A}) < q - \kappa_0$, where $\tilde{A}$ is the matrix A with rows and columns $q+1, \ldots, n$ removed. Hence, equation

$$\tilde{A}\tilde{x} + \tilde{a} = 0 \tag{8.14}$$

has at least $\kappa_0 + 1$ linearly independent solutions, where $\tilde{x}$ and $\tilde{a}$ differ from x and a, respectively, by having their rows $q+1, \ldots, n$ removed. As each $x \in \mathcal{E} \cap \mathrm{aff}(\mathcal{S}')$ corresponds to exactly one solution $\tilde{x}$ of (8.14), $\dim(\mathcal{E} \cap \mathrm{aff}(\mathcal{S}')) \geq \kappa_0 + 1$. This is in contradiction with Lemma 8.5. □

We note that one of our triangulation assumptions, (A2), was not strictly necessary for the above theorem, although its use would have made our invocation of Lemma 8.6 more elegant, as assumption (I1) in Lemma 8.5 would have been automatically satisfied. However, assumption (A3) is necessary.

We use the following example, graphically presented in Figure 8.1, to demonstrate that.

Example 8.8. *Consider the open-loop system*

$$\dot{x} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Let $\mathcal{S} = \mathrm{co}\{v_0, \ldots, v_3\}$*, where* $v_0 = 0$*, and* $v_i = e_i$*, the ith Euclidean coordinate vector. One can verify* $Av_i + a \in \mathcal{C}(v_i)$*,* $i = 0, \ldots, 3$*. We compute* $\mathcal{E} = \{x \mid Ax + a = 0\} = \{x \mid x_1 = 3x_2 - 1, x_3 = -2x_2 + 1\}$*. Next we compute* $\mathcal{E}_{\mathcal{S}} = \mathrm{co}\{\varepsilon_1, \varepsilon_2\}$*, where* $\varepsilon_1 = \frac{1}{2}v_1 + \frac{1}{2}v_2$ *and* $\varepsilon_2 = \frac{1}{3}(v_0 + v_2 + v_3)$*. Thus,* $\dim(\mathcal{E}_{\mathcal{S}}) = 1$*, apparently contradicting Theorem 8.7. However* $I(\varepsilon_1) = \{1, 2\}$ *and* $I(\varepsilon_2) = \{0, 2, 3\}$*. Thus,* $\mathcal{E}_{\mathcal{S}}$ *is a line segment that goes from the facet* $\mathcal{F}_1$ *to the edge between* v_1 *and* v_2*. So, (A2) holds. On the other hand, as* ε_1 *is on the edge between* v_1 *and* v_2*,* $\varepsilon_1 \in \mathcal{F}_0$*. Hence, (A3) is violated.*

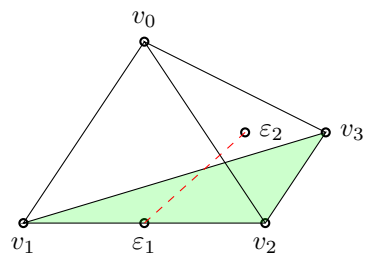

Figure 8.1: An illustration of Example 8.8. $\mathcal{E}_{\mathcal{S}}$ is depicted by a dashed red line, and the exit facet $\mathcal{F}_0$ is painted green.

From Theorem 8.7, we know that, under assumptions (A1), (A3) and (N1), $\mathcal{E}_{\mathcal{S}}$ is a single point. As mentioned, these assumptions represent a significant relaxation of the assumptions made in [55] and [125]. However, assumption (A1) still contains the imposition that $\mathcal{E}_{\mathcal{S}}$ is a simplex. In a particular case of controllable single-input systems, which is the main subject of inquiry of [125], this assumption can be removed as well. We invoke the following well-known result:

Lemma 8.9. *If* (A, B) *is controllable, then* $\mathcal{O}$ *is an affine subspace with* $\dim(\mathcal{O}) = m$*.*

Combined with assumption (A2), we can use Lemma 8.9 to prove the following result for single-input systems.

Proposition 8.10. *Suppose assumptions (A2), (A3) and (N1) hold. Suppose* $m = 1$ *and* (A, B) *is controllable. Then,* $\mathcal{E}_{\mathcal{S}}$ *is either empty or a single point.*

Proof. From Lemma 8.9, we obtain $\dim(\mathcal{O}) = 1$. From (A2), $\dim(\mathcal{O}_{\mathcal{S}}) = \dim(\mathcal{O}) = 1$. Since $\mathcal{O}_{\mathcal{S}}$ is a polytope, the only option is that $\mathcal{O}_{\mathcal{S}}$ is a line segment. Now, if $\mathcal{E}_{\mathcal{S}}$ is not empty, we have $\mathcal{E}_{\mathcal{S}} \subset \mathcal{O}_{\mathcal{S}}$. Since

$\mathcal{E}_{\mathcal{S}}$ is a polytope as well, this implies that $\mathcal{E}_{\mathcal{S}}$ is either a single point or a segment. In both cases, this implies it is a simplex. Hence, (A1) is satisfied as well. Now, by Theorem 8.7, we get that $\mathcal{E}_{\mathcal{S}} = \emptyset$ or $\mathcal{E}_{\mathcal{S}}$ is a single point. □

8.3 Reach Controllability

The notion of reach controllability has been defined for single-input systems in [125]. It provides a way to describe the ability of infinitesimal control actuation in system (3.1) to move the equilibria located on the boundary out of the simplex $\mathcal{S}$. This is similar to the notion of local controllability, which also uses infinitesimal control actuation to reach points in the local neighbourhood. However, in the case of reach controllability it is not the states of the system that we are directly interesting in moving. Instead, the desire is to move the equilibrium set $\{x \in \mathbb{R}^n \mid Ax + Bu(x) + a = 0\}$ out of the simplex.

The definition we provide here is a slight generalization of the definition in [125]. Our definition allows for the vertices of $\mathcal{E}_{\mathcal{S}}$ not to be contained in the vertex set of $\mathcal{O}_{\mathcal{S}}$. We will see such a situation appearing in Example 7.13.

Definition 8.11. *Suppose (N1) and* (5.6) *hold. We say the triple* (A, B, a) *is* reach controllable *if either* $\mathcal{E}_{\mathcal{S}} = \emptyset$*, or* $\mathcal{E}_{\mathcal{S}} = \mathrm{co}\{\varepsilon_1, \ldots, \varepsilon_{\kappa_0+1}\}$ *with* $0 \leq \kappa_0 < \kappa$*, and there exists* $b \in \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}}) \backslash \{0\}$ *such that for each* $\varepsilon_k \in I_{\mathcal{E}_{\mathcal{S}}}$ *there exist* $i \in I(\varepsilon_k)$ *and* $u_i > 0$ *for which* $Av_i + bu_i + a \in \mathcal{C}(v_i)$*.*

An example of a system which satisfies Assumption 8.1, but is not reach controllable was provided as Example 3 in [125]. Along with our Assumption 8.1, results in [125] make use of the following additional assumption:

(A0) $\mathcal{O}_{\mathcal{S}} = \mathrm{co}\{o_1, \cdots, o_{\kappa+1}\}$ is a κ-dimensional simplex with $1 \leq \kappa < n$.

Under assumptions (A0)-(A3) and (N1)-(N2), it has been shown in [125] that the vertices of $\mathcal{E}_{\mathcal{S}}$ in the single-input case are indeed the vertices of $\mathcal{O}_{\mathcal{S}}$. However, this does not hold if we just remove the assumption (N2). This is shown in Example 8.12.

Example 8.12. *Let* $m = 1$ *and let*

$$A = \begin{pmatrix} -4 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & -8 \end{pmatrix}, B = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, a = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Let $\mathcal{S}$ be the standard orthogonal simplex. We note that $\mathcal{C}(v_0) = \{x \in \mathbb{R}^3 \mid x \geq 0\}$ using the definition of a nonnegative vector from Section 2.1, while $\mathcal{C}(v_i) = \{x \in \mathbb{R}^3 \mid x_j \geq 0, j \neq i\}$.

It can be verified by direct calculations that (N1) holds, and we can additionally easily calculate $\mathcal{O} = \{x \mid Ax + a \in \mathcal{B}\}$: it equals $\mathrm{aff}\{v_0/2 + v_1/2, v_0/2 + v_2/4 + v_3/4\}$. Hence, $\mathcal{O}_{\mathcal{S}} = \mathcal{O} \cap \mathcal{S} = \mathrm{co}\{v_0/2 + v_1/2, v_0/2 + v_2/4 + v_3/4\}$. Thus, $\mathcal{O}_{\mathcal{S}}$ clearly satisfies (A0), (A2) and (A3).

Finally, $\mathcal{E}$ is given by $\{x \mid Ax + a = 0\}$, and in this particular case, we can easily calculate $\mathcal{E} = \mathcal{E}_{\mathcal{S}} = \{v_0/2 + v_1/4 + v_2/8 + v_3/8\}$. Clearly, $\mathcal{E}_{\mathcal{S}}$ then satisfies (A1). So, this system satisfies (A0)-(A3) and (N1). However, $\mathcal{E}_{\mathcal{S}}$ is not a vertex of $\mathcal{O}_{\mathcal{S}}$. The resulting picture is given in Figure 8.2.

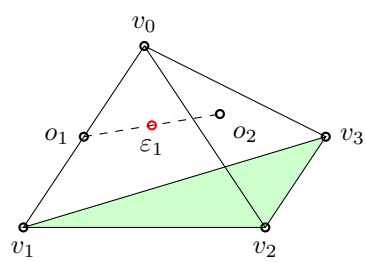

Figure 8.2: An illustration of Example 8.12. $\mathcal{O}_{\mathcal{S}}$ is depicted by a dashed line, and the exit facet $\mathcal{F}_0$ is painted green.

In the remainder of this section, we will be assuming (N2). Thus, by [125], we may assume without loss of generality that $\mathcal{E}_{\mathcal{S}} = \{o_1, \ldots, o_{\kappa_0+1}\}$, and, by Theorem 8.7, we know that $\kappa_0 = 0$.

The main result of [125], which connects reach controllability with solvability of the RCP, contains the assumption that $Ao_{\kappa+1} + a \in \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$. Along with being unintuitive, this assumption seems to significantly constrain the potential values of $Ao_{\kappa+1} + a$, allowing only those values which lie in the ray of $\mathcal{B}$ that points through $\mathrm{cone}(\mathcal{O}_{\mathcal{S}})$. We now show that this assumption is in fact unnecessary. Our proof relies on the use of assumption (N1) to derive the zero structure of matrices A and a, and invoke Lemma 8.6.

Theorem 8.13. *Suppose (A0) and Assumption 8.1 hold. Suppose $m = 1$, and $\mathcal{E}_{\mathcal{S}} = \{o_1, \ldots, o_{\kappa_0+1}\}$. Then, if $\kappa > \kappa_0$, it is impossible that $Ao_k + a \notin \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$ for all $k = \kappa_0 + 2, \ldots, \kappa + 1$.*

Proof. The conditions of Theorem 8.7 hold, so $\kappa_0 = 0$, $\mathcal{E}_{\mathcal{S}} = \{o_1\}$ and $Ao_k + a \neq 0$, $k = 2, \ldots, \kappa + 1$.

Suppose by way of contradiction that $Ao_k + a \notin \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}})$ for all $k = 2, \ldots, \kappa + 1$. Since $0 \neq Ao_k + a \in \mathcal{B}$ for $k = 2, \ldots, \kappa + 1$ and $m = 1$, we have

$$-(Ao_k + a) \in \mathcal{B} \cap \mathrm{cone}(\mathcal{O}_{\mathcal{S}}),\ k = 2, \ldots, \kappa + 1. \tag{8.15}$$

Assume without loss of generality that $\mathcal{S}$ is a standard orthogonal simplex, i.e., $v_0 = 0$, $v_i = e_i$ and $h_i = -e_i$ for all $i \in I$. Also, assume there exists $l \geq 0$ such that $I(o_1) \cap \cdots \cap I(o_{\kappa+1}) = \{0, 1, 2, \ldots, l\}$. By Lemma 8.2, $\text{cone}(\mathcal{O}_{\mathcal{S}}) = \{y \mid h_j \cdot y \leq 0 \text{ for all } j = l+1, \ldots, n\}$. Combining this with (8.15), we obtain

$$h_j \cdot (Ao_k + a) \geq 0,\ k = 2, \ldots, \kappa+1,\ j = l+1, \ldots, n. \tag{8.16}$$

Since we know $I \backslash I(o_k) \subset \{l+1, \ldots, n\}$ for all $k = 2, \ldots, \kappa+1$, (8.16) gives

$$h_j \cdot (Ao_k + a) \geq 0,\ k = 2, \ldots, \kappa+1,\ j \in I \backslash I(o_k). \tag{8.17}$$

From (N1) we also know

$$h_j \cdot (Av_i + a) \leq 0,\ i = 0, \ldots, n,\ j \in I \backslash \{i\}.$$

This implies

$$h_j \cdot (Av_i + a) \leq 0, \quad k = 1, \ldots, \kappa+1, \quad i \in I(o_k),\ j \in I \backslash I(o_k). \tag{8.18}$$

Combining (8.17), (8.18) and convexity, we obtain

$$\begin{aligned} h_j \cdot (Ao_k + a) &= h_j \cdot (A \sum_{i \in I(o_k)} \alpha_i^{o_k} v_i + a) = \\ &= \sum_{i \in I(o_k)} h_j \cdot (Av_i + a) = 0, \quad k = 1, \ldots, \kappa+1,\ j \in I \backslash I(o_k). \end{aligned} \tag{8.19}$$

Note that we include $k = 1$ here because $Ao_1 + a = 0$.

Then, since $\alpha_i^{o_k} > 0$, reapplying (8.18) we get

$$h_j \cdot (Av_i + a) = 0, \quad k = 1, \ldots, \kappa+1, \quad i \in I(o_k),\ j \in I \backslash I(o_k). \tag{8.20}$$

Now, assume that $\{0, 1, \ldots, r_0\}$ is the set of all indices that appear in more than one index set $I(o_k)$. Hence, for any $j = r_0 + 1, \ldots, n$, j is an exclusive member of some $I(o_{k'})$, $k' = 1, \ldots, \kappa+1$. For any $i = 0, \ldots, r_0$, i is not an exclusive member of $I(o_{k'})$. Thus, there exists $k \in \{1, \ldots, \kappa+1\}$ such that $i \in I(o_k)$ and $j \in I \backslash I(o_k)$. From (8.20) we get

$$h_j \cdot (Av_i + a) = 0,\ i = 0, \ldots, r_0,\ j = r_0 + 1, \ldots, n. \tag{8.21}$$

We first invoke Lemma 8.4. By Lemma 5.31 (ii), $\partial\mathcal{O}_{\mathcal{S}} \subset \partial\mathcal{S}$. Thus, the assumptions for Lemma 8.4 are satisfied. Hence, each index set $I(o_k)$, $k = 1, \ldots, \kappa + 1$, has an exclusive member.

We now invoke Lemma 8.6 for $\mathcal{P} = \mathcal{O}_{\mathcal{S}}$. (I1) is satisfied by assumption (A2). (I2) is also satisfied by definition of r_0. (I3) is satisfied by our invocation of Lemma 8.4. Assumption (N1) holds. (E1) holds by (8.21). (E2) is satisfied by (8.19).

We now distinguish between two cases. First, suppose $\kappa + 1 = 2$. (E3) is certainly satisfied for at least $q = 0$, as $Ao_1 + a = 0$. Hence, by Lemma 8.6 $\operatorname{rank}(A) \leq n - 1 = n - \kappa$. This implies $\dim(\mathcal{E}) \geq \kappa$. However, by (A2) $\dim(\mathcal{O}) = \dim(\mathcal{O}_{\mathcal{S}}) = \kappa$, and $\mathcal{E} \subset \mathcal{O}$. As both $\mathcal{E}$ and $\mathcal{O}$ are affine spaces, this implies $\mathcal{E} = \mathcal{O}$, i.e., $\mathcal{O}_{\mathcal{S}} = \mathcal{E}_{\mathcal{S}}$. This means $\kappa = \kappa_0$, and we reach a contradiction.

Suppose now that $\kappa + 1 > 2$. Then, without loss of generality we may assume

$$I(o_2) \cap \cdots \cap I(o_{\kappa+1}) = \{0, 1, 2, \ldots, l'\} \tag{8.22}$$

for some $l' \geq l$.

As $\mathcal{B} \ni Ao_k + a \neq 0$ for $k = 2, \ldots, \kappa + 1$, we know $Ao_k + a = \lambda_k b$ for some $\lambda_k \neq 0$, $k = 2, \ldots, \kappa + 1$. By (8.19) we get

$$h_j \cdot b = 0,\ k = 2, \ldots, \kappa + 1,\ j \in I \backslash I(o_i). \tag{8.23}$$

By (8.22) and (8.23), $h_j \cdot b = 0$ for all $j = l'+1, \ldots, n$. Hence, $h_j \cdot (Ao_k + a) = 0$, for all $k = 2, \ldots, \kappa+1$, $j = l' + 1, \ldots, n$.

By (8.22) and $\kappa+1 > 2$, all exclusive members of $I(o_k)$, $k = 2, \ldots, \kappa+1$ are contained in $\{l'+1, \ldots, n\}$ Thus, $h_j \cdot (Ao_k + a) = 0$ for all $k = 1, \ldots, \kappa + 1$ and j which are exclusive members of $I(o_k)$. We again included $k = 1$ above as $Ao_1 + a = 0$.

We can now take $q = \kappa$ in (E3) of Lemma 8.6. By Lemma 8.6, $\operatorname{rank}(A) \leq n - \kappa - 1$. This implies $\kappa = \dim(\mathcal{O}) \geq \dim(\mathcal{E}) \geq \kappa + 1$. □

A generalization of Theorem 4 of [125] follows immediately.

Theorem 8.14. *Let $m = 1$, and let (A0) and Assumption 8.1 hold. System* (3.1) *is RCP solvable by affine feedback if and only if (A, B, a) is reach controllable.*

Proof. The theorem was proved in [125] under the assumption that $Ao_{\kappa+1} + a \in \mathcal{B} \cap \operatorname{cone}(\mathcal{O}_{\mathcal{S}})$. Clearly, by renaming vertices this is equivalent to stating $Ao_s + a \in \mathcal{B} \cap \operatorname{cone}(\mathcal{O}_{\mathcal{S}})$ for any $s > \kappa_0 + 1$. Assume otherwise: then, $Ao_s + a \notin \mathcal{B} \cap \operatorname{cone}(\mathcal{O}_{\mathcal{S}})$ for all $s > \kappa_0 + 1$. However, this was shown to be impossible by the previous theorem. We are done. □

Chapter 9

Applications of Reach Control

In a finishing segue from Chapters 5-8, this chapter introduces two novel applications of the reach control theory to automated vehicles. While it relates to the theory discussed earlier in the thesis, in particular, the necessary and sufficient conditions for the solvability of the RCP using affine feedback, this chapter is largely self-contained. The application to parallel parking has been developed in [107] and the text in the corresponding section of that chapter is heavily based on that manuscript. Analogously, the application to adaptive cruise control was discussed in [108] and the second part of this chapter is based on that manuscript.

9.1 Parallel Parking

9.1.1 Introduction

Automated parallel parking of a vehicle is a widely studied problem in control theory [29, 37, 75, 92, 110, 115, 129]. The problem seeks to automate the maneuver often used by human drivers to park a car in a limited space between two cars previously parked parallel to the curb (see Figure 9.1).

The control strategies predominantly used in previous literature (e.g., [29, 75, 81]) consist of calculating a feasible path that a vehicle should follow during the parallel parking maneuver, starting from a given initial position, and then computing the appropriate steering control to follow this path. This approach is not robust, as a slight deviation from the initial position or a disturbance during the parallel parking maneuver requires a new path to be calculated, if such a path can be calculated at all. We propose a different method using a reach control approach. We will split the state space into polytopes and then devise a sequence of polytopes leading from the initial state to a desired endpoint. On each

Figure 9.1: An illustration of the initial state in the parallel parking maneuver. Vehicles A and B, marked in dark gray, are previously parked in parallel with the curb (marked in light gray), and at a short distance from it. Vehicle C, marked in blue, seeks to park between vehicle A and vehicle B. The front wheels of all vehicles are coloured red, while the rear wheels are coloured black.

of these polytopes, a closed-loop controller will be devised. It will drive the system state to leave the polytope through the exit facet that connects it to the next polytope in the sequence.

While RCP is amenable to a number of applications mentioned in Chapter 1, the work most relevant to the present problem is [133]. It deals with driving a gantry crane from one area of the state space to another while avoiding an obstacle. The model used by [133] is underactuated. This is the case for the vehicle model in our investigation as well. However, these two works have substantial differences: [133] linearize their nonlinear model around a single point, whereas we seek to ensure the faithfulness of our model by constructing a separate linearization for each region of the partitioned state space. This idea was discussed by [43] but was, to our knowledge, never previously fully investigated in an application. Additionally, while the state space of [133] was four-dimensional, the actual controller design was performed in the 2D output space. In contrast, our design is performed on the full three-dimensional state space.

9.1.2 Model and Problem Statement

We use the standard front-wheel drive car model based on the unicycle described in, e.g., [36]. The model uses an external reference frame (x, y, θ), where the x-axis is parallel to the curb, the y-axis is perpendicular to it and pointing into the road, and θ is the orientation of the car with respect to the curb. The car's position (x, y) in this frame of reference is determined by the midpoint of its rear axle. Let v denote the forward-moving speed of the car, and let φ denote the angle in degrees of the front wheels with respect to the orientation of the car, where $\varphi = 0°$ if the front wheels are aligned with the car; see Figure 9.2.

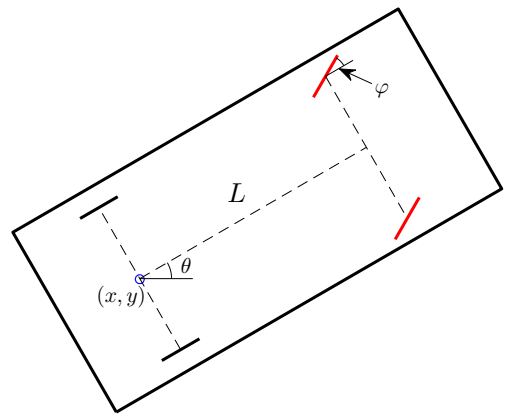

Figure 9.2: The states of the car are (x, y, θ), where x is the horizontal position of the midpoint of the car's rear axle, y is its vertical position, and θ is the orientation of the car with respect to the curb. Also, φ is the angle of the car's front wheels, and L is the distance between the vehicle's front and back wheels.

The model equations are given as follows:

$$\begin{aligned} \dot{x} &= v\cos(\varphi)\cos(\theta) \\ \dot{y} &= v\cos(\varphi)\sin(\theta) \\ \dot{\theta} &= \frac{v}{L}\sin(\varphi), \end{aligned} \tag{9.1}$$

where L is the distance between the front wheels and the back wheels (wheelbase). Our simulation is based on a model of a 2014 Audi TT RS vehicle. The exact dimensions of the car are taken from [6], and are given in Appendix A. We take the maximum steering angle of the front wheels $\varphi_M = 33°$. Because of safety and liveness concerns, we also take the lowest possible speed of the vehicle $v_S = -5\text{km/h} \approx -1.39\text{m/s}$ and the highest speed $v_F = -10\text{km/h} \approx -2.78\text{m/s}$.

Note that the model we are using disregards the Ackermann steering geometry (described, for instance, in [83]). In actual cars, the front wheels do not turn at the same angle in order to avoid slipping. The actual maximum steering angle of a 2014 Audi TT RS can be calculated from the available data [6], using the formulae given in [83]. The true maximum angle of a wheel located inwards when turning is around 33.01°, and the maximum angle of an outer wheel is around 24.76°.

We assume the length of the parking space to be around 1.9 times the length of the car, the distance of the already parked cars from the curb to be 0.05m. The two already parked cars are assumed to be equal in size to the car we are parking, and the roadway is assumed to be clear of any obstacles apart

from the two cars. We note that the length of the parking space allows a comfortable parallel parking maneuver. We will show that our approach does not use the entire area of the parking space.

In order to describe the vehicle's initial position, ending position, and obstacles, we define the following four sets:

- $R(x) \subset \mathbb{R}^2$ is the set of all physical positions covered by a car whose position of the midpoint of its rear axle and orientation are given by x.
- $\mathcal{I} \subset \mathbb{R}^2$ is the set of allowed initial positions of the vehicle. We assume its dimensions to be $0.5\text{m} \times 0.25\text{m}$, and we assume it to be located so that the back of the car is between 0.5m and 0m behind the back of the first parked car, and the side of the car is located between 0.75m and 1m from the side of the parked cars.
- $\mathcal{Q} \subset \mathbb{R}^2$ is the set of all desirable ending positions of the vehicle. It is situated so that no part of the car is colliding with another car or a curb, and that the side of the car is no more than 0.3m from the curb.
- $\mathcal{G} \subset \mathbb{R}^2$ is the set of all obstacles to the vehicle in its path. It consists of the two previously parked vehicles and the sidewalk.

We note that the size of $\mathcal{Q}$ imposes a very tight criterion for correct parking. An illustration of $\mathcal{I}$, $\mathcal{Q}$ and $\mathcal{G}$ is given in Figure 9.3.

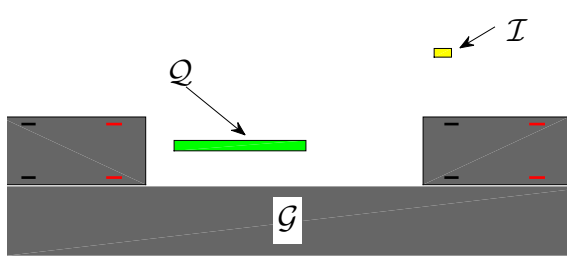

Figure 9.3: An illustration of the initial set, final set, and obstacles in the parallel parking maneuver. The parked cars and the sidewalk are drawn in gray, and are denoted by the set of obstacles $\mathcal{G}$. The initial box $\mathcal{I}$ for the midpoint of the car's rear axle is drawn in yellow. The acceptable ending box $\mathcal{Q}$ is drawn in green.

We now formalize the problem we will be solving.

Problem 9.1. *Let $R(x), \mathcal{I}, \mathcal{Q}, \mathcal{G} \subset \mathbb{R}^2$ be as above. Let $U = [v_F, v_S] \times [-\varphi_M, \varphi_M]$. Consider the model (9.1). Find a connected set $\mathcal{M} \supset \mathcal{I} \times \{0\}$ and a feedback controller $u : \mathcal{M} \to U$ such that for all $x_0 \in \mathcal{I} \times \{0\}$ there exists $T \geq 0$ such that the following holds:*

(i) $\phi(T, x_0) \in \mathcal{Q} \times \{0\}$,

(ii) $\phi(t, x_0) \in \mathcal{M}$ *for all* $t \in [0, T]$,

(iii) $R(\phi(t, x_0)) \cap \mathcal{G} = \emptyset$ *for all* $t \in [0, T]$.

Problem 9.1 is of the reach-avoid type. Note that the initial condition set is $\mathcal{I} \times \{0\}$ and the final condition set is $\mathcal{Q} \times \{0\}$, so the car must be parallel to the curb at the start and the end of the maneuver.

9.1.3 Methodology

In this section, we propose a solution to Problem 9.1 using reach control. We will construct a sequence of polytopes in the state space such that the first polytope contains the initial states of the car, the feedback control on each polytope pushes the system state into the next one, and the last polytope ends in a acceptable ending state for the parking maneuver.

The solution is motivated by the steps human drivers follow to park the car. Using the notation introduced in Figure 9.1, the standard way to execute this maneuver consists of four steps:

(1) Bring vehicle C to a stop in parallel to the curb, and next to vehicle A.

(2) Start going in reverse, with the front wheels rotated as far to the right as possible until vehicle C is roughly at a 45° angle to the curb.

(3) Straighten out the wheels and continue going in reverse until vehicle C is situated almost behind vehicle A.

(4) Continue going in reverse, with the front wheels rotated as far to the left as possible until vehicle C is parallel to the curb.

These steps are illustrated in Figure 9.4.

In the remainder of the section, we assume that a human driver has already completed Step 1 by bringing car to a standstill in state $x_0 \in \mathcal{I} \times \{0\}$. The steps we want to perform are 2–4. Our approach to designing a control strategy which completes this maneuver consists of the following elements:

- Design a sequence of polytopes $\mathcal{P}_i$ constructed around an ideal but possibly infeasible car trajectory, with the goal of having the car remain close to the ideal trajectory by staying within the polytopes $\mathcal{P}_i$.

- Use the necessary conditions for the solvability of the RCP to design appropriate linearization points at each polytope $\mathcal{P}_i$, in order to turn the nonlinear model (9.1) into an affine system amenable to an RCP approach.

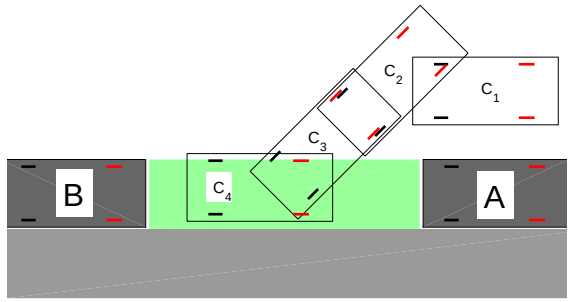

Figure 9.4: An illustration of the parallel parking maneuver. Notation from Figure 9.1 is used. States C_i, $i = 1, \ldots, 4$ denote the position and orientation of vehicle C at the end of Step i of the maneuver described above.

- Use the sufficient conditions for the solvability of the RCP to determine the exact size of each polytope $\mathcal{P}_i$.
- Use reach control theory to design a piecewise affine controller on each of the polytopes $\mathcal{P}_i$.

Polytopes

An intuitive idea of what the position (x, y) and orientation θ of the car "should" look like during steps 2–4 is given in Figure 9.5. The trajectory in Figure 9.5 is not necessarily feasible for the system (9.1) or its linearization. Moreover, it is a trajectory for only one initial condition, whereas our goal is to perform a successful parking maneuver starting from any point in $\mathcal{I} \times \{0\}$. Our approach is to design a control strategy to ensure that actual trajectories of cars starting at any point $x_0 \in \mathcal{I} \times \{0\}$ remain similar to this imaginary trajectory. In order to do that, we will embed the trajectory in Figure 9.5 into a sequence of polytopes $\mathcal{P}_1, \ldots, \mathcal{P}_m$ and ensure that all trajectories starting from initial states in $\mathcal{I} \times \{0\}$ remain inside those polytopes.

Let $\mathcal{I} = [x_{11}, x_{12}] \times [y_{11}, y_{12}]$. We now describe the construction of polytopes $\mathcal{P}_1, \ldots, \mathcal{P}_m$. The process is as follows:

- Grid up the interval $[0°, 45°]$ into intervals $[\theta^0, \theta^1]$, $[\theta^1, \theta^2]$, $\ldots$, $[\theta^{k-1}, \theta^k]$, where $\theta^0 = 0°$ and $\theta^k = 45°$, so that the (x, y) trajectory of the car in Figure 9.5 is nearly linear during the time its orientation is between θ^i and θ^{i+1}. By examining Figure 9.5, we find that $k = 4$ and

$$\theta^i = i \cdot 11.25°, \quad i = 0, 1, \ldots, 4,$$

  are suitable.

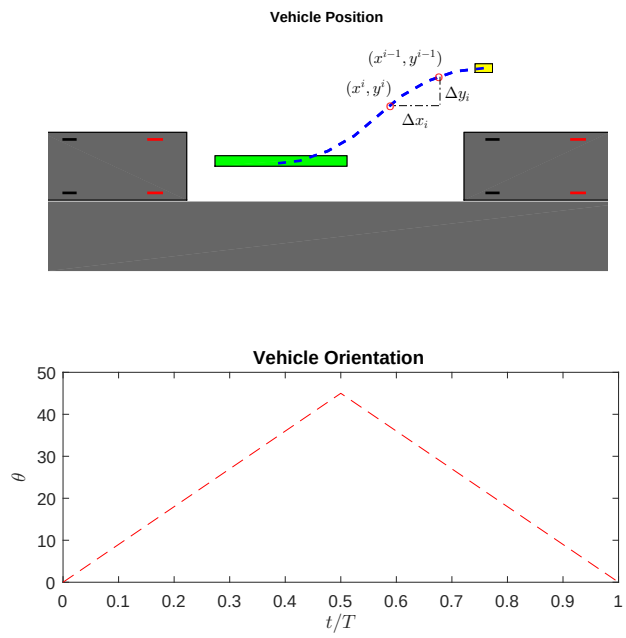

Figure 9.5: An illustration of an imaginary ideal path for the parking maneuver. The top picture contains the (x, y) position of the car during the maneuver indicated by a dashed blue line, while the bottom graph illustrates the orientation θ of the vehicle during the movement, indicated by a dashed red line. In the bottom graph, time t is given as a fraction of the total maneuver time T.

- Based on a measurement from Figure 9.5, let (x^i, y^i) be the position of the car when its orientation is $\theta = \theta^i$. We compute

$$\Delta x_i = x^i - x^{i-1}, \quad \Delta y_i = y^i - y^{i-1}.$$

  For instance, $\Delta x_1 = -0.85$m, $\Delta y_1 = -0.094$m. All subsequent coefficients $\Delta x_i, \Delta y_i$, along with all other relevant measurements and control laws, are available in Appendix A. Note that in the remainder of the procedure, we only work with $\Delta x_i, \Delta y_i$ instead of x^i and y^i, so the choice of origin $(0, 0)$ in Figure 9.5 is not relevant.

- We now start with $\theta \in [0, \theta^1]$. From above we have $\Delta x_1, \Delta y_1$. At $\theta = 0°$, we know that the position of the car is inside $\mathcal{I}$. At $\theta = \theta^1$, the position of the car should be inside $\mathcal{I} + (\Delta x_i, \Delta y_i)$, based on Figure 9.5. Thus, an ideal polytope $\mathcal{P}_1'$ would be constructed by taking the convex hull of faces $\mathcal{F}_1^{\text{in}} = \mathcal{I} \times \{0\}$ and $\mathcal{F}_1^{\text{out}'} = (\mathcal{I} + (\Delta x_1, \Delta y_1)) \times \{\theta^1\}$. Unfortunately, this will not work because the ideal trajectory in Figure 9.5 is not necessarily feasible, and the trajectories are not linear.

- Thus, we introduce *widening coefficients* $w_{1x}, w_{1y} \geq 0$, and we construct a modified polytope $\mathcal{P}_1$ to be the convex hull of $\mathcal{F}_1^{\text{in}} = \mathcal{I} \times \{0\}$ and $\mathcal{F}_1^{\text{out}} = [x_{11} + \Delta x_1 - w_{1x}, x_{12} + \Delta x_1 + w_{1x}] \times [y_{11} +$

$\Delta y_1 - w_{1y}, y_{12} + \Delta x_1 + w_{1y}] \times \{\theta^1\}$. In particular, $\mathcal{F}_i^{\text{out}}$ is now widened.

- We generate $\mathcal{P}_2, \ldots, \mathcal{P}_k$ inductively, where $\mathcal{P}_i$ is the convex hull of the facets given by

$$\mathcal{F}_i^{\text{in}} = \mathcal{F}_{i-1}^{\text{out}},$$

$$\mathcal{F}_i^{\text{out}} = [x_{i+1,1}, x_{i+1,2}] \times [y_{i+1,1}, y_{i+1,2}] \times \{\theta^i\},$$

where

$$x_{i1} = x_{i-1,1} + \Delta x_{i-1} - w_{i-1,x},$$

$$x_{i2} = x_{i-1,2} + \Delta x_{i-1} + w_{i-1,x},$$

$$y_{i1} = y_{i-1,1} + \Delta y_{i-1} - w_{i-1,y},$$

$$y_{i2} = y_{i-1,2} + \Delta y_{i-1} + w_{i-1,y}.$$

Notice that the "in-facet" of $\mathcal{P}_i$ is the "out-facet" of $\mathcal{P}_{i-1}$, and $\mathcal{F}_i^{\text{out}}$ is again widened as above. See Figure 9.6 for an illustration of the above procedure.

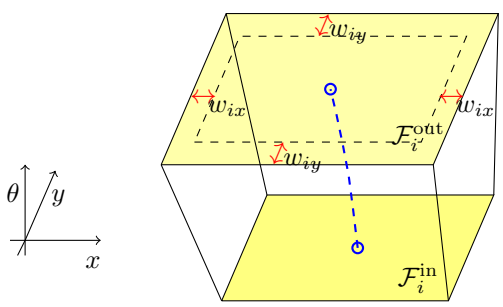

Figure 9.6: An illustration of polytope $\mathcal{P}_i$. The polytope resembles a parallelepiped with $\mathcal{F}_i^{\text{in}}$ and $\mathcal{F}_i^{\text{out}}$ as its bases. The imaginary ideal trajectory of the car while going through polytope $\mathcal{P}_i$ is given by a blue dashed line. Widening coefficients w_{ix}, w_{iy} are represented by red arrows.

The above process corresponds to Step 2 of the maneuver. Step 3, in which the orientation θ of the vehicle does not change significantly, is then interpreted by repeating the same process as above, just with $\theta^i = \theta^{i-1} = 45°$, until some polytope $\mathcal{P}_l$. After that, Step 4 is again performed in the same way as Step 2, just with $0° < \theta^i < \theta^{i-1}$, until $\theta^m = 0°$ for some $m \geq l$. In our case, the entire construction is performed with just eight polytopes: $k = 4$, $l = k = 4$, and $m = 8$. This means that Step 3 is short enough to not make a significant impact in the parking maneuver. Orientations $\theta^5, \ldots, \theta^8$ are chosen by

$$\theta^i = 90° - i \cdot 11.25°, \quad i = 5, 6, 7, 8.$$

We note that we still did not select widening coefficients w_{ix}, w_{iy}. The widening coefficients need to

be selected so that there exists a controller driving all trajectories from $\mathcal{P}_i$ to leave $\mathcal{P}_i$ through the facet $\mathcal{F}_i^{\text{out}}$. Since we will be using reach control theory to construct such a controller, the choice of widening coefficients will follow from the solvability conditions of the RCP. This will be elucidated in Section 9.1.3.

Using the notation of previous sections, it was shown by [54] that, if $\mathcal{O}_{\mathcal{P}} = \mathcal{O} \cap \mathcal{P} = \emptyset$, solvability of the invariance conditions is in fact a sufficient condition for solvability of the RCP by continuous piecewise affine feedback. This is a variation of Corollary 3.7, where condition (ii) is automatically satisfied from $\mathcal{O}_{\mathcal{P}} = \emptyset$. In the following section we will linearize system (9.1) in such a way that $\mathcal{O}_{\mathcal{P}_i} = \emptyset$ for all $i \in \{1, \ldots, m\}$.

Linearization

In order to make (9.1) amenable to an RCP approach, we linearize it around $(x_0, y_0, \theta_0, \varphi_0, v_0)$, and assume $\varphi_0 = 0$. We obtain

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} 0 & 0 & -v_0 \sin(\theta_0) \\ 0 & 0 & v_0 \cos(\theta_0) \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} + \begin{pmatrix} \cos(\theta_0) & 0 \\ \sin(\theta_0) & 0 \\ 0 & v_0/L \end{pmatrix} \begin{pmatrix} v \\ \varphi \end{pmatrix} + \begin{pmatrix} v_0\theta_0 \sin(\theta_0) \\ -v_0\theta_0 \cos(\theta_0) \\ 0 \end{pmatrix}. \tag{9.2}$$

Note that we did not require $(x_0, y_0, \theta_0, \varphi_0, v_0)$ to be an equilibrium for system (9.1). This results in the linearized system having affine form which fits within the setting of reach control.

In order to make the system (9.2) as faithful to (9.1) as possible, we will choose a different linearization point for each $\mathcal{P}_i$, $i \in \{1, \ldots, m\}$. First, we want to choose $(x_0, y_0, \theta_0, v_0, \varphi_0)$ in such a way that $\mathcal{O}_{\mathcal{P}_i} = \emptyset$. The choice of x_0, y_0 is irrelevant as they do not appear on the right side of (9.1). Additionally, as $v_0 \in [v_S, v_F]$ and $\varphi_0 \in [-\varphi_M, \varphi_M]$, it is natural to take $v_0 = (v_S + v_F)/2$, $\varphi_0 = 0$. Going back to the affine model (9.2), we can now calculate $\mathcal{O}_{\mathcal{P}_i}$. It is given by all $(x, y, \theta) \in \mathcal{P}_i$ which satisfy

$$\begin{pmatrix} -v_0 \sin(\theta_0)\theta + v_0\theta_0 \sin(\theta_0) \\ v_0 \cos(\theta_0)\theta - v_0\theta_0 \cos(\theta_0) \\ 0 \end{pmatrix} = \begin{pmatrix} v \cos(\theta_0) \\ v \sin(\theta_0) \\ \varphi v_0/L \end{pmatrix}$$

for some $v, \varphi \in \mathbb{R}$. If $v = 0$, this implies $\sin(\theta_0)(\theta - \theta_0) = 0$ and $\cos(\theta_0)(\theta - \theta_0)$, i.e., $\theta = \theta_0$, and if $v \neq 0$

and $\theta \neq \theta_0$, we have

$$\frac{\sin(\theta_0)(\theta - \theta_0)}{\cos(\theta_0)(\theta - \theta_0)} = -\cot(\theta_0),$$

i.e., $\tan^2(\theta_0) = -1$, which is impossible. Hence,

$$\mathcal{O}_{\mathcal{P}_i} = \{(x, y, \theta) \in \mathcal{P}_i \mid \theta = \theta_0\}.$$

Thus, if we choose θ_0 so that $\mathcal{P}_i$ does not contain any points (x, y, θ) with $\theta = \theta_0$, it is guaranteed that there are no equilibria in $\mathcal{P}$. By construction in Section 9.1.3, the orientations θ in $\mathcal{P}_i$ are between θ^{i-1} and θ^i, with the convention that $\theta^0 = 0$. As we want the linearization (9.2) of the system to be as faithful to (9.1) as possible, we want θ_0 close to $[\theta^{i-1}, \theta^i]$. Thus, we will choose θ_0 for each polytope $\mathcal{P}_i$ so that $\theta_0 < \min(\theta^{i-1}, \theta^i)$, with

$$\min(\theta^{i-1}, \theta^i) - \theta_0 = \varepsilon > 0$$

as small as possible. In our simulations, we chose $\varepsilon = 10^{-10}$.

Choice of Widening Coefficients

Since $\mathcal{O}_{\mathcal{P}_i} = \emptyset$, we know from [54] (i.e., a variation of Corollary 3.7) that if the invariance conditions $Ax + Bu(x) + a \in \mathcal{C}(x)$ are solvable on the vertices $v \in \mathcal{P}_i$, the RCP is solvable using continuous piecewise feedback. The vertices of $\mathcal{P}_i$ were defined in Section 9.1.3, but we did not yet give the exact values for w_{ix}, w_{iy}. It is intuitive to expect that the larger w_{ix}, w_{iy} we choose, the invariance conditions will be more likely to be solvable, as we allow the car to deviate more from our ideal trajectory. The trade-off is that choosing larger w_{ix}, w_{iy} results in larger polytopes $\mathcal{P}_i$. Unfortunately, there is no developed theory for the choices of w_{ix}, w_{iy}. The parameters that worked in our case were given as follows:

$$(w_{ix}, w_{iy}) = \begin{cases} (0, 0.09375\text{m}) & \text{if } i = 1, 8, \\ (0, 0.12\text{m}) & \text{if } 2 \leq i \leq 7. \end{cases} \tag{9.3}$$

We have now fully defined our sequence $\mathcal{P}_i$, $i = 1, \ldots, m$. It can be verified that the invariance conditions are solvable on each polytope $\mathcal{P}_i$. Thus, by Theorem 4.2 in [54], the RCP is solvable on each $\mathcal{P}_i$. What remains is to find a piecewise affine controller on each $\mathcal{P}_i$. This is done using Algorithm 4.11 of [50]. We note that our control set is bounded and given by $v_F \leq v \leq v_S$ and $-\varphi_M \leq \varphi \leq \varphi_M$. This fits the framework of [50], where the set of controls is required to be a polyhedron.

Control Design

Algorithm 4.11 of [50] requires that we first triangulate each polytope $\mathcal{P}_i$ into simplices. This can be performed in multiple ways. We used Delaunay triangulation with the addition of a centroid node (see, e.g., the work by [30] for more on Delaunay triangulation). After splitting polytope $\mathcal{P}_i$ into simplices $\mathcal{S}_1, \ldots, \mathcal{S}_s$, a control vector $u_v \in U$ such that $Av + Bu_v + a \in \mathcal{C}(v)$ is chosen at every point v in the union of vertex sets of $\mathcal{S}_1, \ldots, \mathcal{S}_s$. An affine feedback control $u_j : \mathcal{S}_j \to U$ is then defined by $u_j(v) = u_v$ for every vertex $v \subset \mathcal{S}_j$, and this is affinely extended to the entire $\mathcal{S}_j$. This defines a continuous piecewise affine control $u : \mathcal{P}_i \to U$ by defining the restriction of u on $\mathcal{S}_j$ as $u|_{\mathcal{S}_j} \equiv u_j$. Finally, we can use controls u defined on $\mathcal{P}_1, \ldots, \mathcal{P}_m$ to define $u : \cup_{i=1}^m \mathcal{P}_m \to U$. Note there is no guarantee that u is continuous on the boundary between $\mathcal{P}_i$ and $\mathcal{P}_{i+1}$, $i \in \{1, \ldots, m-1\}$. However, in our assignment of control values, u only has a discontinuity between $\mathcal{P}_4$ and $\mathcal{P}_5$. All simplex vertices, assigned control values, and affine controllers are available in Appendix A, as generated using MATLAB/YALMIP [79], following our discussion above.

This completes our design of a controller for the parallel parking maneuver.

9.1.4 Simulation Results

The procedure in Section 9.1.3 resulted in a hybrid affine system

$$\dot{x} = A_i x + B_i u(x) + a_i, \quad x \in \mathcal{P}_i, \tag{9.4}$$

defined on $\mathcal{M}' = \mathcal{P}_1 \cup \mathcal{P}_2 \ldots \cup \mathcal{P}_8$. The system data (A_i, B_i, a_i), $i \in \{1, \ldots, 8\}$ is obtained from the linearization (9.2) of the system (9.1) around points

$$x_i^L = (x_0, y_0, \min\{\theta^i, \theta^{i-1}\} - \varepsilon, (v_F + v_S)/2, 0).$$

This choice of linearization was justified in Section 9.1.3. Control $u : \mathcal{M} \to U$ is given by defining $u|_{\mathcal{P}_i}$ on each polytope as in Section 9.1.3.

Because of our choice (9.3) of widening parameters w_{iy}, the state space $\mathcal{M}' = \mathcal{P}_1 \cup \mathcal{P}_2 \ldots \cup \mathcal{P}_m$ becomes very large in y-coordinate by the time the trajectories reach $\mathcal{P}_m$. (See Figure 9.7 for a drawing.) The polytopes are placed far from the car parked in front of the parking space to ensure that there is no collision, and can be placed even further, but there is currently no theory available to guarantee that our vehicle will not hit the curb or end up too far from it. Reach control theory merely guarantees that, starting from *any* state in $\mathcal{M}'$, the system state will exit through $\mathcal{F}_m^{\text{out}}$. However, $\mathcal{F}_m^{\text{out}} \not\subset \mathcal{Q} \times \{0\}$. Thus,

there is no theoretical proof that Problem 9.1 is solved by our algorithm. On the other hand, we note that the specifications of our problem do not use reach control theory to the full extent, as the initial position and orientation of the car are actually guaranteed to be inside $\mathcal{I} \times \{0\} = \mathcal{F}_1^{\text{in}}$, which is just one facet of one polytope in $\mathcal{M}'$. Thus, the set of possible system states of the car at the point of leaving $\mathcal{P}_1$ is a proper subset of $\mathcal{F}_1^{\text{out}} = \mathcal{F}_2^{\text{in}}$. Simulations show that it is in fact much smaller than $\mathcal{F}_1^{\text{out}}$. As we then proceed further through subsequent polytopes, this brings us to the conclusion that the final states of the car make up only a small subset of $\mathcal{F}_m^{\text{out}}$. Hence, we achieve much better behaviour than guaranteed by reach control theory.

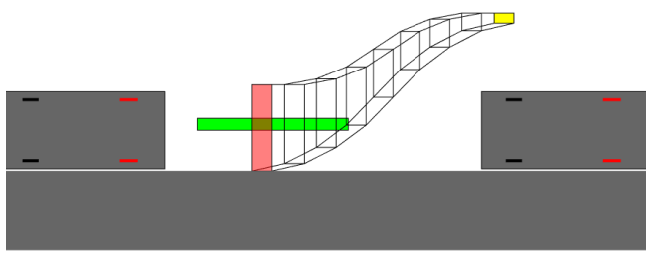

Figure 9.7: A projection of state space $\mathcal{M}'$ designed from parameters in Section 9.1.3 on the (x, y)-space. Polytopes $\mathcal{P}_i \subset \mathcal{M}'$ are shown in black. As in Figure 9.3, the initial box $\mathcal{I}$ is marked in yellow, the ending box $\mathcal{Q}$ in green, while the guaranteed final facet $\mathcal{F}_8^{\text{out}}$ is marked with red.

We now present the results of our simulations. We simulated the behaviour of 10 cars for each of the two models, nonlinear model (9.1) and hybrid model (9.4). The initial states of the cars were chosen at random in $\mathcal{I} \times \{0\}$. In the nonlinear model, to ensure that the system does not leave $\mathcal{M}'$ at the very beginning of the maneuver, we restricted the initial positions to be inside a 0.3m $\times$ 0.15m box in the middle of $\mathcal{I}$. The controller $u : \mathcal{M}' \to U$ remains the same as in the hybrid affine case. Figure 9.8 shows the results.

In the affine case, the only guarantee provided by our control strategy was that all the cars will stop inside $\mathcal{F}_8^{\text{out}}$. However, all the cars in fact stopped in close proximity to each other. In particular, all the x-coordinates of the stopped cars at the end of the maneuvers are almost exactly the same. This was not guaranteed by our construction of the final exit facet $\mathcal{F}_8^{\text{out}}$.

As reach control theory only pertains to the hybrid system (9.4), in the nonlinear case we had no guarantees for the behaviour of system state trajectories. However, Figure 9.8 shows that the linearization performed in Section 9.1.3 worked surprisingly well. In fact, while the trajectories from the linear model tend to end in the lower half of the box $\mathcal{Q}$, that does not occur in the nonlinear case. The differences between the two models could be made smaller by working with a larger number of polytopes m,

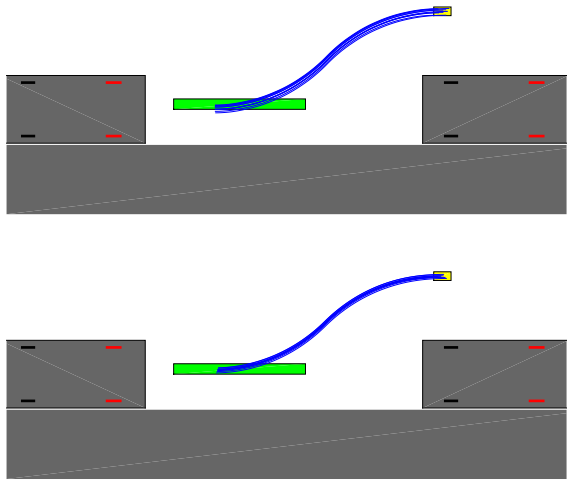

Figure 9.8: The positions of cars during the parallel parking maneuver. In the top image, the movement of cars is simulated using a hybrid affine model (9.4), whereas in the bottom image, the nonlinear model (9.1) is used. The desired ending box $\mathcal{Q}$ is given in green.

at the expense of computational power when calculating the feedback controller. However, we see that already for $m = 8$ the nonlinear system is approximated very well by the sequence of affine systems.

We now focus on a single initial state, where the car starts from the middle of the box $\mathcal{I} \times \{0\}$. This is also the situation from Figure 9.5 which informed our choice of polytopes $\mathcal{P}_i$. Figure 9.9 shows the positions and orientations a car goes through during the parking maneuver when following the nonlinear model (9.1).

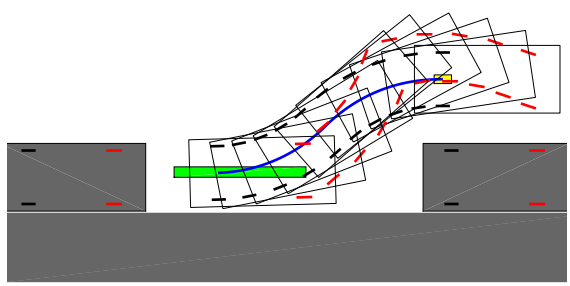

Figure 9.9: The trajectory of a car parking in the desired spot using the nonlinear model (9.1), with the car drawn at multiple instances during its parking procedure.

We note that, in this simulation, as well as in the ones presented in Figure 9.8, the vehicle undergoing the parking procedure does not collide with the vehicles in front or behind.

To illustrate the similarity between the trajectories obtained by the two model, in Figure 9.10 we compare the trajectory of a car from Figure 9.9 with the trajectory of a car starting from the same initial

condition, but obtained using the hybrid model (9.4). We additionally compare this to the intuitive, but not necessarily feasible, trajectory from Figure 9.5, which served as a motivation for our choice of polytopes $\mathcal{P}_i$.

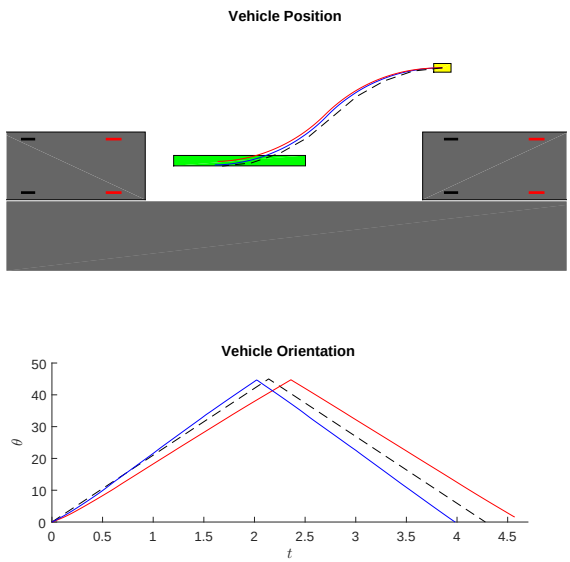

Figure 9.10: A comparison between two trajectories of a car parking in the desired spot using the hybrid model (9.4), with the midpoint of the rear axle drawn in blue, and the nonlinear model (9.1), with the midpoint of the rear axle drawn in red. The trajectory given by the dashed black line is the imaginary trajectory from Figure 9.5. The top picture gives the positional trajectory in (x, y)-coordinates, while the bottom gives the orientation θ of the vehicle with respect to time. The stopping time T of the imaginary trajectory from Figure 9.5 is set to the average of the stopping times (in seconds) of the trajectories obtained from (9.1) and (9.4).

We note that the trajectories are clearly similar. Additionally, it takes the vehicle around 4s to perform the maneuver. Thus, because there is no stopping during the procedure, the controller performs the parking maneuver much faster than a human driver can be expected to.

9.1.5 Conclusion

This work gives a robust control strategy for parallel parking which does not require the car to follow a specific previously calculated path. Our solution is based on a novel application of reach control theory. Extensive simulations show that the constructed feedback control results in the vehicle parking correctly in a desired position. We now list all assumptions that were made to obtain the presented solution.

- Nonlinear model (9.1) was linearized around non-equilibrium points to obtain a hybrid model (9.4). Additionally, the linearization used on each polytope $\mathcal{P}_i$ was obtained by linearizing around a point that is outside of $\mathcal{P}_i$.
- State space $\mathcal{M}'$ was chosen heuristically, based on an imaginary ideal trajectory, and without a guarantee of a correct parking maneuver.
- Controllers $u : \mathcal{P}_i \to U$ were not calculated to ensure the correctness of a parking maneuver. They were not even chosen by hand to match our expectations. The only criterion in place was that u solved the invariance conditions on $\mathcal{P}_i$. During simulations, we also required u to be continuous on $\mathcal{M}'$ except between $\mathcal{P}_4$ and $\mathcal{P}_5$. A controller automatically chosen by MATLAB/YALMIP that satisfied those criteria was used.

The fact that our model went through this many crude assumptions and still produced a correct parking maneuver is surprising and indicative of a deeper theory behind this problem. This work is the first step towards exploring this theory. In particular, the fact that we let an automated optimization tool choose any controller that satisfied the invariance conditions and the maneuver still worked in a desired way suggests a yet unexplored stability property for the RCP. Moreover, in the linearized model, the procedure ended up aligning all the cars to stop at virtually the same x-coordinate, even though their starting x-coordinates were up to 0.5m apart.

The construction of an appropriate state space $\mathcal{M}'$ is another limitation of our current approach. While the $\mathcal{M}'$ that we chose is guaranteed to be invariant under piecewise affine feedback, it is too large to theoretically ensure that the vehicle will stop in a correct position. A potential solution would be dynamic generation of $\mathcal{M}' = \mathcal{P}_1 \cup \ldots \cup \mathcal{P}_m$: when the car reaches a point at an exit facet of polytope $\mathcal{P}_i$, a smaller $\mathcal{P}_{i+1}$ could be constructed on the fly using the information on the current position of the car. This is another future research topic.

Finally, let us briefly discuss an obvious limitation resulting from the assumptions taken in this work. While the approach proposed above is robust to small disturbances in state, we do not claim it is able to directly deal with more significant changes to the maneuver set-up. In particular, a parallel parking maneuver often starts from a less perfect position, i.e., not perfectly straight, or outside the initial box $\mathcal{I}$. In such a case, a common approach of human drivers is to add additional steps in which the vehicle moves back and forth before it finally reaches a desired position. The method proposed in this section clearly does not deal with this scenario. In theory, it would be possible to deal with this situation by an approach similar to the method presented above: design a non-feasible ideal trajectory, and then design a set of boxes around it, and a controller on each box, which finally produce a feasible

trajectory similar to the ideal one. However, this would require a development of a more systematic procedure to perform the above steps. In the approach proposed in this chapter, the ideal trajectory was constructed manually, the number of boxes around it was chosen without any theoretical foundation, and, as discussed in Section 9.1.3, the box sizes were chosen heuristically to be as small as they can, as long as the invariance conditions hold. In the case of a significantly different initial position of the vehicle, significantly different trajectories may be required, and it would be difficult to design by hand different classes of trajectories to cover all initial conditions. Thus, an automated approach may be necessary, once again providing a motivation for further formal development of the underlying theory. If the above generalization was successfully

9.2 Adaptive Cruise Control

9.2.1 Introduction

Automated parallel parking explored in Section 9.1 is a feature still largely under development in commercial vehicles. However, the desire for less dependency on human operation and higher standards in comfort and safety has long motivated the implementation of traditional control techniques in vehicles. This can be evidenced since the early fifties by the emergence of Conventional Cruise Control (CCC) [118], a driver support system designed to track a speed set by the driver. In the mid-nineties, CCC was extended to support Adaptive Cruise Control (ACC) [70], a system which seeks to control a vehicle's speed based on the presence of a preceding car while maintaining a safe distance between the two vehicles. ACC operates by using sensors solely within the actuated vehicle. No outside infrastructure such as beacons, knowledge about the outside environment such as a road map, or vehicle-to-vehicle communication is needed, although such systems are presently being studied as well [4, 113, 131].

While there are multiple proposed implementations of ACC (e.g., [74, 111, 137]), in this introduction we focus on a correct-by-construction ACC design was proposed in [97]. This control was implemented on actual vehicles in [86]. The synthesis proposed in [97] uses Linear Temporal Logic (LTL) (described in, e.g., [114, 132]) to model the desired vehicle dynamics. Two approaches to solving the ACC problem are proposed in [97]: a discrete abstraction of the system with Pessoa [85] and a Model Predictive Control method to compute a chain of sets that reach the desired goal state.

In this section, we propose our own solution to the ACC problem that uses the reach control methodology. Hence, we triangulate the state space into simplices, and then define an affine controller on each of the simplices to satisfy the control specifications. While we follow the ACC model of [97], there are

significant differences between the two approaches. Both methods proposed in [97] rely on discretizing the state space and/or time of the original continuous model for ACC. This can be extremely computationally complex, as witnessed in Section IV of [97]. Additionally, neither of the two approaches in [97] provide guarantees for the behaviour of the continuous system between the sampling times.

In contrast, our model does not contain discretizations, and calculating the desired controller for any speed of the lead vehicle is computationally easy. Unlike the Pessoa-based method described in [97], there is no need for the simplices in our triangulation of the state space to be small. We will show that it is possible to control the system to its desired goal state by using at most 8 simplices, while satisfying the constraints imposed by the ACC specifications throughout the system run. The computational complexity of our method is thus negligible, and we provide guarantees for correct behaviour of the system at all times. As mentioned in Chapter 1, reach control has been used to achieve similar objectives in the past (see, e.g., [82, 107, 133, 134, 135]). However, all of those papers consider a static state space and a triangulation that does not change over time. This work introduces a novel element of reach control theory: both the state space and its triangulation will change over time, depending on the speed of the lead vehicle. Our simulations will show that a reach control approach continues to guarantee safe behaviour of the ACC system. The theory of reach control in the situation when the state space and control specifications change over time has not been developed. The work presented here is the first foray into that question. Additionally, as the triangulations will now be automatically calculated online during the run of the system, instead of a priori by hand, this work is a first step towards the systematization of the state space triangulation, which is a critical bottleneck for reach control [77, 133].

9.2.2 Model and Problem Statement

Model of a Single Vehicle

The model we are using closely follows the model of the dynamics of a one- and two-vehicle system in [97]. The actuated vehicle is modeled as a point mass m moving along a straight line at some speed v. The equations of motion are given by

$$m\dot{v} = F_u - F_f, \tag{9.5}$$

where F_u is the net braking action and engine torque exerted on the vehicle, and F_f is the net friction force. The friction force is modeled as

$$F_f = f_0 + f_1 v + f_2 v^2. \tag{9.6}$$

The engine torque and braking action used by the vehicle are limited. Thus, the input F_u is bounded. We take

$$F_{br} \leq F_u \leq F_{ac}, \tag{9.7}$$

where F_{br} and F_{ac} are maximal braking and acceleration forces, respectively. In the context of cruise control, these are not necessarily the maximal forces that the vehicle can produce, but also take into account the passengers' comfort and convenience.

Model of a Two-Vehicle System

The purpose of adaptive cruise control is to maintain a safe and efficient driving profile in the presence of a moving vehicle ahead. Thus, in addition to the previous single-vehicle dynamics, we need to consider the headway h between the two vehicles: the lead vehicle and the one under our control. We assume the lead vehicle is moving at a known speed v_L. We allow v_L to change over time, i.e. $v_L : [0, +\infty) \to [v_{min}, v_{max}]$, where v_{min} and v_{max} are some reasonable constant minimal and maximal vehicle speeds. These depend on the type of highway, weather conditions, and other characteristics.

The dynamics of the headway in the two-vehicle system can be described by:

$$\dot{h} = v_L - v. \tag{9.8}$$

By combining (9.5) and (9.8) we obtain the following two-vehicle dynamics:

$$\begin{pmatrix} \dot{v} \\ \dot{h} \end{pmatrix} = \begin{pmatrix} -\frac{f_1}{m} & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} v \\ h \end{pmatrix} + \begin{pmatrix} \frac{1}{m} \\ 0 \end{pmatrix} F_u + \begin{pmatrix} \frac{-f_0}{m} \\ v_L \end{pmatrix} + \begin{pmatrix} \frac{-f_2}{m} v^2 \\ 0 \end{pmatrix}. \tag{9.9}$$

We note that, for a constant v_L, the system (9.9) is affine except for the quadratic element $-(f_2/m)v^2$ at the end. However, the value of f_2/m, which we give in Section 9.2.5 is extremely small in practice. Thus, we will approximate the friction $F_f = f_0 + f_1 v + f_2 v^2$ by its linearization F'_f around $v = v_0 := (v_{min} + v_{max})/2$. We obtain $F'_f = (f_0 + f_1 v_0 + f_2 v_0^2) + (2 f_2 v_0 + f_1)(v - v_0)$. This results in the following system

$$\begin{pmatrix} \dot{v} \\ \dot{h} \end{pmatrix} = \begin{pmatrix} -\frac{f_1}{m} - 2\frac{f_2}{m} v_0 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} v \\ h \end{pmatrix} + \begin{pmatrix} \frac{1}{m} \\ 0 \end{pmatrix} F_u + \begin{pmatrix} \frac{-f_0}{m} + \frac{f_2}{m} v_0^2 \\ v_L \end{pmatrix}. \tag{9.10}$$

If v_L is taken to be constant, system (9.10) is an affine control system with the control input F_u.

Thus, it is amenable to a reach control approach. The simulation results presented in Section 9.2.5 will show that there is no substantial difference between models (9.9) and (9.10). Further discussion of this linearization is found in [97], where a method for compensating for the nonlinearity was also developed. In this and the subsequent section, we proceed with our methodology based on system (9.10).

Problem Statement

We again follow [97] for the exposition of the ACC specifications. We first introduce the concept of *time headway*, $\tau = h/v$. Time headway represents the time required for a moving vehicle with velocity v to reach a stationary object at distance h from the vehicle. It is crucial for expressing safety constraints on the system dynamics. We now summarize the ACC specifications from the ISO document [64]:

1) ACC operates in two modes: *no lead car mode* and *lead car mode.*

2) In no lead car mode, a preset desired speed v_{des} eventually needs to be reached and maintained.

3) In lead car mode, a desired lower bound on safe time headway τ_{des} to the lead vehicle and an upper bound on a desired velocity v_{des} eventually need to be reached and maintained.

4) The time headway τ must be greater than $\tau_{min} = 1$s at all times.

5) Independently of the mode, the input F_u satisfies $F_{br} \leq F_u \leq F_{ac}$.

The desired behaviour of the vehicle in the no lead car mode is trivially achievable using a variety of classical cruise control techniques, as well as reach control theory. Thus, we will not be discussing it. Instead, we will impose a soft constraint $h \leq h_{max}$. It is not desirable for a car to fall behind the lead vehicle by more than h_{max}. However, if this does occur, it is permissible as it does not expose the car and its driver to any danger. Essentially, this situation switches the ACC from the lead car mode to the no lead car mode.

Specification 4) above, combined with the specification $h \leq h_{max}$, defines the shape of our constrained state space. It is given by

$$\mathcal{M} = \{(v, h) \mid v_{min} \leq v \leq v_{max}, \tau_{min} v \leq h \leq h_{max}\}. \tag{9.11}$$

Specification 5) defines the space of admissible control inputs by

$$\mathcal{U} = \{F_u \mid F_{br} \leq F_u \leq F_{ac}\}.$$

Finally, specification 3) above defines the desired goal set $\hat{\mathcal{G}} = \{(v,h) \in \mathcal{M} \mid v \leq v_{des}, h \geq \tau_{des} v\}$. If $v_{des} < v_L$, the car will, upon reaching the goal set, start falling behind the lead vehicle and leave $\mathcal{M}$. Thus, we will not be discussing this case, although it can be trivially solved with methods analogous to the ones described in the remainder of the section. Hence, we assume $v_{des} > v_L$ and set the goal set

$$\mathcal{G} = \{(v,h) \in \mathcal{M} \mid v = v_L, h \geq \tau_{des} v\} \subset \hat{\mathcal{G}}.$$

Note that, in the case that v_L is variable, $\mathcal{G}$ varies with time as well. On the other hand, the state and control spaces $\mathcal{M}$ and $\mathcal{U}$ remain the same regardless of the value of $v_L \in [v_{min}, v_{max}]$. Figure 9.11 shows the state space $\mathcal{M}$ and the goal space $\mathcal{G}$ for some value of v_L.

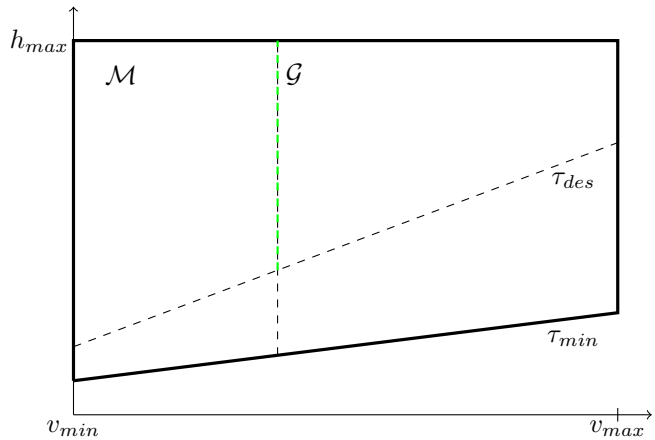

Figure 9.11: An illustration of the state space in the ACC problem. The state space $\mathcal{M}$ is a polytope bounded by lines $v = v_{min}$, $v = v_{max}$, $h/v = \tau_{min}$ and $h = h_{max}$. Desired time headway τ_{des} and lead vehicle speed v_L are marked by dashed lines. The goal set $\mathcal{G}$ is marked in green.

For $x_0 \in \mathcal{M}$ and a control function F_u, let $\phi_{F_u}(\cdot, x_0)$ be the trajectory of the system (9.10) with $\phi_{F_u}(0, x_0) = x_0$. We formulate the ACC problem as follows:

Problem 9.2. *Find a control function F_u taking values in $\mathcal{U}$ such that, for all $x_0 \in \mathcal{M}$, the following holds:*

(i) $\phi_{F_u}(t, x_0) \in \mathcal{M}$ *for all* $t > 0$,

(ii) $\lim_{t \to +\infty} \phi_{F_u}(t, x_0) \in \mathcal{G}$.

9.2.3 Constant Speed Lead Vehicle

Let us examine the state space polytope $\mathcal{M}$ given in Figure 9.11. We would like to pursue a reach control approach to reach the goal set $\mathcal{G}$. Unfortunately, that may not be possible to achieve for the entire $\mathcal{M}$.

Consider the vertex $v^* = (v_{max}, \tau_{min} v_{max})$. This is the point at which the vehicle is moving at a high speed and is very close to the lead vehicle. If the lead vehicle is slow, the maximal braking effort might still not be sufficient to keep the system state within $\mathcal{M}$. In the context of reach control theory, this means that the invariance condition (3.4) would not be solvable at v^* for any control $u \in [F_{br}, F_{ac}]$.

Hence, we need to remove a part of $\mathcal{M}$ where it is impossible to satisfy the invariance conditions. A method for determining which parts of $\mathcal{M}$ need to be removed is discussed in [97]. For simplicity, we will remove a triangle $\mathcal{S}_0$ with vertices $(v_L, \tau_{min} v_L)$, $(v_{max}, \tau_{min} v_{max})$ and $(v_{max}, \tau_{min} v_L + b_{max})$, where

$$
\begin{aligned}
b_{max} &= \frac{m(v_{max} - v_L)^2}{f_0' + f_1' \frac{v_{min}+v_{max}}{2} - F_{br}}, \\
f_0' &= f_0 + f_1 v_0 + f_2 v_0^2, \\
f_1' &= 2 f_2 v_0 + f_1.
\end{aligned}
$$

These values have been calculated so that the invariance conditions (3.4) are made feasible. We assume that $\tau_{min} v_L + b_{max} < \tau_{max} v_{max}$ as it would otherwise be impossible to safely break when the car is at the speed v_{max}.

Having removed $\mathcal{S}_0$, we now triangulate the remainder of the state space $\mathcal{M}$. The triangulation used is driven by our desired control strategy. We distinguish between four cases:

1° if the controlled vehicle is faster than the lead vehicle, and is following it a reasonably large distance, it can slow down to v_L while staying within the desired headway,

2° if the controlled vehicle is faster than the lead vehicle, but the distance between the two vehicles is small, the plan is to first slow down to below v_L, and then gradually increase the headway (see the next case),

3° if the controlled vehicle is slower than the lead vehicle, and the headway is small, it can increase its speed to v_L while also reaching the desired headway,

4° if the controlled vehicle is slower than the lead vehicle, and the distance between the two vehicles is large, it may not be possible to speed up to v_L fast enough without reaching the maximum headway h_{max} first. This is a suboptimal, but safe scenario discussed in Section 9.2.2. In that case, the vehicle should just increase its speed.

We now define vertices $v_1, \ldots, v_9$ to generate a triangulation of the state space $\mathcal{M} \backslash \mathcal{S}_0$. Analogously to b_{max} above, define

$$
b_{min} = \frac{m(v_{min} - v_L)^2}{F_{ac} - f_0' - f_1' \frac{v_{min}-v_{max}}{2}}.
$$

The coordinates of vertices are then given as follows:

$$
\begin{aligned}
v_1 &= (v_{min}, \tau_{min} v_{min}),\\
v_2 &= (v_{min}, h_{max} - b_{min}),\\
v_3 &= (v_{min}, h_{max}),\\
v_4 &= (v_L, \tau_{min} v_L),\\
v_5 &= (v_L, \tau_{des} v_L),\\
v_6 &= (v_L, h_{max}),\\
v_7 &= (v_{max}, \tau_{min} v_L + b_{max}),\\
v_8 &= (v_{max}, \min\{\tau_{des} v_L + b_{max}, h_{max}\}),\\
v_9 &= (v_{max}, h_{max}).
\end{aligned}
$$

We note that vertices v_1, v_3, v_4, v_7, and v_9 are vertices of the trimmed state space $\mathcal{M}\backslash\mathcal{S}_0$. Vertices v_5 and v_6 are vertices of the goal set $\mathcal{G}$. Vertices v_2 and v_8 have been chosen so that the segments $\overline{v_2 v_6}$ and $\overline{v_5 v_8}$ form a boundary between cases 1° and 2° and cases 3° and 4°, respectively. In the context of the RCP, this corresponds to ensuring that the invariance conditions (3.4) are satisfied on all simplices for desired exit facets. Our triangulation of $\mathcal{M}$ is given in Figure 9.12. On each simplex we also denoted an exit facet that the system states are permitted to go through, or an equilibrium point that we want to reach, corresponding to our stated control strategy above. We note that this set-up is clearly similar to the prototypical example of reach control from Figure 1.1.

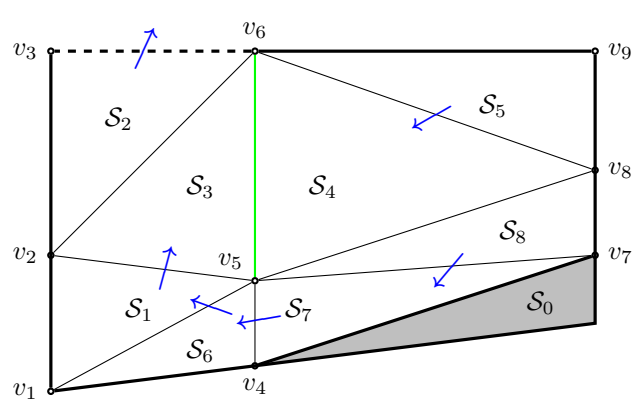

Figure 9.12: A triangulation of the state space $\mathcal{M}$ used in the ACC problem. The desired equilibrium points for our controller, coinciding with the goal set $\mathcal{G}$ are marked in green. Exit facets of simplices are marked by blue arrows.

We note that simplices $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_5, \ldots, \mathcal{S}_8$ contain both an exit facet and an equilibrium point. In those cases, we wish to allow that a trajectory either converges either to an equilibrium point, or leaves the simplex through an exit facet. Thus, our choices of exit facets and equilibria in Figure 9.12 defines the transition graph given in Figure 9.13. This graph contains no cycles. Hence, it shows that the system states starting in $\mathcal{S}_i$, $i = 1, 3, \ldots, 8$ will ultimately converge to an equilibrium in $\mathcal{G}$, while the system states starting in $\mathcal{S}_2$ will either converge to an equilibrium in $\mathcal{G}$ or exit $\mathcal{M}$ in a safe way.

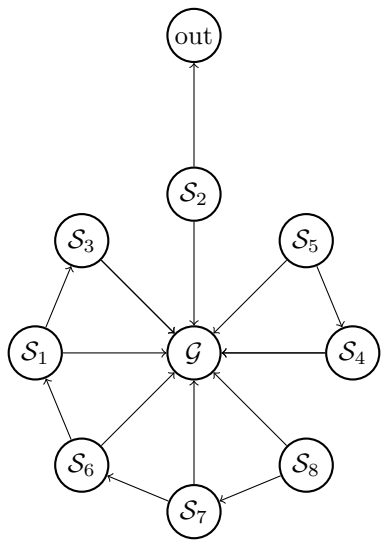

Figure 9.13: An illustration of the transitions between simplices and final states in $\mathcal{M}$. An edge from $\mathcal{S}_i$ and $\mathcal{S}_j$ indicates that, after leaving the interior of $\mathcal{S}_i$, a system state can cross into $\mathcal{S}_j$. An edge from $\mathcal{S}_i$ to $\mathcal{G}$ indicates that a system state can converge to one of the equilibria in $\mathcal{G}$ without ever leaving $\mathcal{S}_i$. An edge from $\mathcal{S}_i$ to out indicates that a system state may exit $\mathcal{S}_i$ by achieving $h > h_{max}$ and hence leaving $\mathcal{M}$.

The final thing that remains is to design controls on each simplex $\mathcal{S}_i$, $i = 1, \ldots, 8$, to ensure the desired state behaviour. We propose these control values F_{ui} at vertices v_i:

$$F_{u1}, F_{u2}, F_{u3} = F_{ac},$$

$$F_{u4}, F_{u7}, F_{u8}, F_{u9} = F_{br},$$

$$F_{u5}, F_{u6} = F'_f,$$

where F'_f is the force which results in the vehicle speed not changing from v_L. The control law $u : \mathcal{M} \backslash \mathcal{S}_0 \to [F_{br}, F_{ac}]$ is then obtained by affinely extended the above controls to each triangle $\mathcal{S}_i$, as in Section 9.1.3. It can be computationally verified that all of these input values indeed satisfy the desired invariance conditions. Direct calculations also show that $Ax + Bu(x) + a = 0$ if and only if $x \in \mathcal{G}$.

Finally, in order to guarantee that our proposed control law indeed satisfies the control objectives, let us discuss a small theoretical detail. The set-up of simplices $\mathcal{S}_i$ is not exactly the same as the RCP set-up in Section 3.1. In Section 3.1, simplex $\mathcal{S}$ had an exit facet and no desired equilibria. In our case, there are two different situations:

- The desired behaviour on each of the simplices $\mathcal{S}_i$, $i = 1, 2, 5, \ldots, 8$ requires the trajectory to either leave a simplex through an facet *or* converge to a single equilibrium $e \in \mathcal{G}$ which is on the boundary of $\mathcal{S}_i$. As mentioned above, it can be shown that the controller on $\mathcal{S}_i$ proposed above solve the invariance conditions (3.4). By Lemma 4.17, if a control input satisfies the invariance conditions, all trajectories that do leave a simplex indeed leave it through a designated facet. Thus, *if* a trajectory leaves a simplex $\mathcal{S}_i$, it will leave it through a desired exit facet.

  As mentioned above, the proposed controller also results in (3.1) on $\mathcal{S}_i$ containing a single equilibrium at e. Hence, because (3.1) is a two-dimensional affine system with a single equilibrium on the boundary of $\mathcal{S}_i$, if a trajectory does not leave $\mathcal{S}_i$, it necessarily converges to this equilibrium. The trajectory cannot diverge because it stays inside the simplex, and it cannot be periodic because that would require it to move around the equilibrium, which is not possible because e is on the boundary of $\mathcal{S}_i$.

- The desired behaviour on simplices $\mathcal{S}_i$, $i = 3, 4$, is to converge towards the facet $\mathcal{G}$ without leaving the simplex. As mentioned, we can show that the control law proposed above solves the invariance conditions on $\mathcal{S}_i$ (with $\mathcal{G}$ defined as the "exit facet"). Thus, by Lemma 4.17, the trajectories will either leave through $\mathcal{G}$ or not leave $\mathcal{S}_i$ at all. However, since $\mathcal{G}$ consists solely of equilibria of the system (3.1), it is impossible to leave through $\mathcal{G}$. Hence, all trajectories remain inside $\mathcal{S}_i$.

  Moreover, since (3.1) on $\mathcal{S}_i$ is a two-dimensional affine system where an entire segment $\mathcal{G}$ is an equilibrium, it can easily be algebraically shown that the trajectories lie on straight lines. Since these trajectories need to remain inside $\mathcal{S}_i$ and there are no equilibria except on $\mathcal{G}$, the only option is that each trajectory converges to a point in $\mathcal{G}$.

The above discussion guarantees the correctness of the proposed control law.

9.2.4 Variable Speed Lead Vehicle

The above construction of the feedback control $u = u(x)$ on the state space $\mathcal{M}$ provides a correct-by-construction solution to the adaptive cruise control problem in the case when the lead vehicle speed v_L is constant. We now generalize this approach to a variable speed $v_L(t)$. We note that the triangulation

defined in Figure 9.12 is with respect to v_L and is valid for any $v_{min} \leq v_L \leq v_{max}$. The control inputs F_{uj} defined at vertices v_j are also parametrized with respect to v_L. Thus, if $v_L = v_L(t)$ is a time-varying function, at every time $t \geq 0$ we can generate a triangulation $\mathbb{T}(t)$ of the state space $\mathcal{M}$, and using $\mathbb{T}(t)$ define a control function $u(x,t)$. The idea is that at every time instant t_0, the control $u(\cdot, t_0)$ will satisfy the invariance conditions imposed by a triangulation $\mathbb{T}(t_0)$ and will be driving system trajectories to converge to the goal set $\mathcal{G}(t_0)$. However, while the correctness of the controller $u(\cdot, t_0)$ is guaranteed for every fixed t_0, there is no guarantee that the time-varying controller $u(\cdot, t)$ is still correct. Such a theory has not yet been developed in reach control; this work is the first step towards developing it. The simulations presented in Section 9.2.5 do show that the above time-varying strategy indeed works.

9.2.5 Simulation Results

In the following simulations, we use the following parameter values from [97]: $m = 1370$ kg, $f_0 = 3.8 \cdot 10^{-3} \cdot mg$ N, $f_1 = 2.6 \cdot 10^{-5} \cdot mg$ Ns/m, $f_2 = 0.4161$ Ns^2/m^2, $F_{br} = -0.3mg$ N, $F_{ac} = 0.2mg$ N. We additionally use $v_{min} = 15$ m/s, $v_{max} = 35$ m/s, $\tau_{min} = 1$s, $\tau_{des} = 2$s, $h_{max} = 300$m.

Vehicle Merging in Front

In this scenario, the vehicle we are controlling starts from a point in $\mathcal{G}$: $(v_0, h_0) = (v_L(0), (v_L(0)\tau_{des} + h_{max})/2) \in \mathcal{G}(0)$. The lead vehicle moves at a constant speed $v_L(t) = 30$, $0 \leq t \leq 10$. At time $T = 10$, a second vehicle merges from a neighbouring lane (see Figure 9.14). Hence, v_L now becomes the speed of the new lead vehicle, and is given by $v_L(t) = 25$, $t > 10$. Additionally, since the new vehicle merged into the lane, it instantaneously reduced the time headway τ at time T to $\tau(T+) = (\tau_{min} + \tau_{des})/2$.

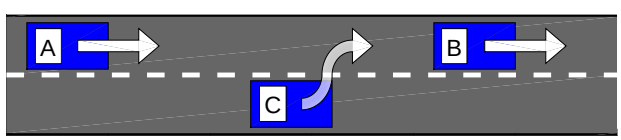

Figure 9.14: An illustration of the maneuver performed by two lead vehicles in Section 9.2.5. Vehicle A which is controlled by our ACC approach is originally following the lead vehicle B. The lead vehicle B speeds up suddenly, and vehicle C merges into the lane, becoming a new lead vehicle.

We note that in this case, both v_L and h are discontinuous, with breaks at time $T = 10$. The change in v_L requires the triangulation $\mathbb{T}(t)$ to be recalculated at time T. Figure 9.15 shows the triangulations $\mathbb{T}(t)$ before and after a new vehicle merged into the lane.

In this simulation our car follows the nonlinear model (9.9). Unlike system (9.10), there are no

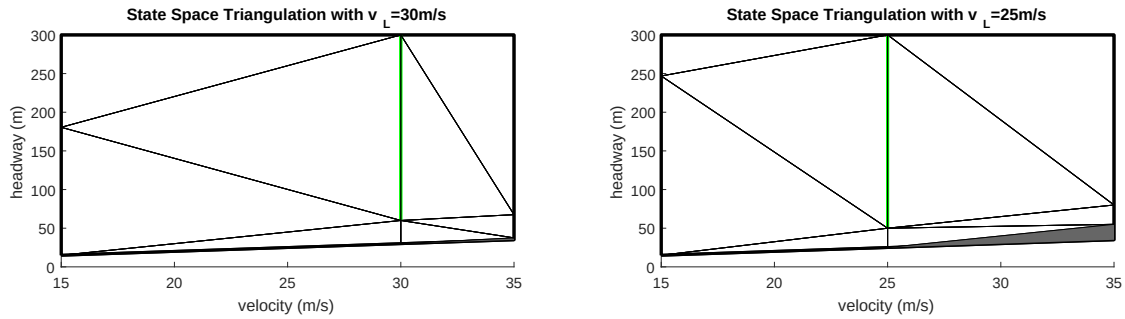

Figure 9.15: Triangulations of the state space $\mathcal{M}$ computed in Section 9.2.5. The left figure shows $\mathbb{T}(t)$ for $0 \leq t \leq 10$. The right figure shows $\mathbb{T}(t)$, $t > 10$. The discarded simplex $\mathcal{S}_0$ is denoted in grey.

guarantees that applying the feedback control $u(x,t)$ developed for the affine system (9.10) will result in a correct behaviour. However, the nonlinear system is well-approximated by an affine model. This is a result of the nonlinear factor $(v - v_0)^2 f_2/m$ being small in magnitude. Since the difference between trajectories produced by systems (9.9) and (9.10) are negligible, we only present the results of the simulations for system (9.9). We again simulate the behaviour of system (9.9) under such a scenario. The results are given in Figure 9.16.

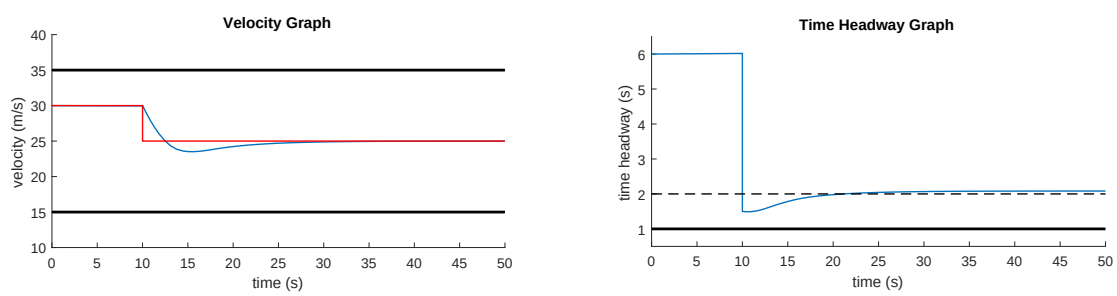

Figure 9.16: The results of a simulation involving a two lead vehicles with constant speed. The left graph shows the lead vehicle speed $v_L(t)$ over time (marked in red), as well as the speed of our vehicle $v(t)$ (marked in blue). The minimal and maximal velocity v_{min} and v_{max} are represented by solid black lines. The right graph shows the time headway $\tau(t)$, with the minimal allowed time headway τ_{min} represented by a solid black line and the minimal desired time headway τ_{des} represented by a dashed line.

We see our controller performs well and tracks the speed of the lead vehicle while staying within the required safety envelope. Apart from minimal nonlinearity issues, this was guaranteed by the reach control theory.

We note that the actions of the merging vehicle in this section did not result in our vehicle being placed in an unsafe position. Hence, our ACC strategy could still be performed, just with new initial conditions $(v(T), v(T)\tau(T+)) \in \mathcal{M}$. Had the new lead vehicle merged into the lane in such a way that

τ became smaller than the minimal safe headway τ_{min}, our controller would not know what action to perform, as the state (v, h) would no longer be in $\mathcal{M}$ after time T. Intuitively, a smart thing to do is to initiate an emergency braking procedure, but this is not covered by the ACC specifications that we are using.

Single Lead Vehicle With Nonconstant Speed

In this scenario, the vehicle we are controlling again follows the nonlinear model (9.9) and starts with the speed $v_0 = 25$, at $h_0 = 37.5$ meters behind the car in front. The car in front behaves according to the following velocity profile:

$$v_L = \begin{cases} 20 + t, & 0 \leq t \leq 15, \\ v_{max}, & 15 \leq t \leq 30, \\ v_{max} - (t - 30), & 30 \leq t \leq 50, \\ v_{min}, & 50 \leq t \leq 65. \end{cases}$$

The results of the simulation are presented in Figure 9.17.

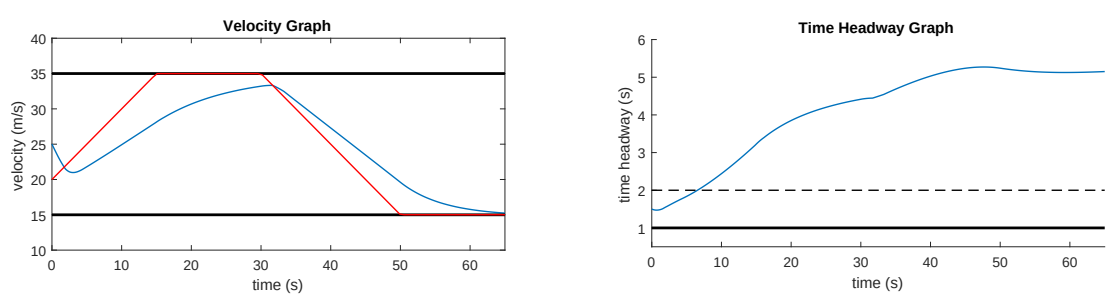

Figure 9.17: The results of a simulation involving a single lead vehicle with nonconstant speed. The left graph shows the lead vehicle speed $v_L(t)$ over time and the speed of our vehicle $v(t)$. The right graph shows the time headway $\tau(t)$. As in Figure 9.16, minimal and maximal allowed velocity v_{min} and v_{max} are represented by solid black lines in the left graph, while in the right graph, the solid line represents the minimal allowed time headway τ_{min}, and the dashed line represents the minimal desired time headway τ_{des}.

As we can see in Figure 9.17, $v(t)$ tracks $v_L(t)$ extremely well. In the segments when v_L is constant, $v(t)$ clearly converges towards v_L. Additionally, the vehicle speed remains between v_{min} and v_{max} at all times, and the time headway τ remains between τ_{min} and τ_{max}, eventually converging to τ_{des}. However, unlike in Section 9.2.5, as the lead vehicle speed is constantly changing, there currently do not exist any

theoretical guarantees for such behaviour.

9.2.6 Conclusion

This section studied the control of a single ACC-enabled vehicle with strict safety specifications. We proposed a correct-by-construction software design approach for an ACC controller using the reach control theory. This design relies on an online calculation of a desired triangulation of the state space $\mathcal{M}$, which is a novel feature not present in previous theory. This is an initial step towards a systematic method of obtaining an appropriate triangulation for reach control. The proposed control strategy allows a lead vehicle with a non-constant speed, as well as a lead vehicle changing lanes. While extensive simulations have shown that the proposed controller indeed meets all the ACC specifications, its correctness is theoretically verified only in the case when the lead vehicle is moving at a constant speed. This is another significant topic of further research.

Chapter 10

Concluding Remarks

In the introductory chapter, we remarked that the contributions of this thesis can be described as belonging to three general areas: results on the existence, uniqueness and behaviour of system trajectories in the RCP, results on the solvability of the RCP, and applications using reach control. At the end of this thesis, perhaps it is more apt to describe a softer division. By their purpose, the problems that this thesis investigates, and the results it presents, can be divided into three parts:

- Building a rigorous mathematical foundation to reach control theory,
- Connecting reach control and RCP theory to previously explored areas in systems control and mathematics, and using these connections for new insights in reach control,
- Motivating new research directions in reach control through applications.

The last of the above elements naturally lends itself into a discussion of future work. For the particular applications introduced in Chapter 9, future theoretical work motivated by these applications has already been discussed in this chapter. Let us, thus, briefly illuminate the bigger picture. It is clear that reach control is not a complete theory. While this thesis makes a substantial contribution towards developing it, let us list some of the challenges still facing reach control theory:

- It is not clear how to triangulate the state space in the best way, i.e., a way that ensures that there is a feasible reach control path to satisfy the control objectives, if such a way exists. There are some insights on how to do it by hand in particular examples (see, e.g., [133] and the applications in Chapter 9), but an algorithmic way of finding an optimal triangulation is still nonexistent. An algorithm for a triangulation of a parametrized state space for the particular example of Section

9.2 is developed in that section, but that algorithm is not extendable to any other scenario. Generally, we note that the assumptions taken when solving the RCP constrain possible triangulations, implying, e.g., that $\mathcal{O}_{\mathcal{S}}$ needs to be contained in a simplex facet (or simplex interior). A possible step forward would be to develop a systematic way of generating a triangulation that is guaranteed to satisfy those constraints. However, such an approach might be overly constraining, because it would be based on currently developed sufficient conditions for the solvability of the RCP. This, thus, clearly motivates the need for further development of the low-level theory.

- It is not even clear how to determine whether a given triangulation contains a feasible path to satisfy the reach control objectives. The solution to this problem banks on determining computationally feasible sufficient and necessary conditions for the solvability of the RCP on a given simplex. While there exist some sufficient and necessary conditions for the solvability of the RCP in the case of affine feedback, and our thesis introduced some novel conditions as well (particularly in Chapter 7), these conditions are largely computationally infeasible, except in limited special cases. In the case of continuous state feedback, there does not even exist a known set of sufficient and necessary conditions for the solvability of the RCP. Necessary conditions, such as the ones in Chapter 5 and Chapter 6, are mathematically valuable in characterizing the nature of an obstacle to solvability, but they are also computationally difficult to verify, and are in no way sufficient for determining the solvability of the RCP.

- While the idea of reach control is applicable to any class of control systems in continuous state space and continuous time, all the previous research on the RCP, as the building block of reach control, has been on affine systems. These systems have a particular structure, and there are generally no guarantees on the behaviour of nonlinear systems when reach control is applied to their linear approximations. There is also no systematic way of deriving these linearizations. This topic was explored to some extent in Section 9.1, but this was done only for a single application, and no theoretical bounds were provided.

Some other research directions were alluded to in the previous chapters (e.g., analysis of reach control in time-varying state spaces). The list above contains three elements which, in the author's opinion, require significant research effort to resolve. While this thesis was mathematically instructive in the sense of problem formalization and exploring the underlying structure of the RCP, we make no attempt to claim that it resolved any of the three above issues. If anything, perhaps the underlying contribution of this entire thesis was exactly to show, as mathematically formally as possible, *why* reach control is so difficult. In hindsight, I believe this should be the framework in which this thesis should be observed.

Bibliography

[1] Alessandro Abate and Ashish Tiwari. Box invariance of hybrid and switched systems. In *2nd IFAC Conference on Analysis and Design of Hybrid Systems*, pages 359–364, 2006.

[2] Aaron Ames and S. Shankar Sastry. Characterization of Zeno behavior in hybrid systems using homological methods. In *2005 American Control Conference*, pages 1160–1165, 2005.

[3] Jan Andres. Topological principles for ordinary differential equations. In Antonio Cañada, Pavel Drábek, and Alessandro Fonda, editors, *Handbook of Differential Equations: Ordinary Differential Equations*, volume 3, pages 1–101. Elsevier, 2006.

[4] Francisco Aparicio, Javier Paez, Felix Moreno, Felipe Jimenez, and Alberto Lopez. Discussion of a new adaptive speed control system incorporating the geometric characteristics of the roadway. *International Journal of Vehicle Autonomous Systems*, 3(1):47–64, 2005.

[5] Graeme Ashford and Mireille E. Broucke. Time-varying affine feedback for reach control on simplices. *Automatica*, 49(5):1365–1369, 2013.

[6] Audi UK. The Audi TT RS pricing and specification guide, edition 3.0 10/13, 2014 model year, 2013.

[7] Ravindra B. Bapat and T. E. S. Reghavan. *Nonnegative Matrices and Applications*. Cambridge University Press, 1997.

[8] Imre Bárány and Victor S. Grinberg. Block partitions of sequences. *Israel Journal of Mathematics*, 206:155–164, 2015.

[9] Matthias Beck and Sinai Robins. *Computing the Continuous Discretely: Integer-point Enumeration in Polyhedra*. Springer, 2007.

[10] Calin Belta, Luc C. G. J. M. Habets, and Vijay Kumar. Control of multi-affine systems on rectangles with applications to biomolecular networks. In *41st IEEE Conference on Decision and Control*, pages 534–539, 2002.

[11] Abraham Berman and Robert J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, 1979.

[12] Dennis S. Bernstein and Anthony N. Michel. A chronological bibliography on saturating actuators. *International Journal of Robust and Nonlinear Control*, 5(5):375–380, 1995.

[13] Anders Björner. Topological methods. In Ronald L. Graham, Martin Grötschel, and László Lovász, editors, *Handbook of Combinatorics*, volume 2, pages 1819–1872. Elsevier / The MIT Press, 1995.

[14] Franco Blanchini. Nonquadratic Lyapunov functions for robust control. *Automatica*, 31(3):451–461, 1995.

[15] Franco Blanchini and Felice Andrea Pellegrino. Relatively optimal control: a static piecewise-affine solution. *SIAM Journal on Control and Optimization*, 46(2):585–603, 2007.

[16] Kim C. Border. *Fixed point theorems with applications to economics and game theory*. Cambridge University Press, 1989.

[17] Karol Borsuk. *Theory of Retracts*. Państwowe Wydawnictwo Naukowe, 1967.

[18] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[19] Michael S. Branicky. Multiple Lyapunov functions and other analysis tools for switched and hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):475–482, 1998.

[20] Glen E. Bredon. *Topology and Geometry*. Springer, 1997.

[21] Haim Brezis. On a characterization of flow-invariant sets. *Communications on Pure and Applied Mathematics*, 23(2):261–263, 1970.

[22] Mireille E. Broucke. On the reach control indices of affine systems on simplices. In *8th IFAC Symposium on Nonlinear Control Systems*, pages 96–101, 2010.

[23] Mireille E. Broucke. Reach control on simplices by continuous state feedback. *SIAM Journal on Control and Optimization*, 48(5):3482–3500, 2010.

[24] Mireille E. Broucke and Marcus Ganness. Reach control on simplices by piecewise affine feedback. *SIAM Journal on Control and Optimization*, 52(5):3261–3286, 2014.

[25] Mireille E. Broucke, Melkior Ornik, and Abdol-Reza Mansouri. A topological obstruction in a control problem. submitted to Systems & Control Letters, 2015.

[26] Kevin Q. Brown. Fast intersection of half spaces. Technical report, Carnegie Mellon University, 1978.

[27] Eduardo F. Camacho and Carlos Bordons Alba. *Model Predictive Control*. Springer, 2007.

[28] M. Kanat Çamlibel. Well-posed bimodal piecewise linear systems do not exhibit Zeno behavior. In *17th IFAC World Congress*, pages 7973–7978, 2008.

[29] Aneesh N. Chand, Michihiro Kawanishi, and Tatsuo Narikiyo. Application of sigmoidal Gompertz curves in reverse parallel parking for autonomous vehicles. *International Journal of Advanced Robotic Systems*, 12, 2015.

[30] Siu-Wing Cheng, Tamal Krishna Dey, and Jonathan Richard Shewchuk. *Delaunay Mesh Generation*. CRC Press, 2013.

[31] Francis H. Clarke, Yuri S. Ledyaev, Ronald J. Stern, and Peter R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer, 1998.

[32] Jorge Cortés. Discontinuous dynamical systems – a tutorial on solutions, nonsmooth analysis, and stability. *IEEE Control Systems Magazine*, 28(3):36–73, 2008.

[33] Richard Courant and David Hilbert. *Methods of Mathematical Physics, Volume 1*. Interscience (Wiley), 1953.

[34] Richard Courant and David Hilbert. *Methods of Mathematical Physics, Volume 2*. Interscience (Wiley), 1962.

[35] Ludwig Danzer, Branko Grünbaum, and Victor Klee. Helly's theorem and its relatives. In Victor Klee, editor, *Convexity*, pages 101–180. American Mathematical Society, 1963.

[36] Alessandro De Luca, Giuseppe Oriolo, and Claude Samson. Feedback control of a nonholonomic car-like robot. In Jean-Paul Laumond, editor, *Robot Motion Planning and Control*, pages 171–253. Springer, 1998.

[37] Kudret Demirli and Mina Khoshnejad. Autonomous parallel parking of a car-like mobile robot by a neuro-fuzzy sensor-based controller. *Fuzzy Sets and Systems*, 160(19):2876–2891, 2009.

[38] Khaled Elbassioni and Hans Raj Tiwary. On a cone covering problem. *Computational Geometry: Theory and Applications*, 44(3):129–134, 2011.

[39] Jelel Ezzine and Abraham H. Haddad. On the controllability and observability of hybrid systems. In *1988 American Control Conference*, pages 41–46, 1988.

[40] Paul R. Garabedian. *Partial Differential Equations*. Wiley, 1964.

[41] Murray Gerstenhaber. Theory of convex polyhedral cones. In *Activity Analysis of Production and Allocation: Proceedings of a Conference*, pages 298–316, 1951.

[42] Elmer G. Gilbert, Ilya Kolmanovsky, and Kok Tin Tan. Discrete-time reference governors and the nonlinear control of systems with state and control constraints. *International Journal of Robust and Nonlinear Control*, 5(5):487–504, 1995.

[43] Antoine Girard and Samuel Martin. Synthesis for constrained nonlinear systems using hybridization and robust controllers on simplices. *IEEE Transactions on Automatic Control*, 57(4):1046–1051, 2012.

[44] Rafal Goebel and Andrew R. Teel. Lyapunov characterization of Zeno behavior in hybrid systems. In *47th IEEE Conference on Decision and Control*, pages 2752–2757, 2008.

[45] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins University Press, 1996.

[46] Per-Olof Gutman and Michael Cwikel. Admissible sets and feedback control for discrete-time linear systems with bounded controls and states. *IEEE Transactions on Automatic Control*, 31(4):373–376, 1986.

[47] Luc C. G. J. M. Habets, Pieter J. Collins, and Jan H. van Schuppen. Reachability and control synthesis for piecewise-affine hybrid systems on simplices. *IEEE Transactions on Automatic Control*, 51(6):938–948, 2006.

[48] Luc C. G. J. M. Habets, Pieter J. Collins, and Jan H. van Schuppen. Control to Facet by Piecewise-Affine Output Feedback. *IEEE Transactions on Automatic Control*, 57(11):2831–2843, 2012.

[49] Luc C. G. J. M. Habets and Jan H. van Schuppen. Control of piecewise-linear hybrid systems on simplices and rectangles. In M. D. Di Benedetto and A. L. Sangiovanni-Vincentelli, editors, *Hybrid Systems: Computation and Control*, pages 261–274. Springer, 2001.

[50] Luc C. G. J. M. Habets and Jan H. van Schuppen. A control problem for affine dynamical systems on a full-dimensional polytope. *Automatica*, 40(1):21–35, 2004.

[51] Jack K. Hale. *Ordinary Differential Equations*. Krieger, 1980.

[52] W. P. M. H. Heemels, M. Kanat Çamlibel, Arjan J. van der Schaft, and J. M. Schumacher. On the existence and uniqueness of solution trajectories to hybrid dynamical systems. In Rolf Johannson and Anders Rantzer, editors, *Nonlinear and hybrid systems in automotive control*, pages 391–422. Springer, 2002.

[53] Mohamed K. Helwa. *Reach Control Problem on Polytopes*. PhD thesis, University of Toronto, 2013.

[54] Mohamed K. Helwa and Mireille E. Broucke. Monotonic reach control on polytopes. *IEEE Transactions on Automatic Control*, 58(10):2704–2709, 2013.

[55] Mohamed K. Helwa and Mireille E. Broucke. Reach control of single-input systems on simplices using multi-affine feedback. In *International Symposium on Mathematical Theory of Networks and Systems*, 2014.

[56] Mohamed K. Helwa and Mireille E. Broucke. Flow functions, control flow functions, and the reach control problem. *Automatica*, 55:108–115, 2015.

[57] Mohamed K. Helwa, Zhiyun Lin, and Mireille E. Broucke. On the necessity of the invariance conditions for reach control on polytopes. *Systems & Control Letters*, 90:16–19, 2016.

[58] Daniel Hershkowitz and Hans Schneider. On the generalized nullspace of M-matrices and Z-matrices. *Linear Algebra and its Applications*, 106:5–23, 1988.

[59] Daniel Hershkowitz and Hans Schneider. Solutions of Z-matrix equations. *Linear Algebra and its Applications*, 106:25–38, 1988.

[60] Michael Heymann, Feng Lin, George Meyer, and Stefan Resmerita. Analysis of Zeno behaviors in a class of hybrid systems. *IEEE Transactions on Automatic Control*, 50(3):376–383, 2005.

[61] Heinz Hopf. Über die abbildungen der dreidimensionalen sphäre auf die kugelfläche. *Mathematische Annalen*, 104(1):637–665, 1931.

[62] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.

[63] Jun-Ichi Imura and Arjan J. van der Schaft. Characterization of well-posedness of piecewise-linear systems. *IEEE Transactions on Automatic Control*, 45(9):1600–1619, 2000.

[64] ISO/TC 204. Intelligent transport systems — adaptive cruise control systems — performance requirements and test procedures. Technical Report 15622:2010, International Organization for Standardization, 2010.

[65] Edmond A. Jonckheere and Jonathan R. Bar-on. Algebraic topology in robust control. In *Control of Uncertain Dynamic Systems (Proceedings of 1991 International Workshop on Robust Control)*, pages 339–358, 1991.

[66] Rudolf E. Kalman. On the general theory of control systems. In *1st International Conference on Automatic Control*, pages 481–492, 1960.

[67] Paul C. Kettler. *Vertex and Path Attributes of the Generalized Hypercube with Simplex Cuts*. University of Oslo, 2008.

[68] Donald E. Knuth. *The Art of Computer Programming: Volume 3: Sorting and Searching*. Addison-Wesley, 1998.

[69] Zachary Kroeze and Mireille E. Broucke. A viability approach to the output reach control problem. In *2016 American Control Conference*, pages 6054–6059, 2016.

[70] Pamela I. Labuhn and William J. Chundrlik, Jr. Adaptive cruise control, 1995. Patent no. 5454442, United States.

[71] Carl W. Lee. Subdivisions and triangulations of polytopes. In Jacob E. Goodman and Joseph O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, pages 383–406. Chapman and Hall/CRC, 2004.

[72] John M. Lee. *Introduction to Topological Manifolds*. Springer, 2010.

[73] Hongbo Li, David Hestenes, and Alyn Rockwood. Spherical conformal geometry with geometric algebra. In Gerald Sommer, editor, *Geometric Computing with Clifford Algebras: Theoretical Foundations and Applications in Computer Vision and Robotics*, pages 61–75. Springer, 2001.

[74] Chi-Ying Liang and Huei Peng. Optimal adaptive cruise control with guaranteed string stability. *Vehicle System Dynamics*, 32(4–5):313–330, 1999.

[75] Zhao Liang, Guoqiang Zheng, and Jishun Li. Automatic parking path optimization based on Bezier curve fitting. In *2012 IEEE International Conference on Automation and Logistics*, pages 583–587, 2012.

[76] Daniel Liberzon. *Switching in Systems and Control*. Springer, 2003.

[77] Zhiyun Lin and Mireille E. Broucke. On a reachability problem for affine hypersurface systems on polytopes. *Automatica*, 47(4):769–775, 2011.

[78] Jesus A. De Loera, Elisha Peterson, and Francis E. Su. A polytopal generalization of Sperner's lemma. *Journal of Combinatorial Theory, Series A*, 100:1–26, 2002.

[79] Johan Löfberg. YALMIP: a toolbox for modeling and optimization in MATLAB. In *IEEE International Symposium on Computer Aided Control System Design*, pages 284–289, 2004.

[80] John Lygeros, Karl Henrik Johansson, Slobodan N. Simić, Jun Zhang, and S. Shankar Sastry. Dynamical properties of hybrid automata. *IEEE Transactions on Automatic Control*, 48(1):2–17, 2003.

[81] Douglas Lyon. Parallel parking with curvature and nonholonomic constraints. In *Intelligent Vehicles '92 Symposium*, pages 341–346, 1992.

[82] Matthew Martino and Mireille E. Broucke. A reach control approach to bumpless transfer of robotic manipulators. In *53rd IEEE Conference on Decision and Control*, pages 25–30, 2014.

[83] Francisco Rovira Más, Qin Zhang, and Alan C. Hansen. *Mechatronics and Intelligent Systems for Off-road Vehicles*. Springer, 2011.

[84] David Q. Mayne, James B. Rawlins, Christopher V. Rao, and Pierre O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

[85] Manuel Mazo, Jr., Anna Davitian, and Paulo Tabuada. Pessoa: a tool for embedded controller synthesis. In *Computer Aided Verification: 22nd International Conference, CAV 2010*, pages 566–569, 2010.

[86] Aakar Mehra, Wen-Loong Ma, Forrest Berg, Paulo Tabuada, Jessy W. Grizzle, and Aaron D. Ames. Adaptive cruise control: experimental validation of advanced controllers on scale-model cars. In *American Control Conference*, pages 1411–1418, 2015.

[87] Krishnaa Mehta. A topological obstruction in a control problem. Master's thesis, University of Toronto, 2012.

[88] Chohong Min. Simplicial isosurfacing in arbitrary dimension and codimension. *Journal of Computational Physics*, 190(1):295–310, 2003.

[89] Miad Moarref, Melkior Ornik, and Mireille E. Broucke. An obstruction to solvability of the reach control problem using affine feedback. *Automatica*, 71:229–236, 2016.

[90] Michael Müger. Topology for the working mathematician, 2016.

[91] James R. Munkres. *Topology*. Prentice Hall, 2000.

[92] Richard M. Murray and S. Shankar Sastry. Nonholonomic motion planning: steering using sinusoids. *IEEE Transactions on Automatic Control*, 38(5):700–716, 1993.

[93] Oleg R. Musin. Homotopy invariants of covers and KKM-type lemmas. *Algebraic & Geometric Topology*, 16(3):1799–1812, 2016.

[94] Gregory L. Naber. *Topological Methods in Euclidean Spaces*. Dover, 2000.

[95] Andrzej Nagórko. Carrier and nerve theorems in the extension theory. *Proceedings of the American Mathematical Society*, 135(2):551–558, 2007.

[96] Hoai-Nam Nguyen. *Constrained Control of Uncertain, Time-Varying, Discrete-Time Systems: An Interpolation Based Approach*. Springer, 2014.

[97] Petter Nilsson, Omar Hussien, Yuxiao Chen, Ayca Balkan, Matthias Rungger, Aaron D. Ames, Jessy W. Grizzle, Necmiye Ozay, Huei Peng, and Paulo Tabuada. Preliminary results on correct-by-construction control software synthesis for adaptive cruise control. In *53rd IEEE Conference on Decision and Control*, pages 816–823, 2014.

[98] Peter Orlik and Volkmar Welker. *Algebraic Combinatorics: Lectures at a Summer School in Nordfjordeid, Norway, June 2003*. Springer, 2007.

[99] Melkior Ornik and Mireille E. Broucke. Chattering in the reach control problem. submitted to Automatica, 2017.

[100] Melkior Ornik and Mireille E. Broucke. A graph-theoretic approach to the reach control problem. submitted to 56th IEEE Conference on Decision and Control, 2017.

[101] Melkior Ornik and Mireille E. Broucke. A nerve-theoretic result on the problem of a topological obstruction in reach control. submitted to 56th IEEE Conference on Decision and Control, 2017.

[102] Melkior Ornik and Mireille E. Broucke. Some results on an affine obstruction to reach control, 2015. arXiv:1507.02957.

[103] Melkior Ornik and Mireille E. Broucke. A topological obstruction to reach control by continuous state feedback. In *54th IEEE Conference on Decision and Control*, pages 2258–2263, 2015.

[104] Melkior Ornik and Mireille E. Broucke. Characterizing equilibria in reach control under affine feedback. In *10th IFAC Symposium on Nonlinear Control Systems*, pages 1078–1083, 2016.

[105] Melkior Ornik and Mireille E. Broucke. On a topological obstruction in the reach control problem. In *Mathematical and Computational Approaches in Advancing Modern Science and Engineering (Proceedings of AMMCS-CAIMS 2015)*, pages 677–687. Springer, 2016.

[106] Melkior Ornik and Mireille E. Broucke. Characterization of a topological obstruction to reach control by continuous state feedback. *Mathematics of Control, Signals, and Systems*, 29(2), 2017.

[107] Melkior Ornik, Miad Moarref, and Mireille E. Broucke. An automated parallel parking strategy using reach control theory. Accepted to 20th IFAC World Congress, 2017.

[108] Melkior Ornik, Mateus S. Moura, Alexander Peplowski, and Mireille E. Broucke. Adaptive cruise control design using reach control. submitted to 56th IEEE Conference on Decision and Control, 2017.

[109] Panos K. Palamides, Gennaro Infante, and Paolamaria Pietramala. Nontrivial solutions of a nonlinear heat flow problem via Sperner's Lemma. *Applied Mathematics Letters*, 22(9):1444–1450, 2009.

[110] Igor E. Paromtchik and Christian Laugier. Autonomous parallel parking of a nonholonomic vehicle. In *1996 IEEE Inteligent Vehicles Symposium*, pages 13–18, 1996.

[111] Mikael Persson, Fredrik Botling, Erik Hesslow, and Rolf Johansson. Stop & go controller for adaptive cruise control. In *1999 IEEE International Conference on Control Applications*, pages 1692–1697, 1999.

[112] Stefan Petterson and Bengt Lennartson. Stability and robustness for hybrid systems. In *35th IEEE Conference on Decision and Control*, pages 1202–1207, 1996.

[113] Jeroen Ploeg, Alex F. A. Serrarens, and Geert J. Heijenk. Connect & drive: design and evaluation of cooperative adaptive cruise control for congestion reduction. *Journal of Modern Transportation*, 19(3):207–213, 2011.

[114] Amir Pnueli. The temporal logic of programs. In *18th Annual Symposium on Foundations of Computer Science*, pages 46–57, 1977.

[115] Jochen Pohl, Magnus Sethsson, Pär Degerman, and Jonas Larsson. A semi-automated parallel parking system for passenger cars. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 220(1):53–65, 2009.

[116] Teodor Przymusiński. Collectionwise normality and absolute retracts. *Fundamenta Mathematicae*, 98(1):61–73, 1978.

[117] Dušan Repovš and Pavel Vladimirovič Semenov. *Continuous Selections of Multivalued Mappings*. Springer, 2013.

[118] Frank J. Riley. Constant speed regulator, 1955. Patent no. 2714880, United States.

[119] Charles V. Robinson. Spherical theorems of Helly type and congruence indices of spherical caps. *American Journal of Mathematics*, 64(1):260–272, 1942.

[120] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University, 1970.

[121] Bartek Roszak and Mireille E. Broucke. Necessary and sufficient conditions for reachability on a simplex. In *44th IEEE Conference on Decision and Control and 2005 European Control Conference*, pages 4706–4711, 2005.

[122] Bartek Roszak and Mireille E. Broucke. Necessary and sufficient conditions for reachability on a simplex. *Automatica*, 42(11):1913–1918, 2006.

[123] Bartek Roszak and Mireille E. Broucke. Reachability of a set of facets for linear affine systems with n-1 inputs. *IEEE Transactions on Automatic Control*, 52(2):359–364, 2007.

[124] Lennart Sandgren. On convex cones. *Mathematica Scandinavica*, 2:19–28, 1954.

[125] Elham Semsar-Kazerooni and Mireille E. Broucke. Reach controllability of single input affine systems on a simplex. *IEEE Transactions on Automatic Control*, 59(3):738–744, 2014.

[126] Steven S. Skiena. *The Algorithm Design Manual*. Springer, 2008.

[127] Michael Spivak. *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*. Addison-Wesley, 1965.

[128] Héctor J. Sussmann. Bounds on the number of switchings for trajectories of piecewise analytic vector fields. *Journal of Differential Equations*, 43(3):399–418, 1982.

[129] Dawn Tilbury, Ole Jakob Sørdalen, Linda Bushnell, and S. Shankar Sastry. A multisteering trailer system: conversion into chained form using dynamic feedback. *IEEE Transactions on Robotics and Automation*, 11(6):807–818, 1995.

[130] Tammo tom Dieck. *Algebraic Topology*. European Mathematical Society, 2008.

[131] Bart van Arem, Cornelie J. G. van Driel, and Ruben Visser. The impact of cooperative adaptive cruise control on traffic-flow characteristics. *IEEE Transactions on Intelligent Transportation Systems*, 7(4):429–436, 2006.

[132] Moshe Y. Vardi. An automata-theoretic approach to linear temporal logic. In Faron Moller and Graham Birtwistle, editors, *Logics for Concurrency: Structure versus Automata*, pages 238–266. Springer, 1996.

[133] Marijan Vukosavljev and Mireille E. Broucke. Control of a gantry crane: a reach control approach. In *53rd IEEE Conference on Decision and Control*, pages 3609–3614, 2014.

[134] Marijan Vukosavljev, Ivo Jansen, Mireille E. Broucke, and Angela P. Schoellig. Safe and robust robot maneuvers based on reach control. In *2016 IEEE International Conference on Robotics and Automation*, pages 5677–5682, 2016.

[135] Marijan Vukosavljev, Zachary Kroeze, Mireille E. Broucke, and Angela P. Schoellig. A framework for multi-vehicle navigation using feedback-based motion primitives. submitted to 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems.

[136] Yuhu Wu and Tielong Shen. Reach control problem for linear differential inclusion systems on simplices. *IEEE Transactions on Automatic Control*, 61(5):1403–1408, 2016.

[137] Yoshinori Yamamura, Yoji Seto, Hikaru Nishira, and Taketoshi Kawabe. An ACC design method for achieving both string stability and ride comfort. *Journal of System Design and Dynamics*, 2(4):979–990, 2008.

Appendices

Appendix A

Numerical Values Used for Simulations in Section 9.1

A.1 Vehicle Size

Dimensions of the car:

wheelbase	2.468m,
width	1.952m,
distance of the front end from the front axle	0.911m,
distance of the back end from the rear axle	0.819m.

A.2 Obstacle Positions

Unless otherwise stated, all values in the remainder of Appendix A are given in metres.

Position of the front parked car:

rear left corner	$(2, 0)$,
rear right corner	$(2, -1.952)$,
front left corner	$(6.198, 0)$,
front right corner	$(6.198, -1.952)$.

Position of the rear parked car:

rear left corner	$(-10.198, 0)$,
rear right corner	$(-10.198, -1.952)$,
front left corner	$(-6, 0)$,
front right corner	$(-6, -1.952)$.

Position of the curb: $\{(x, y) \in \mathbb{R}^2 \mid y < -2.002\}$.

A.3 Positions of the Initial and Ending Box

Coordinates of the initial box $\mathcal{I}$:

top left corner	$(2.319, 1.976)$,
bottom left corner	$(2.319, 1.726)$,
top right corner	$(2.819, 1.976)$,
bottom right corner	$(2.819, 1.726)$.

Coordinates of the ending box $\mathcal{Q}$:

top left corner	$(-5.181, -0.676)$,
bottom left corner	$(-5.181, -0.976)$,
top right corner	$(-1.379, -0.676)$,
bottom right corner	$(-1.379, -0.976)$.

A.4 Polytope Construction

Coefficients Δx_k, Δy_k, w_{kx}, and w_{ky} used in the construction of polytopes $\mathcal{P}_k$, $k \in \{1, \ldots, 8\}$, are as follows:

k	Δx_k	Δy_k	w_{kx}	w_{ky}
1	-0.850	-0.094	0	0.09375
2	-0.834	-0.286	0	0.12
3	-0.785	-0.445	0	0.12
4	-0.707	-0.592	0	0.12
5	-0.707	-0.592	0	0.12
6	-0.785	-0.445	0	0.12
7	-0.834	-0.286	0	0.12
8	-0.850	-0.094	0	0.09375

In total, polytopes $\mathcal{P}_k$ contain 40 vertices, labeled by v_i, $i \in \{1, \ldots, 40\}$. Vertices of polytopes $\mathcal{P}_k$ for $k \in \{1, 2, 3, 4\}$ are $v_{4(k-1)+1}, \ldots, v_{4(k+1)}$, and vertices of polytopes $\mathcal{P}_k$ for $k \in \{5, 6, 7, 8\}$ are $v_{4k+1}, \ldots, v_{4(k+2)}$. Vertices $v_{4k+1}, \ldots, v_{4(k+1)}$ are vertices of exit facets $\mathcal{F}_k^{\text{out}}$ for $k \in \{1, 2, 3, 4\}$, and vertices $v_{4(k+1)+1}, \ldots, v_{4(k+2)}$ are vertices of exit facets $\mathcal{F}_k^{\text{out}}$ for $k \in \{5, 6, 7, 8\}$. Vertices $v_{17}, \ldots, v_{20}$ and $v_{21}, \ldots, v_{24}$, respectively, coincide, but are treated as separate in the remainder of the calculations. The coordinates (x, y, θ) of vertices v_i are as follows:

i	x	y	θ (in rad)
1	2.319	1.976	0.000
2	2.819	1.976	0.000
3	2.819	1.726	0.000
4	2.319	1.726	0.000
5	1.469	1.976	0.196
6	1.969	1.976	0.196

i	x	y	θ (in rad)	i	x	y	θ (in rad)
7	1.969	1.539	0.196	24	-0.857	-0.175	0.785
8	1.469	1.539	0.196	25	-1.563	0.580	0.589
9	0.635	1.820	0.393	26	-1.063	0.580	0.589
10	1.135	1.820	0.393	27	-1.063	-0.897	0.589
11	1.135	1.123	0.393	28	-1.563	-0.897	0.589
12	0.635	1.123	0.393	29	-2.349	0.265	0.393
13	-0.150	1.505	0.589	30	-1.849	0.265	0.393
14	0.350	1.505	0.589	31	-1.849	-1.472	0.393
15	0.350	0.547	0.589	32	-2.349	-1.472	0.393
16	-0.150	0.547	0.589	33	-3.182	0.109	0.196
17	-0.857	1.043	0.785	34	-2.682	0.109	0.196
18	-0.357	1.043	0.785	35	-2.682	-1.888	0.196
19	-0.357	-0.175	0.785	36	-3.182	-1.888	0.196
20	-0.857	-0.175	0.785	37	-4.032	0.109	0.000
21	-0.857	1.043	0.785	38	-3.532	0.109	0.000
22	-0.357	1.043	0.785	39	-3.532	-2.076	0.000
23	-0.357	-0.175	0.785	40	-4.032	-2.076	0.000

A.5 Polytope Triangulation

In addition to vertices $v_1, \ldots, v_{40}$, triangulations of polytopes $\mathcal{P}_k$, $k \in \{1, \ldots, 8\}$, make use of 8 additional points denoted by v_i, $i \in \{41, \ldots, 48\}$. Their coordinates are given as follows:

i	x	y	θ (in rad)
41	2.144	1.804	0.098
42	1.302	1.614	0.295
43	0.493	1.249	0.491
44	-0.253	0.730	0.687
45	-0.960	0.138	0.687
46	-1.706	-0.381	0.491
47	-2.516	-0.747	0.295
48	-3.357	-0.936	0.098

Each polytope $\mathcal{P}_k$ is triangulated into 12 simplices, which are denoted by $\mathcal{S}_{12(k-1)+1}, \ldots, \mathcal{S}_{12k}$. Each

simplex $\mathcal{S}_j$, $j \in \{1, \ldots, 96\}$, is a convex hull of four vertices $v_{j_1}, \ldots, v_{j_4}$, with indices $j_1, \ldots, j_4 \in \{1, \ldots, 48\}$. The vertices of all simplices are given as follows:

j	j_1	j_2	j_3	j_4	j	j_1	j_2	i_3	j_4	j	j_1	j_2	i_3	j_4
1	41	6	1	5	29	43	11	14	15	57	45	27	28	25
2	41	3	6	7	30	43	12	11	15	58	45	26	27	25
3	41	1	6	2	31	43	12	16	13	59	45	24	23	21
4	41	6	3	2	32	43	16	12	15	60	45	21	23	22
5	4	3	41	7	33	43	11	9	10	61	30	46	25	29
6	4	41	8	7	34	43	11	12	9	62	30	25	46	26
7	4	8	41	5	35	43	15	14	13	63	28	32	46	31
8	4	41	1	5	36	43	16	15	13	64	28	46	32	29
9	41	7	6	5	37	44	16	15	19	65	28	25	46	29
10	41	8	7	5	38	44	15	18	19	66	27	28	46	31
11	4	41	3	2	39	44	16	20	17	67	27	46	30	31
12	4	1	41	2	40	44	20	16	19	68	27	30	46	26
13	42	10	5	9	41	44	13	18	14	69	30	32	46	29
14	42	5	8	9	42	44	18	15	14	70	30	46	32	31
15	42	8	12	9	43	44	13	16	17	71	27	46	28	25
16	42	12	8	11	44	44	18	13	17	72	27	46	25	26
17	42	5	10	6	45	44	20	19	17	73	34	47	29	33
18	42	10	7	6	46	44	19	18	17	74	34	29	47	30
19	42	7	10	11	47	44	16	13	14	75	32	36	47	35
20	42	8	7	11	48	44	15	16	14	76	32	47	36	33
21	42	11	10	9	49	45	26	21	22	77	32	29	47	33
22	42	12	11	9	50	45	23	26	22	78	31	32	47	35
23	42	7	5	6	51	45	21	26	25	79	31	47	34	35
24	42	7	8	5	52	45	24	21	25	80	31	34	47	30
25	43	14	11	10	53	45	26	23	27	81	34	36	47	33
26	43	9	14	10	54	45	23	24	27	82	34	47	36	35
27	43	14	9	13	55	45	28	24	25	83	31	47	32	29
28	43	9	12	13	56	45	24	28	27	84	31	47	29	30

j	j_1	j_2	i_3	j_4	j	j_1	j_2	i_3	j_4	j	j_1	j_2	i_3	j_4
85	38	48	33	37	89	36	48	40	37	93	38	40	48	37
86	38	33	48	34	90	36	33	48	37	94	38	48	40	39
87	38	48	35	34	91	36	40	48	39	95	36	48	33	34
88	38	35	48	39	92	36	48	35	39	96	36	35	48	34

A.6 Proposed Control Law

Speeds and front wheel angles assigned at each simplex vertex v_i, $i \in \{1, \ldots, 48\}$, are given as follows:

i	v (in m/s)	φ (in rad)	i	v (in m/s)	φ (in rad)
1	-1.654	-0.484	23	-2.186	0.511
2	-1.391	-0.141	24	-1.767	0.493
3	-2.225	-0.273	25	-1.818	0.504
4	-1.747	-0.499	26	-2.358	0.071
5	-1.773	-0.505	27	-2.227	0.506
6	-2.319	-0.095	28	-1.667	0.516
7	-2.447	-0.509	29	-1.815	0.506
8	-1.711	-0.533	30	-2.348	0.077
9	-1.815	-0.506	31	-2.289	0.488
10	-2.348	-0.077	32	-1.674	0.518
11	-2.289	-0.488	33	-1.773	0.505
12	-1.674	-0.518	34	-2.319	0.095
13	-1.818	-0.504	35	-2.447	0.509
14	-2.358	-0.071	36	-1.711	0.533
15	-2.227	-0.506	37	-1.643	0.495
16	-1.667	-0.516	38	-1.406	-0.195
17	-1.822	-0.459	39	-2.198	0.166
18	-2.198	0.244	40	-1.734	0.502
19	-2.159	-0.504	41	-1.909	-0.380
20	-1.760	-0.493	42	-2.047	-0.404
21	-1.854	0.466	43	-2.025	-0.398
22	-2.264	0.136	44	-2.001	-0.351

i	v (in m/s)	φ (in rad)	i	v (in m/s)	φ (in rad)
45	-2.018	0.400	47	-2.047	0.404
46	-2.025	0.398	48	-1.904	0.326

If $\mathcal{S}_j$ is a simplex with vertices $v_{j_1}, \ldots, v_{j_4}$, the feedback control $u_j(x) = K_j x + g_j$ is obtained by affinely extending the control values assigned to $v_{j_1}, \ldots, v_{j_4}$ above (with x_1 corresponding to speed v, and x_2 corresponding to front wheel angle ϕ). This produces the following feedback controls at each simplex $\mathcal{S}_j$:

j	$u_j(x) = K_j x + g_j$
1	$\begin{pmatrix} -1.092 & -0.456 & -5.332 \\ 0.820 & 0.526 & 3.442 \end{pmatrix} x + \begin{pmatrix} 1.779 \\ -3.425 \end{pmatrix}$
2	$\begin{pmatrix} -1.510 & 0.293 & -7.388 \\ 0.605 & 0.946 & 2.324 \end{pmatrix} x + \begin{pmatrix} 1.526 \\ -3.612 \end{pmatrix}$
3	$\begin{pmatrix} 0.527 & -0.456 & -2.446 \\ 0.685 & 0.526 & 3.202 \end{pmatrix} x + \begin{pmatrix} -1.976 \\ -3.112 \end{pmatrix}$
4	$\begin{pmatrix} -2.081 & 3.338 & -13.736 \\ 0.684 & 0.528 & 3.194 \end{pmatrix} x + \begin{pmatrix} -2.120 \\ -3.112 \end{pmatrix}$
5	$\begin{pmatrix} -0.957 & 1.097 & -4.228 \\ 0.451 & 0.721 & 1.440 \end{pmatrix} x + \begin{pmatrix} -1.419 \\ -2.788 \end{pmatrix}$
6	$\begin{pmatrix} -1.473 & 1.097 & -5.146 \\ 0.048 & 0.721 & 0.721 \end{pmatrix} x + \begin{pmatrix} -0.224 \\ -1.853 \end{pmatrix}$
7	$\begin{pmatrix} -0.621 & -0.142 & -2.643 \\ 0.500 & 0.063 & 2.052 \end{pmatrix} x + \begin{pmatrix} -0.061 \\ -1.767 \end{pmatrix}$
8	$\begin{pmatrix} -0.525 & 0.369 & -2.879 \\ 0.499 & 0.059 & 2.053 \end{pmatrix} x + \begin{pmatrix} -1.165 \\ -1.759 \end{pmatrix}$
9	$\begin{pmatrix} -1.092 & 0.293 & -6.643 \\ 0.820 & 0.946 & 2.707 \end{pmatrix} x + \begin{pmatrix} 0.557 \\ -4.110 \end{pmatrix}$

j	$u_j(x) = K_j x + g_j$
10	$\begin{pmatrix} -1.473 & -0.142 & -8.498 \\ 0.048 & 0.063 & -1.059 \end{pmatrix} x + \begin{pmatrix} 2.339 \\ -0.491 \end{pmatrix}$
11	$\begin{pmatrix} -0.957 & 3.338 & -6.011 \\ 0.451 & 0.528 & 1.593 \end{pmatrix} x + \begin{pmatrix} -5.287 \\ -2.456 \end{pmatrix}$
12	$\begin{pmatrix} 0.527 & 0.369 & -1.003 \\ 0.685 & 0.059 & 2.385 \end{pmatrix} x + \begin{pmatrix} -3.605 \\ -2.190 \end{pmatrix}$
13	$\begin{pmatrix} -1.065 & -0.047 & -4.775 \\ 0.858 & 0.397 & 3.952 \end{pmatrix} x + \begin{pmatrix} 0.823 \\ -3.325 \end{pmatrix}$
14	$\begin{pmatrix} -1.173 & -0.142 & -5.307 \\ 0.478 & 0.063 & 2.077 \end{pmatrix} x + \begin{pmatrix} 1.272 \\ -1.740 \end{pmatrix}$
15	$\begin{pmatrix} -1.189 & -0.203 & -5.287 \\ 0.466 & 0.017 & 2.092 \end{pmatrix} x + \begin{pmatrix} 1.385 \\ -1.654 \end{pmatrix}$
16	$\begin{pmatrix} -1.231 & -0.165 & -5.388 \\ 0.060 & 0.375 & 1.124 \end{pmatrix} x + \begin{pmatrix} 1.410 \\ -1.418 \end{pmatrix}$
17	$\begin{pmatrix} -1.092 & -0.047 & -4.821 \\ 0.820 & 0.397 & 3.889 \end{pmatrix} x + \begin{pmatrix} 0.871 \\ -3.258 \end{pmatrix}$
18	$\begin{pmatrix} -1.478 & 0.293 & -6.190 \\ 0.197 & 0.946 & 1.680 \end{pmatrix} x + \begin{pmatrix} 1.228 \\ -2.682 \end{pmatrix}$
19	$\begin{pmatrix} -1.380 & -0.084 & -5.233 \\ 0.290 & 0.589 & 2.586 \end{pmatrix} x + \begin{pmatrix} 1.427 \\ -2.494 \end{pmatrix}$
20	$\begin{pmatrix} -1.473 & -0.165 & -5.798 \\ 0.048 & 0.375 & 1.103 \end{pmatrix} x + \begin{pmatrix} 1.845 \\ -1.396 \end{pmatrix}$
21	$\begin{pmatrix} -1.065 & -0.084 & -4.699 \\ 0.858 & 0.589 & 3.550 \end{pmatrix} x + \begin{pmatrix} 0.859 \\ -3.517 \end{pmatrix}$

j	$u_j(x) = K_j x + g_j$
22	$\begin{pmatrix} -1.231 & -0.203 & -5.576 \\ 0.060 & 0.017 & -0.670 \end{pmatrix} x + \begin{pmatrix} 1.526 \\ -0.311 \end{pmatrix}$
23	$\begin{pmatrix} -1.092 & 0.293 & -3.568 \\ 0.820 & 0.946 & 5.910 \end{pmatrix} x + \begin{pmatrix} -0.047 \\ -4.739 \end{pmatrix}$
24	$\begin{pmatrix} -1.473 & -0.142 & -5.816 \\ 0.048 & 0.063 & 1.345 \end{pmatrix} x + \begin{pmatrix} 1.813 \\ -0.963 \end{pmatrix}$
25	$\begin{pmatrix} -1.175 & -0.084 & -4.888 \\ 0.323 & 0.589 & 2.272 \end{pmatrix} x + \begin{pmatrix} 1.058 \\ -2.408 \end{pmatrix}$
26	$\begin{pmatrix} -1.065 & -0.150 & -4.555 \\ 0.858 & 0.266 & 3.891 \end{pmatrix} x + \begin{pmatrix} 0.924 \\ -3.063 \end{pmatrix}$
27	$\begin{pmatrix} -1.081 & -0.150 & -4.578 \\ 0.867 & 0.266 & 3.905 \end{pmatrix} x + \begin{pmatrix} 0.943 \\ -3.074 \end{pmatrix}$
28	$\begin{pmatrix} -1.168 & -0.203 & -5.010 \\ 0.455 & 0.017 & 1.857 \end{pmatrix} x + \begin{pmatrix} 1.263 \\ -1.555 \end{pmatrix}$
29	$\begin{pmatrix} -1.162 & -0.137 & -4.728 \\ 0.358 & 0.455 & 2.672 \end{pmatrix} x + \begin{pmatrix} 1.040 \\ -2.454 \end{pmatrix}$
30	$\begin{pmatrix} -1.231 & -0.179 & -5.132 \\ 0.060 & 0.275 & 0.952 \end{pmatrix} x + \begin{pmatrix} 1.325 \\ -1.238 \end{pmatrix}$
31	$\begin{pmatrix} -1.156 & -0.157 & -5.051 \\ 0.454 & 0.013 & 1.860 \end{pmatrix} x + \begin{pmatrix} 1.221 \\ -1.551 \end{pmatrix}$
32	$\begin{pmatrix} -1.119 & -0.179 & -4.969 \\ 0.021 & 0.275 & 0.895 \end{pmatrix} x + \begin{pmatrix} 1.190 \\ -1.191 \end{pmatrix}$
33	$\begin{pmatrix} -1.065 & -0.084 & -4.168 \\ 0.858 & 0.589 & 5.771 \end{pmatrix} x + \begin{pmatrix} 0.651 \\ -4.389 \end{pmatrix}$

j	$u_j(x) = K_j x + g_j$
34	$\begin{pmatrix} -1.231 & -0.203 & -5.102 \\ 0.060 & 0.017 & 1.283 \end{pmatrix} x + \begin{pmatrix} 1.340 \\ -1.078 \end{pmatrix}$
35	$\begin{pmatrix} -1.081 & -0.137 & -4.612 \\ 0.867 & 0.455 & 3.411 \end{pmatrix} x + \begin{pmatrix} 0.943 \\ -3.068 \end{pmatrix}$
36	$\begin{pmatrix} -1.119 & -0.157 & -4.809 \\ 0.021 & 0.013 & -0.975 \end{pmatrix} x + \begin{pmatrix} 1.084 \\ 0.054 \end{pmatrix}$
37	$\begin{pmatrix} -1.119 & -0.162 & -4.278 \\ 0.021 & 0.292 & 1.162 \end{pmatrix} x + \begin{pmatrix} 0.774 \\ -1.357 \end{pmatrix}$
38	$\begin{pmatrix} -0.838 & -0.032 & -2.789 \\ 0.720 & 0.614 & 4.863 \end{pmatrix} x + \begin{pmatrix} -0.273 \\ -3.958 \end{pmatrix}$
39	$\begin{pmatrix} -1.040 & -0.051 & -4.401 \\ 0.554 & 0.028 & 2.215 \end{pmatrix} x + \begin{pmatrix} 0.797 \\ -1.753 \end{pmatrix}$
40	$\begin{pmatrix} -0.798 & -0.162 & -3.940 \\ -0.021 & 0.292 & 1.117 \end{pmatrix} x + \begin{pmatrix} 0.623 \\ -1.338 \end{pmatrix}$
41	$\begin{pmatrix} -1.081 & -0.013 & -3.108 \\ 0.867 & 0.406 & 5.678 \end{pmatrix} x + \begin{pmatrix} -0.130 \\ -4.330 \end{pmatrix}$
42	$\begin{pmatrix} -0.811 & -0.137 & -2.427 \\ 0.762 & 0.455 & 5.413 \end{pmatrix} x + \begin{pmatrix} -0.439 \\ -4.210 \end{pmatrix}$
43	$\begin{pmatrix} -1.067 & -0.157 & -4.233 \\ 0.550 & 0.013 & 2.238 \end{pmatrix} x + \begin{pmatrix} 0.752 \\ -1.760 \end{pmatrix}$
44	$\begin{pmatrix} -0.753 & -0.013 & -2.763 \\ 1.406 & 0.406 & 6.245 \end{pmatrix} x + \begin{pmatrix} -0.284 \\ -4.583 \end{pmatrix}$
45	$\begin{pmatrix} -0.798 & -0.051 & -2.917 \\ -0.021 & 0.028 & -1.320 \end{pmatrix} x + \begin{pmatrix} -0.162 \\ 0.531 \end{pmatrix}$

j	$u_j(x) = K_j x + g_j$
46	$\begin{pmatrix} -0.753 & -0.032 & -2.700 \\ 1.406 & 0.614 & 5.584 \end{pmatrix} x + \begin{pmatrix} -0.313 \\ -4.281 \end{pmatrix}$
47	$\begin{pmatrix} -1.081 & -0.157 & -4.247 \\ 0.867 & 0.013 & 2.572 \end{pmatrix} x + \begin{pmatrix} 0.758 \\ -1.909 \end{pmatrix}$
48	$\begin{pmatrix} -1.119 & -0.137 & -4.324 \\ 0.021 & 0.455 & 0.859 \end{pmatrix} x + \begin{pmatrix} 0.787 \\ -1.268 \end{pmatrix}$
49	$\begin{pmatrix} -0.818 & -0.132 & 3.737 \\ -0.660 & -0.196 & 3.170 \end{pmatrix} x + \begin{pmatrix} -5.353 \\ -2.385 \end{pmatrix}$
50	$\begin{pmatrix} -1.001 & -0.064 & 4.236 \\ -0.359 & -0.308 & 2.348 \end{pmatrix} x + \begin{pmatrix} -5.881 \\ -1.515 \end{pmatrix}$
51	$\begin{pmatrix} -1.081 & -0.132 & 4.014 \\ -0.867 & -0.196 & 3.388 \end{pmatrix} x + \begin{pmatrix} -5.796 \\ -2.733 \end{pmatrix}$
52	$\begin{pmatrix} -0.920 & -0.072 & 3.293 \\ -0.399 & -0.022 & 1.293 \end{pmatrix} x + \begin{pmatrix} -5.154 \\ -0.868 \end{pmatrix}$
53	$\begin{pmatrix} -0.995 & -0.089 & 4.116 \\ -0.362 & -0.295 & 2.411 \end{pmatrix} x + \begin{pmatrix} -5.789 \\ -1.564 \end{pmatrix}$
54	$\begin{pmatrix} -0.838 & -0.031 & 3.338 \\ 0.035 & -0.147 & 0.438 \end{pmatrix} x + \begin{pmatrix} -5.112 \\ 0.154 \end{pmatrix}$
55	$\begin{pmatrix} -0.928 & -0.102 & 3.206 \\ -0.395 & -0.008 & 1.333 \end{pmatrix} x + \begin{pmatrix} -5.098 \\ -0.894 \end{pmatrix}$
56	$\begin{pmatrix} -1.119 & -0.031 & 3.634 \\ -0.021 & -0.147 & 0.497 \end{pmatrix} x + \begin{pmatrix} -5.585 \\ 0.059 \end{pmatrix}$
57	$\begin{pmatrix} -1.119 & -0.102 & 4.383 \\ -0.021 & -0.008 & -0.966 \end{pmatrix} x + \begin{pmatrix} -6.090 \\ 1.045 \end{pmatrix}$

j	$u_j(x) = K_j x + g_j$
58	$\begin{pmatrix} -1.081 & -0.089 & 4.207 \\ -0.867 & -0.295 & 2.943 \end{pmatrix} x + \begin{pmatrix} -5.935 \\ -2.414 \end{pmatrix}$
59	$\begin{pmatrix} -0.838 & -0.072 & 3.207 \\ 0.035 & -0.022 & 0.836 \end{pmatrix} x + \begin{pmatrix} -5.016 \\ -0.137 \end{pmatrix}$
60	$\begin{pmatrix} -0.818 & -0.064 & 3.111 \\ -0.660 & -0.308 & 4.200 \end{pmatrix} x + \begin{pmatrix} -4.933 \\ -3.078 \end{pmatrix}$
61	$\begin{pmatrix} -1.065 & -0.072 & 4.366 \\ -0.858 & -0.134 & 3.635 \end{pmatrix} x + \begin{pmatrix} -6.013 \\ -2.900 \end{pmatrix}$
62	$\begin{pmatrix} -1.081 & -0.072 & 4.388 \\ -0.867 & -0.134 & 3.648 \end{pmatrix} x + \begin{pmatrix} -6.051 \\ -2.923 \end{pmatrix}$
63	$\begin{pmatrix} -1.231 & -0.057 & 5.126 \\ -0.060 & -0.129 & 0.611 \end{pmatrix} x + \begin{pmatrix} -6.664 \\ -0.053 \end{pmatrix}$
64	$\begin{pmatrix} -1.154 & -0.081 & 4.889 \\ -0.454 & -0.007 & 1.828 \end{pmatrix} x + \begin{pmatrix} -6.425 \\ -1.276 \end{pmatrix}$
65	$\begin{pmatrix} -1.160 & -0.102 & 4.789 \\ -0.454 & -0.008 & 1.820 \end{pmatrix} x + \begin{pmatrix} -6.393 \\ -1.273 \end{pmatrix}$
66	$\begin{pmatrix} -1.119 & -0.057 & 4.964 \\ -0.021 & -0.129 & 0.555 \end{pmatrix} x + \begin{pmatrix} -6.393 \\ 0.041 \end{pmatrix}$
67	$\begin{pmatrix} -1.043 & -0.034 & 4.588 \\ -0.365 & -0.236 & 2.245 \end{pmatrix} x + \begin{pmatrix} -6.069 \\ -1.416 \end{pmatrix}$
68	$\begin{pmatrix} -1.028 & -0.089 & 4.204 \\ -0.350 & -0.295 & 1.838 \end{pmatrix} x + \begin{pmatrix} -5.876 \\ -1.213 \end{pmatrix}$
69	$\begin{pmatrix} -1.065 & -0.081 & 4.307 \\ -0.858 & -0.007 & 4.471 \end{pmatrix} x + \begin{pmatrix} -5.987 \\ -3.263 \end{pmatrix}$

j	$u_j(x) = K_j x + g_j$
70	$\begin{pmatrix} -1.231 & -0.034 & 4.862 \\ -0.060 & -0.236 & 1.801 \end{pmatrix} x + \begin{pmatrix} -6.525 \\ -0.678 \end{pmatrix}$
71	$\begin{pmatrix} -1.119 & -0.102 & 4.731 \\ -0.021 & -0.008 & 1.190 \end{pmatrix} x + \begin{pmatrix} -6.295 \\ -0.225 \end{pmatrix}$
72	$\begin{pmatrix} -1.081 & -0.089 & 4.549 \\ -0.867 & -0.295 & 5.223 \end{pmatrix} x + \begin{pmatrix} -6.136 \\ -3.757 \end{pmatrix}$
73	$\begin{pmatrix} -1.092 & -0.021 & 4.440 \\ -0.820 & -0.111 & 3.574 \end{pmatrix} x + \begin{pmatrix} -6.118 \\ -2.794 \end{pmatrix}$
74	$\begin{pmatrix} -1.065 & -0.021 & 4.395 \\ -0.858 & -0.111 & 3.638 \end{pmatrix} x + \begin{pmatrix} -6.038 \\ -2.908 \end{pmatrix}$
75	$\begin{pmatrix} -1.473 & 0.014 & 6.412 \\ -0.048 & -0.117 & 0.374 \end{pmatrix} x + \begin{pmatrix} -7.629 \\ 0.087 \end{pmatrix}$
76	$\begin{pmatrix} -1.303 & -0.031 & 5.787 \\ -0.434 & -0.014 & 1.796 \end{pmatrix} x + \begin{pmatrix} -7.052 \\ -1.227 \end{pmatrix}$
77	$\begin{pmatrix} -1.316 & -0.081 & 5.438 \\ -0.432 & -0.007 & 1.845 \end{pmatrix} x + \begin{pmatrix} -7.020 \\ -1.232 \end{pmatrix}$
78	$\begin{pmatrix} -1.231 & 0.014 & 6.002 \\ -0.060 & -0.117 & 0.395 \end{pmatrix} x + \begin{pmatrix} -6.902 \\ 0.050 \end{pmatrix}$
79	$\begin{pmatrix} -1.046 & 0.064 & 5.107 \\ -0.396 & -0.207 & 2.011 \end{pmatrix} x + \begin{pmatrix} -6.134 \\ -1.338 \end{pmatrix}$
80	$\begin{pmatrix} -1.020 & -0.034 & 4.212 \\ -0.388 & -0.236 & 1.743 \end{pmatrix} x + \begin{pmatrix} -5.879 \\ -1.262 \end{pmatrix}$
81	$\begin{pmatrix} -1.092 & -0.031 & 4.357 \\ -0.820 & -0.014 & 4.418 \end{pmatrix} x + \begin{pmatrix} -6.101 \\ -2.971 \end{pmatrix}$

j	$u_j(x) = K_j x + g_j$
82	$\begin{pmatrix} -1.473 & 0.064 & 5.833 \\ -0.048 & -0.207 & 1.419 \end{pmatrix} x + \begin{pmatrix} -7.422 \\ -0.288 \end{pmatrix}$
83	$\begin{pmatrix} -1.231 & -0.081 & 5.294 \\ -0.060 & -0.007 & 1.212 \end{pmatrix} x + \begin{pmatrix} -6.765 \\ -0.109 \end{pmatrix}$
84	$\begin{pmatrix} -1.065 & -0.034 & 4.520 \\ -0.858 & -0.236 & 4.934 \end{pmatrix} x + \begin{pmatrix} -6.084 \\ -3.383 \end{pmatrix}$
85	$\begin{pmatrix} 0.473 & 0.301 & -2.710 \\ -1.378 & -0.163 & 6.021 \end{pmatrix} x + \begin{pmatrix} 0.230 \\ -5.045 \end{pmatrix}$
86	$\begin{pmatrix} -1.092 & 0.301 & 0.079 \\ -0.820 & -0.163 & 5.026 \end{pmatrix} x + \begin{pmatrix} -5.297 \\ -3.074 \end{pmatrix}$
87	$\begin{pmatrix} -0.104 & 0.064 & -4.200 \\ -0.638 & -0.207 & 4.237 \end{pmatrix} x + \begin{pmatrix} -1.779 \\ -2.425 \end{pmatrix}$
88	$\begin{pmatrix} -0.159 & 0.362 & -0.926 \\ -0.646 & -0.165 & 4.700 \end{pmatrix} x + \begin{pmatrix} -2.009 \\ -2.457 \end{pmatrix}$
89	$\begin{pmatrix} -0.901 & 0.042 & 3.976 \\ -0.750 & -0.004 & 3.404 \end{pmatrix} x + \begin{pmatrix} -5.279 \\ -2.529 \end{pmatrix}$
90	$\begin{pmatrix} -0.914 & -0.031 & 3.294 \\ -0.752 & -0.014 & 3.310 \end{pmatrix} x + \begin{pmatrix} -5.326 \\ -2.536 \end{pmatrix}$
91	$\begin{pmatrix} -0.927 & 0.048 & 4.086 \\ -0.673 & -0.022 & 3.090 \end{pmatrix} x + \begin{pmatrix} -5.373 \\ -2.258 \end{pmatrix}$
92	$\begin{pmatrix} -1.473 & 0.048 & 5.058 \\ -0.048 & -0.022 & 1.974 \end{pmatrix} x + \begin{pmatrix} -7.300 \\ -0.048 \end{pmatrix}$
93	$\begin{pmatrix} 0.473 & 0.042 & -5.465 \\ -1.378 & -0.004 & 7.724 \end{pmatrix} x + \begin{pmatrix} 0.259 \\ -5.063 \end{pmatrix}$

j	$u_j(x) = K_j x + g_j$
94	$\begin{pmatrix} -0.927 & 0.362 & 0.442 \\ -0.673 & -0.165 & 4.749 \end{pmatrix} x + \begin{pmatrix} -4.722 \\ -2.555 \end{pmatrix}$
95	$\begin{pmatrix} -1.092 & -0.031 & 3.611 \\ -0.820 & -0.014 & 3.431 \end{pmatrix} x + \begin{pmatrix} -5.954 \\ -2.777 \end{pmatrix}$
96	$\begin{pmatrix} -1.473 & 0.064 & 5.212 \\ -0.048 & -0.207 & 0.179 \end{pmatrix} x + \begin{pmatrix} -7.300 \\ -0.045 \end{pmatrix}$