Local Controllability of Affine Distributions

by

César Octavio Aguilar

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for the
degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
January 2010

Abstract

In this thesis, we develop a feedback-invariant theory of local controllability for affine distributions. We begin by developing an unexplored notion in control theory that we call *proper small-time local controllability* (PSTLC). The notion of PSTLC is developed for an abstraction of the well-known notion of a control-affine system, which we call an *affine system*. Associated to every affine system is an *affine distribution*, an adaptation of the notion of a distribution. Roughly speaking, an affine distribution is PSTLC if the local behaviour of every affine system that locally approximates the affine distribution is locally controllable in the standard sense. We prove that, under a regularity condition, the PSTLC property can be characterized by studying control-affine systems.

The main object that we use to study PSTLC is a cone of high-order tangent vectors, or *variations*, and these are defined using the vector fields of the affine system. To better understand these variations, we study how they depend on the jets of the vector fields by studying the Taylor expansion of a composition of flows. Some connections are made between labeled rooted trees and the coefficients appearing in the Taylor expansion of a composition of flows. Also, a relation between variations and the formal Campbell–Baker–Hausdorff formula is established.

After deriving some algebraic properties of variations, we define a variational cone for an affine system and relate it to the local controllability problem. We then study the notion of neutralizable variations and give a method for constructing subspaces

of variations.

Finally, using the tools developed to study variations, we consider two important classes of systems: driftless and homogeneous systems. For both classes, we are able to characterize the PSTLC property.

Acknowledgments

I am grateful to my supervisor, Andrew Lewis, for the support he has given me in what now seems a short journey. Thank you for all our one-to-one research sessions, exposing me to rich mathematics through our seminars, for the career advice, and for telling me that I have enough for a thesis.

I am grateful for the staff at the Mathematics and Statistics Department, particularly Jennifer Read, whose behind the scene work made my career as a graduate student a smooth one by ensuring all the supports were in place.

I am grateful to my colleague and friend Bahman Gharesifard for making my experience ever more enjoyable, both academically and personally.

I am grateful to my set of parents, brothers, and sisters for their support and telling me that they are proud.

I am grateful to my wife Dawn and my son Simon for all the unconditional love they have given me throughout this endeavor. I love you both very dearly.

I am grateful for all the wisdom, clarity of thought, patience, confidence, and perspective God has blessed me with to complete this part of my life.

"*I can do everything through him who gives me strength.*" Philippians 4:13

Statement of originality

I herby declare that, except where it is indicated with acknowledgment and citation, the results of this thesis are original.

Contents

Chapter 1

Introduction

1.1 Literature review

Beginning with the work of Chow [14], it has been known for some time that Lie bracket configurations are the key objects to study in the problem of differential geometric controllability. In 1966, Nagano [34] published a result, later generalized by Sussmann [42], that gave a precise reason as to why this is so. To state the Nagano–Sussmann theorem, let us introduce some notation. For a real analytic manifold M, $\Gamma(\mathsf{TM})$ denotes the Lie algebra of real analytic vector fields on M. Given a subset L of $\Gamma(\mathsf{TM})$ and $x \in \mathsf{M}$, let $L(x) = \{\xi(x) \mid \xi \in L\}$. If L is a Lie subalgebra of $\Gamma(\mathsf{TM})$, the set of all $\xi \in L$ such that $\xi(x) = 0_x$, i.e., $\xi(x)$ is the zero vector at x, is called the *isotropy subalgebra* of L and is denoted by L_x. A *transitive subalgebra* of $\Gamma(\mathsf{TM})$ is a Lie subalgebra $L \subset \Gamma(\mathsf{TM})$ such that $\dim L(x) = \dim \mathsf{M}$ for every $x \in \mathsf{M}$. With this notation we state a slightly weaker version of the Nagano–Sussmann theorem.

Theorem 1.1 (Nagano–Sussmann) *Let M and N be connected and simply connected real analytic manifolds. Let $L \subset \Gamma(\mathsf{TM})$ and let $L' \subset \Gamma(\mathsf{TN})$ be transitive Lie subalgebras such that each element of L and L' is complete. Let $\Psi\colon L \to L'$*

be a Lie algebra isomorphism. Assume there exists $x \in \mathsf{M}$ *and* $y \in \mathsf{N}$ *such that* $\Psi(L_x) = L'_y$*. Then there exists a unique diffeomorphism* $\phi\colon \mathsf{M} \to \mathsf{N}$ *such that* $\phi(x) = y$ *and* $\phi_*(\xi) = \Psi(\xi)$ *for every* $\xi \in L$*.*

A local version of the Nagano–Sussmann theorem holds in which the connected, simply connected, and complete assumptions can be dropped, thereby only resulting in the existence of a local diffeomorphism $\phi\colon \Omega \to \Omega'$, with $\phi(x) = y$, such that for every $\xi \in L$, the restriction of ξ to the neighbourhood Ω of x and the restriction of $\Phi(\xi)$ to the neighbourhood Ω' of y correspond under ϕ.

With the knowledge that Lie bracket configurations completely determine the local behaviour of the trajectories of a family of vector fields that generate a transitive Lie subalgebra, a systematic effort to characterize controllability in terms of Lie bracket configurations has resulted in many sufficient conditions for local controllability [31], [46], [18], [30], [19], [43], [20], [44], [7], [45], [5], [8], [9], [2], [22], [28]. A class of systems that has received a lot of the attention in this effort are the so-called *control-affine systems*. These are control systems specified by a family of vector fields $\mathcal{X} = \{X_0, X_1, \dots, X_m\}$ and a subset $U \subset \mathbb{R}^m$, and whose trajectories are absolutely continuous curves $\gamma\colon [0, T] \to \mathsf{M}$ satisfying the differential equation

$$\gamma'(t) = X_0(\gamma(t)) + \sum_{a=1}^{m} u^a(t) X_a(\gamma(t)),$$

for some Lebesgue integrable U-valued function $t \mapsto u(t) = (u^1(t), \dots, u^m(t))$ on $[0, T]$. Assuming that the family of vector fields

$$\mathcal{X}_U = \left\{ X_0 + \sum_{a=1}^{m} u^a X_a \;\middle|\; u \in U \right\}$$

generates a Lie subalgebra that is locally transitive about $x_0 \in \mathsf{M}$, and under mild geometric assumptions on the set U, local controllability from x_0 for the control-affine system $(\mathcal{X}, U)$ is equivalent to studying the local behaviour of the set of trajectories emanating from x_0 of the family of vector fields $\mathcal{X}_U$, i.e., by studying the set of end-points

$$\Phi^{\xi_p}_{t_p} \circ \cdots \circ \Phi^{\xi_1}_{t_1}(x_0), \tag{1.1}$$

for $t_1, \ldots, t_p$ sufficiently small, $\xi_1, \ldots, \xi_p \in \mathcal{X}_U$, $p \geq 1$, and where $(t, x) \mapsto \Phi^{\xi}_t(x)$ denotes the local flow generated by ξ. By the Nagano–Sussmann theorem, the local behaviour of the set of points of the form (1.1) can be described using the isotropy subalgebra of the Lie algebra generated by $\mathcal{X}_U$, the latter we denote by $\mathrm{Lie}(\mathcal{X}_U)$. Assuming that U affinely spans $\mathbb{R}^m$, it is easy to show that $\mathrm{Lie}(\mathcal{X}_U) = \mathrm{Lie}(\mathcal{X})$, and thus much attention has been given to studying the isotropy subalgebra $\mathrm{Lie}(\mathcal{X})_{x_0}$. There are some inherent difficulties that arise, however, by fixing one's attention on $\mathrm{Lie}(\mathcal{X})$ in the way that has been done in the literature. To be precise, and at the same time keep the discussion as simple as possible, let $\Lambda = (\Lambda^a_b)$ be an invertible $m \times m$ matrix, and set $Y_0 = X_0$, set $Y_b = \sum_{a=1}^m \Lambda^a_b X_a$, for $b = 1, \ldots, m$, set $\mathcal{Y} = \{Y_0, Y_1, \ldots, Y_m\}$, and finally set $V = \Lambda(U)$. Then, it is easily seen that the trajectories of the control-affine system $(\mathcal{Y}, V)$ are the same as those of $(\mathcal{X}, U)$. Currently, many sufficient conditions for local controllability for control-affine systems, for example those of [45] which generalize many known results, are not invariant under the Lie algebra isomorphism $\Psi\colon \mathrm{Lie}(\mathcal{X}) \to \mathrm{Lie}(\mathcal{Y})$ that is induced by the mapping $X_j \mapsto Y_j$, for $j \in \{0, 1, \ldots, m\}$. In other words, the obtained results are not invariant under feedback transformations. Let us illustrate this with a simple example.

Example 1.2 Consider the following data:

$$\mathsf{M} = \mathbb{R}^3,\ x_0 = (0,0,0),\ X_0 = ((x^1)^2 - 2(x^2)^2)\frac{\partial}{\partial x^3},\ X_1 = \frac{\partial}{\partial x^1},\ X_2 = \frac{\partial}{\partial x^2}.$$

Let $\Sigma = (\{X_0, X_1, X_2\}, U)$, where U is the unit cube in $\mathbb{R}^2$ centered at the origin. By

Theorem 7.3 in [45], Σ is locally controllable from x_0 if

$$[X_1, [X_1, X_0]](x_0) + [X_2, [X_2, X_0]](x_0) \in \operatorname{span}\{X_1(x_0), X_2(x_0)\}.$$

One can check that this condition does not hold, and so the theorem is inconclusive. Consider the matrix $\Lambda = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}$, and let $\mathcal{Y} = \{Y_0, Y_1, Y_2\}$ be defined as above. One checks that

$$[Y_1, [Y_1, Y_0]](x_0) + [Y_2, [Y_2, Y_0]](x_0) = 0_{x_0},$$

and thus, by Theorem 7.3 in [45], the control-affine system $(\mathcal{Y}, V)$ is locally controllable from x_0 and, therefore, so is Σ. In this case, the Lie algebra isomorphism induced by the mapping $X_j \mapsto Y_j$, for $j \in \{0, 1, 2\}$, does not preserve the isotropy subalgebras $\mathrm{Lie}(\mathcal{X})_{x_0}$ and $\mathrm{Lie}(\mathcal{Y})_{x_0}$ because

$$[X_1, [X_1, X_0]](x_0) + [X_2, [X_2, X_0]](x_0) \neq 0_{x_0}.$$

With the above example in mind, one sees that, in order to obtain a feedback-invariant result for local controllability using the current methods, one must verify that the obtained conditions are satisfied by all families representing the same control-affine system, i.e., by all families that generate the same trajectories. The work of Elkin [16] on the equivalence of control-affine systems can be used to start such an approach. Instead, one can consider the affine distribution generated by a control-affine system. Indeed, a control-affine system $\Sigma = (\{X_0, X_1, \dots, X_m\}, U)$ on M generates the affine distribution $\mathsf{A}_\Sigma \subset \mathsf{TM}$ defined by

$$\mathsf{A}_\Sigma(x) = X_0(x) + \operatorname{span}\{X_1(x), \dots, X_m(x)\},$$

and two control-affine systems have the same trajectories if and only if the affine distributions they generate are the same [16, pg. 117]. Hence, a feedback-invariant theory can be developed by studying the local controllability of affine distributions. To obtain a practical theory, however, one should consider affine distributions with some regularity properties, e.g., possessing local generators. But all constructions should be developed in a generator independent way.

1.2 Contribution of thesis

In this thesis, we propose a feedback-invariant theory of local controllability for affine distributions. The main approach is to use the jets of sections of the affine distribution to study high-order tangent vectors to the reachable set. Below we outline the contents and contributions of the thesis.

- In Chapter 2, we establish our notation and review some basic material from jet bundle theory and set-valued maps. We also prove a technical result regarding the high-order derivatives of an integral curve with respect to a parameter.

- In Chapter 3, we begin by laying a basic foundation for the study of a generator independent theory of local controllability for affine distributions. We start by defining the notion of an *affine system*, which can be seen as a generalization of a control-affine system. With affine systems in hand, we are then able to give a definition of local controllability for affine distributions that we call *proper small-time local controllability* (PSTLC). We then prove that, in the regular case, our notion of PSTLC can be characterized by studying control-affine systems.

- In Chapter 4, we define a type of high-order tangent vector, which we call an *end-time variation*. These tangent vector variations are constructed by concatenating flows of vector fields and parameterizing the switching time between the integral curves of the flows. To better understand these variations, we study how they depend on the jets of the vector fields by studying the Taylor expansion of a composition of flows. This study leads to a theorem which asserts the existence of a linear map on an appropriate jet space of the tangent bundle whose image describes the set of variations. Some connections are made between labeled rooted trees and the coefficients appearing in the Taylor expansion of a composition of flows. We end the chapter by relating variations to the formal Campbell–Baker–Hausdorff formula.

- Using the tools developed in Chapter 4, in Chapter 5 we study a variational cone and its connection with the local controllability problem. We then study the variational cone at low orders and give a method for constructing subspaces in the variational cone.

- In Chapter 6, we consider two important classes of systems, namely, driftless systems and homogeneous systems. For driftless systems, we prove that, under the standard regularity assumptions, there are no obstructions to local controllability. Also, we give a simple proof using our methods to show that, for driftless systems, any Lie bracket direction is realizable as a variation. We then move onto homogeneous systems, which play a key role in many known sufficient conditions for local controllability. We prove that, for homogeneous systems, the variational cone contains all the information needed for the characterization of local controllability. The proof of this result is constructive in the sense that it gives a method for determining the directions that will verify the local controllability, or lack thereof, of the system. Furthermore, for these systems we are able to answer an open question in control theory regarding whether it is possible to determine local controllability in a finite number of differentiations.

- We end the thesis with a summary of the main results and list some natural problems to study using our methods. We also describe how our methods can be used to study Kawski's fast-switching example [24].

Chapter 2

Preliminaries

In this section we establish some of our notation and review some material from jet bundle theory and set-valued maps. We also prove a proposition on the high-order derivatives of the solution of an ODE with respect to a parameter.

2.1 Notation and conventions

If f is a mapping, its domain is denoted by $\mathrm{dom}(f)$ and its image by $\mathrm{img}(f)$.

Let V be a finite-dimensional vector space. Most of our notation regarding vector spaces and linear maps can be found in [12]. The convex hull, affine hull, cone hull, and interior of a set S are denoted by $\mathrm{co}(S)$, $\mathrm{aff}(S)$, $\mathrm{cone}(S)$, and $\mathrm{int}(S)$, respectively. The interior of S relative to W is denoted by $\mathrm{int}_W(S)$. Given a linear map f, $\ker(f)$ will denote its kernel. We identify the kth tensor power of V^*, denoted by $T^k(V^*)$, with the set of k-multilinear maps from V to $\mathbb{R}$, denoted by $L^k(V;\mathbb{R})$. Similarly, we identify $S^k(V^*)$, the kth symmetric power of V^*, with the set of symmetric k-multilinear maps from V to $\mathbb{R}$, denoted by $L^k_{\mathrm{sym}}(V;\mathbb{R})$. With these identifications we have that $T^k(V^*)\otimes W \cong L^k(V;W)$ and that $S^k(V^*)\otimes W \cong L^k_{\mathrm{sym}}(V;W)$. The symbol $\odot$ will denote the symmetric product in the symmetric algebra $S(V^*) = \oplus_{\ell=1}^{\infty} S^{\ell}(V^*)$, that is,

$$\alpha \odot \beta = \mathrm{Sym}(\alpha \otimes \beta),$$

where Sym: $T(V^*) \to S(V^*)$ denotes the symmetrization operator given by

$$\mathrm{Sym}(\alpha)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}),$$

where $\mathfrak{S}_k$ denotes the permutation group on k symbols. For each integer $\ell \geq 1$, define $\delta_\ell \colon V \to S^\ell(V)$ by $\delta_\ell(v) = v \otimes v \otimes \cdots \otimes v$. The proof of the following can be found in Propositions 13 and 15, pg. 54-56, [11].

Lemma 2.1 ([11]) *Let V and W be $\mathbb{R}$-vector spaces with V finite-dimensional and let $f \colon V \to W$ be a homogeneous polynomial mapping of degree $\ell \geq 1$. Then there is a unique mapping $h \in L(S^\ell(V); W)$ such that $f(v) = h(\delta_\ell(v))$ for all $v \in V$.*

By a *manifold* we mean a Hausdorff, second countable, connected, smooth manifold. When not explicitly stated, all maps between manifolds will be assumed to be smooth. We will frequently employ the summation convention in which summation is implied over repeated indices. For a manifold M, TM and $\mathsf{T}^*\mathsf{M}$ denote its tangent and cotangent bundle, respectively, and $\mathsf{T}_x\mathsf{M}$ and $\mathsf{T}^*_x\mathsf{M}$ denote its tangent and cotangent space at x, respectively. The zero vector in $\mathsf{T}_x\mathsf{M}$ is denoted by 0_x. If $f \colon \mathsf{M} \to \mathsf{N}$ is a map differentiable at x, $T_x f$ denotes its derivative at x.

For a complete vector field ξ on a manifold M, $\Phi^\xi \colon \mathbb{R} \times \mathsf{M} \to \mathsf{M}$ will denote its flow. For fixed $t \in \mathbb{R}$, Φ^ξ_t is the diffeomorphism $x \mapsto \Phi^\xi(t, x)$, and for fixed $x \in \mathsf{M}$, Φ^ξ_x denotes the curve $t \mapsto \Phi^\xi(t, x)$, i.e., the integral curve of ξ through x.

Let ξ and η be vector fields of M, defined locally about a common point $x_0 \in \mathsf{M}$. We say that ξ and η are *equivalent at* x_0 if there is a neighbourhood Ω of x_0 such that $\xi(x) = \eta(x)$ for all $x \in \Omega$. This defines an equivalence relation on the set of vector fields defined locally about x_0. The *germ* of ξ at x_0 is the equivalence class of

ξ and is denoted by $[\xi_{x_0}]$. The set of germs of vector fields at x_0 has a natural Lie bracket structure inherited from the Lie bracket structure of $\Gamma(\mathsf{TM})$. Indeed, given germs $[\xi_{x_0}]$ and $[\eta_{x_0}]$, we define their Lie bracket $[[\xi_{x_0}], [\eta_{x_0}]]$ by computing the vector field Lie bracket of the vector fields ξ and η on some neighbourhood of x_0 and letting $[[\xi_{x_0}], [\eta_{x_0}]]$ be the germ of the local vector field $[\xi, \eta]$. In this way, if $\mathcal{X}$ is a family of vector fields defined locally about a common point, we can talk about the Lie algebra generated by the family of vector fields $\mathcal{X}$ by passing to germs.

Given a smooth function $f\colon \Omega \to \mathbb{R}^m$, on the open set $\Omega \subset \mathbb{R}^n$, the derivative of f will be denoted $\mathbf{D}f$, i.e., the $\mathbb{R}^{n\times m}$-valued map on $\mathbb{R}^n$ whose ij-entry is $\frac{\partial f^i}{\partial x^j}$. The higher-order derivatives of f are denoted by $\mathbf{D}^{(k)}f$, which is a map from $\mathbb{R}^n$ to $L^k_{\mathrm{sym}}(\mathbb{R}^n; \mathbb{R}^m)$. Also, we denote by $\mathbf{D}^{(k)}_{\ell} f$ the $\mathbb{R}^n$-valued map on $\mathbb{R}^n$ whose ith entry is $\frac{\partial^k f^i}{(\partial x^\ell)^k}$. The zero vector in $\mathbb{R}^p$ will sometimes be denoted by 0_p, to avoid possible confusion with the zero vector in a different Euclidean space.

Finally, the symbol $*$ will denote concatenation. For example, if $x = (x_1, \dots, x_p)$ and $y = (y_1, \dots, y_q)$, then $x * y = (x_1, \dots, x_p, y_1, \dots, y_q)$. We will sometimes write $x * y = (x, y)$. For maps $f_j\colon X_j \to Y_j$, $j = 1, 2$, the symbol $f_1 * f_2$ denotes the map $(f_1 * f_2)(x_1, x_2) = (f_1(x_1), f_2(x_2))$.

2.2 Jets

In this section we review some basic notions from jet bundle theory, all taken from [38, 27].

Given a vector bundle $\pi\colon \mathsf{E} \to \mathsf{M}$, $\Gamma(\mathsf{E})$ will denote its smooth sections. Given $\xi, \eta \in \Gamma(\mathsf{E})$, we say that ξ and η are *k-equivalent at x* if $\xi(x) = \eta(x)$ and if, in some adapted coordinate system around x, the partial derivatives of ξ and η at x agree up to order k. This defines an equivalence relation on the sections of π. The equivalence class of ξ at x of order k is denoted by $j^k_x\xi$ and is called the *k-jet* of ξ at x. The set of all k-jets at x is denoted by $\mathsf{J}^k_x\mathsf{E}$ and the set of all k-jets is denoted by $\mathsf{J}^k\mathsf{E}$. We will sometimes find it convenient to denote these sets as $\mathsf{J}^k_x\pi$ and $\mathsf{J}^k\pi$, respectively. The set $\mathsf{J}^k\mathsf{E}$ can be given the structure of a smooth manifold by using vector bundle coordinates for E to assign the coordinates of $j^k_x\xi$ as the derivatives of ξ up to order k at x. Note that $\mathsf{J}^0\pi$ is naturally identified with E. The map $\pi_k\colon \mathsf{J}^k\mathsf{E} \to \mathsf{M}$ that takes $j^k_x\xi$ to x defines a vector bundle. In $\pi_k^{-1}(x) = \mathsf{J}^k_x\mathsf{E}$, addition and scalar multiplication are defined as $j^k_x\xi + j^k_x\eta = j^k_x(\xi+\eta)$ and $\lambda \cdot j^k_x\xi = j^k_x(\lambda\xi)$, respectively, where $\lambda \in \mathbb{R}$. For non-negative integers $\ell \leq k$, there is a canonical projection $\pi^k_\ell\colon \mathsf{J}^k\mathsf{E} \to \mathsf{J}^\ell\mathsf{E}$, that maps $j^k_x\xi$ to $j^\ell_x\xi$. When $\ell = k-1$, the map $\pi^k_{k-1}\colon \mathsf{J}^k\mathsf{E} \to \mathsf{J}^{k-1}\mathsf{E}$ can be given an affine bundle structure modeled on the pull-back of the vector bundle $S^k(\mathsf{T}^*\mathsf{M}) \otimes \mathsf{E} \to \mathsf{M}$ to $\mathsf{J}^{k-1}\mathsf{E}$. Explicitly, for smooth functions $f_1, \ldots, f_k$ vanishing at x, the action of $(\mathrm{d}f_1(x) \odot \cdots \odot \mathrm{d}f_k(x)) \otimes \eta(x) \in S^k(\mathsf{T}^*_x\mathsf{M}) \otimes \mathsf{E}_x$ on $j^k_x\xi \in \mathsf{J}^k_x\mathsf{E}$ is given by

$$j^k_x\xi + (\mathrm{d}f_1(x) \odot \cdots \odot \mathrm{d}f_k(x)) \otimes \eta(x) = j^k_x(\xi + (f_1 \cdots f_k)\eta).$$

The affine structure can be represented via the following exact sequence of vector bundles over M,

$$0 \longrightarrow S^k(\mathsf{T}^*\mathsf{M}) \otimes \mathsf{E} \xrightarrow{\epsilon_k} \mathsf{J}^k\mathsf{E} \xrightarrow{\pi^k_{k-1}} \mathsf{J}^{k-1}\mathsf{E} \longrightarrow 0$$

where $\epsilon_k\colon S^k(\mathsf{T}^*\mathsf{M}) \otimes \mathsf{E} \to \mathsf{J}^k\mathsf{E}$ is the injection defined as

$$\epsilon_k((\mathrm{d}f_1(x) \odot \cdots \odot \mathrm{d}f_k(x)) \otimes \eta(x)) = j^k_x((f_1 \cdots f_k)\eta),$$

for smooth functions $f_1, \ldots, f_k$ vanishing at x. The elements of

$$\mathrm{img}(\epsilon_k) \cap \mathsf{J}^k_x\mathsf{E} = \ker(\pi^k_{k-1}) \cap \mathsf{J}^k_x\mathsf{E}$$

are the k-jets of sections of E that vanish up to order $k-1$ at x; that is, elements whose $(k-1)$-jet at x agrees with the $(k-1)$-jet at x of the zero section. Given $j^k_{x_0}\zeta \in \ker(\pi^k_{k-1})$, the corresponding element in $S^k(\mathsf{T}^*_{x_0}\mathsf{M}) \otimes \mathsf{E}_{x_0}$ will be denoted by B^k_ζ. In a coordinate system $(x^1, \ldots, x^n)$ about x_0, B^k_ζ is given by

$$B^k_\zeta = \sum_{j=1}^{p} \sum_{I} \frac{\partial^k \zeta^j}{\partial x^I}(x_0)\, \mathrm{d}x^I(x_0) \otimes e_j, \tag{2.1}$$

where $e_1, \ldots, e_p$ is a basis for E_{x_0}, the inner sum runs through all multi-indices $I = (i_1, \ldots, i_k) \subset \{1, \ldots, n\}^k$ of length k,

$$\frac{\partial^k}{\partial x^I} = \frac{\partial^k}{\partial x^{i_1} \partial x^{i_2} \cdots \partial x^{i_k}} \quad \text{and} \quad \mathrm{d}x^I = \mathrm{d}x^{i_1} \otimes \mathrm{d}x^{i_2} \otimes \cdots \otimes \mathrm{d}x^{i_k}.$$

That (2.1) is indeed in $S^k(\mathsf{T}^*_{x_0}\mathsf{M}) \otimes \mathsf{E}_{x_0}$ follows from the symmetry of the derivative.

We will say that $\xi \in \Gamma(\mathsf{E})$ is of *order k at x* if $j^k_x\xi \in \ker(\pi^k_{k-1})$, but $j^k_x\xi$ is not the zero vector. In other words, $\xi \in \Gamma(\mathsf{E})$ is of order k at x if the first non-zero derivatives of ξ at x are of order k.

Given manifolds M and N, define the trivial bundle $\pi_{\mathsf{M}} \colon \mathsf{M} \times \mathsf{N} \to \mathsf{M}$ by $\pi_M(x, y) = x$. A section of π_{M} is naturally identified with a mapping from M to N. The jet space $\mathsf{J}^k(\mathsf{M} \times \mathsf{N})$ can be defined in the same way as was done for vector bundles. Explicitly, $\mathsf{J}^k_{(x,y)}(\mathsf{M} \times \mathsf{N})$ is the space of equivalence classes of mappings from M to N that map x to y and whose derivatives at x agree up to order k. The set $\mathsf{J}^k(\mathsf{M} \times \mathsf{N})$, which we prefer to denote by $\mathsf{J}^k(\mathsf{M}; \mathsf{N})$, is the set of all such equivalence classes. The space $\mathsf{J}^k_{(x,y)}(\mathsf{M}; \mathsf{N})$ can be given an algebraic structure in the following way. We give the vector space

$$\mathsf{T}^{*k}_x\mathsf{M} := \mathsf{J}^k_{(x,0)}(\mathsf{M}; \mathbb{R})$$

a $\mathbb{R}$-algebra structure by defining multiplication as $j^k_x\phi \cdot j^k_x\psi = j^k_x(\phi\psi)$, for smooth functions ϕ, ψ on M. The set $\mathsf{J}^k_{(x,y)}(\mathsf{M}; \mathsf{N})$ can then be identified with the $\mathbb{R}$-algebra homomorphisms $\mathrm{Hom}(\mathsf{T}_y^{*k}\mathsf{N};\mathsf{T}_x^{*k}\mathsf{M})$ by defining, for $j_x^k f \in \mathsf{J}^k_{(x,y)}(\mathsf{M};\mathsf{N})$ and $j_y^k\phi \in \mathsf{T}_y^{*k}\mathsf{N}$,

$$j_x^k f(j_y^k\phi) = j_x^k(\phi\circ f).$$

Similarly to the vector bundle case, we have an exact sequence of vector spaces

$$0 \longrightarrow S^k(\mathsf{T}_x^*\mathsf{M})\otimes\mathsf{T}_y\mathsf{N} \xrightarrow{\epsilon_k} \mathsf{J}^k_{(x,y)}(\mathsf{M};\mathsf{N}) \xrightarrow{\pi^k_{k-1}} \mathsf{J}^{k-1}_{(x,y)}(\mathsf{M};\mathsf{N}) \longrightarrow 0 \tag{2.2}$$

where $\epsilon_k\colon S^k(\mathsf{T}_x^*\mathsf{M})\otimes\mathsf{T}_y\mathsf{N}\to\mathsf{J}^k_{(x,y)}(\mathsf{M};\mathsf{N})$ is the mapping defined by

$$\epsilon_k((\mathrm{d}f_1(x)\odot\ \cdots\ \odot\mathrm{d}f_k(x))\otimes v_y) = j_x^k(\gamma_{v_y}\circ(f_1f_2\cdots f_k)),$$

for smooth functions $f_1,\dots,f_k$ on M that vanish at x, and where $\gamma_{v_y}\colon\mathbb{R}\to\mathsf{N}$ is any curve at y such that $\gamma'_{v_y}(0)=v_y$. A case that will be of interest to us is when $\mathsf{M}=\mathbb{R}$. In this case, since $S^k(\mathsf{T}_x^*\mathbb{R})$ is canonically isomorphic to $\mathbb{R}$, it follows that $S^k(\mathsf{T}_x^*\mathbb{R})\otimes\mathsf{T}_y\mathsf{N}\cong\mathsf{T}_y\mathsf{N}$. Hence, if $\gamma\colon\mathbb{R}\to\mathsf{N}$ is a curve at y such that $j_0^k\gamma\in\ker(\pi^k_{k-1})$, then $j_0^k\gamma$ can be canonically identified with a tangent vector in $\mathsf{T}_y\mathsf{N}$. Hence, the sequence (2.2) becomes

$$0 \longrightarrow \mathsf{T}_y\mathsf{N} \xrightarrow{\epsilon_k} \mathsf{J}^k_{(0,y)}(\mathbb{R};\mathsf{N}) \xrightarrow{\pi^k_{k-1}} \mathsf{J}^{k-1}_{(0,y)}(\mathbb{R};\mathsf{N}) \longrightarrow 0 \tag{2.3}$$

Another important case of the exact sequence (2.2) is when $\mathsf{N}=\mathbb{R}$. In this case, for $y=0$, the sequence (2.2) becomes

$$0 \longrightarrow S^k(\mathsf{T}_x^*\mathsf{M}) \xrightarrow{\epsilon_k} \mathsf{T}_x^{*k}\mathsf{M} \xrightarrow{\pi^k_{k-1}} \mathsf{T}_x^{*(k-1)}\mathsf{M} \longrightarrow 0 \tag{2.4}$$

When $\mathsf{M}=\mathbb{R}^p$, the set $(\mathbb{R}^p)^{*k} := \mathsf{J}^k_{(0_p,0)}(\mathbb{R}^p;\mathbb{R})$ can be canonically identified with polynomial functions of order k with zero constant term via Taylor's expansion and, therefore, we have the isomorphism

$$(\mathbb{R}^p)^{*k} \cong (\mathbb{R}^p)^{*(k-1)} \oplus S^k((\mathbb{R}^p)^*).$$

Explicitly, for a function $h\colon \mathbb{R}^p \to \mathbb{R}$ vanishing at the origin, $j^k_{0_p} h$ as a polynomial is defined as

$$(j^k_{0_p} h)(\boldsymbol{t}) = \sum_{|I|=1}^{k} \frac{\partial^{|I|} h}{\partial \boldsymbol{t}^I}(0_p) \frac{\boldsymbol{t}^I}{I!},$$

where $|I| := i_1 + \cdots + i_p$, $I! = i_1! i_2! \cdots i_p!$, for a multi-index $I = (i_1, \ldots, i_p) \in \mathbb{Z}^p_{\geq 0}$, and $\boldsymbol{t}^I = t_1^{i_1} t_2^{i_2} \cdots t_p^{i_p}$ for $\boldsymbol{t} = (t_1, \ldots, t_p) \in \mathbb{R}^p$.

2.3 Set-valued maps

In this section we review some basic material from set-valued analysis following [4].

Let A and B be sets. A set-valued map F from A to B, denoted by $F : A \rightrightarrows B$, is a rule that assigns to each $a \in A$ a subset of B, possibly empty. We say that $F : A \rightrightarrows B$ is compact (convex, when B is a vector space) if $F(a)$ is a compact (convex, when B is a vector space) set for each $a \in A$. Suppose that A and B are Hausdorff topological spaces and that $F : A \rightrightarrows B$ has non-empty values. We say that F is *upper semi-continuous* (usc) at $a_0 \in A$ if, for any open set V containing $F(a_0)$, there exists a neighbourhood Ω of a_0 such that $F(\Omega) \subset V$. We say that F is *lower semi-continuous* (lsc) at a_0 if, for any $b_0 \in F(a_0)$ and any neighbourhood V of b_0, there exists a neighbourhood Ω of a_0 such that $F(a) \cap V \neq \emptyset$ for all $a \in \Omega$. We call F *continuous* at a_0 if it is usc and lsc at a_0. If F is usc (lsc, continuous) at each point in A then we say that it is usc (lsc, continuous). A set-valued map is completely determined by its *graph*, that is, the set

$$\operatorname{gph}(F) = \{ (a, b) \in A \times B \mid \; b \in F(a) \}.$$

Given a non-empty subset $G \subset A \times B$, there is a $F : A \rightrightarrows B$ such that $\mathrm{gph}(F) = G$, namely, $F(a) = \{b \in B \mid (a, b) \in G\}$. The proof of the following result can be found in [37, Theorem 5.9].

Theorem 2.2 *Suppose that $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ has non-empty values, is convex and $\mathrm{int}(F(a_0)) \neq \emptyset$. Then F is lsc at a_0 if and only if, for all $b_0 \in \mathrm{int}(F(a_0))$, there exists neighbourhoods $\Omega \ni a_0$ and $V \ni b_0$ such that $\Omega \times V \subset \mathrm{gph}(F)$, that is, $V \subset F(a)$ for all $a \in \Omega$.*

The following will be useful.

Proposition 2.3 *Let A be a topological space and let B a topological vector space. Let F_1 be a set-valued map from A to B and define $F_2 : A \rightrightarrows B$ by $F_2(a) = \mathrm{co}(F_1(a))$. If F_1 is lsc at a_0 then so is F_2.*

Proof: Let $b \in F_2(a_0)$ and let V be a neighbourhood of b. We can write that

$$b = \sum_{j=1}^{m} \lambda^j b_j$$

for some $b_1, \ldots, b_m \in F_1(a_0)$ and some $\lambda^1, \ldots, \lambda^m \geq 0$ with $\sum_{j=1}^{m} \lambda^j = 1$. Consider the map $\rho \colon B^m \to B$ defined by

$$\rho(x_1, \ldots, x_m) = \sum_{j=1}^{m} \lambda^j x_j.$$

Then $\rho(b_1, \ldots, b_m) = b \in V$. Because ρ is continuous, for each $j \in \{1, \ldots, m\}$, there is a neighbourhood V_j of b_j such that $\rho(V_1 \times \cdots \times V_m) \subset V$. By lower semi-continuity of F_1 at a_0, there is a neighourhood Ω of a_0 such that $W_j(a) := F_1(a) \cap V_j \neq \emptyset$ for all $a \in \Omega$ and all $j \in \{1, \ldots, m\}$. Then, by the definition of ρ,

$$\rho(W_1(a) \times \cdots \times W_m(a)) \subset F_2(a),$$

for all $a \in \Omega$. This proves that $F_2(a) \cap V \neq \emptyset$ for all $a \in \Omega$, and, therefore, F_2 is lsc at a_0. ■

A set-valued map $F : A \rightrightarrows B$ is *locally C^r selectionable* at a_0 if, for each $b_0 \in F(a_0)$, there exist a neighbourhood Ω of a_0 and a C^r map $f\colon \Omega \to B$ such that $f(a_0) = b_0$ and $f(a) \in F(a)$ for all $a \in \Omega$. We say that F is *locally C^r selectionable* if it is locally C^r selectionable at each point in A. The following is straightforward to show [4].

Proposition 2.4 *A locally continuously selectionable set-valued map is lower semi-continuous.*

2.4 Derivatives of the solution of an ODE with respect to a parameter

In this section we prove a technical result regarding the high-order derivatives of the solution of an ODE with respect to a parameter.

Proposition 2.5 *Let $I \subset \mathbb{R}$ be an open interval containing the origin and suppose that $\zeta\colon I \times \mathsf{M} \to \mathsf{TM}$ is a smooth map with the following properties:*

(i) $\zeta_s = \zeta(s, \cdot)$ *is a smooth vector field on* M *for each* $s \in I$*;*

(ii) $\zeta(0, x) = 0_x$ *for all* $x \in \mathsf{M}$*;*

(iii) the curve $\zeta_{x_0}\colon I \to \mathsf{T}_{x_0}\mathsf{M}$ *defined by* $\zeta_{x_0}(s) = \zeta(s, x_0)$ *has vanishing derivatives of orders* $1, 2, \dots, \ell - 1$ *at* $s = 0$*.*

If v_s denotes the integral curve of the vector field ζ_s through x_0, then the curve $s \mapsto \gamma(s) = v_s(1)$ satisfies $j_0^\ell \gamma = j_0^\ell \zeta_{x_0} \in \mathsf{T}_{x_0}\mathsf{M}$.

Remark 2.6 Let $T > 0$. From the continuous dependence on the parameters of the solution of a smooth differential equation and the fact that the integral curves of the zero vector field $\zeta_0 = \zeta(0, \cdot)$ exist on any compact interval $[-T, T]$, the integral curves of the vector field ζ_s also exist on the interval $[-T, T]$ provided s is sufficiently close to zero.

The proof of Proposition 2.5 will follow from the next two lemmas.

Lemma 2.7 *Let $\zeta\colon I \times \Omega \to \mathbb{R}^n$ be a smooth map, where $I \subset \mathbb{R}$ is an interval containing the origin and where $\Omega \subset \mathbb{R}^n$ is an open set containing the origin, and suppose that $\zeta(0,x) = 0$ for all $x \in \Omega$. Let $v\colon (-\delta,\delta) \times [-T,T] \to \mathbb{R}^n$ be the smooth map such that $v_s = v(s,\cdot)$ is the integral curve of ζ_s through $x_0 = 0 \in \Omega$. Then, for each integer $\ell \geq 1$, it holds that*

$$\frac{\partial}{\partial t}\left(\frac{\partial^\ell v}{\partial s^\ell}\right) = \mathbf{D}_1^{(\ell)}\zeta(s,v) + \mathbf{D}_2\zeta(s,v)\cdot\frac{\partial^\ell v}{\partial s^\ell} + G_\ell\left(s, v, \frac{\partial v}{\partial s}, \dots, \frac{\partial^{\ell-1} v}{\partial s^{\ell-1}}\right), \tag{2.5}$$

where $G_\ell\colon W \to \mathbb{R}^n$ is a smooth map on a neighbourhood $W \subset \mathbb{R}\times(\mathbb{R}^n)^\ell$ of the origin such that $G_\ell(s,y_1,0,\dots,0) = 0$, for all $(s,y_1) \in \mathbb{R}\times\mathbb{R}^n$. Consequently,

$$\frac{\partial^\ell v}{\partial s^\ell}(0,t) = \mathbf{D}_1^{(\ell)}\zeta(0,x_0)t + \int_0^t G_\ell\left(0, x_0, \frac{\partial v}{\partial s}(0,\sigma), \dots, \frac{\partial^{\ell-1} v}{\partial s^{\ell-1}}(0,\sigma)\right)\mathrm{d}\sigma. \tag{2.6}$$

Proof: The proof is by induction on ℓ. By definition of v,

$$v(s,t) = v_s(t) = x_0 + \int_0^t \zeta(s, v_s(\sigma))\,\mathrm{d}\sigma,$$

and, therefore,

$$\frac{\partial}{\partial t}\left(\frac{\partial v}{\partial s}\right) = \mathbf{D}_1\zeta(s,v) + \mathbf{D}_2\zeta(s,v)\cdot\frac{\partial v}{\partial s}.$$

The claim holds for $\ell = 1$ by setting $G_1(s,y_1) = 0$. Assume the claim holds for $\ell \geq 1$. Let $\mathbf{D}_k G_\ell$ denote the derivative of G_ℓ with respect to y_k for $k \in \{1,\dots,\ell\}$. By commutativity of partial differentiation,

$$\frac{\partial}{\partial t}\left(\frac{\partial^{\ell+1} v}{\partial s^{\ell+1}}\right) = \frac{\partial}{\partial s}\left(\frac{\partial}{\partial t}\left(\frac{\partial^\ell v}{\partial s^\ell}\right)\right).$$

Hence, by the induction hypothesis and the chain-rule (we omit evaluation at (s,v)

for the derivatives of ζ for compactness of notation),

$$\frac{\partial}{\partial t}\left(\frac{\partial^{\ell+1} v}{\partial s^{\ell+1}}\right) = \mathbf{D}_1^{(\ell+1)}\zeta + \mathbf{D}_2\mathbf{D}_1^{(\ell)}\zeta\cdot\frac{\partial v}{\partial s} + \mathbf{D}_1\mathbf{D}_2\zeta\cdot\frac{\partial^\ell v}{\partial s^\ell} + \mathbf{D}_2^{(2)}\zeta\cdot\left(\frac{\partial v}{\partial s},\frac{\partial^\ell v}{\partial s^\ell}\right) + \mathbf{D}_2\zeta\cdot\frac{\partial^{\ell+1} v}{\partial s^{\ell+1}}$$
$$+ \frac{\partial G_\ell}{\partial s}\left(s, v, \frac{\partial v}{\partial s}, \dots, \frac{\partial^{\ell-1} v}{\partial s^{\ell-1}}\right) + \sum_{k=1}^{\ell}\mathbf{D}_k G_\ell\left(s, v, \frac{\partial v}{\partial s}, \dots, \frac{\partial^{\ell-1} v}{\partial s^{\ell-1}}\right)\cdot\frac{\partial^k v}{\partial s^k}.$$

Let

$$G_{\ell+1}(s, y_1, \dots, y_{\ell+1}) = \mathbf{D}_2\mathbf{D}_1^{(\ell)}\zeta(s, y_1)\cdot y_2 + \mathbf{D}_1\mathbf{D}_2\zeta(s, y_1)\cdot y_{\ell+1} + \mathbf{D}_2^{(2)}\zeta(s, y_1)\cdot(y_2, y_{\ell+1})$$
$$+ \frac{\partial G_\ell}{\partial s}(s, y_1, \dots, y_\ell) + \sum_{k=1}^{\ell}\mathbf{D}_k G_\ell(s, y_1, \dots, y_\ell)\cdot y_{k+1}.$$

Then $G_{\ell+1}$ is clearly smooth, and, moreover,

$$G_{\ell+1}(s, y_1, 0, \dots, 0) = \frac{\partial G_\ell}{\partial s}(s, y_1, 0, \dots, 0) = 0,$$

and the latter vanishes by the induction hypothesis. This proves (2.5). Equation (2.6) now follows because $\mathbf{D}_2\zeta(0, x)$ is the zero matrix for all $x \in \Omega$ and v_0 is constant and equal to x_0. This completes the proof. ■

Lemma 2.8 *Let ζ and v be given as in Lemma 2.7. Let $\ell \geq 1$ be an integer and suppose that $\mathbf{D}_1^{(j)}\zeta(0, x_0) = 0$ for $j \in \{0, \dots, \ell-1\}$. Then, for all $t \in (-T, T)$ and for all $j \in \{0, \dots, \ell-1\}$, $\frac{\partial^j v}{\partial s^j}(0, t) = 0$ and*

$$\frac{\partial^\ell v}{\partial s^\ell}(0, t) = \mathbf{D}_1^{(\ell)}\zeta(0, x_0)t.$$

Proof: The proof is by induction on $\ell \geq 1$. For $\ell = 1$, we have that $G_1 \equiv 0$, and, therefore, by (2.5) it holds that $\frac{\partial v}{\partial s}(0, t) = \mathbf{D}_1\zeta(0, x_0)t$ for all $t \in (-T, T)$. The proves the claim for $\ell = 1$.

Now suppose that $\mathbf{D}_1^{(j)}\zeta(0, x_0) = 0$ for $j \in \{0, 1, \dots, \ell\}$ and assume the claim for $\ell \geq 1$. By the induction hypothesis, $\frac{\partial^j v}{\partial s^j}(0, t) = 0$ for $j \in \{0, 1, \dots, \ell-1\}$ and $\frac{\partial^\ell v}{\partial s^\ell}(0, t) = \mathbf{D}_1^{(\ell)}\zeta(0, x_0)t$, for all $t \in (-T, T)$. By assumption, $\mathbf{D}_1^{(\ell)}\zeta(0, x_0) = 0$, and thus also $\frac{\partial^\ell v}{\partial s^\ell}(0,t) = 0$, for all $t \in (-T,T)$. Then by (2.6), for all $t \in (-T,T)$,

$$\begin{aligned}\frac{\partial^{\ell+1} v}{\partial s^{\ell+1}}(0,t) &= \mathbf{D}_1^{(\ell+1)}\zeta(0,x_0)t + \int_0^t G_{\ell+1}\left(0, x_0, \frac{\partial v}{\partial s}(0,\sigma), \ldots, \frac{\partial^\ell v}{\partial s^\ell}(0,\sigma)\right) \mathrm{d}\sigma \\ &= \mathbf{D}_1^{(\ell+1)}\zeta(0,x_0)t,\end{aligned}$$

and this completes the proof. ■

Proof of Proposition 2.5: Choose a coordinate neighbourhood of x_0, mapping x_0 to the origin, and let, by abuse of notation, $\zeta\colon I\times\Omega \to \mathbb{R}^n$ be the coordinate representation of $\zeta\colon I \times \mathsf{M} \to \mathsf{TM}$ about x_0. By assumption, $\mathbf{D}_1^{(j)}\zeta(0,x_0) = 0_{x_0}$ for $j \in \{0,1,\ldots,\ell-1\}$, and, therefore, by Lemma 2.8, $\frac{\partial^j v}{\partial s^j}(0,t) = 0$ for all $j \in \{0,1,\ldots,\ell-1\}$ and

$$\frac{\partial^\ell v}{\partial s^\ell}(0,t) = \mathbf{D}_1^{(\ell)}\zeta(0,x_0)t,$$

for all $t \in (-T,T)$. Choosing $T > 1$ and setting $t = 1$ completes the proof. ■

Chapter 3

Local controllability of affine distributions

The purpose of this chapter is to lay the basic foundation for a theory of local controllability for affine distributions. We start by reviewing the notion of an affine distribution and prove some basic results regarding their local structure. We then introduce an unexplored notion in control theory called an *affine system* and use it to give a definition of local controllability for an affine distribution. A special type of affine system is a control-affine system, and we prove that, in the regular case, it is sufficient to study control-affine systems to prove controllability of affine distributions. Even with this being the case, the setting of affine systems has the advantage of forcing one's viewpoint to be feedback-invariant.

3.1 Affine distributions

By an *affine distribution* on M we mean a subset $\mathsf{A} \subset \mathsf{TM}$ such that, for each $x \in \mathsf{M}$, $\mathsf{A}_x := \mathsf{A} \cap \mathsf{T}_x\mathsf{M}$ is an affine subspace of $\mathsf{T}_x\mathsf{M}$. We say that A is *smooth* if, for each $x_0 \in \mathsf{M}$, there exists a neighbourhood Ω of x_0 and smooth vector fields $X_0, X_1, \ldots, X_m$

on Ω such that

$$\mathsf{A}_x = \{X_0(x)\} + \operatorname{span}\{X_1(x), \ldots, X_m(x)\}$$

for all $x \in \Omega$. The set of vector fields $\{X_0, X_1, \ldots, X_m\}$ is called a *local frame* for A at x_0. Henceforth, we deal exclusively with smooth affine distributions.

An affine distribution will be called a *distribution* if A_x is a subspace for each x. Distributions will typically be denoted with the symbol D. A vector field ξ is said to belong to A, if $\xi(x) \in \mathsf{A}_x$ for each x in the domain of ξ. We let $\Gamma(\mathsf{A})$ denote the set of smooth A-vector fields and let $\Gamma_x(\mathsf{A})$ denote the set of smooth A-vector fields containing x in their domain. The *linear part* of A at x is denoted by $\mathsf{L}(\mathsf{A})_x$, and $\mathsf{L}(\mathsf{A}) \subset \mathsf{TM}$ denotes the corresponding distribution on M. Explicitly,

$$\mathsf{L}(\mathsf{A})_x = \{\, \xi_2(x) - \xi_1(x) \mid \; \xi_1, \xi_2 \in \Gamma_x(\mathsf{A})\} \,.$$

We say that x_0 is a *regular point* of A if there is a neighbourhood of x_0 in which the dimension of the subspace $\mathsf{L}(\mathsf{A})_x$ is constant, and we call A *regular* if it is regular at every point. We say that A is *singular* at x_0 if it is not regular at x_0. The following three lemmas, describing some local properties of affine distributions, will prove to be useful in subsequent analysis.

Lemma 3.1 *Let* $k = \dim(\mathsf{L}(\mathsf{A})_{x_0})$. *There is a local frame* $\{X_0, X_1, \ldots, X_m\}$ *for* A *at* x_0 *such that* $X_{k+j}(x_0) = 0_{x_0}$ *for* $j = 1, \ldots, m-k$. *Moreover, if* $0_{x_0} \in \mathsf{A}_{x_0}$, *then* X_0 *can be chosen to satisfy* $X_0(x_0) = 0_{x_0}$.

Proof: To prove the first statement, let $r = m - k$ and assume that $r > 0$; otherwise there is nothing to prove. Let $\{X_0, X_1, \ldots, X_k, Y_1, \ldots, Y_r\}$ be a local frame for A on Ω about x_0 such that $\mathsf{L}(\mathsf{A})_{x_0} = \operatorname{span}\{X_1(x_0), \ldots, X_k(x_0)\}$. For each $j \in \{1, \ldots, r\}$, there exist $\lambda_j \in \mathbb{R}^k$ such that $Y_j(x_0) = \lambda_j^b X_b(x_0)$ (we are employing the summation convention). Let $X_{k+j} = Y_j - \lambda_j^b X_b$. Then $X_{k+j} \in \mathsf{L}(\mathsf{A})$ and $X_{k+j}(x_0) = 0_{x_0}$, for

$j \in \{1, \ldots, r\}$. Let $\Lambda = [\lambda_1 \; \lambda_2 \quad \cdots \quad \lambda_r] \in \mathbb{R}^{k \times r}$, i.e., Λ's jth column is λ_j. Now let $x \in \Omega$ and $v_x \in \mathsf{L}(\mathsf{A})_x$ be arbitrary. Then $v_x = \alpha_1^b X_b(x) + \alpha_2^c Y_c(x)$ for some $(\alpha_1, \alpha_2) \in \mathbb{R}^k \times \mathbb{R}^r$. Set $\mu_1 = \alpha_1 + \Lambda\alpha_2$ and $\mu_2 = \alpha_2$. Then,

$$\begin{aligned}\mu_1^b X_b(x) + \mu_2^c X_{k+c}(x) &= \left(\alpha_1^b + \lambda_c^b \alpha_2^c\right) X_b(x) + \alpha_2^c \left(Y_c(x) - \lambda_c^b X_b(x)\right) \\ &= \alpha_1^b X_b(x) + \alpha_2^c Y_c(x) + (\lambda_c^b \alpha_2^c - \lambda_c^b \alpha_2^c) X_b(x) = v_x.\end{aligned}$$

This proves that $\{X_1, \ldots, X_m\}$ is a local frame for $\mathsf{L}(\mathsf{A})$ on Ω, and, therefore, $\{X_0, X_1, \ldots, X_m\}$ is a local frame for A on Ω.

To prove the second statement, suppose that $0_{x_0} \in \mathsf{A}_{x_0}$ and let $\{Y_0, X_1, \ldots, X_m\}$ be a local frame for A on Ω about x_0. There is a $\lambda \in \mathbb{R}^m$ such that $0_{x_0} = Y_0(x_0) + \lambda^b X_b(x_0)$. Let $X_0 = Y_0 + \lambda^b X_b$, so that $X_0(x_0) = 0_{x_0}$, and, moreover $X_0 \in \mathsf{A}$. Let $x \in \Omega$ and $v_x \in \mathsf{A}_x$ be arbitrary. We can write $v_x = Y_0(x) + \alpha^b X_b(x)$ for some $\alpha \in \mathbb{R}^m$. If $\mu = \alpha - \lambda$, then

$$X_0(x) + \mu^b X_b(x) = Y_0(x) + (\mu^b + \lambda^b) X_b(x) = Y_0(x) + \alpha^b X_b(x) = v_x.$$

Thus $\{X_0, X_1, \ldots, X_m\}$ is a local frame for A on Ω. This completes the proof. ■

Lemma 3.2 *Let* A *be regular at* x_0*. Suppose that* $\xi_1, \ldots, \xi_p$ *are* A*-vector fields such that* $\operatorname{aff}(\{\xi_1(x_0), \ldots, \xi_p(x_0)\}) = \mathsf{A}_{x_0}$*. Then there exists a neighbourhood* Ω *of* x_0 *such that*

$$\operatorname{aff}(\{\xi_1(x), \ldots, \xi_p(x)\}) = \mathsf{A}_x$$

for all $x \in \Omega$*.*

Proof: It is straightforward to show that, for any $i \in \{1, \ldots, p\}$,

$$\operatorname{aff}(\{\xi_1(x_0), \ldots, \xi_p(x_0)\}) = \xi_i(x_0) + \operatorname{span}\{\xi_j(x_0) - \xi_i(x_0) \mid j \neq i\},$$

so that

$$\operatorname{span}\{\xi_j(x_0) - \xi_i(x_0) \mid j \neq i\} = \mathsf{L}(\mathsf{A})_{x_0}.$$

By lower semi-continuity of the rank and regularity of A at x_0, there exists a neighbourhood Ω of x_0 such that $\operatorname{span}\{\xi_j(x) - \xi_i(x) \mid j \neq i\} = \mathsf{L}(\mathsf{A})_x$ for all $x \in \Omega$. Therefore,

$$\mathsf{A}_x = \xi_i(x) + \operatorname{span}\{\xi_j(x) - \xi_i(x) \mid j \neq i\} = \operatorname{aff}(\{\xi_1(x), \dots, \xi_p(x)\})$$

for all $x \in \Omega$. This completes the proof. ■

The next lemma asserts the existence of a particularly nice frame for a regular distribution.

Lemma 3.3 *Let M be an n-manifold and let D be a smooth distribution that is regular at x_0 and of rank m at x_0. Then there is a coordinate system $(x^1, \dots, x^n)$ about x_0 and a local frame $\{X_1, \dots, X_m\}$ for D such that, in these coordinates,*

$$X_j = \frac{\partial}{\partial x^j} + \sum_{\ell=1}^{n-m} c_j^\ell \frac{\partial}{\partial x^{m+\ell}},$$

for smooth functions c_j^ℓ vanishing at x_0.

Proof: Without loss of generality assume that $\mathsf{M} = \mathbb{R}^n$ and that $x_0 = 0$. Let $\{\tilde{X}_1, \dots, \tilde{X}_m\}$ be a local frame for D on a neighbourhood Ω of x_0. Thinking of the $\tilde{X}_j$'s as column vectors, define the matrix $A(x) = [\tilde{X}_1(x)\ \tilde{X}_2(x)\ \cdots\ \tilde{X}_m(x)]$. By permuting the coordinates, if necessary, we can assume that the uppermost $m \times m$ submatrix in $A(x)$ is invertible on Ω. Call this submatrix $\tilde{A}(x)$. Multiplying $A(x)$ on the right by $\tilde{A}(x)^{-1}$, we obtain a new frame $\{X_1, \dots, X_m\}$ for D on Ω of the form

$$X_j = \frac{\partial}{\partial x^j} + \sum_{\ell=1}^{n-m} b_j^\ell \frac{\partial}{\partial x^{m+\ell}}, \qquad j \in \{1, \dots, m\},$$

for smooth functions b_j^ℓ. Define the coordinate change $\Psi \colon \mathbb{R}^n \to \mathbb{R}^n$ by

$$\tilde{x} = \Psi(x^1, \dots, x^n) = (x^1, \dots, x^m, x^{m+1} - \Sigma_{j=1}^m b_j^1(x_0)x^j, \dots, x^n - \Sigma_{j=1}^m b_j^{n-m}(x_0)x^j).$$

A direct computation gives that

$$\Psi_*(X_j)(\tilde{x}) = \frac{\partial}{\partial \tilde{x}^j} + \sum_{\ell=1}^{n-m} (b_j^\ell(\tilde{x}) - b_j^\ell(x_0)) \frac{\partial}{\partial \tilde{x}^{m+\ell}},$$

which completes the proof. ■

Let ζ be a vector field of order k at x_0 and let

$$B_\zeta^k = \sum_{j=1}^{n} \sum_{I} \frac{\partial^k \zeta^j}{\partial x^I}(x_0)\, \mathrm{d}x^I(x_0) \otimes \frac{\partial}{\partial x^j}(x_0) \tag{3.1}$$

denote the associated symmetric k-multilinear map on $\mathsf{T}_{x_0}\mathsf{M}$, in the coordinates $(x^1, \ldots, x^n)$ about x_0, where $n = \dim(\mathsf{M})$. The next proposition states that, in the regular case, if ζ belongs to a distribution D then B_ζ^k will take its values in D_{x_0}.

Proposition 3.4 *Let D be a smooth distribution, on the n-dimensional manifold M, that is regular at x_0. If ζ is a D-vector field that is of order k at x_0, then* $\mathrm{img}(B_\zeta^k) \subset \mathsf{D}_{x_0}$.

Proof: Let m be the rank of D at x_0. By Lemma 3.3, there is a local frame $\{X_1, \ldots, X_m\}$ for D about x_0 of the form

$$X_j = \frac{\partial}{\partial x^j} + \sum_{\ell=1}^{n-m} c_j^\ell \frac{\partial}{\partial x^{m+\ell}},$$

for smooth functions c_j^ℓ vanishing at x_0, and, thus,

$$\mathsf{D}_{x_0} = \mathrm{span}\left\{ \frac{\partial}{\partial x^1}(x_0), \ldots, \frac{\partial}{\partial x^m}(x_0) \right\}. \tag{3.2}$$

There exist smooth functions $u^1, \ldots, u^m$ such that, locally, $\zeta = u^j X_j$. Now since ζ is of order k at x_0, the partial derivatives of the functions u^j of order less than k vanish at x_0. That is, $\frac{\partial^{|I|} u^j}{\partial x^I}(x_0) = 0$ for multi-indices $0 \leq |I| < k$. Therefore,

$$\frac{\partial^{|I|}(c_j^\ell u^j)}{\partial x^I}(x_0) = 0$$

for all multi-indices I such that $0 \leq |I| \leq k$ because $c^\ell_j(x_0) = 0$. It therefore follows from (3.1) that

$$B^k_\zeta = \sum_{j=1}^m \sum_I \frac{\partial^{|I|} u^j}{\partial x^I}(x_0)\,\mathrm{d}x^I(x_0) \otimes \frac{\partial}{\partial x^j}(x_0),$$

and, therefore, by (3.2), $\operatorname{img}(B^k_\zeta) \subset \mathsf{D}_{x_0}$. ■

The following example shows that the conclusion of Proposition 3.4 is no longer true in the singular case.

Example 3.5 Let $\mathsf{M} = \mathbb{R}^2$, write a typical point $x \in \mathbb{R}^2$ as $x = (x^1, x^2)$, and let $x_0 = (0,0) \in \mathbb{R}^2$ denote the origin. Consider the distribution D given by

$$\mathsf{D}_x = \begin{cases} \operatorname{span}\left\{\frac{\partial}{\partial x^1}(x)\right\}, & x^1 = 0, \\ \operatorname{span}\left\{\frac{\partial}{\partial x^1}(x), \frac{\partial}{\partial x^2}(x)\right\}, & x^1 \neq 0. \end{cases}$$

The vector fields $X_1 = \frac{\partial}{\partial x^1}$ and $X_2 = x^1 \frac{\partial}{\partial x^2}$ form a global frame for D and so D is smooth. The vector X_2 is of order $k = 1$ at x_0. In the canonical coordinates on $\mathbb{R}^2$ we have

$$B^1_{X_2} = \mathrm{d}x^1(x_0) \otimes \frac{\partial}{\partial x^2}(x_0),$$

which does not take its values in $\mathsf{D}_{x_0} = \operatorname{span}\left\{\frac{\partial}{\partial x^1}(x_0)\right\}$.

3.2 Affine systems

In this section we introduce affine systems and their trajectories. To begin, we will say that $\mathcal{F} : \mathsf{M} \rightrightarrows \mathsf{TM}$ is a *multi-valued vector field* if $\mathcal{F}(x) \subset \mathsf{T}_x\mathsf{M}$ for each $x \in \mathsf{M}$.

Definition 3.6 Let A be an affine distribution. An *affine system* in A is a multi-valued vector field $\mathcal{A} : \mathsf{M} \rightrightarrows \mathsf{TM}$ such that $\operatorname{aff}(\mathcal{A}(x)) = \mathsf{A}_x$ for each $x \in \mathsf{M}$.

If $\mathcal{A}$ is an affine system in A, then, necessarily, $\mathcal{A}(x) \neq \emptyset$ and $\mathcal{A}(x) \subset \mathsf{A}_x$ for each $x \in \mathsf{M}$. Henceforth, when the affine distribution A is understood or it is not important in what is to follow, we will simply refer to $\mathcal{A}$ as an affine system, without mentioning A. The restriction of $\mathcal{A}$ to an open set $\Omega \subset \mathsf{M}$ will be denoted by $\mathcal{A}|_\Omega$. Given an open set $\Omega \subset \mathsf{M}$ and a multi-valued vector field $\mathcal{A} : \Omega \rightrightarrows \mathsf{TM}$ such that $\operatorname{aff}(\mathcal{A}(x)) = \mathsf{A}_x$ for each $x \in \Omega$, we will call $\mathcal{A}$ a *local affine system* in A or an affine system in A on Ω. Given two affine systems $\mathcal{A}_1$ and $\mathcal{A}_2$ in A with $\operatorname{dom}(\mathcal{A}_1) \cap \operatorname{dom}(\mathcal{A}_2) \neq \emptyset$, we will write that $\mathcal{A}_1 \subset \mathcal{A}_2$ if $\mathcal{A}_1(x) \subset \mathcal{A}_2(x)$ for all $x \in \operatorname{dom}(\mathcal{A}_1) \cap \operatorname{dom}(\mathcal{A}_2)$.

By an *$\mathcal{A}$-vector field* on an open set $\Omega \subset \mathsf{M}$ we will mean a vector field $\xi \colon \Omega \to \mathsf{TM}$ such that $\xi(x) \in \mathcal{A}(x)$ for each $x \in \Omega$. We say that $\mathcal{A}$ is *smooth at x* if, for each $v_x \in \mathcal{A}(x)$, there exists a smooth $\mathcal{A}$-vector field ξ such that $\xi(x) = v_x$, i.e., $\mathcal{A}$ is smoothly selectionable at x. The set of all smooth $\mathcal{A}$-vector fields will be denoted by $\Gamma(\mathcal{A})$ and $\Gamma_x(\mathcal{A})$ will denote the set of smooth $\mathcal{A}$-vector fields containing x in their domain. Finally, we say that $\mathcal{A}$ is *proper at x_0* if

$$0_{x_0} \in \operatorname{int}_{\mathsf{A}_{x_0}}(\operatorname{co}(\mathcal{A}(x_0))).$$

Definition 3.7 Let $\mathcal{A} : \mathsf{M} \rightrightarrows \mathsf{TM}$ be an affine system. An $\mathcal{A}$-*trajectory* is a locally absolutely continuous curve $\gamma \colon I \to \mathsf{M}$ such that $\gamma'(t) \in \mathcal{A}(\gamma(t))$ a.e., where $I \subset \mathbb{R}$ is an interval.

There are various sufficient conditions for an affine system, or more generally a differential inclusion, to possess trajectories under the above definition of a trajectory [4]. The conditions are of two types, namely, continuity conditions and geometric or topological conditions (convexity, compactness). For example, if $\mathcal{A}$ is a smooth affine system, then it trivially has (smooth) trajectories through each point $x \in \mathsf{M}$ with any given initial velocity vector $v_x \in \mathcal{A}(x)$. To see this, let $\xi \colon \Omega \to \mathsf{TM}$ be a $\mathcal{A}$-vector field with $\xi(x) = v_x$. Then the differential equation $\gamma'(t) = \xi(\gamma(t))$ has a (unique) smooth solution through x with $\gamma'(0) = v_x$. More generally, we have the following.

Theorem 3.8 ([4]) *Let $\Omega \subset \mathbb{R}^n$ be an open subset containing x_0 and let $\mathcal{F} : \Omega \rightrightarrows \mathbb{R}^n$ be continuous with non-empty compact images. Then there exists $T > 0$ and an absolutely continuous curve $\gamma\colon [0,T] \to \Omega$ such that $\gamma'(t) \in \mathcal{F}(\gamma(t))$ and $\gamma(0) = x_0$.*

This theorem can be applied locally to multi-valued vector fields. We include the details for completeness.

Theorem 3.9 *Let $\mathcal{F} : \mathsf{M} \rightrightarrows \mathsf{TM}$ be a continuous multi-valued vector field with non-empty compact images. Then, for any $x_0 \in \mathsf{M}$, there exists $T > 0$ and an absolutely continuous curve $\gamma\colon [0,T] \to \mathsf{M}$ such that $\gamma'(t) \in \mathcal{F}(\gamma(t))$ and $\gamma(0) = x_0$.*

Proof: Let (Ω, φ) be a coordinate chart for x_0 and set $n = \dim(\mathsf{M})$. Let $f\colon \varphi(\Omega) \to \mathbb{R}^n$ be the map defined by $f(y) = T_{\varphi^{-1}(y)}\varphi(\mathcal{F}(\varphi^{-1}(y)))$. This map is well-defined since $\mathcal{F}(x) \subset \mathsf{T}_x\mathsf{M}$ for each $x \in \mathsf{M}$ and it is continuous because it is a composition of continuous maps. Moreover, its images are non-empty compact subsets of $\mathbb{R}^n$. By Theorem 3.8, the differential inclusion $y'(t) \in f(y(t))$ with initial condition $y(0) = \varphi(x_0)$ has a solution $y\colon [0,T] \to \varphi(\Omega)$ for some $T > 0$. It follows then that $t \mapsto \gamma(t) = \varphi^{-1}(y(t))$ is such that $\gamma'(t) \in \mathcal{F}(\gamma(t))$ and $\gamma(0) = x_0$, and γ is absolutely continuous. ■

Corollary 3.10 *A continuous and compact affine system contains trajectories through any point in its domain.*

In this thesis, we will not be concerned with existence issues of solutions of affine systems since we will be focusing on smooth affine systems.

3.3 The Lie algebra rank condition

Given a set $\mathcal{X}$ of smooth vector fields, we let $\mathcal{X}(x) = \{\, X(x) \mid \ X \in \mathcal{X}\}$, and denote by $\mathrm{Lie}(\mathcal{X})$ the smallest Lie subalgebra of vector fields that contains $\mathcal{X}$. If $\mathcal{X}$ consists of vector fields that are defined locally about a common point x_0, then $\mathcal{X}$ generates a set of germs at x_0 of vector fields, which by abuse of notation we denote by the same symbol $\mathcal{X}$. In this case, $\mathrm{Lie}(\mathcal{X})$ will denote the smallest Lie subalgebra of germs at x_0 of vector fields generated by the set of germs $\mathcal{X}$.

Definition 3.11 A set $\mathcal{X}$ of smooth vector fields is said to satisfy the *Lie algebra rank condition* (LARC) at x_0 if $\mathrm{Lie}(\mathcal{X})(x_0) = \mathsf{T}_{x_0}\mathsf{M}$.

Let A be a smooth affine distribution. If $\mathcal{X} = \{X_0, X_1, \ldots, X_m\}$ is a local frame for A at x_0, then it is clear that $\mathrm{Lie}(\mathcal{X}) \subset \mathrm{Lie}(\Gamma_{x_0}(\mathsf{A}))$, and hence a way to test if the LARC holds for $\Gamma_{x_0}(\mathsf{A})$ at x_0, is to compute $\mathrm{Lie}(\mathcal{X})(x_0)$. However, since it is generally not true that $\mathrm{Lie}(\Gamma_{x_0}(\mathsf{A})) \subset \mathrm{Lie}(\mathcal{X})$, a *bad* choice of a local frame can lead to an inconclusive result. Here is a simple example to demonstrate this.

Example 3.12 As in Example 3.5, let $\mathsf{M} = \mathbb{R}^2$, and let $x_0 = (0,0) \in \mathbb{R}^2$ denote the origin. The distribution D is

$$\mathsf{D}_x = \begin{cases} \mathrm{span}\left\{\frac{\partial}{\partial x^1}(x)\right\}, & x^1 = 0, \\ \mathrm{span}\left\{\frac{\partial}{\partial x^1}(x), \frac{\partial}{\partial x^2}(x)\right\}, & x^1 \neq 0. \end{cases}$$

The vector fields $X_1 = \frac{\partial}{\partial x^1}$ and $X_2 = x^1\frac{\partial}{\partial x^2}$ form a global frame for D and so D is smooth. One computes that $[X_1, X_2] = \frac{\partial}{\partial x^2}$ and therefore $\mathrm{Lie}(\{X_1, X_2\})(x_0) = \mathsf{T}_{x_0}\mathsf{M}$. Now let $\tilde{X}_2 = \varphi\frac{\partial}{\partial x^2}$, where $\varphi\colon \mathbb{R} \to \mathbb{R}$ is a smooth function whose derivatives of all orders (including the zeroth derivative) vanish at $x^1 = 0$ and $\varphi(x^1) > 0$ for $x^1 \neq 0$. Then X_1 and $\tilde{X}_2$ also form a global frame for D, but direct computations show that $\mathrm{Lie}(\{\tilde{X}_1, X_2\})(x_0) = \mathrm{span}\left\{\frac{\partial}{\partial x^1}\right\}$.

Remark 3.13 Example 3.12 can be used to show that, if D is a smooth distribution that is singular at x_0 and $\{X_1, X_2\}$ is a local frame for D about x_0, then, for $\xi \in \mathsf{D}$, there may not exist smooth functions u^1, u^2, locally defined about x_0, such that $\xi = u^1 X_1 + u^2 X_2$. For example, and referring to Example 3.12, if $X_1 = \frac{\partial}{\partial x^1}$ but now $X_2 = (x^1)^2 \frac{\partial}{\partial x^2}$, then $\{X_1, X_2\}$ is still a global frame for D. It is clear that $\xi = x^1 \frac{\partial}{\partial x^2}$ is a D-vector field, but there is no smooth function u, defined in a neighbourhood of x_0, such that $\xi = uX_2$.

The previous example is typical of what can happen in the case that A is singular at x_0. Let us now prove a lemma regarding the regular case.

Lemma 3.14 *Let A be an affine distribution that is regular at x_0. Then for any frame $\mathcal{X} = \{X_0, X_1, \ldots, X_m\}$ of A at x_0, it holds that* $\mathrm{Lie}(\Gamma_{x_0}(\mathsf{A}))(x_0) = \mathrm{Lie}(\mathcal{X})(x_0)$.

Proof: It is clear that we only need to show that $\mathrm{Lie}(\Gamma_{x_0}(\mathsf{A}))(x_0) \subset \mathrm{Lie}(\mathcal{X})(x_0)$. To this end, let $\eta_1, \eta_2 \in \Gamma_{x_0}(\mathsf{A})$. In a neighbourhood Ω of x_0, we can write that $\eta_1 = X_0 + u^a X_a$ and that $\eta_2 = X_0 + v^b X_b$, for smooth functions u^a, v^b on Ω. By the properties of the Lie bracket, we can write that

$$[\eta_1, \eta_2] = f^a [X_0, X_a] + g^b X_b + h^{cd} [X_c, X_d]$$

for some smooth functions f^a, g^b, h^{cd}, where $1 \leq a, b, c, d \leq m$. Let

$$X = f^a(x_0)[X_0, X_a] + g^b(x_0) X_b + h^{cd}(x_0)[X_c, X_d].$$

Then $X \in \mathrm{Lie}(\mathcal{X})$ and $[\eta_1, \eta_2](x_0) = X(x_0)$. This procedure can be repeated for any vector field of the form $[\eta_k, [\eta_{k-1}, [\cdots, [\eta_2, \eta_1]] \cdots]$, for any $\eta_1, \ldots, \eta_k \in \Gamma_{x_0}(\mathsf{A})$ and $k \geq 1$. Using the Jacobi-identity, one can show [35, Proposition 3.8] that any element of $\mathrm{Lie}(\Gamma_{x_0}(\mathsf{A}))$ can be written as a linear combination of Lie brackets of the form $[\eta_k, [\eta_{k-1}, [\cdots, [\eta_2, \eta_1]] \cdots]$, and so the claim follows. ■

The LARC plays an important role in controllability theory. For example, in the analytic case, the LARC at x_0 is a necessary condition for controllability [46]. Moreover, if one assumes the LARC at x_0, the type of trajectories that characterize local controllability take a relatively simple form [17]. With this in mind we give the following definition.

Definition 3.15 We say that an affine system $\mathcal{A}$ satisfies the LARC at x_0 if the family $\Gamma_{x_0}(\mathcal{A})$ satisfies the LARC at x_0.

3.4 Local controllability definitions

In this section we define the notion of local controllability for affine distributions that we will study in this thesis.

Let $\mathcal{A}$ be an affine system and let $T > 0$. The *reachable set of $\mathcal{A}$ from x_0 in time T* is

$$\mathcal{R}_{\mathcal{A}}(x_0, T) = \{\gamma(T) \mid \gamma\colon [0,T] \to \mathsf{M} \text{ is a } \mathcal{A}\text{-trajectory such that } \gamma(0) = x_0\},$$

and the *reachable set of $\mathcal{A}$ from x_0 in time at most T* is

$$\mathcal{R}_{\mathcal{A}}(x_0, \leq T) = \bigcup_{0 \leq t \leq T} \mathcal{R}_{\mathcal{A}}(x_0, t).$$

The *reachable set of $\mathcal{A}$ from x_0* is $\mathcal{R}_{\mathcal{A}}(x_0) = \cup_{t \geq 0} \mathcal{R}_{\mathcal{A}}(x_0, t)$.

An affine system $\mathcal{A}$ is called *small-time locally controllable (STLC)* from x_0 if, for each $T > 0$, it holds that $x_0 \in \operatorname{int}(\mathcal{R}_{\mathcal{A}}(x_0, \leq T))$.

Remark 3.16 Let $\mathcal{F} : \mathsf{M} \rightrightarrows \mathsf{TM}$ be a multi-valued vector field. Then the sets $\mathcal{R}_{\mathcal{F}}(x_0, T)$, $\mathcal{R}_{\mathcal{F}}(x_0, \leq T)$, the definition of the LARC at x_0 (if $\mathcal{F}$ is smoothly selectionable at x_0), and the property of STLC from x_0, can all be defined in the same way for $\mathcal{F}$ as was done for an affine system.

We now give our local controllability definitions for an affine distribution.

Definition 3.17 Let A be an affine distribution on M and let $x_0 \in \mathsf{M}$.

(i) We say that A is *properly small-time locally controllable (PSTLC)* from x_0 if every affine system $\mathcal{A}$ in A that (1) is proper and smooth at x_0 and (2) satisfies the LARC at x_0, is STLC from x_0.

(ii) We say that A is *small-time locally uncontrollable (STLUC)* from x_0 if every affine system $\mathcal{A}$ in A that (1) is upper semi-continuous at x_0 and (2) for which $\mathcal{A}(x_0)$ is compact, is not STLC from x_0.

(iii) We say that A is *conditionally small-time locally controllable (CSTLC)* from x_0 if it is neither PSTLC nor STLUC from x_0.

The following two examples illustrate some of the motivations for studying PSTLC. In the first, we give an example that shows that controllability of the linearization is not invariant under feedback transformations.

Example 3.18 Let $\mathsf{M} = \mathbb{R}^3$, let $x_0 = (0,0,0) \in \mathsf{M}$, and consider the affine distribution

$$\mathsf{A}_x = X_0(x) + \operatorname{span}\{X_1(x), X_2(x)\},$$

where $X_0 = x^2 \frac{\partial}{\partial x^1}$, $X_1 = x^3 \frac{\partial}{\partial x^2}$, and $X_2 = \frac{\partial}{\partial x^3}$. Consider the affine system $\mathcal{A}$ given by

$$\mathcal{A}(x) = \left\{ X_0(x) + u^1 X_1(x) + u^2 X_2(x) \;\middle|\; (u^1, u^2) \in [-2,2]^2 \subset \mathbb{R}^2 \right\}.$$

The linearization of $\mathcal{A}$ satisfies the equations

$$\dot{x}^1 = x^2; \qquad \dot{x}^2 = 0; \qquad \dot{x}^3 = u^2,$$

which is not STLC from x_0 since $\dot{x}^2 = 0$. Consider the transformation

$$(u^1, u^2) \mapsto (v^1, v^2) = (u^1 - 1, u^2).$$

The corresponding new frame for A is given by $Y_0 = x^2 \frac{\partial}{\partial x^1} + x^3 \frac{\partial}{\partial x^2}$, $Y_1 = x^3 \frac{\partial}{\partial x^2}$, and $Y_2 = \frac{\partial}{\partial x^3}$. In the new frame, the affine system $\mathcal{A}$ is given by

$$\mathcal{A}(x) = \left\{ Y_0(x) + v^1 Y_1(x) + v^2 Y_2(x) \;\middle|\; (v^1, v^2) \in [-3, 1] \times [-2, 2] \subset \mathbb{R}^2 \right\}.$$

The linearization of $\mathcal{A}$ in this frame satisfies the equations

$$\dot{x}^1 = x^2; \qquad \dot{x}^2 = x^3; \qquad \dot{x}^3 = v^2.$$

Using the standard Kalman rank test, this system is STLC from x_0.

In the next example, we show how the size of the control set can affect controllability.

Example 3.19 Let $\mathsf{M} = \mathbb{R}^3$, let $x_0 = (0, 0, 0) \in \mathbb{R}^3$, and consider the affine distribution

$$\mathsf{A}_x = X_0(x) + \operatorname{span}\{X_1(x), X_2(x)\},$$

where $X_0 = (x^1)^2 \frac{\partial}{\partial x^3}$, $X_1 = \frac{\partial}{\partial x^1}$, and $X_2 = \frac{\partial}{\partial x^2} + \frac{(x^1)^2}{2} \frac{\partial}{\partial x^3}$. Consider the affine system $\mathcal{A}_1$ given by

$$\mathcal{A}_1(x) = \left\{ X_0(x) + u^1 X_1(x) + u^2 X_2(x) \;\middle|\; (u^1, u^2) \in [-1, 1]^2 \subset \mathbb{R}^2 \right\}.$$

It can be shown that $\mathcal{A}_1$ is proper and smooth at x_0 (Proposition 3.21) and, moreover, $\mathcal{A}_1$ satisfies the LARC at x_0. An $\mathcal{A}_1$-trajectory will satisfy the differential equation

$$\dot{x}^1 = u^1; \qquad \dot{x}^2 = u^2; \qquad \dot{x}^3 = (x^1)^2(1 + \tfrac{1}{2}u^2).$$

It is clear that $\mathcal{A}_1$ is not STLC from x_0 because $\dot{x}^3 \geq 0$ for all $\mathcal{A}_1$-trajectories. Hence, A is not PSTLC from x_0. Roughly speaking, the set $U_1 = [-1, 1] \times [-1, 1]$ is not "big enough" to counteract the effects of X_0. Now let $\epsilon > 0$ and consider the affine system

$$\mathcal{A}_2(x) = \left\{ X_0(x) + u^1 X_1(x) + u^2 X_2(x) \;\middle|\; (u^1, u^2) \in [-1, 1] \times [-2 - \epsilon, 1] \right\}.$$

It is clear that $\mathcal{A}_2(x_0)$ is compact, and it can be shown that $\mathcal{A}_2$ is upper semi-continuous at x_0 (Proposition 3.21). With the set $U_2 = [-1, 1] \times [-2 - \epsilon, 1]$, one suspects that $\mathcal{A}_2$ is STLC from x_0 since now $\dot{x}^3$ can be made positive and negative. The proof that this is indeed true will have to wait until Example 5.8. Thus, $\mathcal{A}_2$ is STLC from x_0. Hence, A is CSTLC from x_0. We note that a direct computation shows that

$$\operatorname{span}\{X_1(x_0), X_2(x_0), [X_1, [X_1, X_2]](x_0)\} = \mathsf{T}_{x_0}\mathsf{M},$$

i.e., the value of the Lie algebra generated by $\mathsf{L}(\mathsf{A})$ is $\mathsf{T}_{x_0}\mathsf{M}$, but one cannot conclude STLC for $\mathcal{A}_2$ from the results of, say, [9] since U_2 is bounded.

Hence, a motivation for the notion of PSTLC is that one would like a property of controllability that does not depend on the size of the control set but only on the local differential geometry of the affine distribution.

As given, the definition of PSTLC is difficult to check. In this regard, the next proposition states that, in the regular case, PSTLC can be determined by considering affine systems that are generated by a finite set of vector fields. To state the proposition, we need some notation. For a finite family of vector fields $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ defined on an open set $\Omega \subset \mathsf{M}$, let $\mathcal{A}_{\boldsymbol{\xi}} : \Omega \rightrightarrows \mathsf{TM}$ be defined as $\mathcal{A}_{\boldsymbol{\xi}}(x) = \{\xi_1(x), \ldots, \xi_p(x)\}$. A set-valued map of the form $\mathcal{A}_{\boldsymbol{\xi}}$ is clearly smoothly selectionable.

Proposition 3.20 *Let A be an affine distribution that is regular at x_0. Then A is PSTLC from x_0 if and only if, for every finite collection $\boldsymbol{\xi}$ of smooth A-vector fields such that $\mathcal{A}_{\boldsymbol{\xi}}$ is proper at x_0 and $\boldsymbol{\xi}$ satisfies the LARC at x_0, $\mathcal{A}_{\boldsymbol{\xi}}$ is STLC from x_0.*

Proof: Assume that, for every finite collection $\boldsymbol{\xi}$ of smooth A-vector fields such that

$\mathcal{A}_{\boldsymbol{\xi}}$ is proper at x_0 and $\boldsymbol{\xi}$ satisfies the LARC at x_0, $\mathcal{A}_{\boldsymbol{\xi}}$ is STLC from x_0, and let us prove that A is PSTLC from x_0. To this end, let $\mathcal{A}$ be an affine system in A that is proper and smooth at x_0 and satisfies the LARC at x_0. By smoothness of $\mathcal{A}$ at x_0, there exists a neighbourhood Ω of x_0, and a finite family of $\mathcal{A}$-vector fields $\boldsymbol{\xi}$ defined on Ω such that $0_{x_0} \in \operatorname{int}_{\mathsf{A}_{x_0}}(\operatorname{co}(\mathcal{A}_{\boldsymbol{\xi}}(x_0)))$. By augmenting a finite number of $\mathcal{A}$-vector fields to the family $\boldsymbol{\xi}$, if necessary, we can assume that $\boldsymbol{\xi}$ satisfies the LARC at x_0. Now since $\operatorname{aff}(\mathcal{A}_{\boldsymbol{\xi}}(x_0)) = \mathsf{A}_{x_0}$, by Lemma 3.2, we can assume by shrinking Ω if necessary, that $\operatorname{aff}(\mathcal{A}_{\boldsymbol{\xi}}(x)) = \mathsf{A}_x$ for all $x \in \Omega$, and, therefore, $\mathcal{A}_{\boldsymbol{\xi}}$ is an affine system in A. By construction, $\mathcal{A}_{\boldsymbol{\xi}} \subset \mathcal{A}$ and, therefore, since $\mathcal{A}_{\boldsymbol{\xi}}$ is STLC from x_0, then so is $\mathcal{A}$. Since $\mathcal{A}$ was arbitrary, this proves that A is PSTLC from x_0.

Now assume that A is PSTLC from x_0 and let $\boldsymbol{\xi}$ be a finite collection of smooth A-vector fields. If $\mathcal{A}_{\boldsymbol{\xi}}$ is proper at x_0 then again, by Lemma 3.2, $\mathcal{A}_{\boldsymbol{\xi}}$ is a local affine system in A. Therefore, if $\mathcal{A}_{\boldsymbol{\xi}}$ satisfies the LARC at x_0, then $\mathcal{A}_{\boldsymbol{\xi}}$ is STLC from x_0 because A is PSTLC from x_0. This completes proof. ■

3.5 Control-affine systems

In this section, we describe an important class of affine systems called control-affine systems. After proving some basic properties of control-affine systems, we will show that, in the regular case, it is enough to consider control-affine systems to study the PSTLC property.

A *control-affine system* is a triple $\Sigma = (\mathsf{M}, \{X_0, X_1, \ldots, X_m\}, U)$, where M is a manifold, $\{X_0, X_1, \ldots, X_m\}$ is a set of vector fields on M, and $\operatorname{aff}(U) = \mathbb{R}^m$. The set U is called the *control set* of Σ, and we say that U (or Σ) is *proper* if $0 \in \operatorname{int}(\operatorname{co}(U))$. One can associate to Σ an affine distribution A_Σ and an affine system $\mathcal{A}_\Sigma$ in A_Σ by

$$(\mathsf{A}_\Sigma)_x = \left\{ X_0(x) + u^a X_a(x) \;|\; u = (u^1, \ldots, u^m) \in \mathbb{R}^m \right\}$$

and

$$\mathcal{A}_\Sigma(x) = \left\{ X_0(x) + u^a X_a(x) \mid \; u = (u^1, \dots, u^m) \in U \right\},$$

respectively. We will say that Σ is convex (compact, STLC from x_0, etc.) if the associated affine system $\mathcal{A}_\Sigma$ is convex (compact, STLC from x_0, etc.). If $\gamma\colon [0,T] \to \mathsf{M}$ is a trajectory of $\mathcal{A}_\Sigma$, i.e., γ is absolutely continuous and $\gamma'(t) \in \mathcal{A}_\Sigma(\gamma(t))$ a.e., then there is an integrable map $u\colon [0,T] \to \mathbb{R}^m$ such that $u(t) = (u^1(t), \dots, u^m(t)) \in U$ and

$$\gamma'(t) = X_0(\gamma(t)) + \sum_{a=1}^{m} u^a(t) X_a(\gamma(t)).$$

The next proposition gives some basic properties of control-affine systems.

Proposition 3.21 *Let $\Sigma = (\mathsf{M}, \{X_0, X_1, \dots, X_m\}, U)$ be a control-affine system. Then the following statements hold:*

(i) $\mathcal{A}_\Sigma$ is smooth and, in particular, lower semi-continuous;

(ii) if U is compact then $\mathcal{A}_\Sigma$ is upper semi-continuous;

(iii) if U is proper and $X_0(x_0) = 0_{x_0}$ then $\mathcal{A}_\Sigma$ is proper at x_0.

Proof: To prove (i), let $x \in \mathsf{M}$ and let $v_x \in \mathcal{A}_\Sigma(x)$. Then there exists $u \in U$ such that $v_x = X_0(x) + u^a X_a(x)$. Let $\xi = X_0 + u^a X_a$. Then ξ is a $\mathcal{A}_\Sigma$-vector field and $\xi(x) = v_x$. Hence, $\mathcal{A}_\Sigma$ is smooth, and by Proposition 2.4, $\mathcal{A}_\Sigma$ is lower semi-continuous.

To prove (ii), fix $x_0 \in \mathsf{M}$. Since upper semi-continuity is a local property, we can work locally and consider a coordinate representation of $\mathcal{A}_\Sigma$ on some neighbourhood of x_0. Hence, we can think of $\mathcal{A}_\Sigma$ as a map $\mathcal{A}_\Sigma : \Omega \rightrightarrows \mathbb{R}^n$, where $\Omega \subset \mathbb{R}^n$ is an open set containing x_0. Let $f\colon \Omega \times \mathbb{R}^m \to \mathbb{R}^n$ be defined by $f(x,u) = X_0(x) + u^a X_a(x)$. Then $\mathcal{A}_\Sigma(x) = f(x, U)$ for all $x \in \Omega$. Let W be an open set containing $f(x_0, U)$. Then, for each $u_0 \in U$, $f(x_0, u_0)$ is contained in W. Since f is continuous, there exists a neighbourhood Ω_0 of x_0 and a neighbourhood U_0 of u_0 such that $f(x,u) \in W$ for all $(x,u) \in \Omega_0 \times U_0$. By compactness of U, there exists a finite number of neighbourhoods $\Omega_{0,1}, \dots, \Omega_{0,N}$ containing x_0 such that, if $x \in \Omega' = \cap_{j=1}^{N} \Omega_{0,j}$, then $f(x,u) \in W$ for all

$u \in U$. In other words, for $x \in \Omega'$, $f(x, U) \subset W$, i.e., $\mathcal{A}_\Sigma(x) \subset W$ for all $x \in \Omega'$. This proves that $\mathcal{A}_\Sigma$ is upper semi-continuous at x_0.

To prove (iii), let $f : \mathsf{M} \times \mathbb{R}^m \rightrightarrows \mathsf{TM}$ be defined by $f(x, u) = X_0(x) + u^a X_a(x)$. Then $\mathrm{co}(\mathcal{A}_\Sigma(x)) = f(x, \mathrm{co}(U))$ for each $x \in \mathsf{M}$. Now, since $X_0(x_0) = 0_{x_0}$, $(\mathsf{A}_\Sigma)_{x_0}$ is a subspace and, because $\dim((\mathsf{A}_\Sigma)_{x_0}) \leq m$, $f(x_0, \cdot)$ is a linear map onto $(\mathsf{A}_\Sigma)_{x_0}$, and is therefore an open mapping. Thus, $0 \in \mathrm{int}(\mathrm{co}(U))$, implies that $0_{x_0} \in \mathrm{int}_{(\mathsf{A}_\Sigma)_{x_0}}(\mathrm{co}(\mathcal{A}_\Sigma(x_0)))$. ■

Because of their relatively simple structure in the setting of affine systems, one would like to use control-affine systems to study the local controllability of affine distributions. To this end we give the following definition.

Definition 3.22 Let A be an affine distribution. We say that the control-affine system $\Sigma = (\Omega, \{X_0, X_1, \dots, X_r\}, U)$ is a *local realization* for A at x_0 if Ω is a neighbourhood of x_0 and $\mathsf{A}_\Sigma = \mathsf{A}|_\Omega$.

The next proposition asserts the possibility of including a control-affine system in an arbitrary convex affine system.

Proposition 3.23 *Suppose that x_0 is a regular point of A and let $m = \dim(\mathsf{L}(\mathsf{A})_{x_0})$. Let $\mathcal{A}$ be a convex affine system in A that is lower semi-continuous at x_0. Then there is a neighbourhood Ω of x_0, a convex set $U \subset \mathbb{R}^m$, and a control-affine system $\Sigma = (\Omega, \{X_0, X_1, \dots, X_m\}, U)$ that is a local realization for A at x_0 and $\mathcal{A}_\Sigma \subset \mathcal{A}$. Moreover, if $\mathcal{A}$ is proper at x_0, then Σ can be chosen so that $\mathcal{A}_\Sigma$ is proper at x_0.*

Proof: Let $\{X_0, X_1, \dots, X_m\}$ be a local frame for A on a coordinate neighbourhood Ω of x_0, and let $f \colon \Omega \times \mathbb{R}^m \to \mathsf{A}|_\Omega$ be defined as $f(x, u) = X_0(x) + u^a X_a(x)$. Then, for $x \in \Omega$, $f(x, \cdot)$ is a bijective affine map. Define $\mathcal{F} : \Omega \rightrightarrows \mathbb{R}^m$ as $\mathcal{F}(x) = \mathrm{pr}_2 \circ f^{-1}(\mathcal{A}(x))$, where pr_2 is the canonical projection onto the second factor. Then $\mathcal{F}$ is convex because $\mathcal{A}$ is convex, and $\mathrm{int}(\mathcal{F}(x)) \neq \emptyset$ for each $x \in \Omega$ since $\mathrm{int}_{\mathsf{A}_x}(\mathcal{A}(x)) \neq \emptyset$ for each x. Let

$u_0 \in \mathrm{int}(\mathcal{F}(x_0))$. By Theorem 2.2, there exists a convex open set $U \subset \mathrm{int}(\mathcal{F}(x_0))$ containing u_0 and a neighbourhood $\Omega_0 \ni x_0$ such that $U \subset \mathrm{int}(\mathcal{F}(x))$ for all $x \in \Omega_0$. The first part of the claim follows by letting $\Sigma = (\Omega_0, \{X_0, X_1, \ldots, X_m\}, U)$.

If $\mathcal{A}$ is proper at x_0, by Lemma 3.1 we can assume that $X_0(x_0) = 0_{x_0}$. The linear independence of $X_1(x_0), \ldots, X_m(x_0)$ and properness of $\mathcal{A}$ at x_0 imply that $0 \in \mathrm{int}(\mathcal{F}(x_0))$. Hence we can choose $u_0 = 0 \in \mathbb{R}^m$, and, therefore, U contains the origin in its interior. Therefore, $\mathcal{A}_\Sigma$ is proper at x_0. ■

The method of proof of Proposition 3.23 can be used to prove the following useful lemma.

Lemma 3.24 *Let* A *be regular at* x_0 *and let* $\mathcal{A}$ *be an affine system in* A *that is lower semi-continuous at* x_0*. If* $v_{x_0} \in \mathrm{int}_{\mathsf{A}_{x_0}}(\mathrm{co}(\mathcal{A}(x_0)))$ *and* ξ *is an* A*-vector field such that* $\xi(x_0) = v_{x_0}$*, then* ξ *is a locally* $\mathrm{co}(\mathcal{A})$*-vector field. In fact, there exists a neighbourhood* Ω *of* x_0 *such that, for* $x \in \Omega$*,*

$$\xi(x) \in \mathrm{int}_{\mathsf{A}_x}(\mathrm{co}(\mathcal{A}(x))).$$

Proof: By Proposition 2.3, lower semi-continuity of $\mathcal{A}$ at x_0 implies that $\mathrm{co}(\mathcal{A})$ is lower semi-continuous at x_0. Using the notation of the proof of Proposition 3.23 applied to the convex affine system $\mathrm{co}(\mathcal{A})$, let $u_0 \in \mathrm{int}(\mathcal{F}(x_0))$ be such that $f(x_0, u_0) = v_{x_0}$. Now $\xi(x) = X_0(x) + u^a(x)X_a(x)$ for some smooth functions $u^1, \ldots, u^m$ on Ω such that $(u^1(x_0), \ldots, u^m(x_0)) = u_0$. By shrinking Ω if necessary, and by lower semi-continuity, Theorem 2.2 implies that there is a neigbourhood U_0 of u_0 such that $U_0 \subset \mathrm{int}(\mathcal{F}(x))$ for all $x \in \Omega$. By continuity of $u^1, \ldots, u^m$, and shrinking Ω if necessary, $(u^1(x), \ldots, u^m(x)) \in U_0$ for all $x \in \Omega$. In other words, $\xi(x) \in \mathrm{int}_{\mathsf{A}_x}(\mathrm{co}(\mathcal{A}(x)))$ for all $x \in \Omega$. In particular, ξ is a $\mathrm{co}(\mathcal{A})$-vector field. ■

We are now ready to state the result we have been eluding to, namely that, in the regular case, we can consider control-affine systems to study the PSTLC property.

Theorem 3.25 *Suppose that x_0 is a regular point for A and that $0_{x_0} \in \mathsf{A}_{x_0}$. Then A is PSTLC from x_0 if and only if every control-affine system that*

(i) satisfies the LARC at x_0,

(ii) is a local realization for A at x_0, and

(iii) has a proper and convex control set,

is STLC from x_0.

Proof: Assume that (i)-(iii) hold and let us prove that A is PSTLC from x_0. By Proposition 3.20, it is enough to show that, for every finite family $\boldsymbol{\xi}$ of A-vector fields such that $\mathcal{A}_{\boldsymbol{\xi}}$ is proper and satisfies the LARC at x_0, $\mathcal{A}_{\boldsymbol{\xi}}$ is STLC from x_0. By Proposition 3.23, there exists a control-affine system Σ that is a local realization for A at x_0, with a control set that is proper and convex, such that $\mathcal{A}_\Sigma \subset \mathrm{co}(\mathcal{A}_{\boldsymbol{\xi}})$ on some neighbourhood of x_0. Moreover, by Lemma 3.14, $\mathcal{A}_\Sigma$ satisfies the LARC at x_0. Hence, $\mathcal{A}_\Sigma$ is STLC from x_0, and, therefore, so is $\mathrm{co}(\mathcal{A}_{\boldsymbol{\xi}})$. By Proposition 3.30 below, $\mathcal{A}_{\boldsymbol{\xi}}$ is also STLC from x_0. This proves that A is PSTLC from x_0. The converse statement is obvious. ■

The following corollary to Theorem 3.25 will be useful.

Corollary 3.26 *Suppose that x_0 is a regular point for A with $m = \dim(\mathsf{A}_{x_0})$ and that $0_{x_0} \in \mathsf{A}_{x_0}$. Let $\{X_0, X_1, \ldots, X_m\}$ be a local frame for A at x_0 that satisfies the LARC at x_0. Then A is PSTLC from x_0 if and only if, for every convex set U that is proper, the control-affine system $\Sigma = (\Omega, \{X_0, X_1, \ldots, X_m\}, U)$ is STLC from x_0, where Ω is a neighbourhood of x_0.*

Proof: Let Σ' be a control-affine system that is a local realization for A at x_0, satisfies the LARC at x_0, and has a proper and convex control set. The affine system $\mathcal{A}_{\Sigma'}$ is clearly convex and, by Proposition 3.21, is also lower semi-continuous. Therefore, by Proposition 3.23, there is a neighbourhood Ω of x_0 and a convex and proper control set $U \subset \mathbb{R}^m$, such that control-affine system $\Sigma = (\Omega, \{X_0, X_1, \dots, X_m\}, U)$ satisfies $\mathcal{A}_\Sigma(x) \subset \mathcal{A}_{\Sigma'}(x)$ for each $x \in \Omega$. Properness of U implies that $\text{aff}(U) = \mathbb{R}^m$ and, therefore by Lemma 3.14, $\mathcal{A}_\Sigma$ satisfies the LARC at x_0. If $\mathcal{A}_\Sigma$ is STLC, then so is $\mathcal{A}_{\Sigma'}$. Since Σ' is arbitrary, by Theorem 3.25, this proves that A is PSTLC from x_0. The converse statement is obvious. ■

To prove Theorem 3.25, a key ingredient that was used was that if $\boldsymbol{\xi}$ is a finite family of vector fields that satisfies the LARC at x_0, then $\mathcal{A}_{\boldsymbol{\xi}}$ is STLC from x_0 if and only if $\text{co}(\mathcal{A}_{\boldsymbol{\xi}})$ is STLC from x_0. This fact can be shown to be a corollary of a proposition in [45, Proposition 2.3]. However, to prove Proposition 3.30, we will use some results from [32] and [6] that relate the reachable set of a multi-valued vector field with the reachable set of its smooth selections. To begin, for an arbitrary family $\boldsymbol{\xi}$ of vector fields on M, $x_0 \in \mathsf{M}$, and $T > 0$, we let

$$\mathcal{R}_{\boldsymbol{\xi}}(x_0, T) = \left\{ \Phi^{\xi_p}_{t_p} \circ \cdots \circ \Phi^{\xi_1}_{t_1}(x_0) \;\middle|\; \Sigma\, t_i = T,\, t_i > 0,\, \xi_i \in \boldsymbol{\xi},\, p \geq 0 \right\},$$

and let $\mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T) = \cup_{0 \leq t \leq T} \mathcal{R}_{\boldsymbol{\xi}}(x_0, t)$. Next, we recall that, if (M, d) is a metric space and A and B are subsets of M, then the Hausdorff distance of A and B is defined by

$$d_H(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(a, b) \right\}.$$

It is well-known [33] that, when d_H is restricted to the non-empty, closed, and bounded subsets of M, we obtain a metric space. Let $\Omega \subset \mathbb{R}^n$ be an open set and let $\mathcal{F} : \Omega \rightrightarrows \mathbb{R}^n$ be a Lipschitzean map with respect to the Hausdorff metric, that is, there exists a constant L such that $d_H(F(x), F(y)) \leq L d(x, y)$ for all $x, y \in \Omega$. Suppose that $\mathcal{F}$ admits Lipschitz selections about any point in Ω, that is, for any $x_0 \in \Omega$ and $y_0 \in \mathcal{F}(x_0)$, there is a Lipschitz function $f \colon \Omega_0 \to \mathbb{R}^n$, where $\Omega_0 \subset \Omega$ is a neighbourhood of x_0, such that $f(x_0) = y_0$ and $f(x) \in \mathcal{F}(x)$ for all $x \in \Omega_0$. Let $\Gamma_{x_0}(\mathcal{F})$ denote the set of all Lipschitz selections of $\mathcal{F}$ containing x_0 in their domain. It is clear that

$$\mathcal{R}_{\Gamma_{x_0}(\mathcal{F})}(x_0, T) \subset \mathcal{R}_{\mathcal{F}}(x_0, T) \subset \mathcal{R}_{\overline{\mathrm{co}}(\mathcal{F})}(x_0, T),$$

where $\overline{\mathrm{co}}(S)$ denotes the closure of the convex hull of S. In fact, we have the following.

Lemma 3.27 ([32]) *Let $\mathcal{F} : \Omega \rightrightarrows \mathbb{R}^n$ be a Lipschitzean map that admits Lipschtiz selections about any point in Ω. Then, for each $T > 0$, $\mathcal{R}_{\Gamma_{x_0}(\mathcal{F})}(x_0, T)$ is dense in $\mathcal{R}_{\overline{\mathrm{co}}(\mathcal{F})}(x_0, T)$. Consequently,*

$$\mathrm{cl}\, \mathcal{R}_{\Gamma_{x_0}(\mathcal{F})}(x_0, \leq T) = \mathrm{cl}\, \mathcal{R}_{\overline{\mathrm{co}}(\mathcal{F})}(x_0, \leq T)$$

Proof: The proof of the first statement is the contents of [32]. To prove the second statement, we have that

$$\mathrm{cl}\, \mathcal{R}_{\Gamma_{x_0}(\mathcal{F})}(x_0, \leq T) = \mathrm{cl} \bigcup_{0 \leq t \leq T} \mathcal{R}_{\Gamma_{x_0}(\mathcal{F})}(x_0, t) \supset \bigcup_{0 \leq t \leq T} \mathrm{cl}\, \mathcal{R}_{\Gamma_{x_0}(\mathcal{F})}(x_0, t)$$

$$= \bigcup_{0 \leq t \leq T} \mathcal{R}_{\overline{\mathrm{co}}(\mathcal{F})}(x_0, t) = \mathcal{R}_{\overline{\mathrm{co}}(\mathcal{F})}(x_0, \leq T).$$

It follows that

$$\mathrm{cl}\, \mathcal{R}_{\overline{\mathrm{co}}(\mathcal{F})}(x_0, \leq T) \subset \mathrm{cl}\, \mathcal{R}_{\Gamma_{x_0}(\mathcal{F})}(x_0, \leq T).$$

The reverse inclusion is obvious. ■

The following result is Theorem 3.1 in [6].

Theorem 3.28 ([6]) *Let $f \colon \mathsf{M} \times \mathbb{R}^m \to \mathsf{TM}$ be a smooth map such that, for each $u \in \mathbb{R}^m$, $x \mapsto f(x, u)$ is a smooth vector field on M. Let $U \subset \mathbb{R}^m$ and let $\Gamma(f_U) = \{ f(\cdot, u) \mid u \in U \}$ and suppose that $\mathrm{Lie}(\Gamma(f_U))(x_0) = \mathsf{T}_{x_0}\mathsf{M}$. Then*

$$\mathrm{int}\, \mathrm{cl}\, \mathcal{R}_{\Gamma(f_U)}(x_0, \leq T) = \mathrm{int}\, \mathcal{R}_{\Gamma(f_U)}(x_0, \leq T).$$

The following is an easy consequence of Theorem 3.28.

Lemma 3.29 *Let* $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ *be a family of smooth vector fields on* M *that satisfies the LARC at* x_0*. Then*

$$\operatorname{int} \operatorname{cl} \mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T) = \operatorname{int} \mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T).$$

Proof: Define $f\colon \mathsf{M} \times \mathbb{R}^m \to \mathsf{TM}$ by $f(x,u) = \sum_{a=1}^{p} u^a \xi_a(x)$. Then $f(\cdot, u)$ is a smooth vector field for each $u \in \mathbb{R}^m$. Let U be the standard basis in $\mathbb{R}^p$. Then $\Gamma(f_U) = \boldsymbol{\xi}$, and, therefore, $\operatorname{Lie}(\Gamma(f_U)) = \operatorname{Lie}(\boldsymbol{\xi})$. We then apply Theorem 3.28 to the map f to conclude the proof. ■

We finally have the following proposition, whose proof is an adaptation of the proof of Proposition 2.2 in [10], the difference being that we are not assuming analyticity, but instead assume the LARC and use the standard proof of accessibility [29].

Proposition 3.30 *Let* $\boldsymbol{\xi}$ *be a finite family of smooth vector fields defined on an open set* $\Omega \subset \mathbb{R}^n$*, let* $\mathcal{A}_{\boldsymbol{\xi}} : \Omega \rightrightarrows \mathbb{R}^n$ *be the associated set-valued map, and let* $x_0 \in \Omega$*. If* $\boldsymbol{\xi}$ *satisfies the LARC on* Ω*, then, for each* $T > 0$ *and* $\epsilon > 0$*,*

$$\operatorname{int} \mathcal{R}_{\operatorname{co}(\mathcal{A}_{\boldsymbol{\xi}})}(x_0, \leq T) \subset \operatorname{int} \mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T + \epsilon).$$

Consequently, $\mathcal{A}_{\boldsymbol{\xi}}$ *is STLC from* x_0 *if and only if* $\operatorname{co}(\mathcal{A}_{\boldsymbol{\xi}})$ *is STLC from* x_0*.*

Proof: We first note that, since $\boldsymbol{\xi}$ is finite and consists of smooth vector fields, $\mathcal{A}_{\boldsymbol{\xi}}$ is Lipschitzean with respect to the Hausdorff metric and admits Lipschitz selections. Let $T > 0$ and let $y \in \operatorname{int} \mathcal{R}_{\operatorname{co}(\mathcal{A}_{\boldsymbol{\xi}})}(x_0, \leq T)$. Let $\epsilon > 0$ and consider $\mathcal{R}_{-\boldsymbol{\xi}}(y, \leq \epsilon)$, where $-\boldsymbol{\xi} = \{-\xi \mid \ \xi \in \boldsymbol{\xi}\}$. Because $n = \dim \operatorname{Lie}(\boldsymbol{\xi})(y) = \dim \operatorname{Lie}(-\boldsymbol{\xi})(y)$, it follows that $\mathcal{R}_{-\boldsymbol{\xi}}(y, \leq \epsilon)$ has non-empty interior [29]. Let $V_y \subset \mathcal{R}_{\operatorname{co}(\mathcal{A}_{\boldsymbol{\xi}})}(x_0, \leq T)$ be a neighbourhood of y. By the well-known accessibility theorem [39, page 156, Theorem 9], there exists a sequence $\xi_1, \ldots, \xi_n$ of elements of $\boldsymbol{\xi}$ and $\boldsymbol{t}^* = (t_1^*, \ldots, t_n^*) \in \mathbb{R}^n_{>0}$, with $t_i^* < \epsilon/n$, such that the map

$$(t_1, \ldots, t_n) \mapsto \varphi(t_1, \ldots, t_n) = \Phi_{t_1}^{-\xi_1} \circ \cdots \circ \Phi_{t_n}^{-\xi_n}(y)$$

is defined for $0 \leq t_i < \epsilon/n$, its image belongs to V_y, and it has rank n at $\boldsymbol{t}^*$. Hence there is a neighbourhood $W_{\boldsymbol{t}^*}$ of $\boldsymbol{t}^*$ such that $\varphi(W_{\boldsymbol{t}^*})$ is an open set and is contained in $V_y \cap \operatorname{int} \mathcal{R}_{-\boldsymbol{\xi}}(y, \leq \epsilon)$. Hence, there is a $z \in \varphi(W_{\boldsymbol{t}^*})$ and a neighbourhood V_z of z such that

$$V_z \subset V_y \subset \mathcal{R}_{\operatorname{co}(\mathcal{A}_{\boldsymbol{\xi}})}(x_0, \leq T) \subset \operatorname{cl} \mathcal{R}_{\operatorname{co}(\mathcal{A}_{\boldsymbol{\xi}})}(x_0, \leq T) = \operatorname{cl} \mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T),$$

where the last equality follows from Lemma 3.27. By Lemma 3.29, it holds that $\operatorname{int} \operatorname{cl} \mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T) = \operatorname{int} \mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T)$, and hence $V_z \subset \mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T)$. Now by definition of z, it holds that $y = \Phi_{t_n}^{\xi_n} \circ \cdots \circ \Phi_{t_1}^{\xi_1}(z)$ for some $0 < t_i < \epsilon/n$. Therefore, the open set $\Phi_{t_n}^{\xi_n} \circ \cdots \circ \Phi_{t_1}^{\xi_1}(V_z)$, which is contained in $\mathcal{R}_{\boldsymbol{\xi}}(x_0, \leq T + \epsilon)$, is a neighbourhood of y. This completes the proof. ■

We end this chapter with a corollary to Proposition 3.30 which states that, in the regular case, there is no loss in generality that affine systems are convex.

Corollary 3.31 *Suppose that x_0 is a regular point for A and that $0_{x_0} \in \mathsf{A}_{x_0}$. Then A is PSTLC from x_0 if and only if every smooth convex affine system in A that is proper and satisfies the LARC at x_0 is STLC from x_0.*

Proof: First assume that every convex affine system in A that is proper, smooth, and satisfies the LARC at x_0, is STLC from x_0, and let us prove that A is PSTLC from x_0. By Proposition 3.20, it is enough to consider affine systems of the form $\mathcal{A}_{\boldsymbol{\xi}}$, where $\boldsymbol{\xi}$ is a finite family of vector fields that satisfies the LARC at x_0 and such that $\mathcal{A}_{\boldsymbol{\xi}}$ is proper at x_0. By Proposition 3.23, there exists a control-affine system Σ that is a local realization for A at x_0 and has a proper and convex control set, and $\mathcal{A}_\Sigma \subset \operatorname{co}(\mathcal{A}_{\boldsymbol{\xi}})$.

By Lemma 3.14, Σ satisfies the LARC at x_0. Moreover, $\mathcal{A}_\Sigma$ is clearly convex. Hence, $\mathcal{A}_\Sigma$ is STLC from x_0 and thus so is $\mathrm{co}(\mathcal{A}_{\boldsymbol{\xi}})$. By Proposition 3.30, $\mathcal{A}_{\boldsymbol{\xi}}$ is also STLC from x_0. This proves that A is PSTLC. The converse is obvious. ■

Chapter 4

Composition of flows and related computational tools

The content of this chapter is a set of computational tools for the study of high-order tangent vectors constructed using compositions of flows of vector fields. We start by defining a type of high-order tangent vector that we call an *end-time variation*. We then proceed to describe how these variations depend on the jets of the vector fields used to construct them. We describe a connection between the coefficients of the Taylor series and labeled rooted trees, in a similar way as Butcher [13] relates the coefficients of the Taylor series of the solution of an ODE to rooted trees. We end the chapter with the relationship between a variation and the continuous Campbell–Baker–Hausdorff formula.

4.1 End-time variations

In this section, for ease of presentation and without loss of generality, all vector fields are assumed to be complete. If $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ is a family of vector fields on M, we define the map $\Phi^{\boldsymbol{\xi}}\colon \mathbb{R}^p \times \mathsf{M} \to \mathsf{M}$ by

$$\Phi^{\boldsymbol{\xi}}(\boldsymbol{t}, x) = \Phi^{\xi_p}_{t_p} \circ \Phi^{\xi_{p-1}}_{t_{p-1}} \circ \ \cdots \ \circ \Phi^{\xi_1}_{t_1}(x).$$

The map $\Phi^{\boldsymbol{\xi}}_{\boldsymbol{t}}$ is defined as $x \mapsto \Phi^{\boldsymbol{\xi}}_{\boldsymbol{t}}(x) = \Phi^{\boldsymbol{\xi}}(\boldsymbol{t}, x)$ and $\Phi^{\boldsymbol{\xi}}_x$ is the map defined as $\boldsymbol{t} \mapsto \Phi^{\boldsymbol{\xi}}_x(\boldsymbol{t}) = \Phi^{\boldsymbol{\xi}}(\boldsymbol{t}, x)$.

For a positive integer p, an *end-time* is a smooth map $\boldsymbol{\tau}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^p_{\geq 0}$ such that $\boldsymbol{\tau}(0) = 0_p$. The set of all such maps is denoted by ET_p. Given a family of vector fields $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ and $\boldsymbol{\tau} \in \mathrm{ET}_p$, $\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau}\colon \mathbb{R} \to \mathsf{M}$ is a curve at x_0 whose image consists of points obtained by following concatenations of the integral curves of $\xi_1, \ldots, \xi_p$. The *order* of the pair $(\boldsymbol{\xi}, \boldsymbol{\tau})$ at x_0, denoted $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau})$, is the smallest integer k such that

$$j^k_0(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau}) \neq 0_{x_0},$$

provided such an integer exists, and we set $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) = \infty$ if no such integer exists. If $k = \mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau})$, we call

$$V_{\boldsymbol{\xi},\boldsymbol{\tau}} := j^k_0(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau})$$

the $(\boldsymbol{\xi}, \boldsymbol{\tau})$-*end-time variation* or just *variation* when $(\boldsymbol{\xi}, \boldsymbol{\tau})$ is understood. Recall that, from the exact sequence (2.3), $V_{\boldsymbol{\xi},\boldsymbol{\tau}}$ can be canonically identified with a tangent vector at x_0, and this is how we will view $V_{\boldsymbol{\xi},\boldsymbol{\tau}}$.

4.2 A linear map describing variations

To better understand how a variation $V_{\boldsymbol{\xi},\boldsymbol{\tau}}$ depends on the jets of $\boldsymbol{\xi}$, we note that, as $\mathbb{R}$-algebra homomorphisms,

$$j^k_0(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau}) = j^k_0\boldsymbol{\tau} \circ j^k_{0_p}\Phi^{\boldsymbol{\xi}}_{x_0},$$

where we think of

$$j_0^k \boldsymbol{\tau} \in \mathrm{Hom}((\mathbb{R}^p)^{*k}; (\mathbb{R})^{*k}) \quad \text{and} \quad j_{0_p}^k \Phi_{x_0}^{\boldsymbol{\xi}} \in \mathrm{Hom}(\mathsf{T}_{x_0}^{*k}\mathsf{M}; (\mathbb{R}^p)^{*k}),$$

and where we recall that $(\mathbb{R}^p)^{*k} = \mathsf{J}_{(0_p,0)}^k(\mathbb{R}^p;\mathbb{R})$. Thinking of jets as Taylor polynomials, it is easy to understand what $j_0^k \boldsymbol{\tau}$ is. On the other hand, it is not easy to see how $j_{0_p}^k \Phi_{x_0}^{\boldsymbol{\xi}}$ depends on the jets of $\boldsymbol{\xi}$. To understand this dependence, we first introduce some multi-index notation. For a family of vector fields $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$, a multi-index $I = (i_1, \ldots, i_p)$, and a smooth function $f \colon \mathsf{M} \to \mathbb{R}$, let $\boldsymbol{\xi}^I f$ be the function defined by $(\boldsymbol{\xi}^I f)(x) = (\xi_1^{i_1} \cdots \xi_p^{i_p} f)(x)$, where we think of vector fields as differential operators. Similarly, for $\boldsymbol{t} = (t_1, \ldots, t_p) \in \mathbb{R}^p$, we set $\boldsymbol{t}^I = t_1^{i_1} \cdots t_p^{i_p}$. With this notation we have the following.

Theorem 4.1 *Let $f \colon \mathsf{M} \to \mathbb{R}$ be a smooth function, let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ be a family of smooth vector fields on M, and let $x_0 \in \mathsf{M}$. The Taylor expansion of the function $f \circ \Phi_{x_0}^{\boldsymbol{\xi}} : \mathbb{R}^p \to \mathbb{R}$ at the origin is*

$$\sum_{|I|=0}^{\infty} (\boldsymbol{\xi}^I f)(x_0) \frac{\boldsymbol{t}^I}{I!}. \tag{4.1}$$

Proof: We must show that, for any multi-index $I = (i_1, \ldots, i_p)$,

$$\frac{\partial^{|I|}}{\partial x^I}(f \circ \Phi_{x_0}^{\boldsymbol{\xi}})(0_p) = (\boldsymbol{\xi}^I f)(x_0).$$

Since the order of differentiation does not matter, we begin by differentiating with respect to t_p. It is clear that

$$\frac{\partial}{\partial t_p}(f \circ \Phi_{x_0}^{\boldsymbol{\xi}})(\boldsymbol{t}) = (\xi_p f)(\Phi_{t_p}^{\xi_p} \circ \cdots \circ \Phi_{t_1}^{\xi_1}(x_0))$$

and, therefore,

$$\frac{\partial^{i_p}}{\partial t_p^{i_p}}(f \circ \Phi_{x_0}^{\boldsymbol{\xi}})(\boldsymbol{t}) = (\xi_p^{i_p} f)(\Phi_{t_p}^{\xi_p} \circ \cdots \circ \Phi_{t_1}^{\xi_1}(x_0)),$$

from which it follows that

$$\frac{\partial^{i_p}}{\partial t_p^{i_p}}(f \circ \Phi_{x_0}^{\boldsymbol{\xi}})(t_1, \cdots, t_{p-1}, 0) = (\xi_p^{i_p} f)(\Phi_{t_{p-1}}^{\xi_{p-1}} \circ \cdots \circ \Phi_{t_1}^{\xi_1}(x_0)).$$

We repeat this procedure with t_{p-1} to obtain

$$\frac{\partial^{i_{p-1}+i_p}}{\partial t_{p-1}^{i_{p-1}} \partial t_p^{i_p}}(f \circ \Phi_{x_0}^{\boldsymbol{\xi}})(t_1, \cdots, t_{p-2}, 0, 0) = (\xi_{p-1}^{i_{p-1}} \xi_p^{i_p} f)(\Phi_{t_{p-2}}^{\xi_{p-2}} \circ \cdots \circ \Phi_{t_1}^{\xi_1}(x_0)).$$

This procedure is done repeatedly to obtain the desired result. ■

We now introduce some notation that will be used throughout the rest of the thesis. Given a smooth mapping $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$, written $\phi(x) = (\phi^1(x), \ldots, \phi^n(x))$, and a smooth vector field ξ on $\mathbb{R}^n$, let $\xi\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ be the smooth mapping whose ith component is $\xi\phi^i$. Let $\mathrm{id}_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n$ denote the identity mapping. For a multi-index $I = (i_1, \ldots, i_p)$, let $\boldsymbol{\xi}^I : \mathbb{R}^n \to \mathbb{R}^n$ denote the mapping $(\boldsymbol{\xi}^I)(x) = (\xi_1^{i_1} \cdots \xi_{i_p}^{i_p} \, \mathrm{id}_{\mathbb{R}^n})(x)$. We this notation we have the following corollary to Theorem 4.1.

Corollary 4.2 *Let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ be a family of smooth vector fields on $\mathbb{R}^n$ and let $x_0 \in \mathbb{R}^n$. The Taylor series of the mapping $\Phi_{x_0}^{\boldsymbol{\xi}} : \mathbb{R}^p \to \mathbb{R}^n$ at the origin is*

$$\sum_{|I|=0}^{\infty} \boldsymbol{\xi}^I(x_0) \frac{\boldsymbol{t}^I}{I!}. \tag{4.2}$$

Using Theorem 4.1, we now want to describe how $j_{0_p}^k \Phi_{x_0}^{\boldsymbol{\xi}}$ depends on the jets of $\boldsymbol{\xi}$. We do this in the following theorem which gives the existence of a linear map on an appropriate jet space of the tangent bundle and whose image, on a suitable subset, determines $j_{0_p}^k \Phi_{x_0}^{\boldsymbol{\xi}}$ for every family $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ of smooth vector fields. To state the theorem, let $\pi_{\mathsf{TM}}^p\colon \oplus_{j=1}^p \mathsf{TM} \to \mathsf{M}$ denote the vector bundle over M whose total space is the p-fold direct sum of TM. By abuse of notation, a family of p-vector fields $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ will be identified with a section of π_{TM}^p.

Theorem 4.3 *Let M be a manifold and let $x_0 \in \mathsf{M}$. For positive integers k and p,* *there exists a unique linear map*

$$\mathscr{T}^k_{x_0} \colon \bigoplus_{\ell=1}^{k} S^\ell(\mathsf{J}^{\ell-1}_{x_0}(\pi^p_{\mathsf{TM}})) \to L(\mathsf{T}^{*k}_{x_0}\mathsf{M}; (\mathbb{R}^p)^{*k})$$

such that, for every family of smooth vector fields $\boldsymbol{\xi} = (\xi_1, \dots, \xi_p)$ *on* M*,*

$$\mathscr{T}^k_{x_0}(\oplus^k_{\ell=1}\delta_\ell(j^{\ell-1}_{x_0}\boldsymbol{\xi})) = j^k_{0_p}\Phi^{\boldsymbol{\xi}}_{x_0}.$$

Moreover, the diagram

$$\begin{array}{ccccccc}
\delta_1(\mathsf{J}^0_{x_0}\pi^p_{\mathsf{TM}}) & \longleftarrow & \bigoplus^2_{\ell=1}\delta_\ell(\mathsf{J}^{\ell-1}_{x_0}\pi^p_{\mathsf{TM}}) & \longleftarrow & \bigoplus^3_{\ell=1}\delta_\ell(\mathsf{J}^{\ell-1}_{x_0}\pi^p_{\mathsf{TM}}) & \longleftarrow & \cdots \\
\downarrow{\scriptstyle \mathscr{T}^1_{x_0}} & & \downarrow{\scriptstyle \mathscr{T}^2_{x_0}} & & \downarrow{\scriptstyle \mathscr{T}^3_{x_0}} & & \\
\mathrm{Hom}(\mathsf{T}^{*1}_{x_0}\mathsf{M}; (\mathbb{R}^p)^{*1}) & \longleftarrow & \mathrm{Hom}(\mathsf{T}^{*2}_{x_0}\mathsf{M}; (\mathbb{R}^p)^{*2}) & \longleftarrow & \mathrm{Hom}(\mathsf{T}^{*3}_{x_0}\mathsf{M}; (\mathbb{R}^p)^{*3}) & \longleftarrow & \cdots
\end{array}$$

commutes, where the horizontal arrows are the canonical projections.

Proof: We first recall that $(\mathbb{R}^p)^{*k}$ can be canonically identified with polynomial functions of order k with zero constant term via Taylor's expansion, and thus we will think of elements in $(\mathbb{R}^p)^{*k}$ in this way.

The proof is by induction on k. Let $k = 1$ and let $\boldsymbol{\xi}$ be a family of p-vector fields. We define $\mathscr{T}^1_{x_0} \colon \mathsf{J}^0_{x_0}\pi^p_{\mathsf{TM}} \to L(\mathsf{T}^{*1}_{x_0}\mathsf{M}; (\mathbb{R}^p)^{*1})$ by asking that $\mathscr{T}^1_{x_0}(\boldsymbol{\xi}(x_0)) \in L(\mathsf{T}^{*1}_{x_0}\mathsf{M}; (\mathbb{R}^p)^{*1})$ be defined as

$$\mathscr{T}^1_{x_0}(\boldsymbol{\xi}(x_0))(j^1_{x_0}f)(\boldsymbol{t}) = j^1_{0_p}(f \circ \Phi^{\boldsymbol{\xi}}_{x_0})(\boldsymbol{t}) = \sum_{|I|=1}^{1}(\boldsymbol{\xi}^I f)(x_0)\frac{\boldsymbol{t}^I}{I!},$$

where the last equality follows from Theorem 4.1. If $\boldsymbol{\eta}$ is another family of p-vector fields, then

$$\begin{aligned}\mathscr{T}^1_{x_0}(\boldsymbol{\xi}(x_0)+\boldsymbol{\eta}(x_0))(j^1_{x_0}f)(\boldsymbol{t}) &= \mathscr{T}^1_{x_0}((\boldsymbol{\xi}+\boldsymbol{\eta})(x_0))(j^1_{x_0}f)(\boldsymbol{t})\\ &= j^1_{0_p}(f\circ\Phi^{\boldsymbol{\xi}+\boldsymbol{\eta}}_{x_0})(\boldsymbol{t})\\ &= \sum_{|I|=1}^{1}((\boldsymbol{\xi}+\boldsymbol{\eta})^I f)(x_0)\frac{\boldsymbol{t}^I}{I!}\\ &= j^1_{0_p}(f\circ\Phi^{\boldsymbol{\xi}}_{x_0})(\boldsymbol{t})+j^1_{0_p}(f\circ\Phi^{\boldsymbol{\eta}}_{x_0})(\boldsymbol{t})\\ &= \mathscr{T}^1_{x_0}(\boldsymbol{\xi}(x_0))(j^1_{x_0}f)(\boldsymbol{t})+\mathscr{T}^1_{x_0}(\boldsymbol{\eta}(x_0))(j^1_{x_0}f)(\boldsymbol{t}).\end{aligned}$$

This proves that $\mathscr{T}^1_{x_0}$ is linear. Moreover, by definition, $\mathscr{T}^1_{x_0}$ is the unique map such that $\mathscr{T}^1_{x_0}(\boldsymbol{\xi}(x_0)) = j^1_{0_p}\Phi^{\boldsymbol{\xi}}_{x_0}$. Hence, the claim holds for $k=1$.

By induction, assume the claim for $k\geq 1$ and let $\boldsymbol{\xi}$ be a family of p-vector fields on M. From Theorem 4.1 and the induction hypothesis,

$$\begin{aligned}(j^{k+1}_{0_p}\Phi^{\boldsymbol{\xi}}_{x_0})(j^{k+1}_{x_0}f)(\boldsymbol{t}) &= \sum_{|I|=1}^{k}(\boldsymbol{\xi}^I f)(x_0)\frac{\boldsymbol{t}^I}{I!}+\sum_{|I|=k+1}(\boldsymbol{\xi}^I f)(x_0)\frac{\boldsymbol{t}^I}{I!}\\ &= (j^k_{0_p}\Phi^{\boldsymbol{\xi}}_{x_0})(j^k_{x_0}f)(\boldsymbol{t})+\sum_{|I|=k+1}(\boldsymbol{\xi}^I f)(x_0)\frac{\boldsymbol{t}^I}{I!}\\ &= \mathscr{T}^k_{x_0}(\oplus^k_{\ell=1}\delta_\ell(j^{\ell-1}_{x_0}\boldsymbol{\xi}))(j^k_{x_0}f)(\boldsymbol{t})+\sum_{|I|=k+1}(\boldsymbol{\xi}^I f)(x_0)\frac{\boldsymbol{t}^I}{I!}\\ &= \mathscr{T}^k_{x_0}(\oplus^k_{\ell=1}\delta_\ell(j^{\ell-1}_{x_0}\boldsymbol{\xi}))(j^k_{x_0}f)(\boldsymbol{t})+\psi^{k+1}_{x_0}(j^{k+1}_{x_0}\boldsymbol{\xi})(j^{k+1}_{x_0}f)(\boldsymbol{t}),\end{aligned}$$

where $\psi^{k+1}_{x_0}\colon \mathsf{J}^{k+1}_{x_0}\pi^p_{\mathsf{TM}}\to L(\mathsf{T}^{*(k+1)}_{x_0}\mathsf{M};(\mathbb{R}^p)^{*(k+1)})$ is defined as

$$\psi^{k+1}_{x_0}(j^{k+1}_{x_0}\boldsymbol{\xi})(j^{k+1}_{x_0}f)(\boldsymbol{t}) = \sum_{|I|=k+1}(\boldsymbol{\xi}^I f)(x_0)\frac{\boldsymbol{t}^I}{I!}.$$

If $\lambda\in\mathbb{R}$, then

$$\begin{aligned}\psi^{k+1}_{x_0}(\lambda j^{k+1}_{x_0}\boldsymbol{\xi})(j^{k+1}_{x_0}f)(\boldsymbol{t}) = \psi^{k+1}_{x_0}(j^{k+1}_{x_0}\lambda\boldsymbol{\xi})(j^{k+1}_{x_0}f)(\boldsymbol{t}) &= \sum_{|I|=k+1}((\lambda\boldsymbol{\xi})^I f)(x_0)\frac{\boldsymbol{t}^I}{I!}\\ = \sum_{|I|=k+1}\lambda^{|I|}(\boldsymbol{\xi}^I f)(x_0)\frac{\boldsymbol{t}^I}{I!} &= \lambda^{k+1}\psi^{k+1}_{x_0}(j^{k+1}_{x_0}\boldsymbol{\xi})(j^{k+1}_{x_0}f)(\boldsymbol{t}).\end{aligned}$$

Hence, $\psi_{x_0}^{k+1}$ is a homogeneous polynomial mapping of degree $k+1$. By Lemma 2.1, there exists a unique linear mapping

$$\Psi_{x_0}^{k+1} \colon S^{k+1}(\mathsf{J}_{x_0}^{k+1}\pi_{\mathsf{TM}}^p) \to L(\mathsf{T}_{x_0}^{*(k+1)}\mathsf{M}; (\mathbb{R}^p)^{*(k+1)})$$

such that

$$\Psi_{x_0}^{k+1}(\delta_{k+1}(j_{x_0}^{k+1}\boldsymbol{\xi})) = \psi_{x_0}^{k+1}(j_{x_0}^{k+1}\boldsymbol{\xi}).$$

Now define the mapping

$$\mathscr{T}_{x_0}^{k+1} \colon \bigoplus_{\ell=1}^{k+1} S^{\ell}(\mathsf{J}_{x_0}^{\ell-1}\pi_{\mathsf{TM}}^p) \to L(\mathsf{T}_{x_0}^{*(k+1)}\mathsf{M}; (\mathbb{R}^p)^{*(k+1)})$$

by asking that, for

$$v_1 \oplus \cdots \oplus v_k \oplus v_{k+1} \in \bigoplus_{\ell=1}^{k+1} S^{\ell}(\mathsf{J}_{x_0}^{\ell-1}\pi_{\mathsf{TM}}^p),$$

we have

$$\mathscr{T}_{x_0}^{k+1}(v_1 \oplus \cdots \oplus v_k \oplus v_{k+1})(j_{x_0}^{k+1}f) = \mathscr{T}_{x_0}^{k}(v_1 \oplus \cdots \oplus v_k)(j_{x_0}^{k}f) + \Psi_{x_0}^{k+1}(v_{k+1})(j_{x_0}^{k+1}f), \quad (4.3)$$

where $\mathscr{T}_{x_0}^{k} \colon \bigoplus_{\ell=1}^{k} S^{\ell}(\mathsf{J}_{x_0}^{\ell-1}\pi_{\mathsf{TM}}^p) \to L(\mathsf{T}_{x_0}^{*k}\mathsf{M}; (\mathbb{R}^p)^{*k})$ is the unique linear mapping whose existence is ensured by the induction hypothesis. Because $\mathscr{T}_{x_0}^{k}$ and $\Psi_{x_0}^{k+1}$ are linear, it follows that $\mathscr{T}_{x_0}^{k+1}$ is linear. Moreover,

$$\begin{aligned}\mathscr{T}_{x_0}^{k+1}(\oplus_{\ell=1}^{k+1}\delta_{\ell}(j_{x_0}^{\ell-1}\boldsymbol{\xi}))(j_{x_0}^{k+1}f) &= \mathscr{T}_{x_0}^{k}(\oplus_{\ell=1}^{k}\delta_{\ell}(j_{x_0}^{\ell-1}\boldsymbol{\xi}))(j_{x_0}^{k}f) + \Psi_{x_0}^{k+1}(\delta_{k+1}(j_{x_0}^{k+1}\boldsymbol{\xi}))(j_{x_0}^{k+1}f)\\ &= (j_{0_p}^{k}\Phi_{x_0}^{\boldsymbol{\xi}})(j_{x_0}^{k}f) + \psi_{x_0}^{k+1}(j_{x_0}^{k+1}\boldsymbol{\xi}) = (j_{0_p}^{k+1}\Phi_{x_0}^{\boldsymbol{\xi}})(j_{x_0}^{k+1}f).\end{aligned}$$

Now since $\mathscr{T}_{x_0}^{k+1}$ is uniquely determined by $\mathscr{T}_{x_0}^{k}$ and $\Psi_{x_0}^{k+1}$, it follows that $\mathscr{T}_{x_0}^{k+1}$ is the unique mapping satisfying the claim of the theorem for $k+1$. This proves the first statement. Commutativity of the diagram follows from (4.3). ■

Since we are only interested in the image of $\mathscr{T}_{x_0}^{k}$ on $\bigoplus_{\ell=1}^{k}\delta_{\ell}(\mathsf{J}_{x_0}^{\ell-1}\pi_{\mathsf{TM}}^p)$, we define the map $\boldsymbol{\mathscr{T}}_{x_0}^{k} \colon \mathsf{J}_{x_0}^{k-1}\pi_{\mathsf{TM}}^p \to L(\mathsf{T}_{x_0}^{*k}\mathsf{M}; (\mathbb{R}^p)^{*k})$ by asking that

$$\boldsymbol{\mathscr{T}}_{x_0}^{k}(j_{x_0}^{k-1}\boldsymbol{\xi}) = \mathscr{T}_{x_0}^{k}(\oplus_{\ell=1}^{k}\delta_{\ell}(j_{x_0}^{\ell-1}\boldsymbol{\xi})).$$

Hence, with this notation, for $\boldsymbol{\xi} = (\xi_1, \dots, \xi_p)$ and $\boldsymbol{\tau} \in \mathrm{ET}_p$,

$$j_0^k(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau}) = j_0^k \boldsymbol{\tau} \circ \mathscr{T}^k_{x_0}(j^{k-1}_{x_0}\boldsymbol{\xi}).$$

For $k \in \mathbb{Z}_{\geq 0}$ and a smooth function f, define the polynomial function $e_k^{\boldsymbol{\xi} f} : \mathbb{R}^p \to \mathbb{R}$ by

$$e_k^{\boldsymbol{\xi} f}(\boldsymbol{t}) = \sum_{|I|=0}^{k} (\boldsymbol{\xi}^I f)(x_0) \frac{\boldsymbol{t}^I}{I!}, \tag{4.4}$$

i.e., by Theorem 4.1 $e_k^{\boldsymbol{\xi} f}(\boldsymbol{t})$ is the Taylor polynomial of $f \circ \Phi^{\boldsymbol{\xi}}_{x_0}$ of order k. It will be important for us to know how the Taylor polynomials (4.4) decompose when we view $\boldsymbol{\xi} = (\xi_1, \dots, \xi_p)$ as being a concatenation of two families of vector fields. The following lemma will be key.

Lemma 4.4 *Let $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ be families of smooth vector fields on M of length p and q, respectively, and let $f : \mathsf{M} \to \mathbb{R}$ be a smooth function that vanishes at x_0. Let $\boldsymbol{\xi} = \boldsymbol{\xi}_1 * \boldsymbol{\xi}_2$. Then, for each positive integer k and $(\boldsymbol{t}_1, \boldsymbol{t}_2) \in \mathbb{R}^p \times \mathbb{R}^q$,*

$$e_k^{(\boldsymbol{\xi}_1 * \boldsymbol{\xi}_2) f}(\boldsymbol{t}_1, \boldsymbol{t}_2) = e_k^{\boldsymbol{\xi}_1 f}(\boldsymbol{t}_1) + e_k^{\boldsymbol{\xi}_2 f}(\boldsymbol{t}_2) + m_k^{\boldsymbol{\xi} f}(\boldsymbol{t}_1, \boldsymbol{t}_2),$$

where

$$m_k^{\boldsymbol{\xi} f}(\boldsymbol{t}_1, \boldsymbol{t}_2) = \sum_{|J|=1}^{k-1} \frac{\boldsymbol{t}_2^J}{J!} e_{k-|J|}^{\boldsymbol{\xi}_1 (h_J)}(\boldsymbol{t}_1) \qquad \text{and} \qquad h_J = \boldsymbol{\xi}_2^J f - \boldsymbol{\xi}_2^J f(x_0).$$

Proof: From (4.4),

$$e_k^{(\boldsymbol{\xi}_1 * \boldsymbol{\xi}_2) f}(\boldsymbol{t}_1, \boldsymbol{t}_2) = e_k^{\boldsymbol{\xi}_1 f}(\boldsymbol{t}_1) + e_k^{\boldsymbol{\xi}_2 f}(\boldsymbol{t}_2) + \sum_{\substack{|I|+|J|=2 \\ |I|,|J| \geq 1}}^{k} (\boldsymbol{\xi}_1^I \boldsymbol{\xi}_2^J f)(x_0) \frac{\boldsymbol{t}_1^I \boldsymbol{t}_2^J}{I! J!}.$$

Now, directly,

$$\sum_{\substack{|I|+|J|=2\\|I|,|J|\geq 1}}^{k}(\boldsymbol{\xi}_1^I\boldsymbol{\xi}_2^J f)(x_0)\frac{\boldsymbol{t}_1^I\boldsymbol{t}_2^J}{I!J!} = \sum_{|J|=1}^{k-1}\sum_{|I|=1}^{k-|J|}(\boldsymbol{\xi}_1^I\boldsymbol{\xi}_2^J f)(x_0)\frac{\boldsymbol{t}_1^I\boldsymbol{t}_2^J}{I!J!}$$

$$= \sum_{|J|=1}^{k-1}\frac{\boldsymbol{t}_2^J}{J!}\sum_{|I|=1}^{k-|J|}\boldsymbol{\xi}_1^I(\boldsymbol{\xi}_2^J f - \boldsymbol{\xi}_2^J f(x_0))(x_0)\frac{\boldsymbol{t}_1^I}{I!}$$

$$= \sum_{|J|=1}^{k-1}\frac{\boldsymbol{t}_2^J}{J!}e_{k-|J|}^{\boldsymbol{\xi}_1(\boldsymbol{\xi}_2^J f - \boldsymbol{\xi}_2^J f(x_0))}(\boldsymbol{t}_1),$$

where the last equality follows because the function $\boldsymbol{\xi}_2^J f - \boldsymbol{\xi}_2^J f(x_0)$ vanishes at x_0. This proves the claim. ■

The following lemma will also be useful.

Lemma 4.5 *Let $\boldsymbol{\xi}$ be a family of smooth vector fields of length p and let $\boldsymbol{\tau} \in \mathrm{ET}_p$. Suppose that $k = \mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq 2$ and let $\boldsymbol{\rho}\colon \mathbb{R} \to \mathbb{R}^q$ be a smooth map such that $\boldsymbol{\rho}(0) = 0_q$. For any smooth function $f\colon \mathsf{M} \to \mathbb{R}$ and any multi-index $J = (j_1, \ldots, j_q)$ with $1 \leq |J| \leq k-1$, the derivatives of the function $s \mapsto \boldsymbol{\rho}^J(s)e_{k-|J|}^{\boldsymbol{\xi}f}(\boldsymbol{\tau}(s))$ of orders $0, 1, \ldots, k$ vanish at $s = 0$, where $\boldsymbol{\rho}^J(s) = (\rho^1(s))^{j_1}\cdots(\rho^q(s))^{j_q}$.*

Proof: Suppose that $1 \leq |J| \leq k-1$. By the Leibniz rule, the derivatives of the function $s \mapsto \boldsymbol{\rho}^J(s)$ of orders $0, 1, \ldots, |J|-1$ all vanish at $s = 0$. By definition of ord_{x_0}, the derivatives of the function $s \mapsto e_{k-|J|}^{\boldsymbol{\xi}f}(\boldsymbol{\tau}_1(s))$ of orders $1, \ldots, k-|J|$ all vanish at $s = 0$. Therefore, by the Leibniz rule, the derivatives of the function $s \mapsto \boldsymbol{\rho}^J(s)e_{k-|J|}^{\boldsymbol{\xi}f}(\boldsymbol{\tau}(s))$ of orders $0, 1, \ldots, k$ all vanish at $s = 0$. ■

4.3 The Taylor series of $\Phi_{x_0}^{\boldsymbol{\xi}}$ and rooted trees

In this section, following the work of Butcher [13], we make a graph-theoretic connection between the coefficients in the Taylor series (4.1) and labeled rooted trees. To state Butcher's formula and our extension of it, we introduce some basic notions from graph theory [15]. If G is a graph, we let $|G|$ denote the number of vertices of G. A *tree* is a connected graph with no cycles. A *rooted tree* is a tree with a distinguished vertex called the *root*, which we denote by r. The set of all rooted trees is denoted by $\mathcal{T}$ and a typical rooted tree will be denoted by T, V, or W. The rooted tree with a single vertex is denoted by τ. If v_1, v_2 are vertices of the rooted tree T, let $v_1 T v_2$ be the unique path from v_1 to v_2. There is a natural ordering imposed on the vertices of a rooted tree: $v_1 \leq v_2$ if $v_2 \in rTv_1$. In other words, $v_1 \leq v_2$ if v_2 is closer to the root than v_1, and it is easy to see that the root r is the greatest vertex. If u and v are adjacent vertices in a rooted tree and $u \leq v$, then we say that v is the *parent* of u and u the *child* of v. A *leaf* of a rooted tree is a vertex with no children. For a finite subset $K \subset \mathbb{N}$ with k elements, let $\mathcal{T}_K$ denote the set of rooted trees with k vertices and labeled with the elements of K such that the root is labeled $\max(K)$ and for each $a \in K \backslash \{\max(K)\}$, the labels of the nodes on the unique path from the root to the node labeled a forms a *decreasing* sequence. If $K = \{1, \ldots, k\}$, then we write $\mathcal{T}_k$ for $\mathcal{T}_K$. If $T \in \mathcal{T}_K$ and $T_1, T_2, \ldots, T_m$ denote the rooted trees obtained by removing the root of T and its incident edges, the labeling of T induces a set partition $K_1, K_2, \ldots, K_m$ of $K \backslash \{\max(K)\}$ satisfying $T_a \in \mathcal{T}_{K_a}$ for $a = 1, \ldots, m$. For this reason, we write $T = [T_1, T_2, \ldots, T_m]$. We now introduce the notion of an elementary differential corresponding to a labeled rooted tree, which is a generalization of the elementary differentials considered by Butcher [13].

Definition 4.6 Let $K = \{a_1, \ldots, a_k\} \subset \mathbb{N}$, let $\boldsymbol{\eta} = \{\eta_{a_1}, \ldots, \eta_{a_k}\}$ be a set of smooth vector fields on $\mathbb{R}^n$, and let $K' \subset K$ be a non-empty subset. The *elementary differential of* $\boldsymbol{\eta}$ *corresponding to* $T \in \mathcal{T}_{K'}$ is the map $\boldsymbol{\eta}_T \colon \mathbb{R}^n \to \mathbb{R}^n$ defined as $\boldsymbol{\eta}_T(x) = \eta_{\max(K')}(x)$, if $|K'| = 1$, and

$$\boldsymbol{\eta}_T(x) = (\mathbf{D}^{(m)} \eta_{\max(K')}(x))(\boldsymbol{\eta}_{T_1}(x), \ldots, \boldsymbol{\eta}_{T_m}(x)),$$

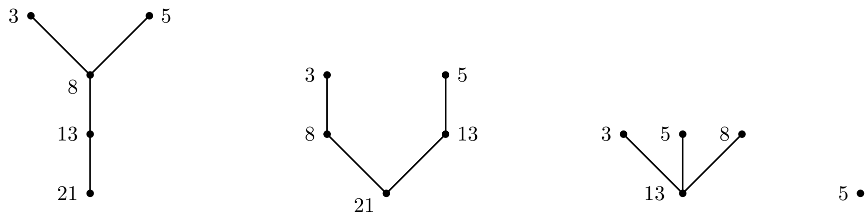

Figure 4.1: Some labeled rooted trees.

if $|K'| \geq 2$, where $T_1, \ldots, T_m$ are the rooted trees obtained by removing the root of T and its incident edges.

The following example illustrates the idea of an elementary differential for labeled rooted trees.

Example 4.7 The elementary differentials corresponding to the labeled trees in Figure 4.1 are (from left to right) $\mathbf{D}\eta_{21}\mathbf{D}\eta_{13}\mathbf{D}^2\eta_8(\eta_3, \eta_5)$, $\mathbf{D}^2\eta_{21}(\mathbf{D}\eta_8(\eta_3), \mathbf{D}\eta_{13}(\eta_5))$, $\mathbf{D}^3\eta_{13}(\mathbf{D}\eta_3, \mathbf{D}\eta_5, \mathbf{D}\eta_8)$, and η_5, respectively. □

Let there be given a multi-index $I = (i_1, \ldots, i_p)$ such that $i_1 + \cdots + i_p = k$ and let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ be a family of vector fields. Define a new family of vector fields $\boldsymbol{\xi}_I$ by

$$\boldsymbol{\xi}_I = (\underbrace{\xi_1, \ldots, \xi_1}_{i_1-\text{times}}, \ldots, \underbrace{\xi_p, \ldots, \xi_p}_{i_p-\text{times}}).$$

For each $T \in \mathfrak{T}_k$, let $[\boldsymbol{\xi}_I]_T \colon \mathbb{R}^n \to \mathbb{R}^n$ be the elementary differential of $\boldsymbol{\xi}_I$ corresponding to T. With this notation we have the following formula.

Theorem 4.8 *Let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ be family of smooth vector fields on $\mathbb{R}^n$ and let $x_0 \in \mathbb{R}^n$. Then, for a multi-index $I = (i_1, \ldots, i_p)$ such that $|I| = k$, it holds that*

$$\frac{\partial^k}{(\partial t_1)^{i_1}(\partial t_2)^{i_2} \cdots (\partial t_p)^{i_p}} \Phi^{\boldsymbol{\xi}}_{x_0}(0_p) = \sum_{T \in \mathfrak{T}_k} [\boldsymbol{\xi}_I]_T(x_0).$$

The case $p = 1$ in Theorem 4.8 is Butcher's formula [13]. Let us recall Butcher's construction. If T is a rooted tree, let $\alpha(T)$ denote the number of ways of labeling T with the set $\{1, 2, \ldots, |T|\}$ such that the root is labeled 1 and for each $2 \leq a \leq |T|$, the labels of the nodes on the unique path from the root to the node labeled a forms an increasing sequence. Let $\xi\colon \mathbb{R}^n \to \mathbb{R}^n$ be a smooth vector field. To each rooted tree T, Butcher associates an *elementary differential* $\xi_T\colon \mathbb{R}^n \to \mathbb{R}^n$ defined as $\xi_T(x) = \xi(x)$ if $|T| = 1$, and

$$\xi_T(x) = (\mathbf{D}^{(m)}\xi)(x)(\xi_{T_1}(x), \ldots, \xi_{T_m}(x))$$

if $|T| \geq 2$, where $T_1, \ldots, T_m$ are the rooted trees obtained by removing the root of T and its incident edges. With this notation we state Butcher's formula, which is an immediate corollary of Theorem 4.8.

Theorem 4.9 ([13]) *If $\xi\colon \mathbb{R}^n \to \mathbb{R}^n$ is $(k-1)$ times differentiable at x_0 and $\boldsymbol{\gamma}'(t) = \xi(\boldsymbol{\gamma}(t))$ with $\boldsymbol{\gamma}(0) = x_0$, then*

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}\boldsymbol{\gamma}(0) = \sum_{\substack{T \text{ a rooted tree} \\ |T|=k}} \alpha(T)\xi_T(x_0).$$

To prove Theorem 4.8 we will first need to prove the following.

Theorem 4.10 *Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)$ be a family of smooth vector fields on $\mathbb{R}^n$. Then, for all $x \in \mathbb{R}^n$,*

$$(\eta_1\eta_2\cdots\eta_k)(x) = \sum_{T \in \mathscr{T}_k} \boldsymbol{\eta}_T(x).$$

Let us prove Theorem 4.8 using Theorem 4.10

Proof of Theorem 4.8: By Corollary 4.2, it follows that, for a multi-index $I = (i_1, \ldots, i_p)$ satisfying $|I| = k$,

$$\frac{\partial^k}{(\partial t_1)^{i_1}(\partial t_2)^{i_2}\cdots(\partial t_p)^{i_p}}\Phi^{\boldsymbol{\xi}}_{x_0}(0_p) = (\xi_1^{i_1}\xi_2^{i_2}\cdots\xi_p^{i_p})(x_0).$$

Define k vector fields $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)$ via the equation

$$\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_k) = (\underbrace{\xi_1, \ldots, \xi_1}_{i_1-\text{times}}, \ldots, \underbrace{\xi_p, \ldots, \xi_p}_{i_p-\text{times}}).$$

Then, for any $T \in \mathcal{T}_k$ and multi-index $I = (i_1, \ldots, t_p)$ with $|I| = k$, by definition $\boldsymbol{\eta}_T = [\boldsymbol{\xi}_I]_T$. Therefore,

$$(\xi_1^{i_1}\xi_2^{i_2}\cdots\xi_p^{i_p})(x_0) = (\eta_1\eta_2\cdots\eta_k)(x_0) = \sum_{T\in\mathcal{T}_k} \boldsymbol{\eta}_T(x_0) = \sum_{T\in\mathcal{T}_k} [\boldsymbol{\xi}_I]_T(x_0),$$

where the second equality follows from Theorem 4.10 and the third equality follows from the definition of $[\boldsymbol{\xi}_I]_T$. This completes the proof. ■

The rest of this section is devoted to proving Theorem 4.10. Let $K = \{a_1, \ldots, a_k\} \subset \mathbb{N}$ and let $K_0 = \{a_0, a_1, \ldots, a_k\} \subset \mathbb{N}$, where $a_0 < \min(K)$. If $V \in \mathcal{T}_{K_0}$, and the vertex labeled with a_0 (which is necessarily a leaf) is removed along with its incident edge, then the resulting labeled rooted tree is an element of $\mathcal{T}_K$. Conversely, every element of $\mathcal{T}_{K_0}$ can be obtained by adding a leaf to a vertex of some element of $\mathcal{T}_K$ and labeling it with a_0. Hence, if $T \in \mathcal{T}_K$ and $\mathscr{D}_T \subset \mathcal{T}_{K_0}$ denotes the subset whose elements are obtained by adding a leaf to a vertex of T and labeling with a_0, we have $\mathcal{T}_{K_0} = \bigcup_{T\in\mathcal{T}_K} \mathscr{D}_T$. The next lemma states that the operation of adding a leaf to each vertex of a rooted tree is equivalent to taking the derivative of the associated elementary differential.

Lemma 4.11 *Let $K = \{a_1, \ldots, a_k\}$, let $K' \subset K$ be a non-empty subset, and let $K_0' = \{a_0\} \cup K'$, where $a_0 < \min(K)$. Let $T \in \mathcal{T}_{K'}$ and let $\boldsymbol{\zeta} = \{\zeta_{a_0}, \zeta_{a_1}, \ldots, \zeta_{a_k}\}$ be a set of smooth vector fields on $\mathbb{R}^n$. Then*

$$\mathbf{D}(\boldsymbol{\zeta}_T)(\zeta_{a_0}) = \sum_{V\in\mathscr{D}_T} \boldsymbol{\zeta}_V, \tag{4.5}$$

where $\mathbf{D}(\boldsymbol{\zeta}_T)(\zeta_{a_0})\colon \mathbb{R}^n \to \mathbb{R}^n$ denotes the smooth mapping $x \mapsto \mathbf{D}(\boldsymbol{\zeta}_T)(x)(\zeta_{a_0}(x))$.

Proof: Let τ_{a_0} denote the rooted tree with one vertex and labeled with a_0. Let $P(\mathcal{T}_{K_0})$ denote the power set of $\mathcal{T}_{K_0}$. The set $\mathscr{D}_T$ is the image of T under the map $\mathrm{d}\colon \mathcal{T}_K \to P(\mathcal{T}_{K_0})$ defined recursively as $\mathrm{d}(T) = \{[\tau_{a_0}]\}$ if $|T| = 1$, and

$$\mathrm{d}(T) = \{[\tau_{a_0}, T_1, T_2, \ldots, T_m]\} \bigcup \left(\bigcup_{j=1}^{m} \bigcup_{W \in \mathrm{d}(T_j)} \{[T_1, \ldots, T_{j-1}, W, T_{j+1}, \ldots, T_m]\} \right), \tag{4.6}$$

if $|T| \geq 2$. It is straightforward to show that (4.5) holds for $T \in \mathcal{T}_{K'}$ whenever $|K| = 1$. Assume by induction that the claim holds for $K' \subset K$ and $T \in \mathcal{T}_{K'}$, whenever $|K| = k - 1$. To prove the induction step, let $K \subset \mathbb{N}$ have k elements, $K' \subset K$ is non-empty and $T \in \mathcal{T}_{K'}$. Write $T = [T_1, \ldots, T_m]$, and let $K_1', \ldots, K_m'$ be the induced partition on $K' \backslash \{\max(K')\}$ such that $T_j \in \mathcal{T}_{K_j'}$. Then each K_j' is a subset of $K \backslash \{\max(K)\}$, which has $k-1$ elements, and so the claim holds for each T_j by the induction hypothesis. Now, by the definition of an elementary differential $\boldsymbol{\zeta}_T$, and letting $r := \max(K') = \max(K_0')$, we have $\boldsymbol{\zeta}_T = \mathbf{D}^{(m)}\zeta_r(\boldsymbol{\zeta}_{T_1}, \ldots, \boldsymbol{\zeta}_{T_m})$. Therefore, from the chain rule, it follows that

$$\begin{aligned}\mathbf{D}(\boldsymbol{\zeta}_T)(\zeta_{a_0}) = {} & \mathbf{D}^{(m+1)}\zeta_r(\boldsymbol{\zeta}_{T_1}, \ldots, \boldsymbol{\zeta}_{T_m}, \zeta_{a_0}) \\ & + \sum_{j=1}^{m} \mathbf{D}^{(m)}\zeta_r(\boldsymbol{\zeta}_{T_1}, \ldots, \boldsymbol{\zeta}_{T_{j-1}}, \mathbf{D}(\boldsymbol{\zeta}_{T_j})(\zeta_{a_0}), \boldsymbol{\zeta}_{T_{j+1}}, \ldots, \boldsymbol{\zeta}_{T_m}).\end{aligned}$$

On the other hand, from the definition of $\mathscr{D}_T = \mathrm{d}(T)$ given in (4.6),

$$\begin{aligned}\sum_{V\in\mathscr{D}_T}\boldsymbol{\zeta}_V &= \mathbf{D}^{(m+1)}\zeta_r(\boldsymbol{\zeta}_{T_1},\ldots,\boldsymbol{\zeta}_{T_m},\zeta_{a_0})\\ &\quad+\sum_{j=1}^{m}\sum_{W\in\mathrm{d}(T_j)}\mathbf{D}^{(m)}\zeta_r(\boldsymbol{\zeta}_{T_1},\ldots,\boldsymbol{\zeta}_{T_{j-1}},\boldsymbol{\zeta}_W,\boldsymbol{\zeta}_{T_{j+1}},\ldots,\boldsymbol{\zeta}_{T_m})\\ &= \mathbf{D}^{(m+1)}\zeta_r(\boldsymbol{\zeta}_{T_1},\ldots,\boldsymbol{\zeta}_{T_m},\zeta_{a_0})\\ &\quad+\sum_{j=1}^{m}\mathbf{D}^{(m)}\zeta_r(\boldsymbol{\zeta}_{T_1},\ldots,\boldsymbol{\zeta}_{T_{j-1}},\sum_{W\in\mathrm{d}(T_j)}\boldsymbol{\zeta}_W,\boldsymbol{\zeta}_{T_{j+1}},\ldots,\boldsymbol{\zeta}_{T_m})\\ &= \mathbf{D}^{(m+1)}\zeta_r(\boldsymbol{\zeta}_{T_1},\ldots,\boldsymbol{\zeta}_{T_m},\zeta_{a_0})\\ &\quad+\sum_{j=1}^{m}\mathbf{D}^{(m)}\zeta_r(\boldsymbol{\zeta}_{T_1},\ldots,\boldsymbol{\zeta}_{T_{j-1}},\mathbf{D}(\boldsymbol{\zeta}_{T_j})(\zeta_{a_0}),\boldsymbol{\zeta}_{T_{j+1}},\ldots,\boldsymbol{\zeta}_{T_m}),\end{aligned}$$

where the second equality follows from multilinearity and the last equality follows from the induction hypothesis. This completes the proof. ■

The following lemma is an easy consequence of our notation and the definition of the Lie derivative.

Lemma 4.12 *For any collection of smooth vector fields $\eta_1,\ldots,\eta_k$ on $\mathbb{R}^n$, it holds that*

$$(\eta_1\eta_2\cdots\eta_k)(x) = \mathbf{D}(\eta_2\cdots\eta_k)(x)(\eta_1(x)),$$

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

We now prove Theorem 4.10.

Proof of Theorem 4.10: The proof is a modification of Lemma 302B in [13]. The proof is by induction on k. The case $k=1$ is trivial. Assume it holds for $k\geq 1$. Let $\zeta_0=\eta_1,\zeta_1=\eta_2,\ldots,\zeta_k=\eta_{k+1}$, $\boldsymbol{\zeta}=\{\zeta_0,\zeta_1,\ldots,\zeta_k\}$, and $K_0=\{0,1,\ldots,k\}$. Then, from Lemma 4.12,

$$\eta_1\eta_2\cdots\eta_{k+1} = \mathbf{D}(\eta_2\cdots\eta_{k+1})(\eta_1) = \mathbf{D}(\zeta_1\cdots\zeta_k)(\zeta_0) = \mathbf{D}\left(\sum_{T\in\mathfrak{T}_k}\boldsymbol{\zeta}_T\right)(\zeta_0)$$
$$= \sum_{T\in\mathfrak{T}_k}\mathbf{D}(\boldsymbol{\zeta}_T)(\zeta_0) = \sum_{V\in\mathfrak{T}_{K_0}}\boldsymbol{\zeta}_V = \sum_{V\in\mathfrak{T}_{k+1}}\boldsymbol{\eta}_V,$$

where the third equality follows from the induction hypothesis and the penultima equality follows from Lemma 4.11. ■

To give an application of some of the ideas in this section, we will give a Lie bracket interpretation of the vector-valued symmetric k-multilinear map B^k_ζ, for a vector field ζ of order k at x_0. Recall that, in a coordinate system $(x^1,\ldots,x^n)$ about x_0, $B^k_\zeta \in S^k(\mathsf{T}^*_{x_0}\mathsf{M})\otimes\mathsf{T}_{x_0}\mathsf{M}$ is given by

$$B^k_\zeta = \sum_{j=1}^n\sum_I \frac{\partial^k\zeta^j}{\partial x^I}(x_0)\,\mathrm{d}x^I(x_0)\otimes\frac{\partial}{\partial x^j}(x_0).$$

Hence, identifying $\mathsf{T}_{x_0}\mathsf{M}$ with $\mathbb{R}^n$ via the basis $\{\frac{\partial}{\partial x^1}(x_0),\ldots,\frac{\partial}{\partial x^n}(x_0)\}$, the action of B^k_ζ on $v_1,\ldots,v_k\in\mathbb{R}^n$ is given by

$$B^k_\zeta(v_1,\ldots,v_k) = \mathbf{D}^k\zeta(x_0)(v_1,\ldots,v_k). \tag{4.7}$$

With (4.7) in mind we have the following.

Proposition 4.13 *Let ζ be a vector field of order k at x_0. Then $B^k_\zeta\in S^k(\mathsf{T}^*_{x_0}\mathsf{M})\otimes\mathsf{T}_{x_0}\mathsf{M}$ is given by*

$$B^k_\zeta(v_1,\ldots,v_k) = [X_1,[X_2,[\cdots,[X_k,\zeta]]\cdots](x_0),$$

where $X_j(x_0)=v_j$, for vector fields $X_1,\ldots,X_k$.

Proof: By a simple induction, one can show that, for vector fields $X_1,\ldots,X_{k+1}$, as differential operators,

$$[X_1, [X_2, [\cdots, [X_k, X_{k+1}]] \cdots] = X_1 X_2 \cdots X_{k+1} + \Psi(X_1, \ldots, X_{k+1}),$$

where $\Psi(X_1, \ldots, X_{k+1})$ is a linear combination of monomials of the form $X_{i_1} X_{i_2} \cdots X_{i_{k+1}}$ with $i_{k+1} \neq k+1$. That is, X_{k+1} does not appear as the right-most factor in any monomial in the sum $\Psi(X_1, \ldots, X_{k+1})$. From Lemma 4.12, the coordinate expression of $X_{i_1} X_{i_2} \cdots X_{i_{k+1}}(x_0)$ is a sum of terms involving the derivatives at x_0, of $X_{i_{k+1}}$ up to order k, the derivatives of X_{i_k} up to order $k-1$, etc., and the zeroth derivative of X_{i_1}. Therefore, if X_{k+1} is of order k at x_0, $\Psi(X_1, \ldots, X_{k+1})(x_0)$ vanishes. Hence, by Lemma 4.12, if X_{k+1} is of order k at x_0, then

$$X_1 X_2 \cdots X_k X_{k+1}(x_0) = \mathbf{D}^k X_{k+1}(x_0)(X_1(x_0), X_2(x_0), \ldots, X_k(x_0)),$$

which proves the claim. ■

4.4 The continuous Campbell–Baker–Hausdorff formula

In this section, we establish a connection between a variation and the formal Campbell–Baker–Hausdorff formula. The algebraic material that follows can all be found in [23].

Let J be a set. The *free* $\mathbb{R}$*-vector space* generated by J will be denoted by $V(J)$. By definition, $V(J)$ is the set of maps $\phi\colon J \to \mathbb{R}$ such that $\phi(j) = 0$ for all but finitely many $j \in J$ and the vector space operations on $V(J)$ are the usual pointwise definitions. A basis for $V(J)$ is the set of maps $e_j\colon J \to \mathbb{R}$ defined by $e_j(j') = 1$ if $j = j'$ and zero otherwise. We let $A(J) = \bigoplus_{k=0}^{\infty} T^k(V(J))$ and $\hat{A}(J) = \prod_{k=0}^{\infty} T^k(V(J))$ denote the *free associative* $\mathbb{R}$*-algebra* and the $\mathbb{R}$*-algebra of non-commutative power* *series* generated by J, respectively. We have the canonical projections

$$\pi_k\colon A(J) \to \bigoplus_{j=0}^{k} T^j(V(J)), \quad \hat{\pi}_k\colon \hat{A}(J) \to \bigoplus_{j=0}^{k} T^j(V(J)).$$

Using the usual commutator definition, $A(J)$ and $\hat{A}(J)$ become Lie algebras. We let $L(J) \subset A(J)$ and $\hat{L}(J) \subset \hat{A}(J)$ denote the Lie subalgebras generated by the basis $(e_j)_{j\in J}$. We say that $b \in L(J)$ is of *degree* k if it belongs to $T^k(V(J))$, and the set of elements in $L(J)$ of degree k is denoted by $L^k(J)$. Let $\hat{A}_0(J) = \prod_{k=1}^{\infty} T^k(V(J))$, which is easily seen to be a subalgebra of $\hat{A}(J)$, and that $\hat{L}(J) \subset \hat{A}_0(J)$. For $a \in \hat{A}_0(J)$, $\exp(a) = \sum_{k=0}^{\infty} \frac{a^k}{k!}$ and $\log(1+a) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{a^k}{k}$ are well-defined elements of $\hat{A}_0(J)$. Moreover, a direct calculation shows that $\exp(\log(1+a)) = 1+a$ and $\log(\exp(a)) = a$. The Campbell–Baker–Hausdorff formula gives a map $\mathrm{CBH}\colon \hat{L}(J)\times\hat{L}(J) \to \hat{L}(J)$ such that $\exp(b_1)\cdot\exp(b_2) = \exp(\mathrm{CBH}(b_1, b_2))$ [23]. The formula for $\mathrm{CBH}(b_1, b_2)$ up to order three is

$$\mathrm{CBH}(b_1, b_2) = b_1 + b_2 + \frac{1}{2}[b_1, b_2] + \frac{1}{12}[b_1, [b_1, b_2]] + \frac{1}{12}[b_2, [b_2, b_1]] + \cdots$$

Given $b_1, \ldots, b_p \in \hat{L}(J)$ and applying the CBH formula recursively, there exists $\mathrm{CBH}(b_1, \ldots, b_p) \in \hat{L}(J)$ such that

$$\exp(b_1)\cdot\exp(b_2)\cdot\;\cdots\;\cdot\exp(b_p) = \exp(\mathrm{CBH}(b_1, \ldots, b_p)).$$

The formula for $\mathrm{CBH}(b_1, \ldots, b_p)$ up to order three is

$$\begin{aligned}\mathrm{CBH}(b_1, \ldots, b_p) = \sum b_i + \frac{1}{2}\sum_{i<j}[b_i, b_j] + \frac{1}{12}\sum_{i<j}([b_i, [b_i, b_j]] + [b_j, [b_j, b_i]]) \\ + \frac{1}{6}\sum_{i<j<k}([b_i, [b_j, b_k]] + [b_k, [b_i, b_j]]) + \cdots .\end{aligned}$$

Now let $\psi\colon J \to \Gamma(\mathsf{TM})$ be a map, and recall that the set of vector fields on a manifold M has the structure of a Lie algebra. By the universal property of the free Lie algebra $L(J)$, there exists a unique Lie algebra homomorphism $\mathrm{Ev}_\psi\colon L(J) \to \Gamma(\mathsf{TM})$ such that $\mathrm{Ev}_\psi(e_j) = \psi(j)$ for each $j \in J$. One cannot in general extend Ev_ψ to $\hat{L}(J)$, and so for example, the expression $\mathrm{Ev}_\psi(\mathrm{CBH}(b_1, \ldots, b_p))$ does not generally make sense. However, it is possible to use the CBH formula to relate the flows of a family of vector fields $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ and the vector field obtained by truncating the formal CBH formula. Explicitly, let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ be a family of vector fields, let $J = \{1, \ldots, p\}$, and define $\psi\colon J \to \Gamma(\mathsf{TM})$ by $\psi(j) = \xi_j$. Let $\{e_1, \ldots, e_p\}$ be the canonical basis of $V(J)$. For $k \in \mathbb{Z}_{>0}$, let

$$\mathrm{CBH}_k(\xi_1, \ldots, \xi_p) = \mathrm{Ev}_\psi(\hat{\pi}_k(\mathrm{CBH}(e_1, \ldots, e_p))),$$

that is, $\mathrm{CBH}_k(\boldsymbol{\xi})$ is the vector field obtained by "plugging in" ξ_j for e_j in the kth-order truncated Lie series $\mathrm{CBH}(e_1, \ldots, e_p)$. With this notation we can state the following result.

Theorem 4.14 ([41]) *Let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ be a family of smooth vector fields. Then*

$$\Phi^{\boldsymbol{\xi}}_{x_0}(t_1, \ldots, t_p) = \Phi_1^{\mathrm{CBH}_k(t_1\xi_1, \ldots, t_p\xi_p)}(x_0) + O((t_1 + \cdots + t_p)^{k+1}),$$

as $(t_1, \ldots, t_p) \to 0_p$ in $\mathbb{R}^p_{>0}$.

Combining Proposition 2.5 and Theorem 4.14, we obtain the following.

Proposition 4.15 *Let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ be a family of smooth vector fields and let $\boldsymbol{\tau} \in \mathrm{ET}_p$, and suppose that $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) = k$. Then*

$$j_0^k(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau}) = \frac{\mathrm{d}^k}{\mathrm{d}s^k}\bigg|_{s=0} \mathrm{CBH}_k(\tau^1(s)\xi_1, \ldots, \tau^p(s)\xi_p)(x_0).$$

Proposition 4.15 and uniqueness of the CBH formula, immediately gives the existence and uniqueness of a "universal" Lie bracket at each order k such that, when evaluated at x_0, gives the variation of every pair $(\boldsymbol{\xi}, \boldsymbol{\tau})$ with $k = \mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau})$, on every manifold M.

Theorem 4.16 *Let p be a positive integer and let $J = \{1, \ldots, p\}$. Let $\psi\colon J \to \Gamma(\mathsf{TM})$ be defined as $\psi(j) = \xi_j$. Then, for each positive integer k, there exists a unique map $\beta_p^k\colon J^k_{(0,0_p)}(\mathbb{R}; \mathbb{R}^p) \to L^k(J)$ such that*

(i) for every manifold M and every $x_0 \in \mathsf{M}$ and

(ii) for every family of smooth vector fields $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ on M and every $\boldsymbol{\tau} \in \mathrm{ET}_p$, with $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) = k$,

it holds that

$$j_0^k(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau}) = \mathrm{Ev}_\psi(\beta_p^k(j_0^k \boldsymbol{\tau}))(x_0).$$

.

Chapter 5

A variational cone for affine systems

In this chapter we describe a class of high-order tangent vectors to the reachable set of an affine system and relate them to the local controllability of the affine system. We then discuss the notion of a neutralizable variation and show that variations of orders $k = 1$ and $k = 2$ are always neutralizable, provided the affine system is proper. We then give a method for construction subspaces of variations for affine systems.

5.1 A variational cone

Fix a smooth affine system $\mathcal{A}$ and $x_0 \in \mathsf{M}$. Let $\Gamma_{x_0}(\mathcal{A}^p)$ denote the set of p-tuples of elements of $\Gamma_{x_0}(\mathcal{A})$ and let $\mathsf{V}_{\mathcal{A}} = \cup_{p\geq 1}\Gamma_{x_0}(\mathcal{A}^p) \times \mathrm{ET}_p$. For a positive integer k, let $\mathsf{V}_{\mathcal{A}}^k$ denote the elements of $\mathsf{V}_{\mathcal{A}}$ of order k at x_0, and let

$$\mathcal{V}_{x_0}^k\mathcal{A} = \left\{ V_{\boldsymbol{\xi},\boldsymbol{\tau}} \mid \ (\boldsymbol{\xi}, \boldsymbol{\tau}) \in \mathsf{V}_{\mathcal{A}}^k \right\} \cup \{0_{x_0}\}$$

and let

$$\mathcal{V}_{x_0}\mathcal{A} = \bigcup_{k\geq 1} \mathcal{V}^k_{x_0}\mathcal{A}.$$

By definition, $\mathcal{V}_{x_0}\mathcal{A}$ is a set of high-order tangent vectors at x_0 to the reachable set of $\mathcal{A}$ from x_0. It is well-known that a curve $\gamma\colon [0,\epsilon] \to \mathsf{M}$ is of order k at 0 if and only if for any smooth function $f\colon \mathsf{M} \to \mathbb{R}$, the derivatives at 0 of the function $f \circ \gamma$ vanish up to order $k-1$, and in this case

$$\frac{\mathrm{d}^k}{\mathrm{d}s^k}(f \circ \gamma)(0) = Vf,$$

where $V = \gamma^{(k)}(0) \in S^k(\mathsf{T}^*_0\mathbb{R}) \otimes \mathsf{T}_{\gamma(0)}\mathsf{M} \cong \mathsf{T}_{\gamma(0)}\mathsf{M}$. Therefore, if $k = \mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau})$, then, for any function $f\colon \mathsf{M} \to \mathbb{R}$, the derivatives of the function $e^{\boldsymbol{\xi} f}_k \circ \boldsymbol{\tau}\colon [0,\epsilon] \to \mathbb{R}$ vanish up to order $k-1$ at $s = 0$, and

$$\frac{\mathrm{d}^k}{\mathrm{d}s^k}(e^{\boldsymbol{\xi} f}_k \circ \boldsymbol{\tau})(0) = V_{\boldsymbol{\xi},\boldsymbol{\tau}} f.$$

With this in mind, let us state and prove the most important property of $\mathcal{V}^k_{x_0}\mathcal{A}$.

Proposition 5.1 *The set $\mathcal{V}^k_{x_0}\mathcal{A}$ is a convex cone.*

Proof: We first prove that $\mathcal{V}^k_{x_0}\mathcal{A}$ is closed under addition. Let $(\boldsymbol{\xi}_1, \boldsymbol{\tau}_1), (\boldsymbol{\xi}_2, \boldsymbol{\tau}_2) \in \mathsf{V}^k_{\mathcal{A}}$, set $\boldsymbol{\xi} = \boldsymbol{\xi}_1 * \boldsymbol{\xi}_2$, and set $\boldsymbol{\tau} = \boldsymbol{\tau}_1 * \boldsymbol{\tau}_2$. We will show that $(\boldsymbol{\xi}, \boldsymbol{\tau}) \in \mathsf{V}^k_{\mathcal{A}}$ and that $V_{\boldsymbol{\xi},\boldsymbol{\tau}} = V_{\boldsymbol{\xi}_1,\boldsymbol{\tau}_1} + V_{\boldsymbol{\xi}_2,\boldsymbol{\tau}_2}$. We can assume that $V_{\boldsymbol{\xi}_1,\boldsymbol{\tau}_1} \neq -V_{\boldsymbol{\xi}_2,\boldsymbol{\tau}_2}$; if not, then $V_{\boldsymbol{\xi}_1,\boldsymbol{\tau}_1} + V_{\boldsymbol{\xi}_2,\boldsymbol{\tau}_2} = 0_{x_0} \in \mathcal{V}^k_{x_0}\mathcal{A}$. Let $f\colon \mathsf{M} \to \mathbb{R}$ be a smooth function that vanishes at x_0. By Lemma 4.4,

$$e^{\boldsymbol{\xi} f}_k(\boldsymbol{\tau}(s)) = e^{\boldsymbol{\xi}_1 f}_k(\boldsymbol{\tau}_1(s)) + e^{\boldsymbol{\xi}_2 f}_k(\boldsymbol{\tau}_2(s)) + m^{\boldsymbol{\xi} f}_k(\boldsymbol{\tau}_1(s), \boldsymbol{\tau}_2(s)),$$

where

$$m^{\boldsymbol{\xi} f}_k(\boldsymbol{\tau}_1(s), \boldsymbol{\tau}_2(s)) = \sum_{|J|=1}^{k-1} \frac{\boldsymbol{\tau}^J_2(s)}{J!} e^{\boldsymbol{\xi}_1 (h_J)}_{k-|J|}(\boldsymbol{\tau}_1(s)),$$

and $h_J = \boldsymbol{\xi}^J_2 f - \boldsymbol{\xi}^J_2 f(x_0)$. By Lemma 4.5, the first k derivatives of the function $s \mapsto m^{\boldsymbol{\xi} f}_k(\boldsymbol{\tau}_1(s), \boldsymbol{\tau}_2(s))$ at $s = 0$ vanish. This proves that $\mathcal{V}^k_{x_0}\mathcal{A}$ is closed under addition.

To prove that $\mathcal{V}^k_{x_0}\mathcal{A}$ is closed under $\mathbb{R}_{>0}$-multiplication, let $(\boldsymbol{\xi},\boldsymbol{\tau}) \in \mathsf{V}^k_{\mathcal{A}}$, let $\alpha \in \mathbb{R}_{>0}$, and define $\boldsymbol{\tau}_\alpha$ by $\boldsymbol{\tau}_\alpha(s) = \boldsymbol{\tau}(\alpha^{1/k}s)$. By the chain-rule, for all $\ell \in \mathbb{Z}_{>0}$,

$$\frac{\mathrm{d}^\ell}{\mathrm{d}s^\ell}(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau}_\alpha)(0) = \alpha^{\ell/k}\frac{\mathrm{d}^\ell}{\mathrm{d}s^\ell}(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau})(0).$$

Therefore, $(\boldsymbol{\xi},\boldsymbol{\tau}_\alpha) \in \mathsf{V}^k_{\mathcal{A}}$ and $V_{\boldsymbol{\xi},\boldsymbol{\tau}_\alpha} = \alpha V_{\boldsymbol{\xi},\boldsymbol{\tau}}$. This completes the proof. ■

Lemma 5.2 ([30]) *For positive integers k and m, $\mathcal{V}^k_{x_0}\mathcal{A} \subseteq \mathcal{V}^{km}_{x_0}\mathcal{A}$.*

Proof: If $(\boldsymbol{\xi},\boldsymbol{\tau}) \in \mathsf{V}^k_{\mathcal{A}}$, then, for any function f vanishing at x_0,

$$e^{\boldsymbol{\xi} f}_k(\boldsymbol{\tau}(s)) = (V_{\boldsymbol{\xi},\boldsymbol{\tau}} f)\frac{s^k}{k!} + o(s^k).$$

Therefore,

$$e^{\boldsymbol{\xi} f}_k(\boldsymbol{\tau}((k!/(km)!)^{1/k}s^m)) = (V_{\boldsymbol{\xi},\boldsymbol{\tau}} f)\frac{s^{km}}{(km)!} + o(s^{km}).$$

It follows that if

$$\boldsymbol{\rho}(s) = \boldsymbol{\tau}((k!/(km)!)^{1/k}s^m),$$

then $(\boldsymbol{\xi},\boldsymbol{\rho}) \in \mathsf{V}^{km}_{\mathcal{A}}$ and $V_{\boldsymbol{\xi},\boldsymbol{\rho}} = V_{\boldsymbol{\xi},\boldsymbol{\tau}}$. ■

Corollary 5.3 *$\mathcal{V}_{x_0}\mathcal{A}$ is a convex cone.*

Proof: The set $\mathcal{V}_{x_0}\mathcal{A}$ is a cone because it is a union of cones. By Lemma 5.2, if $V_1,\dots,V_r \in \mathcal{V}_{x_0}\mathcal{A}$, with $V_j \in \mathcal{V}^{k_j}_{x_0}\mathcal{A}$ and $k = \operatorname{lcm}(k_1,\dots,k_r)$, then $V_1,\dots,V_r \in \mathcal{V}^k_{x_0}\mathcal{A}$. By Proposition 5.1, $\mathcal{V}^k_{x_0}\mathcal{A}$ is convex and, therefore, any convex combination of $V_1,\dots,V_r$ is an element of $\mathcal{V}^k_{x_0}\mathcal{A} \subset \mathcal{V}_{x_0}\mathcal{A}$. This proves that $\mathcal{V}_{x_0}\mathcal{A}$ is convex. This completes the proof. ■

Remark 5.4 Our definition of a variation uses *smooth* functions $\boldsymbol{\tau}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^p_{\geq 0}$, so that in general we do not have $\mathcal{V}^k_{x_0}\mathcal{A} \subseteq \mathcal{V}^{k+1}_{x_0}\mathcal{A}$. If the end-times $\boldsymbol{\tau}$ are allowed to be C^r at $s = 0$ for $r \geq 1$, then a variation of order k can be realized as a variation of order $\ell > k$. However, one then needs to keep track of the order of differentiability of the end-times $\boldsymbol{\tau}$ to be able to work with high-order jets. For this reason we choose to work with smooth end-times, and Lemma 5.2 ensures that essentially nothing is lost by doing so. The use of smooth end-times are employed for instance in [30], whereas [19] uses end-times that are C^r, $r \geq 1$.

The following theorem relates $\mathcal{V}_{x_0}\mathcal{A}$ and STLC of $\mathcal{A}$ at x_0.

Theorem 5.5 *Let $\mathcal{A}$ be a smooth affine system in $\mathsf{A} \subset \mathsf{TM}$ and let $x_0 \in \mathsf{M}$. If $\mathcal{V}_{x_0}\mathcal{A} = \mathsf{T}_{x_0}\mathsf{M}$ then $\mathcal{A}$ is STLC from x_0.*

Proof: Let $T > 0$. By assumption, there exists $V_{\boldsymbol{\xi}_1,\boldsymbol{\tau}_1}, \dots, V_{\boldsymbol{\xi}_r,\boldsymbol{\tau}_r} \in \mathcal{V}_{x_0}\mathcal{A}$ such that

$$0_{x_0} \in \mathrm{int}(\mathrm{co}\{V_{\boldsymbol{\xi}_1,\boldsymbol{\tau}_1}, \dots, V_{\boldsymbol{\xi}_r,\boldsymbol{\tau}_r}\}).$$

By Lemma 5.2, we can assume that $V_{\boldsymbol{\xi}_i,\boldsymbol{\tau}_i} \in \mathcal{V}^k_{x_0}\mathcal{A}$ for some $k \in \mathbb{Z}_{>0}$, for all $i = 1, \dots, r$. Consider the map $\mu \colon \Omega \cap \mathbb{R}^r_{\geq 0} \to \mathsf{M}$ defined by

$$\mu(s_1, \dots, s_r) = \Phi^{\boldsymbol{\xi}_1}_{\boldsymbol{\tau}_1((k!s_1)^{1/k})} \circ \cdots \circ \Phi^{\boldsymbol{\xi}_r}_{\boldsymbol{\tau}_r((k!s_r)^{1/k})}(x_0),$$

where Ω is a neighbourhood of the origin in $\mathbb{R}^r$ such that if $(s_1, \dots, s_r) \in \Omega \cap \mathbb{R}^r_{\geq 0}$, then $\sum_{i,j} \tau_{j,i}((k!s_j)^{1/k})) \leq T$. By construction, μ is continuously differentiable at the origin, $\mu(0) = x_0$, and the image of μ consists of points reachable from x_0 in time at most T. The theorem now follows by Lemma 5.6 below. ■

Lemma 5.6 ([3]) *Let $\mu \colon \mathbb{R}^r \to \mathbb{R}^n$ be Lipschitzean, $\mu(0) = 0$, and differentiable at 0. Assume that $\mathbf{D}\mu(0)(\mathbb{R}^r_{\geq 0}) = \mathbb{R}^n$. Then, for any neighbourhood Ω of the origin in $\mathbb{R}^r$,*

$$0 \in \mathrm{int}\, \mu(\Omega \cap \mathbb{R}^r_{\geq 0}).$$

Let us state an immediate corollary to Theorem 5.5.

Corollary 5.7 *Let A be an affine distribution on M and let $x_0 \in \mathsf{M}$. If, for every smooth affine system $\mathcal{A}$ in A that is proper and satisfies the LARC x_0, it holds that $\mathcal{V}_{x_0}\mathcal{A} = \mathsf{T}_{x_0}\mathsf{M}$, then A is PSTLC from x_0.*

The following example illustrates the usage of Theorem 5.5 and at the same time proves the claim made in Example 3.19.

Example 5.8 As in Example 3.19, let $\mathsf{M} = \mathbb{R}^3$, let $x_0 = (0,0,0)$, and consider the affine distribution $\mathsf{A}_x = X_0(x) + \operatorname{span}\{X_1(x), X_2(x)\}$, where

$$X_0 = (x^1)^2 \frac{\partial}{\partial x^3}, \; X_1 = \frac{\partial}{\partial x^1}, \text{ and } X_2 = \frac{\partial}{\partial x^2} + \frac{(x^1)^2}{2} \frac{\partial}{\partial x^3}.$$

A trajectory of an affine system in A satisfies

$$\dot{x}^1 = u^1; \qquad \dot{x}^2 = u^2; \qquad \dot{x}^3 = x_1^2(1 + \tfrac{1}{2}u^2).$$

Let $U = [-1,-1] \times [-2-\epsilon, 1]$, for $\epsilon > 0$, and let $\Sigma = (\mathsf{M}, \{X_0, X_1, X_2\}, U)$ be the associated control-affine system. It is clear that $\operatorname{span}\left\{\frac{\partial}{\partial x^1}(x_0), \frac{\partial}{\partial x^2}(x_0)\right\} \subset \mathcal{V}^1_{x_0}\mathcal{A}_\Sigma$. Hence, to prove that Σ is STLC from x_0 using Theorem 5.5, we need only construct variations in the $\pm\frac{\partial}{\partial x^3}$ directions. Let $\xi_j = X_0 + u_j^1 X_1 + u_j^2 X_2$, for $j = 1,2,3$, let $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$, and let $\boldsymbol{\tau}(s) = (s,s,s)$. Suppose that $u_1^1 + u_2^1 + u_3^1 = u_1^2 + u_2^2 + u_3^2 = 0$. Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau})(0) = \dot{\tau}^j(0)\xi_j(x_0) &= (u_1^1 + u_2^1 + u_3^1)\frac{\partial}{\partial x^1}(x_0) + (u_1^2 + u_2^2 + u_3^2)\frac{\partial}{\partial x^2}(x_0) \\ &= 0_{x_0},\end{aligned}$$

and, therefore, $\operatorname{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq 2$. Moreover, since $\xi_i\xi_j(x_0)$ vanish at x_0, for any $i,j \in \{1,2,3\}$, by Theorem 4.1 we actually have that $\operatorname{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq 3$. For ease of notation, set $a_j = u_j^1$ and $b_j = u_j^2$, for $j = 1,2,3$. Then a direct computation gives that

$$\frac{\mathrm{d}^3}{\mathrm{d}s^3}(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau})(0) = (a_1^2(10 + 2b_2) + a_1a_2(10 - 2b_1 + b_2) + a_2^2(4 - b_1))\frac{\partial}{\partial x^3}(x_0).$$

If $b_1 = b_2 = 0$, then, for all a_1, a_2, the above produces a variation in the $\frac{\partial}{\partial x^3}$ direction.

Let $b_1 = b_2$ so that

$$\frac{\mathrm{d}^3}{\mathrm{d}s^3}(\Phi^{\boldsymbol{\xi}}_{x_0} \circ \boldsymbol{\tau})(0) = (a_1^2(10 + 2b_2) + a_1 a_2(10 - b_2) + a_2^2(4 - b_2))\frac{\partial}{\partial x^3}.$$

The determinant of the matrix associated with the quadratic form

$$Q(a_1, a_2) = a_1^2(10 + 2b_2) + a_1 a_2(10 - b_2) + a_2^2(4 - b_2)$$

is $\chi(b_2) = 60 + 12b_2 - 9b_2^2 = 3(2 + b_2)(10 - 3b_2)$. Therefore, the quadratic form Q can be made to have a saddle at the origin by choosing $-2 - \epsilon \leq b_2 < -2$. Hence, for such b_2, there are values of a_1, a_2 arbitrarily close to the origin that produce a variation in the $-\frac{\partial}{\partial x^3}$ direction. By Theorem 5.5, $\mathcal{A}_\Sigma$ is STLC from x_0.

5.2 Neutralizable variations

Given $(\boldsymbol{\xi}_1, \boldsymbol{\tau}_1) \in \mathsf{V}^k_{\mathcal{A}}$, under what conditions does there exist a $(\boldsymbol{\xi}_2, \boldsymbol{\tau}_2) \in \mathsf{V}^k_{\mathcal{A}}$ such that $V_{\boldsymbol{\xi}_1,\boldsymbol{\tau}_1} = -V_{\boldsymbol{\xi}_2,\boldsymbol{\tau}_2}$? Motivated by this question we give the following definition.

Definition 5.9 We say that $(\boldsymbol{\xi}_1, \boldsymbol{\tau}_1) \in \mathsf{V}^k_{\mathcal{A}}$ is *neutralizable* if there exists $(\boldsymbol{\xi}_2, \boldsymbol{\tau}_2) \in \mathsf{V}^k_{\mathcal{A}}$ such that $V_{\boldsymbol{\xi}_1,\boldsymbol{\tau}_1} = -V_{\boldsymbol{\xi}_2,\boldsymbol{\tau}_2}$.

The following result is a trivial consequence of Proposition 5.1

Proposition 5.10 *Every element of $\mathsf{V}^k_{\mathcal{A}}$ is neutralizable if and only if $\mathcal{V}^k_{x_0}\mathcal{A}$ is a subspace.*

For orders $k = 1$ and $k = 2$, we will show that neutralizability is ensured by properness of the affine system. In fact, for $k = 1$ we have the following.

Proposition 5.11 *Let $\mathcal{A}$ be an affine system in A that is proper at x_0. Then*

$$\mathcal{V}^1_{x_0}\mathcal{A} = \mathsf{A}_{x_0}.$$

Proof: Given $(\boldsymbol{\xi}, \boldsymbol{\tau}) \in \mathsf{V}_{\mathcal{A}}$, a direct calculation yields that

$$\frac{\mathrm{d}}{\mathrm{d}s}(\Phi^{\boldsymbol{\xi}} \circ \boldsymbol{\tau})(0) = \dot{\tau}^j(0)\xi_j(x_0).$$

Since $\dot{\tau}^j(0) \geq 0$, it follows that $\mathcal{V}^1_{x_0}\mathcal{A} = \mathrm{cone}(\mathrm{co}(\mathcal{A}(x_0)))$, and the result follows by properness of $\mathcal{A}$ at x_0. ■

To treat the $k = 2$ case, we derive the expression for a second order variation. If ξ, η are vector fields, then, as differential operators,

$$\xi\eta = \frac{1}{2}(\xi\eta + \eta\xi) + \frac{1}{2}[\xi, \eta].$$

Therefore, by Theorem 4.1, if $f \colon \mathsf{M} \to \mathbb{R}$ is a smooth function, $\boldsymbol{\xi} = (\xi_1, \dots, \xi_p)$ is a family of vector fields, and $\boldsymbol{\tau} \in \mathrm{ET}_p$, then

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}s^2}(e_2^{\boldsymbol{\xi}f} \circ \boldsymbol{\tau})(0) &= \sum_{j=1}^{p}(\xi_j f)(x_0)\ddot{\tau}^j(0) + \sum_{i \leq j}(\xi_i\xi_j f)(x_0)\dot{\tau}^i(0)\dot{\tau}^j(0) \\ &= \sum_{j=1}^{p}(\xi_j f)(x_0)\ddot{\tau}^j(0) + (\dot{\tau}^1(0)\xi_1 + \dots + \dot{\tau}^p(0)\xi_p)^2(f)(x_0) \\ &\quad + \sum_{i<j}[\xi_i, \xi_j](f)(x_0)\dot{\tau}^i(0)\dot{\tau}^j(0).\end{aligned}$$

Hence, if $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) = 2$, then $(\dot{\tau}^1(0)\xi_1 + \dots + \dot{\tau}^p(0)\xi_p)^2(f)(x_0) = 0$, and, therefore,

$$V_{\boldsymbol{\xi},\boldsymbol{\tau}} = \sum_{j=1}^{p}\xi_j(x_0)\ddot{\tau}^j(0) + \sum_{i<j}[\xi_i, \xi_j](x_0)\dot{\tau}^i(0)\dot{\tau}^j(0). \tag{5.1}$$

Lemma 5.12 *Let $\mathcal{A}$ be an affine system in A that is proper at x_0. If $\xi, \eta \in \Gamma_{x_0}(\mathcal{A})$, then $[\xi, \eta](x_0) \in \mathcal{V}^2_{x_0}\mathcal{A}$.*

Proof: Set $\xi_0 := \xi$ and $\eta_0 := \eta$. By properness of $\mathcal{A}$ at x_0, there are positive constants $\alpha_0, \alpha_1, \dots, \alpha_p$ and $\mathcal{A}$-vector fields $\xi_1, \dots, \xi_p$ such that $\sum_{j=0}^{p}\alpha_j\xi_j(x_0) = 0_{x_0}$. Similarly, there are positive constants $\beta_0, \beta_1, \dots, \beta_q$ and $\mathcal{A}$-vector fields $\eta_1, \dots, \eta_q$ such that

$\sum_{\ell=0}^{q}\beta_\ell\eta_\ell(x_0) = 0_{x_0}$. Let $\boldsymbol{\xi} = (\xi_0, \eta_0, \xi_1, \dots, \xi_p, \eta_1, \dots, \eta_q)$ and let

$$\boldsymbol{\tau}(s) = (\alpha_0 s, \beta_0 s, \alpha_1 s, \dots, \alpha_p s, \beta_1 s, \dots, \beta_q s).$$

Then $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq 2$ and, if $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) = 2$, then, by (5.1) (we will suppress evaluation of the Lie brackets at x_0 to simplify the notation),

$$\begin{aligned} V_{\boldsymbol{\xi},\boldsymbol{\tau}} = \alpha_0\beta_0[\xi_0, \eta_0] + \sum_{0\leq i<j\leq p} \alpha_i\alpha_j[\xi_i, \xi_j] + \sum_{\ell=1}^{q} \alpha_0\beta_\ell[\xi_0, \eta_\ell] + \sum_{j=1}^{p} \beta_0\alpha_j[\eta_0, \xi_j] \\ + \sum_{0\leq k<\ell\leq q} \beta_k\beta_\ell[\eta_k, \eta_\ell] + \sum_{j=1}^{p}\sum_{\ell=1}^{q} \alpha_j\beta_\ell[\xi_j, \eta_\ell]. \end{aligned}$$

Now let $\tilde{\boldsymbol{\xi}} = (\eta_q, \dots, \eta_1, \xi_p, \dots, \xi_1, \xi_0, \eta_0)$ and let

$$\tilde{\boldsymbol{\tau}}(s) = (\beta_q s, \dots, \beta_1 s, \alpha_p s, \dots, \alpha_1 s, \alpha_0 s, \beta_0 s).$$

Then $\mathrm{ord}_{x_0}(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\tau}}) \geq 2$ and, by (5.1),

$$\begin{aligned} V_{\tilde{\boldsymbol{\xi}},\tilde{\boldsymbol{\tau}}} = \sum_{0\leq k<\ell\leq q} \beta_k\beta_\ell[\eta_\ell, \eta_k] + \sum_{\ell=1}^{q}\sum_{j=1}^{p} \beta_\ell\alpha_j[\eta_\ell, \xi_j] + \sum_{\ell=1}^{q} \alpha_0\beta_\ell[\eta_\ell, \xi_0] \\ + \sum_{0\leq i<j\leq p} \alpha_i\alpha_j[\xi_j, \xi_i] + \sum_{j=1}^{p} \alpha_j\beta_0[\xi_j, \eta_0] + \alpha_0\beta_0[\xi_0, \eta_0]. \end{aligned}$$

One computes that $V_{\boldsymbol{\xi},\boldsymbol{\tau}} + V_{\tilde{\boldsymbol{\xi}},\tilde{\boldsymbol{\tau}}} = 2\alpha_0\beta_0[\xi_0, \eta_0](x_0)$, and, therefore, $[\xi_0, \eta_0](x_0) \in \mathcal{V}^2_{x_0}\mathcal{A}$ because $\mathcal{V}^2_{x_0}\mathcal{A}$ is a cone. This completes the proof. ■

For an affine system $\mathcal{A}$ define

$$\mathcal{D}^{(2)}_{\mathcal{A}} = \{\, [\xi, \eta] \mid \ \xi, \eta \in \Gamma(\mathcal{A}) \} \,.$$

With this notation we have the following.

Proposition 5.13 *Let $\mathcal{A}$ be an affine system in A that is proper at x_0. Then*

$$\mathcal{V}^2_{x_0}\mathcal{A} = \mathsf{A}_{x_0} + \mathrm{span}\{\mathcal{D}^{(2)}_{\mathcal{A}}(x_0)\}.$$

In particular, $\mathcal{V}^2_{x_0}\mathcal{A}$ is a subspace.

Proof: By (5.1) it is clear that $\mathcal{V}^2_{x_0}\mathcal{A} \subseteq \mathsf{A}_{x_0} + \mathrm{span}\{\mathcal{D}^{(2)}_{\mathcal{A}}(x_0)\}$. Now let

$$w \in \mathsf{A}_{x_0} + \mathrm{span}\{\mathcal{D}^{(2)}_{\mathcal{A}}(x_0)\},$$

and write

$$w = \sum_{i=1}^{p} v_i + \sum_{j=1}^{r} \alpha_j [\xi_{j,1}, \xi_{j,2}](x_0),$$

for $v_i \in \mathsf{A}_{x_0}$ and $\xi_{j,k} \in \Gamma_{x_0}(\mathcal{A})$. Since $\mathcal{V}^1_{x_0}\mathcal{A} = \mathsf{A}_{x_0}$, it follows that $v_i \in \mathcal{V}^1_{x_0}\mathcal{A} \subset \mathcal{V}^2_{x_0}\mathcal{A}$. By Lemma 5.12, $[\xi_{j,1}, \xi_{j,2}](x_0) \in \mathcal{V}^2_{x_0}\mathcal{A}$, and we can assume without loss of generality that $\alpha_j > 0$. Hence, w is a sum of elements of $\mathcal{V}^2_{x_0}\mathcal{A}$, and thus $w \in \mathcal{V}^2_{x_0}\mathcal{A}$ since $\mathcal{V}^2_{x_0}\mathcal{A}$ is closed under addition. This proves the reverse inclusion. ■

The next step would be to consider the cases $k \geq 3$, but this seems to be a difficult task. Notwithstanding, some results have been obtained by fixing a set of local generators for the affine distribution and identifying "bad Lie brackets" that are potentially non-neutralizable [9, 45]. These results, unfortunately, are not invariant under a change of local generators for the affine distribution.

5.3 Subspaces of variations

In this section, using a technique from Krener [30, Section 4], we construct subspaces of variations. The idea of this section is to obtain linear approximations, i.e., subspaces, to the variational cone $\mathcal{V}_{x_0}\mathcal{A}$.

Let ζ be a vector field on M that vanishes at x_0. Then, by Proposition 4.13, ζ induces a linear map $B_\zeta \colon \mathsf{T}_{x_0}\mathsf{M} \to \mathsf{T}_{x_0}\mathsf{M}$. Explicitly, for $v_{x_0} \in \mathsf{T}_{x_0}\mathsf{M}$, $B_\zeta(v_{x_0}) = [X, \zeta](x_0)$, where X is any vector field extending v_{x_0}.

Let $\mathcal{A}$ be an affine system and set

$$\mathcal{Z}_{x_0}\mathcal{A} = \{\, \zeta \in \Gamma_{x_0}(\mathcal{A}) \mid \ \zeta(x_0) = 0_{x_0} \,\}\,.$$

We identify $\mathcal{Z}_{x_0}\mathcal{A}$ with the corresponding subset of $\mathrm{End}(\mathsf{T}_{x_0}\mathsf{M})$, which we still denote by $\mathcal{Z}_{x_0}\mathcal{A}$. For a subspace $W \subseteq \mathsf{T}_{x_0}\mathsf{M}$, let $\langle \mathcal{Z}_{x_0}\mathcal{A}; W\rangle$ denote the smallest subspace containing W and that is invariant under the linear maps in $\mathcal{Z}_{x_0}\mathcal{A}$. It is not hard to show that

$$\langle \mathcal{Z}_{x_0}\mathcal{A}; W\rangle = \mathrm{span}\left\{B_{\zeta_1}B_{\zeta_2}\ \cdots\ B_{\zeta_r}(w_{x_0}) \mid w_{x_0} \in W,\ \zeta_i \in \mathcal{Z}_{x_0}\mathcal{A},\ r \in \mathbb{Z}_{\geq 0}\right\}. \quad (5.2)$$

Theorem 5.14 *Let $\mathcal{A}$ be a smooth affine system in A and let $x_0 \in M$. For any subspace $W \subseteq \mathcal{V}_{x_0}\mathcal{A}$, it holds that $\langle \mathcal{Z}_{x_0}\mathcal{A}; W\rangle \subseteq \mathcal{V}_{x_0}\mathcal{A}$.*

Proof: To prove the theorem, it is enough to show that, if $w_{x_0} \in W$ and $\zeta \in \mathcal{Z}_{x_0}\mathcal{A}$, then $B_\zeta(w_{x_0}) \in \mathcal{V}_{x_0}\mathcal{A}$.

Let $w_{x_0} \in W$ and let $\zeta \in \mathcal{Z}_{x_0}\mathcal{A}$. By Lemma 5.2, we can assume that there exist an integer $k \geq 1$ and $(\boldsymbol{\xi}_i, \boldsymbol{\tau}_i) \in \mathsf{V}^k_{\mathcal{A}}$ such that $V_{\boldsymbol{\xi}_i,\boldsymbol{\tau}_i} = (-1)^{i+1}w_{x_0}$ for $i = 1, 2$. Let $\tilde{\boldsymbol{\tau}}_i(s) = \boldsymbol{\tau}_i((k!/(2k)!)^{1/k}s^2)$, for $i = 1, 2$. Then, by Lemma 5.2, $\mathrm{ord}_{x_0}(\boldsymbol{\xi}_i, \tilde{\boldsymbol{\tau}}_i) = 2k$ and $V_{\boldsymbol{\xi}_1,\tilde{\boldsymbol{\tau}}_1} = (-1)^{i+1}w_{x_0}$, for $i = 1, 2$. Now, since $\zeta(x_0) = 0_{x_0}$ and $V_{\boldsymbol{\xi}_1,\tilde{\boldsymbol{\tau}}_2} = -V_{\boldsymbol{\xi}_2,\tilde{\boldsymbol{\tau}}_2}$, we have $\mathrm{ord}_{x_0}(\boldsymbol{\xi}_1 * \zeta * \boldsymbol{\xi}_2, \tilde{\boldsymbol{\tau}}_1 * s * \tilde{\boldsymbol{\tau}}_2) \geq 2k + 1$. By definition and then expanding,

$$\begin{aligned} e^{(\boldsymbol{\xi}_1*\zeta*\boldsymbol{\xi}_2)f}_{2k+1}(\tilde{\boldsymbol{\tau}}_1(s), s, \tilde{\boldsymbol{\tau}}_2(s)) &= e^{\boldsymbol{\xi}_1 f}_{2k+1}(\tilde{\boldsymbol{\tau}}_1(s)) + e^{\zeta f}_{2k+1}(s) + e^{\boldsymbol{\xi}_2 f}_{2k+1}(\tilde{\boldsymbol{\tau}}_2(s)) \\ &\quad + \sum_{\substack{|I_1|+j=2\\ |I_1|,j\geq 1}}^{2k+1} (\boldsymbol{\xi}_1^{I_1}\zeta^j f)(x_0)\frac{s^j\tilde{\boldsymbol{\tau}}_1^{I_1}(s)}{j!I_1!} + \sum_{\substack{|I_2|+j=2\\ |I_2|,j\geq 1}}^{2k+1} (\zeta^j\boldsymbol{\xi}_2^{I_2} f)(x_0)\frac{s^j\tilde{\boldsymbol{\tau}}_2^{I_2}(s)}{j!I_2!} \\ + \sum_{\substack{|I_1|+|I_2|=2\\ |I_1|,|I_2|\geq 1}}^{2k+1} (\boldsymbol{\xi}_1^{I_1}\boldsymbol{\xi}_2^{I_2} f)(x_0)\frac{\tilde{\boldsymbol{\tau}}_1^{I_1}(s)\tilde{\boldsymbol{\tau}}_2^{I_2}(s)}{I_1!I_2!} &+ \sum_{\substack{|I_1|+j+|I_2|=3\\ |I_1|,j,|I_2|\geq 1}}^{2k+1} (\boldsymbol{\xi}_1^{I_1}\zeta^j\boldsymbol{\xi}_2^{I_2} f)(x_0)\frac{\tilde{\boldsymbol{\tau}}_1^{I_1}s^j\tilde{\boldsymbol{\tau}}_2^{I_2}}{I_1!j!I_2!}. \end{aligned} \quad (5.3)$$

Using the fact that $\zeta(x_0) = 0_{x_0}$ and letting h_j, for each $j \in \{1, \ldots, 2k\}$, be the smooth function $x \mapsto h_j(x) = (\zeta^j f)(x) - (\zeta^j f)(x_0)$, we can rewrite (5.3) as

$$
\begin{aligned}
e_{2k+1}^{(\boldsymbol{\xi}_1 * \zeta * \boldsymbol{\xi}_2)f}(\tilde{\boldsymbol{\tau}}_1(s), s, \tilde{\boldsymbol{\tau}}_2(s)) &= e_{2k+1}^{(\boldsymbol{\xi}_1 * \boldsymbol{\xi}_2)f}(\tilde{\boldsymbol{\tau}}_1(s), \tilde{\boldsymbol{\tau}}_2(s)) + \sum_{\substack{|I_1|+j=2 \\ |I_1|, j \geq 1}}^{2k+1} (\boldsymbol{\xi}_1^{I_1} \zeta^j f)(x_0) \frac{s^j \tilde{\boldsymbol{\tau}}_1^{I_1}(s)}{j! I_1!} \\
&\quad + \sum_{\substack{|I_1|+j+|I_2|=3 \\ |I_1|, j, |I_2| \geq 1}}^{2k+1} (\boldsymbol{\xi}_1^{I_1} \zeta^j \boldsymbol{\xi}_2^{I_2} f)(x_0) \frac{\tilde{\boldsymbol{\tau}}_1^{I_1} s^j \tilde{\boldsymbol{\tau}}_2^{I_2}}{I_1! j! I_2!} \\
&= e_{2k+1}^{(\boldsymbol{\xi}_1 * \boldsymbol{\xi}_2)f}(\tilde{\boldsymbol{\tau}}_1(s), \tilde{\boldsymbol{\tau}}_2(s)) + \sum_{j=1}^{2k} \frac{s^j}{j!} e_{(2k+1)-j}^{\boldsymbol{\xi}_1(h_j)}(\tilde{\boldsymbol{\tau}}_1(s)) \\
&\quad + \sum_{\substack{|I_1|+j+|I_2|=3 \\ |I_1|, j, |I_2| \geq 1}}^{2k+1} (\boldsymbol{\xi}_1^{I_1} \zeta^j \boldsymbol{\xi}_2^{I_2} f)(x_0) \frac{\tilde{\boldsymbol{\tau}}_1^{I_1} s^j \tilde{\boldsymbol{\tau}}_2^{I_2}}{I_1! j! I_2!}. \qquad (5.4)
\end{aligned}
$$

Now, $\mathrm{ord}_{x_0}(\boldsymbol{\xi}_1 * \boldsymbol{\xi}_2, \boldsymbol{\tau}_1 * \boldsymbol{\tau}_2) \geq k+1$ because $V_{\boldsymbol{\xi}_1,\boldsymbol{\tau}_1} + V_{\boldsymbol{\xi}_2,\boldsymbol{\tau}_2} = w_{x_0} - w_{x_0} = 0_{x_0}$, and, therefore, $\mathrm{ord}_{x_0}(\boldsymbol{\xi}_1 * \boldsymbol{\xi}_2, \tilde{\boldsymbol{\tau}}_1 * \tilde{\boldsymbol{\tau}}_2) \geq 2(k+1) = 2k+2$. Hence, the derivatives of $e_{2k+1}^{(\boldsymbol{\xi}_1 * \boldsymbol{\xi}_2)f}(\tilde{\boldsymbol{\tau}}_1(s), \tilde{\boldsymbol{\tau}}_2(s))$ of orders $1, \ldots, 2k+1$ all vanish at $s = 0$. By Lemma 4.4, the term (5.4) can be written as

$$
\sum_{j+|I_2|=2}^{2k} \frac{s^j \tilde{\boldsymbol{\tau}}_2^{I_2}(s)}{j! I_2!} \, e_{(2k+1)-(j+|I_2|)}^{\boldsymbol{\xi}_1(H_{j,I_2})}(\tilde{\boldsymbol{\tau}}_1(s)), \qquad (5.5)
$$

where H_{j,I_2} is the smooth function $x \mapsto H_{j,I_2}(x) = Z^j \boldsymbol{\xi}_2^{I_2} f(x) - Z^j \boldsymbol{\xi}_2^{I_2} f(x_0)$. By Lemma 4.5, the derivatives of (5.5) up to order $2k+1$ vanish at $s = 0$. Hence $V_{\boldsymbol{\xi}_1 * \zeta * \boldsymbol{\xi}_2, \tilde{\boldsymbol{\tau}}_1 * s * \tilde{\boldsymbol{\tau}}_2}$ is determined by the $(2k+1)$st derivative of the $\mathbb{R}$-valued function

$$
s \mapsto g(s) := \sum_{j=1}^{2k} \frac{s^j}{j!} f_j(s),
$$

where, for each $j \in \{1, \ldots, 2k\}$,

$$
f_j(s) = e_{(2k+1)-j}^{\boldsymbol{\xi}_1(h_j)}(\tilde{\boldsymbol{\tau}}_1(s)).
$$

Now since $\mathrm{ord}_{x_0}(\boldsymbol{\xi}_1, \tilde{\boldsymbol{\tau}}_1) = 2k$, if $j \in \{2, \ldots, 2k\}$, then the $(2k+1-j)$th derivative of f_j at $s = 0$ vanishes and, therefore, the $(2k+1)$st derivative of $s \mapsto s^j f_j(s)$ vanishes at $s = 0$. Thus the $(2k+1)$st derivative of g at $s = 0$ is equal to the $(2k+1)$st derivative of $s \mapsto sf_1(s)$ at $s = 0$. The $2k$th derivative of f_1 at $s = 0$ is precisely $w_{x_0}(\zeta f - \zeta f(x_0)) = B_\zeta(w_{x_0})(f)$, and therefore, the $(2k+1)$st derivative of $s \mapsto sf_1(s)$ is $(2k+1)\zeta(w_{x_0})(f)$. Therefore, $(2k+1)B_\zeta(w_{x_0}) \in \mathcal{V}_{x_0}\mathcal{A}$ and since $\mathcal{V}_{x_0}\mathcal{A}$ is a cone, $B_\zeta(w_{x_0}) \in \mathcal{V}_{x_0}\mathcal{A}$. This completes the proof. ■

In the case that A is regular at x_0 and $\mathsf{A}_{x_0} \subset W$, computing $\langle \mathcal{Z}_{x_0}\mathcal{A}; W\rangle$ is greatly simplified in the following sense.

Theorem 5.15 *Suppose that* A *is regular at* x_0 *and let* $\mathcal{A}$ *be an affine system in* A*. If* $W \subset \mathcal{V}_{x_0}\mathcal{A}$ *is a subspace containing* A_{x_0}*, then* $\langle \mathcal{Z}_{x_0}\mathcal{A}; W\rangle = \langle B_\zeta; W\rangle$ *for any* $\zeta \in \mathcal{Z}_{x_0}\mathcal{A}$*.*

Proof: Fix $\zeta \in \mathcal{Z}_{x_0}\mathcal{A}$ and let $\zeta_1 \in \mathcal{Z}_{x_0}\mathcal{A}$ be arbitrary. Then $Y_1 = \zeta_1 - \zeta$ is a $\mathsf{L}(\mathsf{A})$-vector field and, moreover, $Y_1(x_0) = 0_{x_0}$. Therefore, for $w_{x_0} \in \mathsf{T}_{x_0}\mathsf{M}$,

$$B_{\zeta_1}(w_{x_0}) = B_\zeta(w_{x_0}) + B_{Y_1}(w_{x_0}),$$

and, therefore, by Proposition 3.4,

$$B_{\zeta_1}(w_{x_0}) - B_\zeta(w_{x_0}) \in \mathsf{A}_{x_0}.$$

If $\zeta_2 = \zeta + Y_2 \in \mathcal{Z}_{x_0}\mathcal{A}$, then

$$B_{\zeta_2}B_{\zeta_1}(w_{x_0}) = (B_\zeta)^2(w_{x_0}) + B_\zeta(B_{Y_1}(w_{x_0})) + B_{Y_2}(B_\zeta(w_{x_0}) + B_{Y_1}(w_{x_0})),$$

and, therefore, by Proposition 3.4,

$$B_{\zeta_2}B_{\zeta_1}(w_{x_0}) - (B_\zeta)^2(w_{x_0}) - B_\zeta(B_{Y_1}(w_{x_0})) \in \mathsf{A}_{x_0}.$$

By induction, if $\zeta_i = \zeta + Y_i \in \mathcal{Z}_{x_0}\mathcal{A}$ for $i \in \{1, \dots, k\}$, then

$$B_{\zeta_k} \cdots B_{\zeta_1}(w_{x_0}) - (B_\zeta)^k(w_{x_0}) - (B_\zeta)^{k-1}(b_{x_0,1}) - \cdots - B_\zeta(b_{x_0,k-1}) \in \mathsf{A}_{x_0},$$

where $b_{x_0,\ell} \in \mathsf{A}_{x_0}$ for $\ell \in \{1, \dots, k-1\}$. Hence, if W is a subspace containing A_{x_0}, for $w_{x_0} \in W$ and any collection $\zeta_1, \dots, \zeta_k \in \mathcal{Z}_{x_0}\mathcal{A}$, it holds that $B_{\zeta_k} \cdots B_{\zeta_1}(w_{x_0}) \in$

$\langle B_\zeta; W\rangle$. By (5.2), this proves that $\langle \mathcal{Z}_{x_0}\mathcal{A}; W\rangle \subset \langle B_\zeta; W\rangle$ if W contains A_{x_0}. The reverse inclusion holds regardless of whether $\mathsf{A}_{x_0} \subset W$ or not. This completes the proof. ■

Proposition 5.13 and Theorem 5.14 imply the following.

Theorem 5.16 *Let $\mathcal{A}$ be an affine system that is proper at x_0. If*

$$\langle \mathcal{Z}_{x_0}\mathcal{A}; \mathsf{A}_{x_0}\rangle + \mathrm{span}\{\mathcal{D}^{(2)}_{\mathcal{A}}(x_0)\} = \mathsf{T}_{x_0}\mathsf{M}$$

then $\mathcal{A}$ is STLC from x_0.

Example 5.17 On $\mathsf{M} = \mathbb{R}^n$, let $\mathcal{A}$ be the linear control system $\dot{x} = Ax + Bu$, where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and u lies in the unit cube in $\mathbb{R}^m$. By Proposition 5.11, $\mathcal{V}^1_{x_0}\mathcal{A} = \mathrm{img}(B)$. The set $\mathcal{Z}_{x_0}\mathcal{A}$ contains the vector field $x \mapsto Ax$, i.e., the drift vector field. Hence, by Theorem 5.14, the smallest subspace containing $\mathrm{img}(B)$ and invariant under the linear vector field $x \mapsto Ax$ is a subspace of variations. In other words, the image of the classical Kalman controllability matrix $[B\ AB\ \cdots A^{n-1}B]$ is a subspace of variations.

Chapter 6

Driftless and homogeneous systems

In this chapter, we give two applications of the methods developed in this thesis for two important classes of systems: driftless and homogeneous systems. For driftless systems, we show that the LARC at x_0 is sufficient for PSTLC from x_0. This result is well-known, of course, dating to the work of Chow and Rashevski [14, 36]. We then consider homogeneous systems which are central to proving the well-known sufficient conditions of Sussmann [45] and Bianchini-Stefani [8]; see Hermes [21] for an excellent survey. We give a necessary and sufficient condition, in terms of the variational cone, for local controllability for homogeneous systems. Moreover, for these systems, we are able to give a positive answer to an open problem in control theory regarding whether local controllability can be determined in a finite number of differentiations.

6.1 Driftless Systems

When the affine distribution A is a distribution, affine systems become what are commonly called driftless systems. Here is the definition.

Definition 6.1 Let D be a smooth distribution. A *driftless system* in D is a multi-valued vector field $\mathcal{D} : \mathsf{M} \rightrightarrows \mathsf{TM}$ such that $\operatorname{span} \mathcal{D}(x) = \mathsf{D}_x$ for each $x \in \mathsf{M}$.

We begin by showing that all variations can be neutralized for a distribution.

Proposition 6.2 *Let* D *be a smooth distribution and let* $x_0 \in \mathsf{M}$*. For each positive integer* k*,* $\mathcal{V}^k_{x_0}\mathsf{D}$ *is a subspace.*

Proof: By Proposition 5.1, $\mathcal{V}^k_{x_0}\mathsf{D}$ is a convex cone, so to prove that it is a subspace we need only prove that it is closed under multiplication by -1. Let $(\boldsymbol{\xi}, \boldsymbol{\tau}) \in \mathsf{V}^k_{\mathsf{D}}$, let $\tilde{\boldsymbol{\xi}} = (-\xi_p, \ldots, -\xi_1)$, and let $\tilde{\boldsymbol{\tau}} = (\tau^p, \ldots, \tau^1)$. We note that, for each $i \in \{1, \ldots, p\}$, $-\xi_i$ is a D-vector field, and thus $(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\tau}}) \in \mathsf{V}_{\mathsf{D}}$. It is clear that

$$\Phi^{\tilde{\boldsymbol{\xi}}}_{\tilde{\boldsymbol{\tau}}(s)} \circ \Phi^{\boldsymbol{\xi}}_{\boldsymbol{\tau}(s)}(x_0) = x_0$$

for all s in a neighbourhood of zero. Thus, for all $\ell \geq 0$ and any smooth function f vanishing at x_0, it holds that $s \mapsto e^{(\boldsymbol{\xi} * \tilde{\boldsymbol{\xi}})f}_{\ell}(\boldsymbol{\tau}(s), \tilde{\boldsymbol{\tau}}(s)) = 0$. Hence, by Lemma 4.4,

$$e^{\tilde{\boldsymbol{\xi}}f}_{k}(\tilde{\boldsymbol{\tau}}(s)) = -e^{(\boldsymbol{\xi}f)}_{k}(\boldsymbol{\tau}(s)) - m^{(\boldsymbol{\xi} * \tilde{\boldsymbol{\xi}})f}_{k}(\boldsymbol{\tau}(s), \tilde{\boldsymbol{\tau}}(s)). \tag{6.1}$$

By Lemma 4.5, the derivatives of $m^{(\boldsymbol{\xi} * \tilde{\boldsymbol{\xi}})f}_{k}(\boldsymbol{\tau}, \tilde{\boldsymbol{\tau}})$ at $s = 0$ of orders $1, \ldots, k$ vanish because $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) = k$. Therefore, differentiating (6.1) and evaluating at $s = 0$ through orders $1, \ldots, k$, we obtain that $\mathrm{ord}_{x_0}(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\tau}}) = k$ and $V_{\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\tau}}} = -V_{\boldsymbol{\xi}, \boldsymbol{\tau}}$. This proves that $-V_{\boldsymbol{\xi}, \boldsymbol{\tau}} \in \mathcal{V}^k_{x_0}\mathsf{D}$. ■

The next proposition states that, in the regular case, the variational cones of convex driftless systems agree with those of the distribution.

Proposition 6.3 *Let* D *be a smooth distribution that is regular at* x_0 *and let* $\mathcal{D}$ *be a smooth driftless system in* D*. If* $\mathcal{D}$ *is proper at* x_0 *then* $\mathcal{V}^k_{x_0}\mathsf{D} = \mathcal{V}^k_{x_0}\mathrm{co}(\mathcal{D})$ *for each* $k \geq 1$*. Consequently,* $\mathcal{V}_{x_0}\mathsf{D} = \mathcal{V}_{x_0}\mathrm{co}(\mathcal{D})$*.*

Proof: Let $(\boldsymbol{\xi}, \boldsymbol{\tau}) \in \mathsf{V}^k_{\mathsf{D}}$, where $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ and $\boldsymbol{\tau} \in \mathrm{ET}_p$. There exists $\lambda > 0$ such that $\lambda\xi_j(x_0) \in \mathrm{int}_{\mathsf{D}_{x_0}}(\mathrm{co}(\mathcal{D}(x_0)))$ for all $j \in \{1, \ldots, p\}$. By Lemma 3.24, there exists a neighbourhood Ω of x_0 such that, for all $j \in \{1, \ldots, p\}$ and all $x \in \Omega$, it holds that

$\lambda\xi_j(x) \in \mathrm{co}(\mathcal{D}(x))$, i.e., the $\lambda\xi_j$'s are $\mathrm{co}(\mathcal{D})$-vector fields. Define $\boldsymbol{\xi}_\lambda = (\lambda\xi_1, \ldots, \lambda\xi_p)$ and $\boldsymbol{\tau}_\lambda(s) = \frac{1}{\lambda}\boldsymbol{\tau}(s)$. Then $(\boldsymbol{\xi}_\lambda, \boldsymbol{\tau}_\lambda) \in \mathsf{V}_{\mathrm{co}(\mathcal{D})}$ and, by Theorem 4.1, for any smooth function f,

$$e_k^{\boldsymbol{\xi}_\lambda f}(\boldsymbol{\tau}_\lambda(s)) = \sum_{|I|=0}^{k} (\boldsymbol{\xi}_\lambda^I f)(x_0)\frac{\boldsymbol{\tau}_\lambda^I(s)}{I!} = \sum_{|I|=0}^{k} (\lambda^{|I|}\boldsymbol{\xi}^I f)(x_0)\frac{\boldsymbol{\tau}^I(s)}{\lambda^{|I|}I!} = e_k^{\boldsymbol{\xi} f}(\boldsymbol{\tau}(s)).$$

Therefore, $V_{\boldsymbol{\xi},\boldsymbol{\tau}} = V_{\boldsymbol{\xi}_\lambda,\boldsymbol{\tau}_\lambda} \in \mathcal{V}^k_{x_0}\mathrm{co}(\mathcal{D})$. This proves that $\mathcal{V}^k_{x_0}\mathsf{D} \subseteq \mathcal{V}^k_{x_0}\mathrm{co}(\mathcal{D})$, and, therefore, $\mathcal{V}_{x_0}\mathsf{D} \subseteq \mathcal{V}_{x_0}\mathrm{co}(\mathcal{D})$. The reverse inclusion is obvious. ■

Combining Proposition 6.3 and Corollary 3.31 we obtain the following.

Theorem 6.4 *Let D be a smooth distribution that is regular at x_0. If $\mathcal{V}_{x_0}\mathsf{D} = \mathsf{T}_{x_0}\mathsf{M}$ then D is PSTLC from x_0.*

Proof: Let $\mathcal{D}$ be a convex driftless system that is proper, smooth, and satisfies the LARC at x_0. By Proposition 6.3, $\mathcal{V}_{x_0}\mathcal{D} = \mathcal{V}_{x_0}\mathsf{D} = \mathsf{T}_{x_0}\mathsf{M}$, which implies that $\mathcal{D}$ is STLC from x_0 by Theorem 5.5. Since $\mathcal{D}$ was arbitrary and by Corollary 3.31, the proof is complete. ■

In the rest of this section, we will construct an explicit type of variation for a driftless system that will lead to the result that, for driftless systems, the LARC at x_0 is sufficient for PSTLC from x_0. The construction is motivated by [27, Theorem 3.16] but we will use our methods to prove the result. For vector fields ξ_1, ξ_2 and $s \in \mathbb{R}$ sufficiently small, define the (local) diffeomorphism

$$[\Phi_s^{\xi_2}, \Phi_s^{\xi_1}] = \Phi_s^{-\xi_1} \circ \Phi_s^{-\xi_2} \circ \Phi_s^{\xi_1} \circ \Phi_s^{\xi_2}.$$

In our notation, if $\boldsymbol{\xi} = (\xi_2, \xi_1, -\xi_2, -\xi_1)$ and $\boldsymbol{\tau}(s) = (s, s, s, s)$, then $\Phi^{\boldsymbol{\xi}}_{\boldsymbol{\tau}(s)} = [\Phi_s^{\xi_2}, \Phi_s^{\xi_1}]$. It is clear that $[\Phi_s^{\xi_2}, \Phi_s^{\xi_1}]^{-1} = [\Phi_s^{\xi_1}, \Phi_s^{\xi_2}]$. If ξ_3 is another vector field, put

$$[\Phi_s^{\xi_3}, [\Phi_s^{\xi_2}, \Phi_s^{\xi_1}]] = [\Phi_s^{\xi_2}, \Phi_s^{\xi_1}]^{-1} \circ \Phi_s^{-\xi_3} \circ [\Phi_s^{\xi_2}, \Phi_s^{\xi_1}] \circ \Phi_s^{\xi_3}.$$

In our notation, this corresponds to $\boldsymbol{\xi} = (\xi_3, \xi_2, \xi_1, -\xi_2, -\xi_1, -\xi_3, \xi_1, \xi_2, -\xi_1, -\xi_2, -\xi_3)$ and $\boldsymbol{\tau}(s) = (s, s, \dots, s) \in \mathbb{R}^{10}$. We can iterate this process to define higher-order commutators of local diffeomorphisms of the form

$$[\Phi_s^{\xi_p}, [\cdots, [\Phi_s^{\xi_2}, \Phi_s^{\xi_1}]] \cdots] \tag{6.2}$$

for vector fields $\xi_1, \dots, \xi_p$. An elementary induction shows that, for each $p \geq 2$, the commutator (6.2) corresponds to a family of $a_p = 3 \cdot 2^p - 2$ vectors fields.

Lemma 6.5 *Let $p \geq 2$ be an integer and let $\xi_1, \dots, \xi_p$ be smooth vector fields on M. Then, for any $x \in \mathsf{M}$, it holds that, for $\ell \in \{1, \dots, p-1\}$,*

$$\frac{\mathrm{d}^\ell}{\mathrm{d}s^\ell}\bigg|_{s=0} [\Phi_s^{\xi_p}, [\cdots, [\Phi_s^{\xi_2}, \Phi_s^{\xi_1}]] \cdots](x) = 0_x$$

and

$$\frac{\mathrm{d}^p}{\mathrm{d}s^p}\bigg|_{s=0} [\Phi_s^{\xi_p}, [\cdots, [\Phi_s^{\xi_2}, \Phi_s^{\xi_1}]] \cdots](x) = p! [\xi_p, [\cdots, [\xi_2, \xi_1]] \cdots](x).$$

Proof: The proof is by induction on p. Let $p = 2$. From Corollary 4.2, the first order term in the Taylor expansion of $[\Phi_s^{\xi_1}, \Phi_s^{\xi_2}](x)$ is $(\xi_2(x) + \xi_1(x) - \xi_2(x) - \xi_1(x))s$, which is identically zero for all $x \in \mathsf{M}$. Again, by Corollary 4.2, the second order term in the Taylor expansion of $[\Phi_s^{\xi_1}, \Phi_s^{\xi_2}](x)$ is (we are suppressing evaluation at x)

$$\xi_2^2 \frac{s^2}{2!} + \xi_2\xi_1 s^2 - \xi_2^2 s^2 - \xi_2\xi_1 s^2 + \xi_1^2 \frac{s^2}{2!} - \xi_1\xi_2 s^2 - \xi_1^2 s^2 + \xi_2^2 \frac{s^2}{2!} + \xi_2\xi_1 s^2 + \xi_1^2 \frac{s^2}{2!},$$

which simplifies to $\xi_2\xi_1 s^2 - \xi_1\xi_2 s^2 = [\xi_2, \xi_1]s^2$. This proves the case $p = 2$. Assume it for $p \geq 2$. Let $\xi_1, \xi_2, \dots, \xi_{p+1}$ be smooth vector fields and let $\boldsymbol{\xi}$ be the a_p-tuple of vector fields and let $\boldsymbol{\tau} \in \mathrm{ET}_{a_p}$ be defined as $\boldsymbol{\tau}(s) = (s, s, \dots, s) \in \mathbb{R}^{a_p}$, so that $\Phi_{\boldsymbol{\tau}(s)}^{\boldsymbol{\xi}} = [\Phi_s^{\xi_p}, [\cdots, [\Phi_s^{\xi_2}, \Phi_s^{\xi_1}]] \cdots]$. Let $\tilde{\boldsymbol{\xi}}$ denote the family obtained by reversing the order of the sequence $\boldsymbol{\xi}$ and multiplying each element by -1. Let $\boldsymbol{\eta} = \xi_{p+1} * \boldsymbol{\xi} * (-\xi_{p+1}) * \tilde{\boldsymbol{\xi}}$

and let

$$e^{\boldsymbol{\eta}}_{p+1}(s) = \sum_{\ell=1}^{p+1} \sum_{i+j+|I|+|J|=\ell} \big((\xi_{p+1})^i \boldsymbol{\xi}^I (-\xi_{p+1})^j \tilde{\boldsymbol{\xi}}^J\big)(x) \frac{s^\ell}{i!j!I!J!}. \tag{6.3}$$

By Corollary 4.2, $e^{\boldsymbol{\eta}}_{p+1}$ is the Taylor expansion of order $p+1$ of the curve

$$s \mapsto [\Phi^{\xi_{p+1}}_s, [\cdots, [\Phi^{\xi_2}_s, \Phi^{\xi_1}_s]] \cdots](x),$$

so that, at $s = 0$, they have the same derivatives up to order $p+1$. From the induction hypothesis, it follows that $\sum_{|I|=a} \frac{1}{I!}\boldsymbol{\xi}^I$ is identically zero for each $a \in \{1, \dots, p-1\}$ and that also

$$\sum_{|I|=p} \tfrac{1}{I!}\boldsymbol{\xi}^I = [\xi_p, [\cdots, [\xi_2, \xi_1]] \cdots].$$

Moreover, by the proof of Lemma 6.2, the same is true for the family $\tilde{\boldsymbol{\xi}}$. Hence, the only coefficients in the polynomial (6.3) that are potentially non-zero are when either $|I| = p$ or $|J| = p$. Hence there are two cases to consider: (i) $|I| = p$ (or $|J| = p$) and $i = j = 0$, or (ii) $|I| = p$ (or $|J| = p$) and $i = 1$ or $j = 1$. In the first case, if $|I| = p$ (so that $|J| = 1$), these coefficients will vanish identically by the induction hypothesis, and the case $|J| = p$ and $|I| = 1$ is identical. In the second case, the four coefficients that remain are when $|I| = p$ and $i = 1$, $|I| = p$ and $j = 1$, $|J| = p$ and $i = 1$, and $|J| = p$ and $j = 1$. Therefore,

$$\begin{aligned}
e^{\boldsymbol{\eta}}_{p+1}(s) &= \left(\xi_{p+1} \sum_{|I|=p} \frac{\boldsymbol{\xi}^I}{I!}\right)(x)s^{p+1} + \left(\sum_{|I|=p} \frac{\boldsymbol{\xi}^I}{I!}(-\xi_{p+1})\right)(x)s^{p+1} \\
&\quad + \left(\sum_{|J|=p} (\xi_{p+1})\frac{\tilde{\boldsymbol{\xi}}^J}{J!}\right)(x)s^{p+1} + \left(\sum_{|J|=p} (-\xi_{p+1})\frac{\tilde{\boldsymbol{\xi}}^J}{J!}\right)(x)s^{p+1} \\
&= \left(\xi_{p+1} \sum_{|I|=p} \frac{\boldsymbol{\xi}^I}{I!}\right)(x)s^{p+1} + \left(\sum_{|I|=p} \frac{\boldsymbol{\xi}^I}{I!}(-\xi_{p+1})\right)(x)s^{p+1} \\
&= [\xi_{p+1}, [\xi_p, [\cdots, [\xi_2, \xi_1]] \cdots](x)s^{p+1}
\end{aligned}$$

and the proof is complete. ∎

Lemma 6.5 gives us a formula for iterative Lie brackets in terms of monomials. To state the formula, we introduce some notation. Given a finite list of indeterminates $\boldsymbol{x} = (x_1, \dots, x_k)$, let $\tilde{\boldsymbol{x}} = (-x_k, \dots, -x_1)$. Given another indeterminate y, define $\mathsf{\rho}_y(\boldsymbol{x}) = (y, \boldsymbol{x}, -y, \tilde{\boldsymbol{x}})$. For example, given x_1, x_2, x_3,

$$\mathsf{\rho}_{x_3}(\mathsf{\rho}_{x_2}(x_1)) = (x_3, \mathsf{\rho}_{x_2}(x_1), -x_3, \widetilde{\mathsf{\rho}_{x_2}(x_1)}) = (x_3, x_2, x_1, -x_2, -x_1, x_3, x_1, x_2, -x_1, -x_2).$$

With this notation and Lemma 6.5, the following formula is immediate.

Corollary 6.6 *Let $\xi_1, \dots, \xi_p$ be vector fields and define the family of vector fields via $\boldsymbol{\xi} = \mathsf{\rho}_{\xi_p}(\mathsf{\rho}_{\xi_{p-1}}(\cdots(\mathsf{\rho}_{\xi_2}(\xi_1))\cdots))$. Then*

$$[\xi_p, [\xi_{p-1}, [\cdots [\xi_2, \xi_1]]\cdots] = \sum_{|I|=p} \boldsymbol{\xi}^I \frac{1}{I!}.$$

The following is an immediate consequence of Lemma 6.5.

Proposition 6.7 *Let D be a smooth distribution and let $p \in \mathbb{Z}_{\geq 2}$. Then, for any family $\boldsymbol{\xi} = (\xi_1, \dots, \xi_p)$ of D-vector fields and $x_0 \in \mathsf{M}$, it holds that*

$$[\xi_p, [\cdots, [\xi_2, \xi_1]]\cdots](x_0) \in \mathcal{V}^p_{x_0}\mathsf{D}.$$

For completeness, we state the following.

Theorem 6.8 *Let D be a smooth distribution and suppose that $\Gamma_{x_0}(\mathsf{D})$ satisfies the LARC at x_0. Then $\mathcal{V}_{x_0}\mathsf{D} = \mathsf{T}_{x_0}\mathsf{M}$. Consequently, if D is regular at x_0 then it is PSTLC from x_0.*

Proof: It is well-known that $\mathrm{Lie}(\Gamma_{x_0}(\mathsf{D}))$ is spanned by elements of the form

$$[\xi_p, [\cdots, [\xi_2, \xi_1]]\cdots](x_0),$$

where $\xi_1, \dots, \xi_p \in \Gamma_{x_0}(\mathsf{D})$ and $p \in \mathbb{Z}_{\geq 1}$, see for example [35, Proposition 3.8]. Hence, if $\Gamma_{x_0}(\mathsf{D})$ satisfies the LARC at x_0, by Proposition 6.7 there is some k such that $\mathcal{V}^k_{x_0}\mathsf{D} = \mathsf{T}_{x_0}\mathsf{M}$ and thus $\mathcal{V}_{x_0}\mathsf{D} = \mathsf{T}_{x_0}\mathsf{M}$. The second statement is Theorem 6.4. ■

6.2 Homogeneous systems

Homogeneous systems have received much attention in the literature with regards to controllability and also stabilizability, see [21] for a survey. One of the basic problems is concerned with constructing homogeneous approximations that preserve the property of interest, for example, STLC or stabilizability. Our aim in this section is to show that, for homogeneous systems, one can characterize the local controllability property with the variational cone.

To define homogeneous systems, we need the notion of a dilation. A *one-parameter family of dilations* on $\mathbb{R}^n$ is a map $\Delta\colon \mathbb{R}_{>0} \to L(\mathbb{R}^n; \mathbb{R}^n)$ of the form

$$\Delta(s)(x^1, \dots, x^n) = (s^{k_1}x^1, s^{k_2}x^2, \dots, s^{k_n}x^n)$$

for positive integers $k_1, \dots, k_n$. We write Δ_s for the linear map $\Delta(s)$.

Given a control-affine system $\Sigma = (\mathbb{R}^n, \{X_0, X_1, \dots, X_m\}, U)$, a *controlled trajectory on* $[0, T]$ of Σ is a pair (γ, u), where $u\colon [0, T] \to \mathbb{R}^m$ is an integrable map such that $u(t) \in U$ and $\gamma\colon [0, T] \to \mathbb{R}^n$ is the absolutely continuous curve satisfying

$$\gamma'(t) = X_0(\gamma(t)) + \sum_{a=1}^{m} u^a(t) X_a(\gamma(t)).$$

The set of all controlled trajectories on $[0, T]$ of Σ is denoted by $\mathrm{Traj}_\Sigma(T)$. Given $(\gamma, u) \in \mathrm{Traj}_\Sigma(T)$ and $s > 0$, we define $(\gamma^{\mathrm{tm}}_s, u^{\mathrm{tm}}_s) \in \mathrm{Traj}_\Sigma(sT)$ by setting $u^{\mathrm{tm}}_s(t) = u(t/s)$, for all $t \in [0, sT]$. Similarly, given $(\gamma, u) \in \mathrm{Traj}_\Sigma(T)$ and $\varepsilon > 0$, we define $(\gamma^{\mathrm{cr}}_\varepsilon, u^{\mathrm{cr}}_\varepsilon) \in \mathrm{Traj}_\Sigma(T)$ by setting $u^{\mathrm{cr}}_\varepsilon(t) = \varepsilon u(t)$, for all $t \in [0, T]$.

Definition 6.9 Let $\Sigma = (\mathbb{R}^n, \{X_0, X_1, \dots, X_m\}, U)$ be a control-affine system.

(i) We say that Σ is *time-homogeneous* with respect to the one-parameter family of dilations Δ^{tm} if, for every $(\gamma, u) \in \mathrm{Traj}_\Sigma(T)$ inducing $(\gamma_s^{\mathrm{tm}}, u_s^{\mathrm{tm}})$, it holds that $\gamma_s^{\mathrm{tm}}(st) = \Delta_s^{\mathrm{tm}}(\gamma(t))$, for all $t \in [0, T]$.

(ii) We say that Σ is *control-homogeneous* with respect to the one-parameter family of dilations Δ^{cr} if, for every $(\gamma, u) \in \mathrm{Traj}_\Sigma(T)$ inducing $(\gamma_\varepsilon^{\mathrm{cr}}, u_\varepsilon^{\mathrm{cr}})$, it holds that $\gamma_\varepsilon^{\mathrm{cr}}(t) = \Delta_\varepsilon^{\mathrm{cr}}(\gamma(t))$, for all $t \in [0, T]$.

Time-homogeneous systems have, naturally, homogeneous reachable sets.

Lemma 6.10 *Let Σ be a control-affine system and suppose that Σ is time-homogeneous with respect to the dilation Δ^{tm}. Then, for each $T > 0$,*

$$\mathcal{R}_\Sigma(x_0, sT) = \Delta_s^{\mathrm{tm}}(\mathcal{R}_\Sigma(x_0, T)).$$

Consequently,

$$\mathcal{R}_\Sigma(x_0, \leq sT) = \Delta_s^{\mathrm{tm}}(\mathcal{R}_\Sigma(x_0, \leq T)).$$

Proof: Let $(\gamma, u) \in \mathrm{Traj}_\Sigma(T)$ and let $(\gamma_s^{\mathrm{tm}}, u_s^{\mathrm{tm}}) \in \mathrm{Traj}_\Sigma(sT)$ be the induced controlled trajectory. By definition of time-homogeneity, $\Delta_s^{\mathrm{tm}}(\gamma(T)) = x_s(sT) \in \mathcal{R}_\Sigma(x_0, sT)$, so that $\Delta_s^{\mathrm{tm}}(\mathcal{R}_\Sigma(x_0, T)) \subset \mathcal{R}_\Sigma(x_0, sT)$. To prove the reverse inclusion, let $(x_s, u_s) \in \mathrm{Traj}_\Sigma(sT)$. Define $u\colon [0, T] \to \mathbb{R}^m$ by $u(t) = u_s(st)$ and let $(\gamma, u) \in \mathrm{Traj}_\Sigma(T)$ be the resulting controlled trajectory. Then, by definition, $(x_s, u_s) = (\gamma_s^{\mathrm{tm}}, u_s^{\mathrm{tm}})$, and, therefore, by time-homogeneity, $x_s(st) = \gamma_s^{\mathrm{tm}}(st) = \Delta_s^{\mathrm{tm}}(\gamma(t))$ for $t \in [0, T]$. Hence, $x_s(sT) = \Delta_s^{\mathrm{tm}}(\gamma(T)) \in \Delta_s^{\mathrm{tm}}(\mathcal{R}_\Sigma(x_0, T))$. This proves the reverse inclusion. ■

The proof of the following is similar to the proof of the previous lemma but we include it for completeness.

Lemma 6.11 *Let $\Sigma = (\mathbb{R}^n, \mathcal{X}, U)$ be a control-affine system and suppose that Σ is control-homogeneous with respect to the dilation Δ^{cr}. For $\varepsilon > 0$, let $U_\varepsilon =$*

$\{\varepsilon u \mid \ u \in U\}$ *and let* $\Sigma_\varepsilon = (\mathbb{R}^n, \mathcal{X}, U_\varepsilon)$. *Then, for each* $T > 0$,

$$\mathcal{R}_{\Sigma_\varepsilon}(x_0, T) = \Delta_\varepsilon^{\mathrm{cr}}(\mathcal{R}_\Sigma(x_0, T)).$$

Consequently,

$$\mathcal{R}_{\Sigma_\varepsilon}(x_0, \leq T) = \Delta_\varepsilon^{\mathrm{cr}}(\mathcal{R}_\Sigma(x_0, \leq T)).$$

Proof: Let $(\gamma, u) \in \mathrm{Traj}_\Sigma(T)$ and let $(\gamma_\varepsilon^{\mathrm{cr}}, u_\varepsilon^{\mathrm{cr}}) \in \mathrm{Traj}_{\Sigma_\varepsilon}(T)$ be the induced controlled trajectory. By definition of control-homogeneity, $\Delta_\varepsilon^{\mathrm{cr}}(\gamma(T)) = \gamma_\varepsilon^{\mathrm{cr}}(T) \in \mathcal{R}_{\Sigma_\varepsilon}(x_0, T)$ so that $\Delta_\varepsilon^{\mathrm{cr}}(\mathcal{R}_\Sigma(x_0, T)) \subset \mathcal{R}_{\Sigma_\varepsilon}(x_0, T)$. To prove the reverse inclusion, let $(x_\varepsilon, u_\varepsilon) \in \mathrm{Traj}_{\Sigma_\varepsilon}(T)$. Define $u\colon [0, T] \to \mathbb{R}^m$ by $u(t) = \frac{1}{\varepsilon} u_\varepsilon(t)$ and let $(\gamma, u) \in \mathrm{Traj}_\Sigma(T)$ be the resulting controlled trajectory. Then by definition, $(x_\varepsilon, u_\varepsilon) = (\gamma_\varepsilon^{\mathrm{cr}}, u_\varepsilon^{\mathrm{cr}})$, and, therefore, by control-homogeneity, $x_\varepsilon(t) = \gamma_\varepsilon^{\mathrm{cr}}(t) = \Delta_\varepsilon^{\mathrm{cr}}(\gamma(t))$. Hence, $x_\varepsilon(T) = \Delta_\varepsilon^{\mathrm{cr}}(\gamma(T)) \in \Delta_\varepsilon^{\mathrm{cr}}(\mathcal{R}_\Sigma(x_0, T))$. This proves the reverse inclusion. ■

Definition 6.12 Let Σ be a control-affine system on $\mathbb{R}^n$. We say that Σ is *controllable from* x_0 if $\mathcal{R}_\Sigma(x_0) = \mathbb{R}^n$.

For a control-affine system Σ, we will use the more compact notation $\mathcal{V}_{x_0}^k \Sigma$ for $\mathcal{V}_{x_0}^k \mathcal{A}_\Sigma$. We are now ready to state the main result of this section.

Theorem 6.13 *Let* Σ *be a control-affine system on* $\mathbb{R}^n$ *that is time-homogeneous for some dilation* $\Delta_s^{\mathrm{tm}}(x) = (s^{k_1} x^1, \ldots, s^{k_n} x^n)$ *and let* $x_0 = 0 \in \mathbb{R}^n$. *Let* $k = \mathrm{lcm}(k_1, k_2, \ldots, k_n)$. *Suppose that* Σ *satisfies the LARC at* x_0. *The following are equivalent:*

(i) Σ *is STLC from* x_0*;*

(ii) $\mathcal{V}_{x_0}^{k_1} \Sigma + \mathcal{V}_{x_0}^{k_2} \Sigma + \cdots + \mathcal{V}_{x_0}^{k_n} \Sigma = \mathbb{R}^n$ *and hence* $\mathcal{V}_{x_0}^k \Sigma = \mathbb{R}^n$*;*

(iii) Σ *is controllable from* x_0.

Proof: We first prove that (i) implies (ii). Suppose that Σ is STLC from x_0 and let $T^* > 0$. Let $\{e_1, \ldots, e_n\}$ be the standard basis in $\mathbb{R}^n$ and let $e_j \in \{e_1, \ldots, e_n\}$

be arbitrary. By hypothesis and by a theorem of Grasse [17, Corollary 4.15], there exists a piecewise-constant control $u\colon [0,T] \to \mathbb{R}^m$ for Σ, where $T < T^*$, and a positive constant $c > 0$ such that the corresponding trajectory $\boldsymbol{\gamma}\colon [0,T] \to \mathbb{R}^n$ satisfies $\boldsymbol{\gamma}(T) = ce_j$. In other words, there exists a family of vector fields $\boldsymbol{\xi} = (\xi_1, \dots, \xi_p) \subset \Gamma_{x_0}(\Sigma)$, times $t_1, \dots, t_p > 0$ satisfying $t_1 + \cdots + t_p = T$, such that

$$\boldsymbol{\gamma}(T) = ce_j = \Phi^{\xi_p}_{t_p} \circ \ \cdots \ \circ \Phi^{\xi_1}_{t_1}(x_0).$$

Consider the curve $\boldsymbol{\nu}\colon [0,1] \to \mathbb{R}^n$ given by

$$\boldsymbol{\nu}(s) = \Phi^{\xi_p}_{t_p s} \circ \ \cdots \ \circ \Phi^{\xi_1}_{t_1 s}(x_0).$$

Then, by construction of $\boldsymbol{\nu}$, for $s \in (0,1]$, it holds that $\boldsymbol{\nu}(s) = \boldsymbol{\gamma}^{\mathrm{tm}}_s(sT)$, where $(\boldsymbol{\gamma}^{\mathrm{tm}}_s, u^{\mathrm{tm}}_s) \in \mathrm{Traj}_\Sigma(sT)$ is induced by $(\boldsymbol{\gamma}, u) \in \mathrm{Traj}_\Sigma(T)$. By time-homogeneity and the fact that $\boldsymbol{\nu}(0) = x_0$, it follows that $\boldsymbol{\nu}(s) = ce_j s^{k_j}$ for all $s \in [0,1]$. By construction of $\boldsymbol{\nu}$ and the definition of $\mathcal{V}^{k_j}_{x_0}\Sigma$, it is clear that $e_j \in \mathcal{V}^{k_j}_{x_0}\Sigma$. An identical procedure shows that also $-e_j \in \mathcal{V}^{k_j}_{x_0}\Sigma$. This proves that $\mathcal{V}^{k_1}_{x_0}\Sigma + \mathcal{V}^{k_2}_{x_0}\Sigma + \cdots + \mathcal{V}^{k_n}_{x_0}\Sigma = \mathbb{R}^n$. By Lemma 5.2, it follows that $\mathcal{V}^{k}_{x_0}\Sigma = \mathbb{R}^n$.

Now we prove that (ii) implies (iii). If (ii) holds then, by Theorem 5.5, (i) holds. Let $x^* \in \mathbb{R}^n$ and $T^* > 0$ be arbitrary. For $\lambda > 0$ sufficiently large, $\Delta^{\mathrm{tm}}_{1/\lambda}(x^*) \in \operatorname{int} \mathcal{R}_\Sigma(x_0, \leq T^*)$. Let $T \leq T^*$ and let $(\boldsymbol{\gamma}, u) \in \mathrm{Traj}_\Sigma(T)$ be such that $\boldsymbol{\gamma}(T) = \Delta^{\mathrm{tm}}_{1/\lambda}(x^*)$. Let $(\boldsymbol{\gamma}^{\mathrm{tm}}_s, u^{\mathrm{tm}}_s) \in \mathrm{Traj}_\Sigma(sT)$ be the induced trajectory. Then, by time-homogeneity, $\boldsymbol{\gamma}^{\mathrm{tm}}_s(sT) = \Delta^{\mathrm{tm}}_s(\Delta^{\mathrm{tm}}_{1/\lambda}(x^*)) = \Delta^{\mathrm{tm}}_{s/\lambda}(x^*)$. Hence, setting $s = \lambda$, we obtain that x^* is reachable from x_0 in time λT using the control u_λ. This proves that $x^* \in \mathcal{R}_\Sigma(x_0)$.

The proof that (iii) implies (i) can be done by first proving that (iii) implies (ii) in the exact same way as was shown that (i) implies (ii). Then we use the fact that (ii) implies (i), by Theorem 5.5. ■

Let us illustrate the procedure in the proof of Theorem 6.13 with an example.

Example 6.14 The following system was considered by Stefani [40]. The state manifold is $\mathsf{M} = \mathbb{R}^3$, $x_0 = (0,0,0)$, and the affine distribution is

$$\mathsf{A}_x = X_0(x) + \mathrm{span}\left\{X_1(x)\right\},$$

where

$$X_0 = x^1\frac{\partial}{\partial x^2} + (x^1)^3x^2\frac{\partial}{\partial x^3} \text{ and } X_1 = \frac{\partial}{\partial x^1}.$$

Consider the control-affine system $\Sigma = (\mathbb{R}^n, \{X_0, X_1\}, [-1,1])$. It is straightforward to show that Σ is control-homogeneous with respect to the dilation $\Delta^{\mathrm{cr}}_{\varepsilon}(x) = (\varepsilon x^1, \varepsilon x^2, \varepsilon^4 x^3)$, and that it is time-homogeneous with respect to the dilation $\Delta^{\mathrm{tm}}_s(x) = (sx^1, s^2x^2, s^6x^3)$. For $u \in U$ let $\xi_u = X_0 + uX_1$. Using Proposition 5.13, one computes that $\mathcal{V}^2_{x_0}\Sigma = \mathrm{span}\left\{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}\right\}$. According to Theorem 6.13, to produce variations in the $\pm\frac{\partial}{\partial x^3}$ directions, we need to look at variations of order six. Following the proof of Theorem 6.13, let $\boldsymbol{\tau}(s) = (a_1s, a_2s, a_3s)$ and let $\boldsymbol{\xi} = (\xi_{u_1}, \xi_{u_2}, \xi_{u_3})$, with $a_1u_1 + a_2u_2 + a_3u_3 = 0$. Then one computes that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\bigg|_{s=0}\Phi^{\boldsymbol{\xi}}_{x_0}(\boldsymbol{\tau}(s)) = (u_1a_1(a_1+2a_2+a_3) + u_2a_2(a_2+a_3))\frac{\partial}{\partial x^3}$$

and so we set $u_2 = -\frac{a_1(a_1+2a_2+a_3)u_1}{a_2(a_2+a_3)}$, so that $\mathrm{ord}(\boldsymbol{\xi},\boldsymbol{\tau}) \geq 3$. Then one computes that the derivatives of $\Phi^{\boldsymbol{\xi}}_{x_0}(\boldsymbol{\tau}(s))$ of orders 3,4, and 5 vanish at $s=0$, and that $\frac{\mathrm{d}^6}{\mathrm{d}s^6}\Big|_{s=0}\Phi^{\boldsymbol{\xi}}_{x_0}(\boldsymbol{\tau}(s))$ is equal to

$$-\frac{30u^4a_1^4(a_1+a_2)(a_1-a_3)(a_1+a_2+a_3)(a_1a_2+2a_1a_3+a_2a_3)}{(a_2+a_3)^3}\frac{\partial}{\partial x^3}.$$

By inspection, the above expression can be made negative and positive for all choices of u for appropriate values of $a_1, a_2, a_3 > 0$. Hence, $\mathcal{V}^6_{x_0}\Sigma = \mathrm{span}\left\{\frac{\partial}{\partial x^3}\right\}$, and, therefore, by Corollary 3.26, the affine distribution spanned by $\{X_0, X_1\}$ is PSTLC from the origin x_0.

In the proof of Theorem 6.13, *linear* end-times were used. This can potentially result in an over estimate for an integer k for which $\mathcal{V}^k_{x_0}\Sigma = \mathbb{R}^n$, i.e., the bound $\mathrm{lcm}(k_1, \dots, k_r)$ is not sharp, as the following example shows.

Example 6.15 The following system was considered by Kawski [25]. The control-affine system is given as $\Sigma = (\mathbb{R}^4, \{X_0, X_1\}, U)$ where

$$X_0 = x^1 \frac{\partial}{\partial x^2} + \frac{1}{6}(x^1)^3 \frac{\partial}{\partial x^3} + (x^2 x^3)\frac{\partial}{\partial x^4}, \quad X_1 = \frac{\partial}{\partial x^1},$$

and $U \subset \mathbb{R}$ is convex and proper. One can easily check that Σ is time-homogeneous with respect to the dilation $\Delta^{\mathrm{tm}}_s(x) = (sx^1, s^2x^2, s^4x^3, s^7x^4)$. Hence, from Theorem 6.13, Σ is STLC from $x_0 = 0$ if and only if $\mathcal{V}^{28}_{x_0}\Sigma = \mathbb{R}^4$. Using the procedure of Theorem 6.13, it is not too difficult to produce variations of orders $1, 2, 4$, and 7, in both positive and negative directions, so that $\mathcal{V}^{28}_{x_0}\Sigma = \mathbb{R}^4$. However, here we will show that actually $\mathcal{V}^{8}_{x_0}\Sigma = \mathbb{R}^4$ by using end-times that are not linear to produce the $\pm\frac{\partial}{\partial x^4}$ directions as eighth order variations. For $u \in U$ let $\xi_u = X_0 + uX_1$.

(i) Using Proposition 5.13, one computes that $\mathcal{V}^2_{x_0}\Sigma = \operatorname{span}\left\{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}\right\}$.

(ii) According to Theorem 6.13, to produce $\pm\frac{\partial}{\partial x^3}$ as variations, we must look at variations of order 4. Let $\boldsymbol{\tau}(s) = (a_1 s, a_2 s, a_3 s)$ and $\boldsymbol{\xi} = (\xi_{u_1}, \xi_{u_2}, \xi_{u_3})$ be such that $a_1u_1 + a_2u_2 + a_3u_3 = 0$. Then $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq 2$ and

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0} \Phi^{\boldsymbol{\xi}}_{x_0}(\boldsymbol{\tau}(s)) = (a_1^2 u_1 + a_1(2a_2 + a_3)u_1 + a_2(a_2 + a_3)u_2)\frac{\partial}{\partial x^2}.$$

Setting $u_2 = -\frac{1}{a_2(a_2+a_3)}(a_1^2u_1 + a_1(2a_2 + a_3)u_1)$ results in $\mathrm{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq 4$ and

$$\frac{\mathrm{d}^4}{\mathrm{d}s^4}\Big|_{s=0} \Phi^{\boldsymbol{\xi}}_{x_0}(\boldsymbol{\tau}(s)) = -\frac{a_1^3(a_1 + a_2)(a_1 - a_3)(a_1 + a_2 + a_3)u_1^3}{(a_2 + a_3)^2}\frac{\partial}{\partial x^3}.$$

We can then easily choose a_1, a_2, and a_3 to produce the variations $\pm\frac{\partial}{\partial x^3}$.

(iii) Now we produce variations in the directions $\pm\frac{\partial}{\partial x^4}$. Producing a variation in the direction $\frac{\partial}{\partial x^4}$ is straightforward but we will treat both cases simultaneously. To this end, let

$$\tau^i(s) = a_i s + b_i \frac{s^2}{2},$$

for $i = 1, 2, 3$, let $\boldsymbol{\tau}(s) = (\tau^1(s), \tau^2(s), \tau^3(s))$, let $\boldsymbol{\xi} = (\xi_{u_1}, \xi_{u_2}, \xi_{u_3})$, let $\tilde{\boldsymbol{\tau}}(s) = (\tau^3(s), \tau^2(s), \tau^1(s))$, and let $\tilde{\boldsymbol{\xi}} = (\xi_{u_3}, \xi_{u_2}, \xi_{u_1})$. If $a_1u_1 + a_2u_2 + a_3u_3 = 0$ then $\operatorname{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq 2$ and

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0} \Phi^{\boldsymbol{\xi}}_{x_0}(\boldsymbol{\tau}(s)) = \left(b_1u_1 + b_2u_2 - \frac{b_3(a_1u_1 + a_2u_2)}{a_3}\right)\frac{\partial}{\partial x^2} \\ + \left(a_1^2u_1 + a_1(2a_2 + a_3)u_1 + a_2(a_2 + a_3)u_2\right)\frac{\partial}{\partial x^3}.$$

If we set $b_3 = \frac{a_3}{a_1u_1 + a_2u_2}(b_1u_1 + b_2u_2)$, then we obtain that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0} \Phi^{\boldsymbol{\xi}}_{x_0}(\boldsymbol{\tau}(s)) = \left[(a_1^2u_1 + a_1(2a_2 + a_3)u_1 + a_2(a_2 + a_3)u_2\right]\frac{\partial}{\partial x^2}. \tag{6.4}$$

It is not hard to choose u_1, u_2, a_1, a_2 to make the tangent vector in (6.4) equal to zero, so that we can continue to produce a higher-order variation. But instead of this, we augment to $(\boldsymbol{\xi}, \boldsymbol{\tau})$ the reverse pair $(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\tau}})$ so that we can keep the variables u_1, u_2, a_1, a_2 free and simultaneously cancel the tangent vector in (6.4). In fact, one computes that, if we continue to use

$$a_1u_1 + a_2u_2 + a_3u_3 = 0 \quad \text{and} \quad b_3 = \frac{a_3}{a_1u_1 + a_2u_2}(b_1u_1 + b_2u_2),$$

then $\operatorname{ord}_{x_0}(\boldsymbol{\xi} * \tilde{\boldsymbol{\xi}}, \boldsymbol{\tau} * \tilde{\boldsymbol{\tau}}) \geq 7$ and

$$\frac{\mathrm{d}^7}{\mathrm{d}s^7}\Big|_{s=0} \Phi^{\boldsymbol{\xi} * \tilde{\boldsymbol{\xi}}}_{x_0}((\boldsymbol{\tau} * \tilde{\boldsymbol{\tau}})(s)) = Q_{\boldsymbol{a}}(u_1, u_2)\frac{\partial}{\partial x^4},$$

where $Q_{\boldsymbol{a}} \colon \mathbb{R}^2 \to \mathbb{R}$ is a homogeneous polynomial of degree 4 whose coefficients are homogeneous polynomials in $\boldsymbol{a} = (a_1, a_2, a_3)$ of degree 7. It is difficult to determine if the homogeneous polynomial $Q_{\boldsymbol{a}}$ can be made to be indefinite by an appropriate choice of the parameter $\boldsymbol{a}$, i.e., have both negative and positive values in its image. Instead, we investigate whether it is possible to choose $\boldsymbol{a}$ so that there exists a non-trivial subspace of $\mathbb{R}^2$ on which $Q_{\boldsymbol{a}}$ vanishes. To this end, set, for example, $a_1 = 1, a_2 = 1/10, a_3 = 5$, and $u_2 = \lambda u_1$, where $\lambda \in \mathbb{R}$ is to be determined. Then, one computes that

$$Q_{\boldsymbol{a}}(u_1, \lambda u_1) = \left[c_0 + c_1\lambda + c_2\lambda^2 + c_3\lambda^3 + c_4\lambda^4\right] u_1^4,$$

where $c_0, \ldots, c_4$ are positive rational numbers. Using a computer algebra system, one can check that the polynomial $c(\lambda) = c_0 + c_1\lambda + c_2\lambda^2 + c_3\lambda^3 + c_4\lambda^4$ has two real roots and they can be computed explicitly. Up to four digits they are given as $\lambda_1 = -15.7499\ldots$ and $\lambda_2 = -13.4544\ldots$. Hence, setting $a_1 = 1, a_2 = 1/10, a_3 = 5$, the subspaces $S_j = \{(u_1, u_2) \in \mathbb{R}^2 \mid \; u_2 = \lambda_j u_1\}$, $j = 1, 2$, are killed by $Q_{\boldsymbol{a}}$. Hence, setting $u_2 = \lambda_1 u_1$ yields that $\mathrm{ord}_{x_0}(\boldsymbol{\xi} * \tilde{\boldsymbol{\xi}}, \boldsymbol{\tau} * \tilde{\boldsymbol{\tau}}) \geq 8$ and one computes that, for these choices of parameters,

$$\frac{\mathrm{d}^8}{\mathrm{d}s^8}\Big|_{s=0} \Phi_{x_0}^{\boldsymbol{\xi} * \tilde{\boldsymbol{\xi}}}((\boldsymbol{\tau} * \tilde{\boldsymbol{\tau}})(s)) = (-r_1 b_1 + r_2 b_2) u_1^4 \frac{\partial}{\partial x^4},$$

where $r_1, r_2 > 0$ are constants. By inspection, one can then easily choose b_1 and b_2 to produce variations in the $\pm\frac{\partial}{\partial x^4}$ directions for any choice of u_1. Moreover, since u_2 and u_3 are directly proportional to u_1, by choosing u_1 sufficient small, we can force $u_1, u_2, u_3 \in U$.

Therefore, by Corollary 3.26, the affine distribution spanned by $\{X_0, X_1\}$ is PSTLC from the origin x_0.

Now we turn to the question of determining PSTLC for homogeneous systems.

Theorem 6.16 *Let $\Sigma = (\mathbb{R}^n, \{X_0, X_1, \dots, X_m\}, [-1,1]^m)$ be a control-affine system that is both time- and control-homogeneous for some dilations Δ^{tm} and Δ^{cr}, respectively. Assume that $X_1, \dots, X_m$ are linearly independent at $x_0 = 0$ and that Σ satisfies the LARC at x_0. Let A be the affine distribution generated by $\{X_0, X_1, \dots, X_m\}$. Then A is PSTLC from x_0 if and only if $\mathcal{V}^k_{x_0}\Sigma = \mathbb{R}^n$, where $k = \mathrm{lcm}(k_1, \dots, k_n)$ and the integers $k_1, \dots, k_n$ are those associated with the dilation Δ^{tm}.*

Proof: If A is PSTLC from x_0 then Σ is STLC from x_0. Therefore, by Theorem 6.13, $\mathcal{V}^k_{x_0}\Sigma = \mathbb{R}^n$.

Now suppose that $\mathcal{V}^k_{x_0}\Sigma = \mathbb{R}^n$ and let U be an arbitrary proper and convex control set. By Theorem 5.5, Σ is STLC from x_0. For $\varepsilon > 0$ let $\Sigma_\varepsilon = (\mathbb{R}^n, \{X_0, X_1, \dots, X_m\}, \varepsilon U)$. By control-homogeneity, Σ_ε is also STLC from x_0. For ε sufficiently small, $[-\varepsilon, \varepsilon]^m \subset U$. Hence, the control-affine system $(\mathbb{R}^n, \{X_0, X_1, \dots, X_m\}, U)$ is also STLC from x_0. By Corollary 3.26, A is PSTLC from x_0. ■

6.2.1 A class of high-order systems

In this section, we will consider the affine distribution A on $\mathbb{R}^n$ generated by the frame $\{\zeta, X_1, \dots, X_m\}$, where

$$\zeta = F^1 \frac{\partial}{\partial x^{m+1}} + F^2 \frac{\partial}{\partial x^{m+2}} + \cdots + F^r \frac{\partial}{\partial x^{m+r}},$$

where the $F^j \colon \mathbb{R}^m \to \mathbb{R}$ are homogeneous polynomials of order k, for $j = 1, \dots, r$, and $X_a = \frac{\partial}{\partial x^a}$ for $a = 1, 2, \dots, m$. In this section, $x_0 \in \mathbb{R}^n$ will denote the origin, and $n = m + r$.

Proposition 6.17 *Let A be the affine distribution on $\mathbb{R}^n$ generated by $\{\zeta, X_1, \dots, X_m\}$. Suppose that $\{\zeta, X_1, \dots, X_m\}$ satisfies the LARC at*

x_0. *Then* A *is PSTLC from* x_0 *if and only if* $\mathcal{V}^{k+1}_{x_0}\Sigma = \mathbb{R}^n$, *where* $\Sigma = (\mathbb{R}^n, \{\zeta, X_1, \dots, X_m\}, [-1,1]^m)$.

Proof: Write a point in $\mathbb{R}^{m+r}$ as (x, y). Let $u\colon [0,T] \to \mathbb{R}^m$ be a control for the system Σ and let $t \mapsto (x(t), y(t))$ the corresponding trajectory. Let $((x_s^{\mathrm{cr}}, y_s^{\mathrm{cr}}), u_s^{\mathrm{cr}})$ denote the induced controlled trajectory. Then

$$x_s^{\mathrm{cr}}(t) = \int_0^t u_s(w)\,\mathrm{d}w = \int_0^t u(w/s)\,\mathrm{d}w = s\int_0^{t/s} u(\sigma)\mathrm{d}\sigma = sx(t/s)$$

and therefore $x_s^{\mathrm{cr}}(sT) = sx(T)$. Let $F = (F^1, \dots, F^r)$. Then,

$$y_s^{\mathrm{cr}}(t) = \int_0^t F(x_s^{\mathrm{cr}}(w))\,\mathrm{d}w = \int_0^t F(sx(w/s))\,\mathrm{d}w = s^k \int_0^t F(x(w/s))\,\mathrm{d}w$$

$$= s^k \cdot s \int_0^{t/s} F(x(\sigma))\,\mathrm{d}\sigma = s^{k+1}y(t/s).$$

Thus, $y_s^{\mathrm{cr}}(sT) = s^{k+1}y(T)$. Thus, Σ is time-homogeneous with respect to the dilation $\Delta_s^{\mathrm{tm}}(x,y) = (sx, s^{k+1}y)$. In like manner, one can show that Σ is control-homogeneous with respect to the dilation $\Delta_\varepsilon^{\mathrm{cr}}((x,y)) = (\varepsilon x, \varepsilon^k y)$. The result now follows by Theorem 6.16. ■

If ξ is a vector field of order k at x_0 and

$$B_\xi^k = \sum_{j=1}^{\dim(\mathsf{M})} \sum_I \frac{\partial^k \xi^j}{\partial x^I}(x_0)\,\mathrm{d}x^I(x_0) \otimes \frac{\partial}{\partial x^j}(x_0)$$

is the associated vector-valued symmetric k-multilinear map on $\mathsf{T}_{x_0}\mathsf{M}$, we define $Q_\xi^k\colon \mathsf{T}_{x_0}\mathsf{M} \to \mathsf{T}_{x_0}\mathsf{M}$ by

$$Q_\xi^k(v_{x_0}) = B_\xi^k(v_{x_0}, \dots, v_{x_0}).$$

Note that when k is odd, $\mathrm{img}(Q_\xi^k)$ is closed under multiplication by -1. Indeed,

$$-Q_\xi^k(v_{x_0}) = -B_\xi^k(v_{x_0}, \dots, v_{x_0}) = B_\xi^k(-v_{x_0}, \dots, -v_{x_0}) = Q_\xi^k(-v_{x_0}).$$

Proposition 6.18 *Let* A *be the affine distribution on* $\mathbb{R}^n$ *generated by* $\{\zeta, X_1, \dots, X_m\}$. *Suppose that* $\{\zeta, X_1, \dots, X_m\}$ *satisfies the LARC at* x_0. *Let* $\mathcal{W}_{x_0} = \operatorname{span}\left\{Q_\zeta^k(\mathsf{A}_{x_0})\right\}$ *and suppose that* $\mathsf{A}_{x_0} + \mathcal{W}_{x_0} = \mathbb{R}^n$. *Then the following hold:*

(i) if k *is odd then* A *is PSTLC from* x_0*;*

(ii) if k *is even and* $0_{x_0} \in \operatorname{int}_{\mathcal{W}_{x_0}}(\operatorname{co}(Q_\zeta^k(\mathsf{A}_{x_0})))$ *then* A *is PSTLC from* x_0.

Proof: Put $\Sigma = (\mathbb{R}^n, \{\zeta, X_1, \dots, X_m\}, U = [-1,1]^m)$. According to Proposition 6.17, it is enough to prove that $\mathcal{V}_{x_0}^{k+1}\Sigma = \mathbb{R}^n$. Let $\xi_1 = \zeta + Y$ and $\xi_2 = \zeta - Y$, where $Y = u^a X_a$ for $u = (u^1, \dots, u^m) \in U$. Let $\boldsymbol{\xi} = (\xi_1, \xi_2)$ and let $\boldsymbol{\tau}(s) = (s, s)$. Then, by Theorem 4.1, for any smooth function f,

$$e_{k+1}^{\boldsymbol{\xi} f}(\boldsymbol{\tau}(s)) = \sum_{\ell=0}^{k+1} \sum_{j=0}^{k+1-\ell} (\xi_1^\ell \xi_2^j f)(x_0) \frac{s^{\ell+j}}{\ell! j!}. \tag{6.5}$$

Since $\xi_1(x_0) + \xi_2(x_0) = 0_{x_0}$, it follows that $\operatorname{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq 2$. In fact, since ζ is of order k at x_0 and Y is a constant vector field, by Lemma 4.12 we actually have that $\operatorname{ord}_{x_0}(\boldsymbol{\xi}, \boldsymbol{\tau}) \geq k+1$. Therefore, (6.5) simplies to

$$e_{k+1}^{\boldsymbol{\xi} f}(\boldsymbol{\tau}(s)) = \sum_{\ell=0}^{k+1} \binom{k+1}{\ell} (\xi_1^\ell \xi_2^{k+1-\ell} f)(x_0) \frac{s^{k+1}}{(k+1)!}.$$

Now, as differential operators,

$$\xi_1^\ell \xi_2^{k+1-\ell} = \begin{cases} \xi_1^\ell \xi_2^{k-\ell} \zeta - X_1^\ell \xi_2^{k-\ell} Y, & \ell = 0, 1, \dots, k, \\ \\ \xi_1^k \zeta + \xi_1^k Y, & \ell = k+1 \end{cases}$$

and, therefore, by Lemma 4.12 and the fact that $Y(x_0) = u$,

$$\xi_1^\ell \xi_2^{k+1-\ell}(x_0) = \begin{cases} B_\zeta^k(\underbrace{u, \dots, u}_{\ell-\text{times}}, \underbrace{-u, \dots, -u}_{(k-\ell)-\text{times}}), & \ell = 0, 1, \dots, k, \\ \\ B_\zeta^k(u, \dots, u), & \ell = k+1, \end{cases}$$

$$= \begin{cases} (-1)^{k-\ell} B_\zeta^k(u, \dots, u), & \ell = 0, 1, \dots, k, \\ \\ B_\zeta^k(u, \dots, u), & \ell = k+1. \end{cases}$$

Therefore,

$$e^{\boldsymbol{\xi}}_{k+1}(\boldsymbol{\tau}(s)) = \left(\sum_{\ell=0}^{k} \binom{k+1}{\ell} (-1)^{k-\ell} + 1\right) B^k_\zeta(u, u, \dots, u) \frac{s^{k+1}}{(k+1)!}.$$

Now

$$\begin{aligned}\sum_{\ell=0}^{k} \binom{k+1}{\ell} (-1)^{k-\ell} &= \sum_{\ell=0}^{k+1} \binom{k+1}{\ell} (-1)^{k-\ell} - (-1)^{-1} \\ &= (-1)^k \sum_{\ell=0}^{k+1} \binom{k+1}{\ell} (-1)^{\ell} + 1 = 1,\end{aligned}$$

and, therefore,

$$e^{\boldsymbol{\xi}}_{k+1}(\boldsymbol{\tau}(s)) = 2Q^k_\zeta(u) \frac{s^{k+1}}{(k+1)!}. \tag{6.6}$$

Using the fact that $\mathcal{V}^{k+1}_{x_0}\Sigma$ is a cone and by properness of U, we have $Q^k_\zeta(\mathsf{A}_{x_0}) \subset \mathcal{V}^{k+1}_{x_0}\Sigma$, and, therefore, the convexity of $\mathcal{V}^{k+1}_{x_0}\Sigma$ implies that $\mathrm{co}(Q^k_\zeta(\mathsf{A}_{x_0})) \subset \mathcal{V}^{k+1}_{x_0}\Sigma$. If k is odd then $\mathrm{co}(Q^k_\zeta(\mathsf{A}_{x_0}))$ is closed under multiplication by -1, and, therefore, $\mathrm{co}(Q^k_\zeta(\mathsf{A}_{x_0})) = \mathcal{W}_{x_0}$. If k is even and $0_{x_0} \in \mathrm{int}_{\mathcal{W}_{x_0}}(\mathrm{co}(Q^k_\zeta(\mathsf{A}_{x_0})))$ then again $\mathrm{co}(Q^k_\zeta(\mathsf{A}_{x_0})) = \mathcal{W}_{x_0}$. Now $\mathsf{A}_{x_0} = \mathcal{V}^1_{x_0}\Sigma \subset \mathcal{V}^{k+1}_{x_0}\Sigma$, and therefore $\mathbb{R}^n = \mathsf{A}_{x_0} + \mathcal{W}_{x_0} \subset \mathcal{V}^{k+1}_{x_0}\Sigma$. Hence, in either case of k, we have that $\mathcal{V}^{k+1}_{x_0}\Sigma = \mathbb{R}^n$, and the result now follows by Proposition 6.17. ■

6.2.2 Determining STLC in a finite number of differentiations

In [26] (see also [1]), the following problem was posed.

Open problem: If the smooth control-affine system $\Sigma = (\mathsf{M}, \{X_0, X_1, \dots, X_m\}, U)$ is STLC from x_0, does there exist an integer k such that every smooth control-affine system $\Sigma' = (\mathsf{M}, \{Y_0, Y_1, \dots, Y_m\}, U)$ is also STLC from x_0 if the Taylor expansions at x_0 of the vector fields of the two systems agree up to order k?

Theorem 6.13 gives a positive answer to this question for the special case of time-homogeneous systems.

Theorem 6.19 *Suppose that* $\Sigma = (\mathbb{R}^n, \{X_0, X_1, \ldots, X_m\}, U)$ *is time-homogeneous with respect to the dilation* $\Delta_s^{\mathrm{tm}}(x) = (s^{k_1}x^1, \ldots, s^{k_n}x^n)$ *and let* $k = \max\{k_1, \ldots, k_n\} - 1$*. If* Σ *STLC from* $x_0 = 0$*, then every control-affine system* $\Sigma' = (\mathbb{R}^n, \{Y_0, Y_1, \ldots, Y_m\}, U)$ *with* $j_{x_0}^k Y_a = j_{x_0}^k X_a$*, for all* $a \in \{0, 1, \ldots, m\}$*, is also STLC from* x_0*.*

Proof: If Σ is STLC from x_0, by Theorem 6.13, $\mathcal{V}_{x_0}^{k_1}\Sigma + \mathcal{V}_{x_0}^{k_2}\Sigma + \cdots + \mathcal{V}_{x_0}^{k_n}\Sigma = \mathbb{R}^n$. By definition, $\mathcal{V}_{x_0}^{\ell}\Sigma$ depends only on the $(\ell - 1)$-jets of $X_0, X_1, \ldots, X_m$ at x_0. Hence, if $\Sigma' = (\mathbb{R}^n, \{Y_0, Y_1, \ldots, Y_m\}, U)$ is a control-affine system such that $j_{x_0}^k Y_a = j_{x_0}^k X_a$, then $\mathcal{V}_{x_0}^{k_j}\Sigma = \mathcal{V}_{x_0}^{k_j}\Sigma'$ for all $j \in \{1, \ldots, n\}$. Hence, $\mathcal{V}_{x_0}^{k_1}\Sigma' + \mathcal{V}_{x_0}^{k_2}\Sigma' + \cdots + \mathcal{V}_{x_0}^{k_n}\Sigma' = \mathbb{R}^n$. Consequently, by Theorem 5.5, Σ' is STLC from x_0. ■

Chapter 7

Conclusions and future work

7.1 Conclusions

In this thesis, we have developed a feedback invariant theory of local controllability for affine distributions. The main geometric notion we have studied is what we call proper small-time local controllability and the main tool used to study this notion is a set of high-order tangent vectors. To better understand these high-order tangent vectors, some computational tools were developed on appropriate jet spaces of the tangent bundle. Using these tools we were able to characterize proper small-time local controllability for driftless and homogeneous systems.

7.2 Future work

The following list of questions and problems are natural avenues of future research from this point.

1. In Chapter 3 it was shown that, for a regular affine distribution, there is no loss of generality by considering convex control-affine systems for the study of PSTLC. A natural question is whether this is still true in the singular case.

2. It would be fruitful to better understand the algebraic properties of the linear map $\mathscr{T}^k_{x_0}$ described in Chapter 4. Having established a solid understanding of this map, it should be possible to obtain new interesting sufficient conditions for local controllability. The connection between the coefficients of the Taylor series of a composition of flows and labeled rooted trees, established in Section 4.3, might be useful for this task.

3. Using the tools developed in this thesis, give a sufficient condition for $\mathcal{V}^k_{x_0}\mathcal{A}$ to be a subspace. In other words, when are all variations of order k neutralizable?

4. For homogeneous systems, it was shown that the variational cone completely characterizes local controllability. It would be natural to explore what new necessary conditions can be obtained in the general case.

5. In Section 6.2.1, a special class of homogeneous systems was considered and we were able to give a geometric sufficient condition for the variational cone to be the whole tangent space. This was done by considering a specific type of concatenation that resulted in the high-order vector-valued form appearing as the variation. A natural generalization of this result would be to do the same for general homogeneous systems.

6. In [24] it was shown that, for the control-affine system $\Sigma = (\mathbb{R}^4, \{X_0, X_1\}, [-\epsilon, \epsilon])$ given by

$$\dot{x}^1 = u; \qquad \dot{x}^2 = x^1; \qquad \dot{x}^3 = (x^1)^3; \qquad \dot{x}^4 = (x^3)^2 - (x^2)^7,$$

if the set of controls is restricted to piecewise constant controls with at most N jumps, then, if $T > 0$ and $\epsilon > 0$ satisfy $N^7 T^{7/2} \epsilon^{3/4} < 1$, then $x^4(T) \geq 0$ if $x^1(T) = 0$. However, Kawski proved that Σ was indeed STLC from the origin $x_0 = 0$ by using variations that were parameterized by a discrete parameter related to the number of switchings in a specific type of variation. The number of switchings grew as the final time tended to zero. Roughly speaking, one can say that Kawski's example satisfies the following behaviour: If the control set is bounded then one needs high frequency to control the system in small-time. One natural question is: If the control set is unbounded, can Kawski's example be controlled using finite jumps in small-time? The tools developed in this thesis can be used to give a positive answer to this question. To show this, for $u \in \mathbb{R}$ let $\xi_u = X_0 + uX_1$. For Σ, it is easy to produce variations in the directions $\pm\frac{\partial}{\partial x^1}, \pm\frac{\partial}{\partial x^2}, \pm\frac{\partial}{\partial x^3}$, and $\frac{\partial}{\partial x^4}$, so that $-\frac{\partial}{\partial x^4}$ is the only missing direction for local controllability. For $s > 0$, let $\boldsymbol{\xi}(s) = (\xi_{u(s)}, \xi_{-u(s)}, \xi_{-u(s)}, \xi_{u(s)})$ where $u(s) = \frac{f(s)}{s^7}$ and the function f is to be determined. Let $\boldsymbol{\tau}(s) = (s, s, s, s)$. We can compute directly, by either integrating the differential equation or using the Taylor series tools developed in Chapter 4, that

$$\Phi_{x_0}^{\boldsymbol{\xi}(s)}(\boldsymbol{\tau}(s)) = \frac{(c_1 s - c_2 f(s))}{c_3 s^{34}} f(s)^6 e_4$$

for constants $c_1, c_2, c_3 > 0$, and where e_4 denotes the 4th standard basis vector in $\mathbb{R}^4$. Let

$$f(s) = \frac{c_1 s}{c_2} + \lambda s^{34},$$

for $\lambda \in \mathbb{R}$. Then direct substitution and simplification yields

$$\Phi_{x_0}^{\boldsymbol{\xi}(s)}(\boldsymbol{\tau}(s)) = -\frac{\lambda s^6 (c_1 + \lambda c_2 s^{33})^6}{c_3 c_2^5} e_4.$$

Hence, we get a *smooth* curve in the reachable of Σ from x_0 approaching the origin and in the $-e_4$ direction provided that $\lambda > 0$, i.e.,

$$\Phi_{x_0}^{\boldsymbol{\xi}(s)}(\boldsymbol{\tau}(s)) \in \mathcal{R}_\Sigma(x_0, 4s) \cap \left\{ x \in \mathbb{R}^n \mid \; x^4 < 0 \right\}.$$

The resulting parameterized control $u(s)$ is given by

$$u(s) = \frac{f(s)}{s^7} = \frac{c_1 s + \lambda c_2 s^{34}}{c_2 s^7} = \frac{c_1 + \lambda c_2 s^{33}}{c_2 s^6}, \tag{7.1}$$

which goes to infinity as $s \to 0$. Using a more general notion of variation than the one used in this thesis, see for example [25], this shows that, if one replaces the control set $U = [-\epsilon, \epsilon]$ with $U = \mathbb{R}$, then Σ is STLC from x_0 using piecewise constant controls with a finite number of switchings. In view of Kawski's example, a natural line of future research would be to build a theory, in the same spirit as was done in this thesis, to deal with this fast-switching phenomenon.

Bibliography

[1] A. Agrachev, *Is it possible to recognize local controllability in a finite number of differentiations?*, in Open Problems in Mathematical Systems and Control Theory, V. Blondel, E. Sontag, M. Vidyasagar, and J. Willems, eds., Springer-Verlag, 1999, pp. 15–18.

[2] A. Agrachev and R. Gamkrelidze, *Local controllability and semigroups of diffeomorphisms*, Acta Applicandae Mathematicae, 32 (1993), pp. 1–57.

[3] A. Agrachev and Y. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer-Verlag, 2004.

[4] J.-P. Aubin and A. Cellina, *Differential Inclusions: Set-Valued Maps and Viability Theory*, Springer-Verlag, 1984.

[5] J. Basto-Gonçalves, *Sufficient conditions for local controllability with unbounded controls*, SIAM Journal on Control and Optimization, 25 (1987), pp. 1371–1378.

[6] R. Bianchini, *Variations of a control process at the initial point*, Journal of Optimization Theory and Applications, 81 (1994), pp. 249–258.

[7] R. Bianchini and G. Stefani, *Normal local controllability of order one*, International Journal of Control, 39 (1984), pp. 701–714.

[8] ——, *Graded approximations and controllability along a trajectory*, SIAM Journal on Control and Optimization, 28 (1990), pp. 903–924.

[9] ——, *Controllability along a trajectory: A variational approach*, SIAM Journal on Control and Optimization, 31 (1993), pp. 900–927.

[10] R. Bianchini and P. Zecca, *Local controllability for autonomous nonlinear systems*, Journal of Optimization Theory and Applications, 31 (1980), pp. 69–83.

[11] N. Bourbaki, *Algebra II*, Springer-Verlag, 1990.

[12] F. Bullo and A. D. Lewis, *Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Simple Mechanical Control Systems*, no. 49 in Texts in Applied Mathematics, Springer-Verlag, 2004.

[13] J. Butcher, *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and general linear methods*, John Wiley & Sons, 1987.

[14] W. Chow, *Über systeme von linearen partiellen differentialgleichungen erster ordnung*, Mathematische Annalen, 117 (1939), pp. 98–105.

[15] R. Diestel, *Graph Theory*, Springer, 2nd ed., 2000.

[16] V. Elkin, *Reduction of Nonlinear Control Systems: A Differential Geometric Approach*, Kluwer Academic Publishers, 1999.

[17] K. Grasse, *On the relation between small-time local controllability and normal self-reachability*, Mathematics of Control, Signals, and Systems, 5 (1992), pp. 41–66.

[18] H. Hermes, *Local controllability and sufficient conditions in singular problems*, Journal of Differential Equations, 20 (1976), pp. 213–232.

[19] ——, *Lie algebras of vector fields and local approximations of attainable sets*, SIAM Journal on Control and Optimization, 16 (1978), pp. 715–727.

[20] ——, *Control systems which generate decomposable Lie algebras*, Journal of Differential Equations, 44 (1982), pp. 166–187.

[21] ——, *Nilpotent and high-order approximations of vector fields systems*, SIAM Review, 33 (1991), pp. 238–264.

[22] R. Hirschorn and A. Lewis, *High-order variations for families of vector fields*, SIAM Journal on Control and Optimization, 43 (2004), pp. 301–324.

[23] N. Jacobson, *Lie Algebras*, Dover Publications, Inc., New York, 1962.

[24] M. Kawski, *Control variations with an increasing number of switchings*, Bulletin of the American Mathematical Society (New Series), 18 (1988), pp. 149–152.

[25] M. Kawski, *High-order small-time local controllability*, in Nonlinear Controllability and Optimal Control, H. Sussmann, ed., Marcel Dekker, Inc., 1990, pp. 431–467.

[26] ——, *On the problem whether controllability is finitely determined*, in 17th International Symposium on Mathematical Theory of Networks and Systems, 2006.

[27] I. Kolář, P. W. Michor, and J. Slovák, *Natural Operations in Differential Geometry*, Springer-Verlag, 1993.

[28] M. I. Krastanov, *A sufficient condition for small-time local controllability*, SIAM Journal on Control and Optimization, 48 (2009), pp. 2296–2322.

[29] A. Krener, *A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems*, SIAM J. Control, 12 (1974), pp. 43–51.

[30] ——, *The high order maximal principle and its application to singular extremals*, SIAM Journal on Control and Optimization, 15 (1977), pp. 256–293.

[31] C. Lobry, *Contrôlabilité des systèmes non-linéaires*, SIAM Journal on Control and Optimization, 8 (1970), pp. 573–605.

[32] ——, *Sur l'ensemble des points atteignables par les solutions d'une equation differentielle multivoque*, Publ. Math. Bordeaux, 1 (1973), pp. 43–60.

[33] J. R. Munkres, *Topology*, Prentice Hall, 2 ed., 2000.

[34] T. Nagano, *Linear differential systems with singularities and an application to transitive Lie algebras*, Journal of the Mathematical Society of Japan, 18 (1966), pp. 398–404.

[35] H. Nijmeijer and A. van der Shaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, 1990.

[36] P. Rashevski, *About connecting two points of complete nonholonomic space by admissible curve*, Uchen. Zap. Ped. Inst. Libknexta, 2 (1938), pp. 83–94.

[37] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, 1998.

[38] D. Saunders, *The Geometry of Jet Bundles*, Cambridge University Press, 1989.

[39] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, no. 6 in Texts in Applied Mathematics, Springer-Verlag, New York, NY, 1998.

[40] G. Stefani, *Local controllability of nonlinear systems: An example*, Systems & Control Letters, 6 (1985), pp. 123–125.

[41] R. Strichartz, *The Campbell–Baker–Hausdorff–Dynkin formula and solutions of differential equations*, Journal of Functional Analysis, 72 (1987), pp. 320–345.

[42] H. Sussmann, *An extension of a theorem of Nagano on transitive Lie algebras*, Proceedings of the American Mathematical Society, 45 (1974), pp. 349–356.

[43] ——, *A sufficient condition for local controllability*, SIAM Journal on Control and Optimization, 16 (1978), pp. 790–802.

[44] ——, *Lie brackets and local controllability: A sufficient condition for scalar-input systems*, SIAM Journal on Control and Optimization, 21 (1983), pp. 686–713.

[45] ——, *A general theorem on local controllability*, SIAM Journal on Control and Optimization, 25 (1987), pp. 158–194.

[46] H. Sussmann and V. Jurdjevic, *Controllability of nonlinear systems*, Journal of Differential Equations, 12 (1972), pp. 95–116.