

Geometric local controllability: second-order conditions*

Ron M. Hirschorn[†] Andrew D. Lewis[‡]

30/06/2002

Last updated: 06/03/2004

Abstract

In a geometric point of view, a nonlinear control system, affine in the controls, is thought of as an affine subbundle of the tangent bundle of the state space. In deriving conditions for local controllability from this point of view, one should describe those properties of the affine subbundle that either ensure or prohibit local controllability. In this paper, second-order conditions of this nature are provided. The techniques involve a fusion of well-established analytical methods with differential geometric ideas.

Contents

1. Introduction	1
1.1 Context and contribution.	1
1.2 A motivating example.	2
1.3 Organisation.	3
2. Basic notation and vector-valued quadratic forms	3
2.1 Basics.	3
2.2 Vector-valued quadratic forms.	3
3. Affine systems	4
3.1 Basic definitions.	4
3.2 Control-affine systems and local control-affine realisations.	5
3.3 Subbundles associated with affine subbundles.	5
3.4 Linear maps associated with affine subbundles.	6
3.5 Vector-valued quadratic forms associated with affine subbundles.	7
4. Statement of geometric results	10
4.1 Existing low-order conditions.	10
4.2 Statement of results for affine systems.	10

*Preprint

[†]PROFESSOR, DEPARTMENT OF MATHEMATICS AND STATISTICS, QUEEN'S UNIVERSITY, KINGSTON, ON K7L 3N6, CANADA

Email: ron@mast.queensu.ca, URL: <http://www.mast.queensu.ca/~ron/>

Research supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

[‡]ASSISTANT PROFESSOR, DEPARTMENT OF MATHEMATICS AND STATISTICS, QUEEN'S UNIVERSITY, KINGSTON, ON K7L 3N6, CANADA

Email: andrew@mast.queensu.ca, URL: <http://www.mast.queensu.ca/~andrew/>

Research supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

5. Statement of results for control-affine systems	11
5.1 Some control-affine system notation.	11
5.2 Statement of results for control-affine systems.	11
6. Examples	12
6.1 Discussion.	12
6.2 The examples.	12
7. Proofs for the control-affine system results	17
7.1 Proof of Theorem 5.2.	17
7.2 Proof of Theorem 5.4.	21
8. Proofs of geometric results for affine systems	26
8.1 Proof of Theorem 4.3.	27
8.2 Proof of Theorem 4.4.	27

1. Introduction

The property of local controllability is one of the most fundamental for a control system. Much effort has been devoted to the determining of useful conditions for ascertaining local controllability of nonlinear systems. A sampling of the extensive literature includes the papers [Basto-Gonçalves 1985, 1998, Bianchini and Stefani 1984, 1993, Hermes 1982, Hermes and Kawski 1987, Kawski 1987, 1992, Knobloch and Wagner 1984, Petrov 1977, Stefani 1986, Sussmann 1978, 1983, 1987, Sussmann and Jurdjevic 1972]. Despite this significant effort, the precise nature of the mechanisms governing local controllability is imperfectly understood. Indeed, some of the most basic properties of the reachable set remain unresolved [Agrachev 1999]. What's more, Sontag [1988] and Kawski [1990] show that general conditions for determining the local controllability of a nonlinear system are likely to be computationally difficult. Nonetheless, the problem is so fundamental that it is of continuing interest.

1.1. Context and contribution. The emphasis in this paper is in coming to grips with some of the basic structural properties of local controllability, rather than providing a broad class of, say, sufficient conditions for local controllability; the latter approach is amply represented in the current literature. The machinery for studying local controllability typically revolves around the selection of suitable control variations which give rise to the so-called variational cone. This point of view is well presented by Bianchini and Stefani [1993]. A crucial ingredient in using control variations to determine sufficient conditions for local controllability is the introduction of a suitable method for approximating a system. Perhaps the most popular approximation technique is the “nilpotent approximation,” studied systematically in, for example, [Hermes 1991, Sussmann 1987]. The application of these techniques has been quite successful in the development of myriad sufficient conditions for local controllability, and, particularly for single-input systems, necessary conditions [Stefani 1986, Sussmann 1983]. Much of the effort in this work is focused on the analytical aspects of the problem.

In this paper, the intent is to focus on the geometric, rather than analytical, features of local controllability, and in this way understand the structure of local controllability. Specifically, for a control-affine system given by

$$\dot{\xi}(t) = f_0(\xi(t)) + \sum_{a=1}^m u_a(t) f_a(\xi(t)), \quad \xi(t) \in M,$$

we are interested in the character of the affine subbundle of TM whose fibre at $x \in M$ is

$$\left\{ f_0(x) + \sum_{a=1}^m u_a f_a(x) \mid u \in \mathbb{R}^m \right\}.$$

While the techniques we use to obtain our results rely on the analytical approach described above, the objective in the end is to describe those properties of an affine subbundle that concern the local controllability of systems whose drift vector field f_0 and control vector fields f_1, \dots, f_m generate that subbundle. Thus we give results of two sorts.

1. Theorems 4.3 and 4.4 have as hypotheses second-order conditions on an affine subbundle. One result (Theorem 4.3) gives conditions that ensure that control-affine systems giving rise to this affine subbundle will be locally controllable, and the other (Theorem 4.4) gives conditions that prohibit control-affine systems giving rise to this affine subbundle from being locally controllable. In essence, what we provide are two second-order characterisations of the “shape” of an affine subbundle. The statements and proofs of these theorems contain the main geometric contribution of the paper.
2. Theorems 5.2 and 5.4 provide results on local controllability in the usual manner, in that they involve conditions on the drift vector field f_0 and the control vector fields f_1, \dots, f_m . These results are novel in their own right as they provide a geometric description, delivered in terms of vector-valued quadratic forms, of the second-order nature of local controllability. The proofs of these new conditions for local controllability contain the majority of the analytical contribution of the paper.

Certainly the key geometric idea in the statement of our results is the employment of a vector-valued quadratic form to phrase our hypotheses. This idea seems to have been explicitly exploited in local controllability first by Basto-Gonçalves [1998], although the authors independently uncovered similar structure in mechanical systems [Hirschorn and Lewis 2001]. As we shall see, a certain property of the vector-valued quadratic form we present is equivalent to there being a basis of the input vector fields so that the second-order obstructions to local controllability, the “bad” brackets of Sussmann [1987], are neutralised. In optimal control, vector-valued quadratic forms have been used by Agrachev [1990] in the description of second-order necessary conditions for optimality. The task of finding second-order necessary conditions for optimality is related to finding second-order sufficient conditions for controllability, so it is unsurprising that vector-valued quadratic forms would also come up in our work. Vector-valued quadratic forms come up in other places in control theory. A review of this and a list of references may be found in the paper of Bullo, Cortés, Lewis, and Martínez [2004].

1.2. A motivating example. In this section we provide motivation for our results by considering an example. This example will reappear in the paper at various points as a tool for illustrating some of the techniques employed. In order to provide some context for the example, we briefly consider some general constructions, referring to Section 3.1 for precise definitions. For an affine subbundle $\mathcal{A} \subset TM$, an **affine system** \mathcal{A} assigns to each $x \in M$ a subset $\mathcal{A}(x) \subset \mathcal{A}_x$. The assignment should have some regularity properties that shall not concern us for now. A **trajectory** for \mathcal{A} is an absolutely continuous curve $\xi: [0, T] \rightarrow M$ satisfying $\dot{\xi}(t) \in \mathcal{A}(\xi(t))$ for each $t \in [0, T]$. Roughly speaking, \mathcal{A} is **small-time locally controllable (STLC)** from $x_0 \in M$ if the set of points that can be reached from x_0 by trajectories of \mathcal{A} contains x_0 in its interior.

The example we consider is really a collection of examples, as there are some free parameters. We take as our state space $M = \mathbb{R}^4$ with coordinates (x_1, x_2, x_3, x_4) . On M we define an affine subbundle \mathcal{A} by asking that its fibre at (x_1, x_2, x_3, x_4) be given by

$$\left\{ (0, 0, 0, x_2^2 + \frac{1}{2}ax_1^2 + bx_1x_2 + \frac{1}{2}cx_2^2) \right\} + \text{span}_{\mathbb{R}} \left\{ (0, 0, 0, dx_1^2), (1, 0, 0, 0), (0, 1, x_1, 0) \right\}.$$

One wishes to determine conditions on the real parameters a , b , c , and d so that an affine system \mathcal{A} defined using \mathcal{A} is STLC from $x_0 = (0, 0, 0, 0)$. Our results provide the following information about \mathcal{A} .

1. When $d = 0$ and when $b^2 - ac > 0$, any affine system \mathcal{A} defined using \mathcal{A} is STLC from x_0 (Theorem 4.3).
2. When $d = 0$ and $b^2 - ac < 0$, any affine system \mathcal{A} defined using \mathcal{A} is not STLC from x_0 (Theorem 4.4).

In the case that $d = 0$ and $b^2 - ac = 0$, the results in our paper do not predict whether an affine system \mathcal{A} defined using \mathcal{A} will or will not be STLC from x_0 , although this example is simple enough that one can ascertain this “by hand.”

The conditions $b^2 - ac > 0$ and $b^2 - ac < 0$ are those that ensure that the matrix $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ will have eigenvalues of mixed sign and equal sign, respectively. In our geometric formulation, these sorts of conditions will arise in terms of quadratic forms taking values in a certain vector space. For this example, the vector space in which the quadratic forms take their value is one-dimensional. We have intentionally presented this example in such a way that its controllability properties, at least when $d = 0$, will be clear to anyone who considers the problem for a few moments. Of course, what one wants is a systematic way of arriving at conditions that capture the essence of this example.

The case left yet unconsidered is an interesting one, that when $d \neq 0$. In this case, Theorem 4.4 in no longer able to predict when the system will not be STLC from x_0 . Indeed, it is no longer true that any affine system \mathcal{A} defined using \mathcal{A} will or will not be STLC from x_0 . Instead, the controllability of such an affine system \mathcal{A} will depend on the exact nature of \mathcal{A} at points away from x_0 . More precisely, and perhaps a little surprisingly, the following circumstance can arise.

There are two affine systems \mathcal{A}_1 and \mathcal{A}_2 defined using \mathcal{A} and satisfying (1) $\mathcal{A}_1(x_0) = \mathcal{A}_2(x_0)$, (2) \mathcal{A}_1 is STLC from x_0 , and (3) \mathcal{A}_2 is not STLC from x_0 .

An interesting question is “What distinguishes the case when $d \neq 0$ from the case when $d = 0$?” The answer we give involves a certain linear map that is zero when $d = 0$, and nonzero otherwise. This linear map is constructed both in general terms and for our example in Section 3.5.

One of the interesting features of certain of the statements we make above is that they are independent of the particular affine system \mathcal{A} , and depend only on the affine subbundle \mathcal{A} . This idea of determining the controllability of a system based on statements about the affine subbundle that it defines is one that has received no attention in the existing literature on nonlinear controllability. This paper takes the first steps towards identifying techniques that may be useful in the addressing of these fundamental geometric notions.

1.3. Organisation. The layout of the paper is as follows. In Section 2 is presented the background and notation. The formal definitions and properties of affine systems are given in Section 3. In Sections 3.4 and 3.5 are made some important geometric constructions associated with an affine subbundle that intersects the zero section in TM . At this point in the paper, it is possible to state the main results. This is done in Section 4.2, after a presentation in Section 4.1, using our terminology and notation, of some known zeroth and first-order conditions for local controllability. To prove our main results, Theorems 4.3 and 4.4, we rely on new results for “standard” control-affine systems that we prove using well-known analytical methods. Since these results are themselves of interest to control theoreticians, they are presented separately in Section 5. The proofs collectively requires some effort, and so are postponed to the end of the paper. We choose to immediately follow the statement of results with a collection of examples in Section 6. These examples, unlike the example of Section 1.2 above, are chosen to illustrate that the “obvious” attempts to weaken the hypotheses of our main results do not succeed. The examples given amply illustrate the subtle character of local controllability. In some of the examples, local controllability is demonstrated using the machinery of [Hirschorn and Lewis 2002]. The final two sections of the paper contain the proofs of the main results, with Theorems 5.2 and 5.4 being proved in Section 7, and Theorems 4.3 and 4.4 being proved in Section 8. While the proofs of Section 7 involve the standard tools for the study of control-affine systems, the proofs of Section 8 are geometric in nature, and follow in part from the development in Sections 3.4 and 3.5.

2. Basic notation and vector-valued quadratic forms

In this section we first provide our basic notation. An important element in the statement of our main results is a pair of vector-valued quadratic forms. In Section 2.2 we provide some generalities associated with such objects.

2.1. Basics. For a set S , 2^S denotes the set of subsets of S .

If V is a \mathbb{R} -vector space with V^* its dual, $\text{ann}(S) \subset V^*$ denotes the annihilator of $S \subset V$. $\text{End}(V)$ denotes the endomorphisms of V . If $\mathcal{L} \subset \text{End}(V)$ and if $W \subset V$ is a subspace, then $\langle \mathcal{L}, W \rangle \subset V$ denotes the smallest subspace of V containing W and which is an invariant subspace for each $L \in \mathcal{L}$. Thus $\langle \mathcal{L}, W \rangle$ is generated by those vectors of the form

$$L_1 \circ \cdots \circ L_k(w), \quad k \in \mathbb{N} \cup \{0\}, \quad L_1, \dots, L_k \in \mathcal{L}, \quad w \in W.$$

Again, let V be a \mathbb{R} -vector space. If $S \subset V$ then $\text{conv}(S) \subset V$ denotes the convex hull of S and $\text{aff}(S) \subset V$ denotes the affine hull of S .

If X is a topological space with subsets $S \subset A \subset X$, then $\text{int}_A(S)$ denotes the interior of S relative to the topology on A induced from X . For $A \subset X$, $\text{cl}(A)$ is the closure of A .

If M is a real-analytic manifold, $0_x \in T_x M$ denotes the zero vector in the tangent space. We always assume objects are real-analytic unless otherwise stated. A subset $E \subset TM$ having the property that $E \cap T_x M \neq \emptyset$ for each $x \in M$ is called a **fibred subset** of TM . We denote by $\Gamma(E)$ the set of real-analytic vector fields on M taking values in the fibred subset E . In particular, $\Gamma(TM)$ denotes the set of real-analytic vector fields on M . Although it will generally be the case that $\Gamma(E) = \emptyset$, the fibred subsets we consider are sufficiently nice that they admit real-analytic sections. The real-analytic functions on M are denoted $\mathcal{F}(M)$. For a vector field X , $\mathbb{R} \times M \ni (t, x) \mapsto e^{tX}(x) \in M$ denotes the flow of X . If $\phi: M \rightarrow N$ is a differentiable map, we denote its derivative by $T\phi: TM \rightarrow TN$, and we denote $T_x \phi = T\phi|_{T_x M}$.

To test controllability, one typically looks at Lie brackets of vector fields. We shall adopt this approach as well, and next we present the notation we use for carrying out such discussions. Let $\mathcal{V} \subset \Gamma(TM)$ be a family of vector fields. For $k \in \mathbb{N}$ let $\mathcal{L}^{(k)}(\mathcal{V})$ be the $\mathcal{F}(M)$ -submodule of $\Gamma(TM)$ generated by those vector fields that are brackets of vector fields from \mathcal{V} of degree up to k . Thus, for example, $\mathcal{L}^{(1)}(\mathcal{V}) = \text{span}_{\mathcal{F}(M)}\{\mathcal{V}\}$. Corresponding to these are the distributions on M given by

$$L^{(k)}(\mathcal{V}) = \{X(x) \mid X \in \mathcal{L}^{(k)}(\mathcal{V}), \quad x \in M\}, \quad k \in \mathbb{N}.$$

With this notation, it is reasonable to denote by $\mathcal{L}^{(\infty)}(\mathcal{V})$ the Lie subalgebra of $\Gamma(TM)$ generated by \mathcal{V} , and by

$$L^{(\infty)}(\mathcal{V}) = \{X(x) \mid X \in \mathcal{L}^{(\infty)}(\mathcal{V}), \quad x \in M\}$$

the corresponding distribution. For none of these distributions do we *a priori* make any assumptions concerning their regularity.

We shall find occasion to make an abuse of the above notation. This will arise when we have a fibred subset $E \subset TM$, and we consider the subset $\Gamma(E)$ of $\Gamma(TM)$ comprised of the E -valued sections. In this case, for brevity we will write $L^{(k)}(E) = L^{(k)}(\Gamma(E))$, $k \in \mathbb{N} \cup \{\infty\}$.

Continue with \mathcal{V} a family of vector fields on M and let $x_0 \in M$. Let $Z_{x_0}(\mathcal{V})$ be those vector fields in \mathcal{V} which vanish at x_0 . If $X \in Z_{x_0}(\mathcal{V})$ and $Y \in \Gamma(TM)$ then $[X, Y](x_0)$ depends only on the value of Y at x_0 , and not on any derivatives of Y at x_0 . Therefore, the map $v \mapsto \text{ad}_X Y(x_0)$ is \mathbb{R} -linear if Y is any vector field extending $v \in T_{x_0} M$. For notational brevity, we will make an abuse of notation like that mentioned above for families of vector fields generated by Lie brackets. That is, if $E \subset TM$ is a fibred subset, then we will write $Z_{x_0}(E)$ for $Z_{x_0}(\Gamma(E))$. We shall also switch freely between regarding $Z_{x_0}(\mathcal{V})$ as a subset of $\Gamma(TM)$ and as a subset of $\text{End}(T_{x_0} M)$ (even though these subsets are not in one-to-one correspondence). Generally, when we wish to think of $X \in Z_{x_0}(\mathcal{V})$ as being in $\text{End}(T_{x_0} M)$, we shall write ad_X .

2.2. Vector-valued quadratic forms. An essential ingredient in our development is a vector-valued quadratic form. Since such objects are not entirely well-understood, at least

in the sense that we shall use them, in this section we present an overview of their salient properties.

Let V and W be finite-dimensional \mathbb{R} -vector spaces and let $\Sigma_2(V; W)$ denote the set of symmetric \mathbb{R} -bilinear maps from $V \times V$ to W . For $B \in \Sigma_2(V; W)$ we define $Q_B: V \rightarrow W$ by $Q_B(v) = B(v, v)$. For $\lambda \in W^*$ we define $\lambda B: V \times V \rightarrow \mathbb{R}$ by $\lambda B(v_1, v_2) = \langle \lambda; B(v_1, v_2) \rangle$.

2.1 DEFINITION: Let $B \in \Sigma_2(V; W)$.

- (i) B is **definite** if there exists $\lambda \in W^*$ so that λB is positive-definite.
- (ii) B is **indefinite** if for each $\lambda \in W^*$, λB is neither positive nor negative-semidefinite.
- (iii) If $W = \{0\}$ then B is declared to be indefinite. \square

The following properties of symmetric bilinear maps will be important for us.

2.2 LEMMA: Let V and W be finite-dimensional \mathbb{R} -vector spaces with $B \in \Sigma_2(V; W)$ nonzero. The following statements hold:

- (i) B is indefinite if and only if

$$0 \in \text{int}_{\text{aff}(\text{image}(Q_B))}(\text{conv}(\text{image}(Q_B))); \quad (2.1)$$

- (ii) B is definite if and only if there exists a hyperplane P through $0 \in W$ so that

- (a) $\text{image}(Q_B)$ lies on one side of P and
- (b) $\text{image}(Q_B) \cap P = \{0\}$.

Proof: First note that (2.1) holds if and only if there is no hyperplane $P \subset W$ through the origin for which $\text{image}(Q_B)$ lies on one side of the hyperplane. Equivalently, (2.1) holds if and only if there is no $\lambda \in W^*$ so that $\langle \lambda; Q_B(v) \rangle \geq 0$ for every $v \in V$.

Now assume that (2.1) does not hold. Then there exists $\lambda \in W^*$ so that $\langle \lambda; Q_B(v, v) \rangle \geq 0$ for every $v \in V$ and so that there is at least one $v \in V$ so that $\langle \lambda; Q_B(v, v) \rangle > 0$. Thus λB is positive-semidefinite, so B is not indefinite. Now assume that B is not indefinite. Then there exists $\lambda \in W^*$ so that λB is positive-semidefinite. Then $\langle \lambda; Q_B(v) \rangle \geq 0$ for all $v \in V$. This proves (i).

Suppose that there is a hyperplane P having properties (ii a) and (ii b). If λ has the property that $P = \ker(\lambda)$ then by changing the sign of λ if necessary, we have $\langle \lambda; Q_B(v) \rangle \geq 0$ for all $v \in V$ by property (ii a). Property (ii b) ensures that $\langle \lambda; Q_B(v) \rangle = 0$ only if $v = 0$. Thus λB is positive-definite. The preceding argument may be reversed to show that if λB is positive-definite, then $P = \ker(\lambda)$ has properties (ii a) and (ii b). \blacksquare

The question of whether it is possible to ascertain, in a computationally efficient manner, whether a given vector-valued symmetric bilinear map is indefinite is unresolved, to the best knowledge of the authors. This is given some consideration by Bullo, Cortés, Lewis, and Martínez [2004].

3. Affine systems

In this section we discuss affine systems in the general, geometric manner that characterises our approach. Basic definitions are given in Section 3.1. The usual notion of a

control-affine system is a special case of what we call an affine system, and this specialisation is given explicitly in Section 3.2. To state our results on local controllability of affine systems, we need some constructions that are outlined in Sections 3.3, 3.4, and 3.5. The results in the last two of these sections contain some of the more important geometric constructions of the paper.

3.1. Basic definitions. Let M be a finite-dimensional real-analytic manifold. By a **real-analytic generalised distribution** on M we shall mean a subset $\mathcal{D} \subset TM$ having the property that for each $x_0 \in M$ there exists a neighbourhood \mathcal{N} of x_0 and real-analytic vector fields $\{X_1, \dots, X_k\}$ on \mathcal{N} so that for each $x \in \mathcal{N}$ we have

$$\mathcal{D}_x \triangleq \mathcal{D} \cap T_x M = \text{span}_{\mathbb{R}} \{X_1(x), \dots, X_k(x)\}.$$

We call the vector fields $\{X_1, \dots, X_k\}$ **local generators** for \mathcal{D} . In like manner, a **real-analytic generalised affine subbundle** on M is a subset $\mathcal{A} \subset TM$ having the property that for each $x_0 \in M$ there exists a neighbourhood \mathcal{N} of x_0 and real-analytic vector fields $\{X_0, X_1, \dots, X_k\}$ on \mathcal{N} so that for each $x \in \mathcal{N}$ we have

$$\mathcal{A}_x \triangleq \mathcal{A} \cap T_x M = \{X_0(x)\} + \text{span}_{\mathbb{R}} \{X_1(x), \dots, X_k(x)\}.$$

The vector fields $\{X_0, X_1, \dots, X_k\}$ as above we call **local linear generators** of \mathcal{A} . If we additionally have

$$\mathcal{A}_x = \text{aff} \{X_0(x) + X_j(x) \mid j \in \{1, \dots, k\}\}, \quad x \in \mathcal{N},$$

then the vector fields $\{X_0, X_1, \dots, X_k\}$ are called **local affine generators** of \mathcal{A} . Corresponding to a real-analytic generalised affine subbundle is a generalised distribution $L(\mathcal{A})$ defined by asking that $L(\mathcal{A})_x$ is the linear part of the affine subspace \mathcal{A}_x .

3.1 DEFINITION: Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle. An **affine system in \mathcal{A}** is a map $\mathcal{A}: M \rightarrow 2^{TM}$ with the following properties:

- (i) $\mathcal{A}(x) \subset \mathcal{A}_x$ for each $x \in M$;
- (ii) $\text{aff}(\mathcal{A}(x)) = \mathcal{A}_x$ for each $x \in M$;
- (iii) \mathcal{A} is continuous;
- (iv) for each $x_0 \in M$ there exists local affine generators $\{X_0, X_1, \dots, X_k\}$ for \mathcal{A} defined on a neighbourhood \mathcal{N} of x_0 , and so that for each $x \in \mathcal{N}$ and $j \in \{1, \dots, k\}$, $X_0(x) + X_j(x) \in \text{conv}(\mathcal{A}(x))$.

An affine system \mathcal{A} in \mathcal{A} is **proper at x_0** if $0_{x_0} \in \text{int}(\text{conv}(\mathcal{A}(x_0)))$. \square

The notion of continuity we employ in property (iii) is the continuity of set-valued maps using the Hausdorff metric [see, e.g., Filippov 1988, page 65]. The property (iv) also bears some explanation. If the rank of the distribution $L(\mathcal{A})$ is not constant, then at points x_0 where the rank of $L(\mathcal{A})$ is less than maximal, we wish to ensure that the sets $\mathcal{A}(x)$ near x_0 grow sufficiently quickly in the directions transverse to $L(\mathcal{A}_{x_0})$ that we may fit real-analytic vector fields inside the sets $\text{conv}(\mathcal{A}(x))$. Other than the conditions imposed by the definition, we shall allow the sets $\mathcal{A}(x)$ to be quite general, only specifying their properties as needed.

A **trajectory** for an affine system \mathcal{A} in \mathcal{A} is an absolutely continuous curve $\xi: [0, T] \rightarrow M$ with the property that $\dot{\xi}(t) \in \mathcal{A}(\xi(t))$ for each $t \in [0, T]$. If, for example, $\mathcal{A}(x)$ is compact for each $x \in M$, the existence of trajectories is ensured [Filippov 1988, §2.7]. One might consider trajectories defined on more general intervals (say infinite intervals, or intervals not containing 0), but this will not be important for us. We denote by $\text{Traj}(\mathcal{A}, T)$ the set of trajectories defined on $[0, T]$ and $\text{Traj}(\mathcal{A}) = \bigcup_{T \geq 0} \text{Traj}(\mathcal{A}, T)$. For $x_0 \in M$ and $T \geq 0$ we define

$$\mathcal{R}_{\mathcal{A}}(x_0, T) = \{\xi(T) \mid \xi \in \text{Traj}(\mathcal{A}, T), \xi(0) = x_0\},$$

and similarly we define $\mathcal{R}_{\mathcal{A}}(x_0, \leq T) = \bigcup_{t \in [0, T]} \mathcal{R}_{\mathcal{A}}(x_0, t)$.

3.2 DEFINITION: Let \mathcal{A} be a real-analytic generalised affine subbundle of TM and let \mathcal{A} be an affine system in \mathcal{A} . If there exists $T > 0$ so that $x_0 \in \text{int}(\mathcal{R}_{\mathcal{A}}(x_0, t))$ for each $t \in]0, T]$, then \mathcal{A} is **small-time locally controllable (STLC)** from x_0 . \square

Some of the results we provide are of a nature that they hold for large classes of affine systems \mathcal{A} . In such a context, the following notions are useful.

3.3 DEFINITION: Let \mathcal{A} be a real-analytic generalised affine subbundle.

- (i) \mathcal{A} is **properly small-time locally controllable (properly STLC)** from x_0 if \mathcal{A} is STLC from x_0 for every affine system \mathcal{A} in \mathcal{A} that is proper at x_0 .
- (ii) \mathcal{A} is **small-time locally uncontrollable** from x_0 (**STLUC**) if \mathcal{A} is not STLC from x_0 for every affine system \mathcal{A} in \mathcal{A} with the property that $\mathcal{A}(x_0)$ is compact. \square

3.4 REMARKS: 1. The compactness condition on $\mathcal{A}(x_0)$ in the definition of STLUC ensures that one does not see behaviour unique to systems with unbounded controls. For example, if $L^{(\infty)}(L(\mathcal{A}))_{x_0} = T_{x_0}M$ and if $\mathcal{A}(x) = \mathcal{A}_x$ for each $x \in M$, then \mathcal{A} is STLC from x_0 [Bianchini and Stefani 1993, Pomet 1999]. However, if $\mathcal{A}(x)$ is bounded, then this is generally not true, as we shall see in Example 6.2.

- 2. This seems to be the first time that definitions of controllability in terms of affine subbundles have appeared in the literature. \square

3.2. Control-affine systems and local control-affine realisations. The notion of an affine system in the preceding section is not the one that is normally seen in the nonlinear control literature. Let us now present this usual notion, and see how we can go back and forth between it and our more geometric notion of an affine system.

3.5 DEFINITION: A **control-affine system** is a triple $\Sigma = (M, \mathcal{F}, U)$ where M is a finite-dimensional real-analytic manifold, $\mathcal{F} = \{f_0, f_1, \dots, f_m\}$ is a set of real-analytic vector fields on M , and $U \subset \mathbb{R}^m$. \square

To a control-affine system $\Sigma = (M, \mathcal{F}, U)$ we associate a real-analytic generalised affine subbundle $\mathcal{A}_{\mathcal{F}}$ with fibre at $x \in M$ given by

$$\mathcal{A}_{\mathcal{F}, x} = \left\{ f_0(x) + \sum_{a=1}^m u_a f_a(x) \mid u \in \mathbb{R}^m \right\}$$

and an affine system \mathcal{A}_{Σ} in $\mathcal{A}_{\mathcal{F}}$ given by

$$\mathcal{A}_{\Sigma}(x) = \left\{ f_0(x) + \sum_{a=1}^m u_a f_a(x) \mid u \in U \right\}.$$

Thus affine systems generalise control-affine systems in this way. However, there are affine systems that cannot be realised, even locally, as control-affine systems. (As an instance of this, let $M = \mathbb{R}$ with (x) the canonical coordinate, and let $\mathcal{A} = TM$. Let $\rho: M \rightarrow \mathbb{R}$ be a strictly positive-valued function on M that is continuous, but not differentiable. If we take $\mathcal{A}(x) = [-\rho(x), \rho(x)]$, then this affine system in \mathcal{A} cannot be represented as a control-affine system.) It will also be convenient to define

$$f_u = f_0 + \sum_{a=1}^m u_a f_a, \quad u \in U \quad \text{and} \quad \mathcal{F}_U = \{f_u \mid u \in U\} \subset \Gamma(TM). \quad (3.1)$$

Let us now give a method for retrieving a control-affine system from an affine system. Since the latter notion is truly more general than the former, there cannot be expected to be any sort of general equivalence between the control-affine system we construct and the affine system used to construct it.

3.6 DEFINITION: Let \mathcal{A} be a real-analytic generalised affine subbundle of TM with \mathcal{A} an affine system in \mathcal{A} . A **local control-affine realisation of \mathcal{A} at x_0** is a control-affine system $\Sigma = (N, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ with the properties

- (i) N is a neighbourhood of x_0 ;
- (ii) \mathcal{F} is a family of vector fields defined on N ;
- (iii) $\mathcal{A}_{\mathcal{F}} = \mathcal{A}|N$,
- (iv) $\text{aff}(U) = \mathbb{R}^m$, and
- (v) $\mathcal{A}(x_0) = \{f_u(x_0) \mid u \in U\}$. \square

Note that this realisation is made with respect to a given point x_0 , and will generally have nothing to do with \mathcal{A} away from x_0 , even locally. However, we shall see that it is possible to prove results about small-time local controllability of \mathcal{A} from x_0 by using a local control-affine realisation of \mathcal{A} at x_0 (Corollaries 4.6 and 4.8).

3.3. Subbundles associated with affine subbundles. Let \mathcal{A} be a real-analytic generalised affine subbundle. We will be interested in characterising the distributions $L^{(k)}(\Gamma(\mathcal{A}))$, $k \in \mathbb{N} \cup \{\infty\}$. As we stated in Section 2.1, we shall abbreviate $L^{(k)}(\Gamma(\mathcal{A}))$ with $L^{(k)}(\mathcal{A})$. The following result associates these and related distributions to the usual distributions one deals with in nonlinear control theory. Although the result is elementary, it is worth making explicit to avoid confusion in the sequel about the meaning of symbols.

3.7 PROPOSITION: Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle, and let $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ be a set of local linear generators for \mathcal{A} defined on a neighbourhood N of $x_0 \in M$. Also denote $\mathcal{X}_1 = \{X_1, \dots, X_k\}$. The following statements hold:

- (i) $L^{(\ell)}(\mathcal{A})|N = L^{(\ell)}(\mathcal{X})$ for each $\ell \in \mathbb{N} \cup \{\infty\}$;
- (ii) $L^{(\ell)}(L(\mathcal{A}))|N = L^{(\ell)}(\mathcal{X}_1)$ for each $\ell \in \mathbb{N} \cup \{\infty\}$.

Proof: For simplicity, and without loss of generality, let us suppose that $N = M$.

(i) Clearly $L^{(\ell)}(\mathcal{X}) \subset L^{(\ell)}(\mathcal{A})$, so it is the opposite inclusion we prove. We use the fact that if \mathcal{V} is a family of real-analytic vector fields, then $\mathcal{L}^{(\ell)}(\mathcal{V})$ is generated as a $\mathcal{F}(M)$ -module by the vector fields

$$[V_1, [V_2, \dots, [V_{\ell-1}, V_\ell]]], \quad V_1, \dots, V_\ell \in \mathcal{V}.$$

Since \mathcal{X} is a set of local linear generators, we have $L^{(1)}(\mathcal{A}) \subset L^{(1)}(\mathcal{X})$. Now suppose that $L^{(\ell)}(\mathcal{A}) \subset L^{(\ell)}(\mathcal{X})$ for some $\ell \in \mathbb{N}$. Let $V_1, \dots, V_\ell \in \Gamma(\mathcal{A})$ and write

$$V_s = X_0 + \sum_{j=1}^k \phi_{sj} X_j, \quad s \in \{1, \dots, \ell\},$$

where $\phi_{sj} \in \mathcal{F}(M)$, $s \in \{1, \dots, \ell\}$, $j \in \{1, \dots, k\}$. Also let $V \in \Gamma(\mathcal{A})$ and write

$$V = X_0 + \sum_{j=1}^k \phi_j X_j$$

for $\phi_1, \dots, \phi_k \in \mathcal{F}(M)$. Now we have

$$\begin{aligned} [V, [V_1, [V_2, \dots, [V_{\ell-1}, V_\ell]]]] &= [X_0, [V_1, [V_2, \dots, [V_{\ell-1}, V_\ell]]]] \\ &\quad + \sum_{j=1}^k \phi_j [X_j, [V_1, [V_2, \dots, [V_{\ell-1}, V_\ell]]]] + H_\ell, \end{aligned}$$

where $H_\ell \in \mathcal{L}^{(\ell)}(\mathcal{X})$. The result now follows from the induction hypothesis.

(ii) follows in exactly the same manner. \blacksquare

3.4. Linear maps associated with affine subbundles. Let \mathcal{A} be a real-analytic generalised affine subbundle with \mathcal{A} an affine system in \mathcal{A} . It will generally be the case that we consider local controllability of \mathcal{A} at a point $x_0 \in M$ for which $0_{x_0} \in \mathcal{A}_{x_0}$. Indeed, if this were not the case, then local controllability from x_0 is an impossibility (cf. Theorem 4.1). In this section we will consider some constructions arising from this assumption that $0_{x_0} \in \mathcal{A}_{x_0}$. Following our notation of Section 2.1, $\Gamma(\mathcal{A})$ denotes the set of sections of \mathcal{A} and $Z_{x_0}(\Gamma(\mathcal{A}))$ denotes those sections that vanish at x_0 ; the assumption that $0_{x_0} \in \mathcal{A}_{x_0}$ ensures that this latter set is nonempty. As promised in Section 2.1, we will adopt an abuse of notation by abbreviating $Z_{x_0}(\Gamma(\mathcal{A}))$ as $Z_{x_0}(\mathcal{A})$. For each $X \in Z_{x_0}(\mathcal{A})$, ad_X can be thought of as defining an endomorphism of $T_{x_0}M$. The following result gives a useful description of certain of the invariant subspaces of $Z_{x_0}(\mathcal{A})$.

3.8 PROPOSITION: *Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle with the property that for $x_0 \in M$, $0_{x_0} \in \mathcal{A}_{x_0}$. Also let $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ be a set of local linear generators for \mathcal{A} defined on a neighbourhood \mathcal{N} of x_0 , and define $\mathcal{X}_{\mathbb{R}^k}$ as per (3.1). If $S_{x_0} \subset T_{x_0}M$ is a subspace containing $L(\mathcal{A}_{x_0})$ then $\langle Z_{x_0}(\mathcal{A}), S_{x_0} \rangle = \langle Z_{x_0}(\mathcal{X}_{\mathbb{R}^k}), S_{x_0} \rangle$.*

Proof: We assume without loss of generality, since all computations occur in a neighbourhood of x_0 , that $N = M$. We begin by modifying the local linear generators \mathcal{X} . Since

$\{X_1(x_0), \dots, X_k(x_0)\}$ generate $L(\mathcal{A}_{x_0})$ and since $L(\mathcal{A}_{x_0}) = \mathcal{A}_{x_0}$, we may write

$$X_0(x_0) = \sum_{j=1}^k \bar{\alpha}_j X_j(x_0)$$

for some $\bar{\alpha} \in \mathbb{R}^k$. We then define

$$\tilde{X}_0 = X_0 - \sum_{j=1}^k \bar{\alpha}_j X_j$$

so that $\tilde{X}_0(x_0) = 0_{x_0}$. We may also rearrange the vector fields $\{X_1, \dots, X_k\}$ so that $\{X_{\ell+1}(x_0), \dots, X_k(x_0)\}$ form a basis for $L(\mathcal{A}_{x_0})$. We then let $\{v_1, \dots, v_\ell\} \subset \mathbb{R}^k$ be a basis for the kernel of the linear map

$$\mathbb{R}^k \ni \alpha \mapsto \sum_{j=1}^k \alpha_j X_j(x_0) \in L(\mathcal{A}_{x_0}).$$

Defining $R \in GL(k; \mathbb{R})$ by

$$R = [v_1 \mid \dots \mid v_\ell \mid e_{\ell+1} \mid \dots \mid e_k],$$

where $e_j \in \mathbb{R}^k$, $j \in \{1, \dots, k\}$, is the j th standard basis vector, and defining $\tilde{X}_1, \dots, \tilde{X}_k$ by

$$\tilde{X}_j = \sum_{s=1}^k R_{js}^s X_s, \quad j \in \{1, \dots, k\},$$

we see that $\tilde{X}_1(x_0) = \dots = \tilde{X}_\ell(x_0) = 0_{x_0}$, and that $\{\tilde{X}_{\ell+1}(x_0), \dots, \tilde{X}_k(x_0)\}$ is linearly independent. If we take $\tilde{\mathcal{X}} = \{\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_k\}$ then clearly, as subsets of $\Gamma(TM)$, we have $\mathcal{X}_{\mathbb{R}^k} = \tilde{\mathcal{X}}_{\mathbb{R}^k}$. Thus we also have the equality $Z_{x_0}(\mathcal{X}_{\mathbb{R}^k}) = Z_{x_0}(\tilde{\mathcal{X}}_{\mathbb{R}^k})$, as subsets of $\text{End}(T_{x_0}M)$.

Now let $S_{x_0} \subset T_{x_0}M$ be as stated in the proposition. Clearly we have $\langle Z_{x_0}(\tilde{\mathcal{X}}_{\mathbb{R}^k}), S_{x_0} \rangle \subset \langle Z_{x_0}(\mathcal{A}), S_{x_0} \rangle$, so it is the opposite inclusion we prove. Let $V \in Z_{x_0}(\mathcal{A})$ and write

$$V = \tilde{X}_0 + \sum_{j=1}^k \phi_j \tilde{X}_j$$

for real-analytic functions ϕ_1, \dots, ϕ_k that satisfy $\phi_{\ell+1}(x_0) = \dots = \phi_k(x_0) = 0$. As differential operators on $\Gamma(TM)$ we have

$$\text{ad}_V = \text{ad}_{\tilde{X}_0} + \sum_{j=1}^k \phi_j \text{ad}_{\tilde{X}_j} - \sum_{j=1}^k \tilde{X}_j \otimes d\phi_j.$$

This shows that as an element of $\text{End}(T_{x_0}M)$ we have $\text{ad}_V = L_1 + L_2$ where $L_1 \in Z_{x_0}(\tilde{\mathcal{X}}_{\mathbb{R}^k})$ and where $\text{image}(L_2) \subset L(\mathcal{A}_{x_0})$.

Finally, to prove the proposition, it suffices to show that every vector in $T_{x_0}M$ of the form

$$\text{ad}_{v_1} \circ \dots \circ \text{ad}_{v_\ell}(v), \quad \ell \in \mathbb{N} \cup \{0\}, \quad V_1, \dots, V_\ell \in Z_{x_0}(\mathcal{A}), \quad v \in S_{x_0},$$

lies in $\langle Z_{x_0}(\tilde{\mathcal{X}}), S_{x_0} \rangle$. This follows by an induction on ℓ , using the decomposition $\text{ad}_V = L_1 + L_2$ given above for $V \in Z_{x_0}(\mathcal{A})$, and using the fact that $L(\mathcal{A}_{x_0}) \subset S_{x_0}$. \blacksquare

Let us consider the example of Section 1.2 in light of the preceding result.

3.9 EXAMPLE: Recall that in Section 1.2 we had defined an affine subbundle \mathcal{A} on $M = \mathbb{R}^4$ whose fibre at (x_1, x_2, x_3, x_4) is

$$\{(0, 0, 0, x_1^2, \frac{1}{2}ax_1^2 + bx_1x_2 + \frac{1}{2}cx_2^2)\} + \text{span}_{\mathbb{R}} \{(0, 0, 0, dx_1^2), (1, 0, 0, 0), (0, 1, x_1, 0)\}.$$

For this subbundle we shall utilise the local (in fact global) linear generators $\mathcal{X} = \{X_0, X_1, X_2, X_3\}$ where

$$\begin{aligned} X_0(x_1, x_2, x_3, x_4) &= (0, 0, 0, x_1^2, \frac{1}{2}ax_1^2 + bx_1x_2 + \frac{1}{2}cx_2^2) \\ X_1(x_1, x_2, x_3, x_4) &= (0, 0, 0, dx_1^2) \\ X_2(x_1, x_2, x_3, x_4) &= (1, 0, 0, 0) \\ X_3(x_1, x_2, x_3, x_4) &= (0, 1, x_1, 0). \end{aligned}$$

Note that $Z_{x_0}(\mathcal{X}_{\mathbb{R}^3}) = \{X_0, X_1\}$. Thus, if $X \in Z_{x_0}(\mathcal{A})$ then $X = X_0 + \sum_{j=1}^3 \phi_j X_j$ for real-analytic functions ϕ_j , $j \in \{1, 2, 3\}$, satisfying $\phi_2(x_0) = \phi_3(x_0) = 0$. Let us also take $x_0 = (0, 0, 0, 0)$ and

$$S_{x_0} = \text{span}_{\mathbb{R}} \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)\},$$

noting that $L(\mathcal{A}_{x_0}) \subset S_{x_0}$. In this case one readily determines that $\langle Z_{x_0}(\mathcal{X}_{\mathbb{R}^3}), S_{x_0} \rangle = S_{x_0}$. From Proposition 3.8 we ascertain that $\langle Z_{x_0}(\mathcal{A}), S_{x_0} \rangle = S_{x_0}$. \square

3.5. Vector-valued quadratic forms associated with affine subbundles. We continue considering the situation where $0_{x_0} \in \mathcal{A}_{x_0}$. Here we define an important class of geometric objects, a set of vector-valued quadratic forms defined on $L(\mathcal{A}_{x_0})$. These quadratic forms are defined relative to a given subspace $S_{x_0} \subset T_{x_0}M$. Along with the general constructions, we also illustrate them by continuing the computations begun in Example 3.9 in the preceding section.

In order for the definitions to make sense, the subspace S_{x_0} should have certain properties that we now begin to define.

3.10 DEFINITION: Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle with the property that for $x_0 \in M$, $0_{x_0} \in \mathcal{A}_{x_0}$. A subspace $S_{x_0} \subset T_{x_0}M$ is **second-order compatible** with \mathcal{A} at x_0 if $L^{(2)}(L(\mathcal{A}))_{x_0} + \langle Z_{x_0}(\mathcal{A}), L(\mathcal{A}_{x_0}) \rangle \subset S_{x_0}$. \square

The notion of second-order compatibility has to do with the well-definedness of a map that we now introduce. Given a subspace $S_{x_0} \subset T_{x_0}M$ we define $S_{x_0}^A: Z_{x_0}(\mathcal{A}) \rightarrow \Sigma_2(L(\mathcal{A}_{x_0}); T_{x_0}M/S_{x_0})$ by

$$S_{x_0}^A(X)(v_1, v_2) = \pi_{S_{x_0}}([V_1, [X, V_2]](x_0)),$$

where $V_1, V_2 \in \Gamma(TM)$ are vector fields extending $v_1, v_2 \in L(\mathcal{A}_{x_0})$, and where $\pi_{S_{x_0}}: T_{x_0}M \rightarrow T_{x_0}M/S_{x_0}$ is the natural projection. Let us show when $S_{x_0}^A$ is well-defined, at the same time understanding the meaning of second-order compatibility.

3.11 PROPOSITION: *Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle satisfying $0_{x_0} \in \mathcal{A}_{x_0}$. For a subspace $S_{x_0} \subset T_{x_0}M$, the following statements hold:*

- (i) if S_{x_0} is second-order compatible with \mathcal{A} at x_0 then the map $S_{x_0}^A$ is well-defined;
- (ii) if $S_{x_0}^A$ is well-defined, then it is an affine map.

Proof: (i) We must show that $S_{x_0}^A(X)(v_1, v_2)$ is independent of the extensions V_1 and V_2 . For real-analytic functions ϕ_1 and ϕ_2 a computation gives

$$\begin{aligned} [\phi_1 V_1, [X, \phi_2 V_2]] &= \phi_1 \phi_2 [V_1, [X, V_2]] + \phi_1 (V_1 \phi_2) [X, V_2] \\ &\quad + \phi_1 (X \phi_2) [V_1, V_2] + \phi_1 (V_1 (X \phi_2)) V_2 - \phi_2 ([X, V_2] \phi_1) V_1 - (X \phi_2) (V_2 \phi_1) V_1. \end{aligned}$$

Any terms on the right-hand side of this expression involving derivatives of ϕ_1 and ϕ_2 lie in S_{x_0} if S_{x_0} is second-order compatible with \mathcal{A} at x_0 . Thus we have

$$\pi_{S_{x_0}}([\phi_1 V_1, [X, \phi_2 V_2]](x_0)) = \pi_{S_{x_0}}(\phi_1(x_0) \phi_2(x_0) [V_1, [X, V_2]](x_0)).$$

This shows that the definition of $S_{x_0}^A$ depends only on the values of V_1 and V_2 at x_0 , and not on their derivatives, meaning that $S_{x_0}^A$ is indeed well-defined.

(ii) We first show that both $Z_{x_0}(\mathcal{A})$ and $\Sigma_2(L(\mathcal{A}); T_{x_0}M/S_{x_0})$ are \mathbb{R} -affine spaces. Since $\Sigma_2(L(\mathcal{A}); T_{x_0}M/S_{x_0})$ is a \mathbb{R} -vector space, it is also a \mathbb{R} -affine space, so we need only consider $Z_{x_0}(\mathcal{A})$. Fix $\bar{X} \in Z_{x_0}(\mathcal{A})$ and define

$$L(Z_{x_0}(\mathcal{A})) = \{X - \bar{X} \mid X \in Z_{x_0}(\mathcal{A})\}.$$

We claim that $L(Z_{x_0}(\mathcal{A}))$ is a \mathbb{R} -vector space and is independent of the choice of \bar{X} . To see that $L(Z_{x_0}(\mathcal{A}))$ is indeed a \mathbb{R} -vector space, one can easily show that the operations

$$\begin{aligned} (X_1 - \bar{X}) + (X_2 - \bar{X}) &= X_1 + X_2 - 2\bar{X}, & X_1, X_2 \in Z_{x_0}(\mathcal{A}) \\ a(X - \bar{X}) &= aX - a\bar{X}, & a \in \mathbb{R}, X \in Z_{x_0}(\mathcal{A}) \end{aligned}$$

are well-defined (i.e., restrict to $L(Z_{x_0}(\mathcal{A}))$) and satisfy the axioms for vector addition and scalar multiplication. If $\bar{X}_1, \bar{X}_2 \in Z_{x_0}(\mathcal{A})$ then we have

$$\begin{aligned} \{X - \bar{X}_1 \mid X \in Z_{x_0}(\mathcal{A})\} &= \{X - (\bar{X}_1 - \bar{X}_2) - \bar{X}_2 \mid X \in Z_{x_0}(\mathcal{A})\} \\ &= \{Y - \bar{X}_2 \mid Y \in Z_{x_0}(\mathcal{A})\}, \end{aligned}$$

showing that the set $L(Z_{x_0}(\mathcal{A}))$ is independent of choice of $\bar{X} \in Z_{x_0}(\mathcal{A})$. One also has

$$\begin{aligned} (X_1 - \bar{X}_1) + (X_2 - \bar{X}_1) &= X_1 + X_2 - 2\bar{X}_1 \\ &= X_1 + X_2 - 2(\bar{X}_1 - \bar{X}_2) - 2\bar{X}_2 \\ &= (X_1 - (\bar{X}_1 - \bar{X}_2) - \bar{X}_2) + (X_2 - (\bar{X}_1 - \bar{X}_2) - \bar{X}_2), \end{aligned}$$

showing that vector addition is also independent of choice of \bar{X} . In like fashion one also shows that scalar addition is independent of choice of \bar{X} , and so as a \mathbb{R} -vector space, $L(Z_{x_0}(\mathcal{A}))$ is independent of the choice of \bar{X} . It is now a simple matter to verify that $Z_{x_0}(\mathcal{A})$ is a \mathbb{R} -affine space modelled on the \mathbb{R} -vector space $L(Z_{x_0}(\mathcal{A}))$.

Now, to show that $S_{x_0}^A$ is affine, one should show that the map

$$L(Z_{x_0}(\mathcal{A})) \ni X - \bar{X} \mapsto S_{x_0}^A(X) - S_{x_0}^A(\bar{X}) \in \Sigma_2(L(\mathcal{A}); T_{x_0}M/S_{x_0})$$

is \mathbb{R} -linear for any $\bar{X} \in Z_{x_0}(\mathcal{A})$. This, however, is obvious. \blacksquare

We illustrate this result with an example.

3.12 EXAMPLE: (Example 3.9 cont'd) We take \mathcal{A} , x_0 , and S_{x_0} as in Example 3.9. A direct computation verifies that S_{x_0} is second-order compatible with \mathcal{A} at x_0 . Let us use $\{(1, 0, 0, 0), (0, 1, 0, 0)\}$ as a basis for $L(\mathcal{A}_{x_0})$ and $(0, 0, 0, 1) + S_{x_0}$ as a basis for the quotient space $T_{x_0}M/S_{x_0}$. In this case, an element of $\Sigma_2(L(\mathcal{A}_{x_0}); T_{x_0}M/S_{x_0})$ is represented by a 2×2 matrix. If $X = X_0 + \sum_{j=1}^3 \phi_j X_j \in Z_{x_0}(\mathcal{A})$, a computation shows that $S_{x_0}^A$ is represented by the matrix

$$-\begin{bmatrix} a + 2d\phi_1(x_0) & b \\ b & c \end{bmatrix}.$$

Note that this does indeed define an affine map from $Z_{x_0}(\mathcal{A})$ to $\Sigma_2(L(\mathcal{A}_{x_0}); T_{x_0}M/S_{x_0})$. \square

We denote the linear part of the affine map $S_{x_0}^A$ by $L(S_{x_0}^A)$.

3.13 DEFINITION: Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle with the property that $0_{x_0} \in \mathcal{A}_{x_0}$, and let S_{x_0} be a subspace of $T_{x_0}M$ that is second-order compatible with \mathcal{A} at x_0 . The subspace S_{x_0} is **second-order invariant** with respect to \mathcal{A} at x_0 if $L(S_{x_0}^A) = 0$. \square

Second-order invariance has to do with the well-definedness of a vector-valued quadratic form which we now introduce. We assume as usual that $0_{x_0} \in \mathcal{A}_{x_0}$. For a subspace $S_{x_0} \subset T_{x_0}M$ we define $B_{\mathcal{A}}(S_{x_0}) \in \Sigma_2(L(\mathcal{A}_{x_0}); T_{x_0}M/S_{x_0})$ by

$$B_{\mathcal{A}}(S_{x_0})(v_1, v_2) = \pi_{S_{x_0}}([V_1, [X, V_2]](x_0))$$

where $X \in Z_{x_0}(\mathcal{A})$ is arbitrary, and, as in the definition of $S_{x_0}^A$, $V_1, V_2 \in \Gamma(TM)$ extend $v_1, v_2 \in L(\mathcal{A}_{x_0})$. The following result tells us when $B_{\mathcal{A}}(S_{x_0})$ is well-defined.

3.14 PROPOSITION: Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle satisfying $0_{x_0} \in \mathcal{A}_{x_0}$. If a subspace $S_{x_0} \subset T_{x_0}M$ is second-order invariant with respect to \mathcal{A} at x_0 then $B_{\mathcal{A}}(S_{x_0})$ is well-defined.

Proof: Since second-order invariance implies second-order compatibility with \mathcal{A} at x_0 , we need only show that $B_{\mathcal{A}}(S_{x_0})$ is independent of the choice of $X \in Z_{x_0}(\mathcal{A})$. This, however, follows since if $L(S_{x_0}^A) = 0$, then $S_{x_0}^A$ must be a constant map, since it is an affine map whose linear part is zero. \blacksquare

Let us now address the matter of actually determining when $L(S_{x_0}^A)$ vanishes. Since $Z_{x_0}(\mathcal{A})$ is generally an infinite-dimensional \mathbb{R} -affine space, one would like find a simple way of checking when $L(S_{x_0}^A) = 0$. To set this up, let $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ be a set of local linear generators for \mathcal{A} defined on a neighbourhood \mathcal{N} of x_0 . For $\alpha \in \mathbb{R}^k$ we denote (consistent with (3.1))

$$X_\alpha = X_0 + \sum_{j=1}^k \alpha_j X_j \in \Gamma(TM).$$

We define a map $L_{x_0}^{\mathcal{X}}: \mathbb{R}^k \rightarrow \mathcal{A}_{x_0}$ by $L_{x_0}^{\mathcal{X}}(\alpha) = X_\alpha(x_0)$. Generally $L_{x_0}^{\mathcal{X}}$ is not linear, but affine. Thus $(L_{x_0}^{\mathcal{X}})^{-1}(0_{x_0})$ will be an affine subspace of \mathbb{R}^k which we denote $\ker(L_{x_0}^{\mathcal{X}})$, making a convenient abuse of notation. We then define a map $S_{x_0}^{\mathcal{X}}: \ker(L_{x_0}^{\mathcal{X}}) \rightarrow \Sigma_2(L(\mathcal{A}_{x_0}); T_{x_0}M/S_{x_0})$ by

$$S_{x_0}^{\mathcal{X}}(\alpha)(v_1, v_2) = \pi_{S_{x_0}}([V_1, [X_\alpha, V_2]](x_0)), \quad (3.2)$$

where, as usual, $V_1, V_2 \in \Gamma(TM)$ extend $v_1, v_2 \in L(\mathcal{A}_{x_0})$. The following result provides a convenient characterisation of when $L(S_{x_0}^A) = 0$.

3.15 PROPOSITION: Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle satisfying $0_{x_0} \in \mathcal{A}_{x_0}$ and let $S_{x_0} \subset T_{x_0}M$ be second-order compatible with \mathcal{A} at x_0 . Also let $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ be local linear generators for \mathcal{A} defined in a neighbourhood of x_0 . Then $L(S_{x_0}^A) = 0$ if and only if $L(S_{x_0}^{\mathcal{X}}) = 0$.

Proof: As usual, we may assume without loss of generality that $\mathcal{N} = M$. For convenience let us further assume without loss of generality (cf. the proof of Proposition 3.8) that the local linear generators \mathcal{X} have the property that $X_0(x_0) = X_1(x_0) = \dots = X_\ell(x_0) = 0_{x_0}$ and that $\{X_{\ell+1}(x_0), \dots, X_k(x_0)\}$ are linearly independent. For $X \in Z_{x_0}(\mathcal{A})$ let us write

$$X = X_0 + \sum_{j=1}^k \phi_j X_j$$

where the real-analytic functions ϕ_1, \dots, ϕ_k have the property that $\phi_{\ell+1}(x_0) = \dots = \phi_k(x_0) = 0$. We then compute

$$\begin{aligned} [V_1, [X, V_2]] &= [V_1, [X_0, V_2]] \\ &\quad + \sum_{j=1}^k (\phi_j [V_1, [X_j, V_2]] + (V_1 \phi_j) [X_j, V_2] + (V_2 \phi_j) [X_j, V_1] - (V_1 V_2 \phi_j) X_j). \end{aligned}$$

Upon evaluation of this expression at x_0 we see that, provided that S_{x_0} is second-order compatible with \mathcal{A} at x_0 , the only terms not generally in S_{x_0} combine to exactly give $[V_1, [X_{\alpha(X)}, V_2]]$ if we take $\alpha(X) = (\phi_1(x_0), \dots, \phi_\ell(x_0), 0, \dots, 0) \in \mathbb{R}^k$. Note that this defines a linear map

$$Z_{x_0}(\mathcal{A}) \ni X \mapsto \alpha(X) \in \ker(L_{x_0}^{\mathcal{X}}). \quad (3.3)$$

Furthermore, note that this map is surjective.

Now let $V_1, V_2 \in \Gamma(TM)$ extend $v_1, v_2 \in L(\mathcal{A}_{x_0})$ and let $X \in Z_{x_0}(\mathcal{A})$. We directly compute

$$\begin{aligned} S_{x_0}^A(X)(v_1, v_2) - S_{x_0}^A(X_0)(v_1, v_2) &= S_{x_0}^{\mathcal{X}}(X_{\alpha(X)})(v_1, v_2) - S_{x_0}^{\mathcal{X}}(X_0)(v_1, v_2) \\ \implies L(S_{x_0}^A)(X - X_0) &= L(S_{x_0}^{\mathcal{X}})(X_{\alpha(X)} - X_0). \end{aligned}$$

The proposition follows from the surjectivity of the map (3.3). \blacksquare

Let us continue with our running example.

3.16 EXAMPLE: (Example 3.9 cont'd) We again use \mathcal{A} , \mathcal{X} , x_0 , and S_{x_0} as defined in Example 3.9. In Example 3.12 we had computed $S_{x_0}^A$, and from our computations there we see that $L(S_{x_0}^A) = 0$ if and only if $d = 0$. Now let us compute $S_{x_0}^{\mathcal{X}}$. For $\alpha \in \mathbb{R}^3$, let $X_\alpha \in Z_{x_0}(\mathcal{X})$, meaning simply that $\alpha_2 = \alpha_3 = 0$. A computation then shows that $S_{x_0}^{\mathcal{X}}$ is represented by the matrix

$$-\begin{bmatrix} a + 2d\alpha_1 & b \\ b & c \end{bmatrix}.$$

We see that $L(S_{x_0}^{\mathcal{X}}) = 0$ if and only if $d = 0$. This is the same condition for the vanishing of $L(S_{x_0}^A)$, consistent with Proposition 3.15. \square

3.17 REMARK: The preceding results may be summarised in the following way. Let \mathcal{A} be a real-analytic generalised affine subbundle and let $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ and $\tilde{\mathcal{X}} = \{\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_k\}$ be two sets of local linear generators for \mathcal{A} defined on a neighbourhood \mathcal{N} of x_0 (restrict each to the intersection of the neighbourhoods of definition of \mathcal{X} and $\tilde{\mathcal{X}}$ in case these neighbourhoods do not agree). Suppose that $X_0(x_0) = \tilde{X}_0(x_0) = 0_{x_0}$. For a subspace $S_{x_0} \subset T_{x_0}M$ that is second-order compatible, let us define $B_{\mathcal{X}}(S_{x_0})$ by

$$B_{\mathcal{X}}(S_{x_0})(v_1, v_2) = \pi_{S_{x_0}}([V_1, [X_0, V_2]](x_0)), \quad (3.4)$$

and define $B_{\tilde{\mathcal{X}}}(S_{x_0})$ similarly. One may then readily verify the following statements:

1. if S_{x_0} is second-order compatible with \mathcal{A} at x_0 then $B_{\mathcal{X}}(S_{x_0})$ is well-defined, in that it is independent of the extensions $V_1, V_2 \in \Gamma(TM)$ of $v_1, v_2 \in L(\mathcal{A}_{x_0})$ (of course, $B_{\tilde{\mathcal{X}}}(S_{x_0})$ is similarly well-defined);
2. to check whether S_{x_0} is second-order invariant with respect to \mathcal{A} at x_0 one may check whether either of the linear maps $L(S_{x_0}^{\mathcal{X}})$ or $L(S_{x_0}^{\tilde{\mathcal{X}}})$ between finite-dimensional \mathbb{R} -vector spaces are zero (this is exactly the content of Proposition 3.15);
3. if S_{x_0} is second-order invariant with respect to \mathcal{A} at x_0 then $B_{\mathcal{A}}(S_{x_0}) = B_{\mathcal{X}}(S_{x_0}) = B_{\tilde{\mathcal{X}}}(S_{x_0})$.

The upshot is that even if S_{x_0} is not second-order invariant but only second-order compatible with \mathcal{A} at x_0 , then we may define $B_{\mathcal{X}}(S_{x_0})$ relative to a given set \mathcal{X} of local linear generators for \mathcal{A} . If S_{x_0} is additionally second-order invariant, then $B_{\mathcal{X}}(S_{x_0})$ is independent of the choice of generators (provided that $X_0(x_0) = 0_{x_0}$), and so merits being denoted $B_{\mathcal{A}}(S_{x_0})$. One way to think of this is that second-order invariance refers to invariance of $B_{\mathcal{X}}(S_{x_0})$ under feedback transformations. \square

Our example helps to clarify the preceding discussion.

3.18 EXAMPLE: (Example 3.9 cont'd) We continue with \mathcal{A} , x_0 , and S_{x_0} as in Example 3.9. As we saw in Example 3.16, S_{x_0} is second-order invariant with respect to \mathcal{A} at x_0 if and only if $d = 0$. Furthermore, in this case the computations of Example 3.12 show that $B_{\mathcal{A}}(S_{x_0})$ is represented by the matrix

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

The upshot of Remark 3.17 is that the same matrix also represents $B_{\mathcal{X}}(S_{x_0})$ for any choice of local linear generators $\mathcal{X} = \{X_0, X_1, X_2, X_3\}$ having the property that $X_0(x_0) = (0, 0, 0, 0)$. In particular, as we see in Example 3.16, if we choose \mathcal{X} as in Example 3.9, then we have $B_{\mathcal{A}}(S_{x_0}) = B_{\mathcal{X}}(S_{x_0})$. However, when $d \neq 0$ then $B_{\mathcal{X}}(S_{x_0})$ will generally depend on the choice of local linear generators \mathcal{X} . What's more, and this is of essential importance for the results we state in Section 4, it is possible when $d \neq 0$ that the signs of the eigenvalues of the matrix representation of $B_{\mathcal{X}}(S_{x_0})$ can change depending on the choice of local linear generators \mathcal{X} . \square

The appearance of the linear map $L(S_{x_0}^{\mathcal{A}})$ in the conditions for controllability stated in Section 4 is one of the novel contributions of our work. The following result tells us

that generically we can expect that $L(S_{x_0}^{\mathcal{A}}) = 0$. Indeed, only at singular points of the generalised distribution $L(\mathcal{A})$ can we expect hypotheses involving the vanishing of $L(S_{x_0}^{\mathcal{A}})$ to be important.

3.19 PROPOSITION: *Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle with the property that $0_{x_0} \in \mathcal{A}_{x_0}$, and let S_{x_0} be a subspace of $T_{x_0}M$ that is second-order compatible with \mathcal{A} at x_0 . If x_0 is a regular point for $L(\mathcal{A})$ then $L(S_{x_0}^{\mathcal{A}}) = 0$.*

Proof: Choose local generators $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ for \mathcal{A} around x_0 having the property that $X_0(x_0) = 0$ and that $\{X_1(q_0), \dots, X_k(q_0)\}$ form a basis for $L(\mathcal{A}_{x_0})$. This is possible since x_0 is a regular point for $L(\mathcal{A})$. Then, by Proposition 3.15, it suffices to show that $L(S_{x_0}^{\mathcal{X}}) = 0$. However, since $\ker(L_{x_0}^{\mathcal{X}}) = \{0\}$ by virtue of the linear independence of $\{X_1(q_0), \dots, X_k(q_0)\}$, the result follows immediately. \blacksquare

Finally, we are interested in cases where $B_{\mathcal{A}}(S_{x_0})$ is symmetric.

3.20 PROPOSITION: *Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle satisfying $0_{x_0} \in \mathcal{A}_{x_0}$, and let $S_{x_0} \subset T_{x_0}M$ be a subspace that is second-order invariant with respect to \mathcal{A} at x_0 . The following statements hold:*

- (i) $B_{\mathcal{A}}(S_{x_0})$ is symmetric if and only if $\langle Z_{x_0}(\mathcal{A}), L^{(2)}(L(\mathcal{A}))_{x_0} \rangle \subset S_{x_0}$;
- (ii) if x_0 is a regular point for the distribution $L^{(\infty)}(L(\mathcal{A}))$ and if $L^{(\infty)}(L(\mathcal{A}))_{x_0} \subset S_{x_0}$, then $B_{\mathcal{A}}(S_{x_0})$ is symmetric.

Proof: (i) Let $X \in Z_{x_0}(\mathcal{A})$. First let us do the essential computation. Since $B_{\mathcal{A}}(S_{x_0})$ is a \mathbb{R} -bilinear map, it may be written as a sum of its symmetric and skew-symmetric parts:

$$B_{\mathcal{A}}(S_{x_0})(v_1, v_2) = \frac{1}{2}(B_{\mathcal{A}}(S_{x_0})(v_1, v_2) + B_{\mathcal{A}}(S_{x_0})(v_2, v_1)) + \frac{1}{2}(B_{\mathcal{A}}(S_{x_0})(v_1, v_2) - B_{\mathcal{A}}(S_{x_0})(v_2, v_1)).$$

Of course, $B_{\mathcal{A}}(S_{x_0})$ is symmetric if and only if its skew-symmetric part is zero. For $V_1, V_2 \in \Gamma(TM)$ extensions of $v_1, v_2 \in L(\mathcal{A}_{x_0})$, the skew-symmetric part of $B_{\mathcal{A}}(S_{x_0})(v_1, v_2)$ is

$$\frac{1}{2}\pi_{S_{x_0}}([V_1, [X, V_2]](x_0) - [V_2, [X, V_1]](x_0)).$$

Now, by the Jacobi identity we have

$$\begin{aligned} & [V_1, [X, V_2]] + [V_2, [V_1, X]] + [X, [V_2, V_1]] = 0 \\ \implies & [V_1, [X, V_2]] - [V_2, [X, V_1]] = [X, [V_1, V_2]]. \end{aligned} \quad (3.5)$$

From this computation, the lemma follows easily.

(ii) Let $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ be a set of local linear generators for \mathcal{A} with the property that $X_0(x_0) = 0_{x_0}$. Let $\mathcal{X}_1 = \{X_1, \dots, X_k\}$ so that $L^{(\infty)}(L(\mathcal{A})) = L^{(\infty)}(\mathcal{X}_1)$ by Proposition 3.7. Since x_0 is a regular point for $L^{(\infty)}(\mathcal{X}_1)$ there exists coordinates (x^1, \dots, x^n) for a neighbourhood \mathcal{N} of x_0 so that

$$L^{(\infty)}(\mathcal{X}_1)_x = \text{span}\left\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}\right\}$$

for some $k \in \{1, \dots, n\}$ and for all $x \in \mathcal{N}$. Define a family of vector fields on \mathcal{N} by

$$\tilde{\mathcal{X}} = \{\tilde{X}_0 = f_0, \tilde{X}_1 = \frac{\partial}{\partial x^1}, \dots, \tilde{X}_k = \frac{\partial}{\partial x^k}\}.$$

Also define $\tilde{S}_{x_0} = L^{(1)}(\tilde{\mathcal{X}}_1)_{x_0} = L^{(\infty)}(\mathcal{X}_1)_{x_0}$. We first claim that $B_{\tilde{\mathcal{X}}}(\tilde{S}_{x_0})$ is symmetric. Indeed, the Jacobi identity shows that

$$B_{\tilde{\mathcal{X}}}(\tilde{S}_{x_0})(\tilde{X}_a(x_0), \tilde{X}_b(x_0)) - B_{\tilde{\mathcal{X}}}(\tilde{S}_{x_0})(\tilde{X}_b(x_0), \tilde{X}_a(x_0)) = \pi_{\tilde{S}_{x_0}}([X_0, [\tilde{X}_a, \tilde{X}_b]](x_0)),$$

for $a, b \in \{1, \dots, k\}$. Since $[\tilde{X}_a, \tilde{X}_b] = 0$, we have shown that $B_{\tilde{\mathcal{X}}}(\tilde{S}_{x_0})$ is symmetric on a basis for $L^{(1)}(\tilde{\mathcal{X}}_1)_{x_0}$, and so is symmetric. Since $L^{(1)}(\mathcal{X})_{x_0} \subset L^{(1)}(\tilde{\mathcal{X}})_{x_0}$, it follows that $B_{\mathcal{X}}(\tilde{S}_{x_0})$ is symmetric. Then, since $\tilde{S}_{x_0} \subset S_{x_0}$, it also follows that $B_{\mathcal{X}}(S_{x_0})$ is symmetric. ■

3.21 REMARK: If S_{x_0} is only second-order compatible with \mathcal{A} at x_0 , and if $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ are local linear generators for \mathcal{A} with $X_0(x_0) = 0_{x_0}$, then the computations in part (i) of the proof of the preceding result show that $B_{\mathcal{X}}(S_{x_0})$ is symmetric if $\text{ad}_{X_0}(L^{(2)}(L(\mathcal{A}))_{x_0}) \subset S_{x_0}$, where $B_{\mathcal{X}}(S_{x_0})$ is defined as in (3.4). □

4. Statement of geometric results

In this section we state the main results of the paper, namely those that provide geometric conditions on affine subbundles that ensure or prohibit their local controllability. To put these results into context, we first state, using our language and notation, some low-order results from the existing literature. The proofs of the new results in this section rely on control-affine system results that we state in Section 5. The proofs of the results from both sections are distributed over the last two sections of the paper, Sections 7 and 8.

4.1. Existing low-order conditions. In this section, for context and completeness, we review the zeroth and first-order conditions for small-time local controllability from x_0 . The results we state will be for affine systems, and so differ slightly from the actual statements in the literature that are given for control-affine systems. The statements we give can be derived from those in the literature using the ideas we develop in this paper. For the zeroth-order case, Theorem 4.1, the relation to the form of the results stated in the literature is clear. For the first-order result, Theorem 4.2, we refer to Proposition 3.8.

Let us first look at conditions which involve *no* differentiations of the input vector fields.

4.1 THEOREM: ([Sussmann 1978]) *Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle. The following statements hold:*

- (i) if $\mathcal{A}_{x_0} = T_{x_0}M$ then \mathcal{A} is properly STLC from x_0 ;
- (ii) if neither of the equivalent conditions

- (a) $0_{x_0} \in \mathcal{A}_{x_0}$ nor
- (b) $L(\mathcal{A}_{x_0}) = \mathcal{A}_{x_0}$

hold, then \mathcal{A} is STLUC from x_0 .

- (iii) more generally, if \mathcal{A} is an affine system in \mathcal{A} and if $0_{x_0} \notin \text{conv}(\mathcal{A}(x_0))$ then \mathcal{A} is not STLC from x_0 ;

Furthermore, the conditions (i) and (ii) are the best possible zeroth-order conditions in the sense that if \mathcal{A} satisfies neither of these conditions, then one cannot determine whether \mathcal{A} is properly STLC or STLUC from x_0 using only the values of sections of \mathcal{A} at x_0 .

These conditions are exactly what one would guess. The sufficient condition demands that the system be fully actuated, and the necessary condition indicates the system is not controllable if, in essence, one cannot counteract the effects of the drift vector field at x_0 .

Now let us turn to first-order conditions. One should, of course, assume that one falls through the gap in the conditions (i) and (ii) of Theorem 4.1. This gap occurs when $0_{x_0} \in \mathcal{A}_{x_0}$ but $\mathcal{A}_{x_0} \subsetneq T_{x_0}M$. In the first-order result we state, we actually restrict to systems for which $0_{x_0} \in \text{int}(\text{conv}(\mathcal{A}(x_0)))$. We do this primarily for convenience, as to consider systems not satisfying this condition changes the character of the investigation.

4.2 THEOREM: ([Bianchini and Stefani 1984]) *Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle satisfying $0_{x_0} \in \mathcal{A}_{x_0}$. Then \mathcal{A} is properly STLC from x_0 if*

$$\langle Z_{x_0}(\mathcal{A}), L^{(2)}(\mathcal{A})_{x_0} \rangle = T_{x_0}M. \quad (4.1)$$

Furthermore, the condition (4.1) is the best possible first-order condition in the sense that if \mathcal{A} does not satisfy this condition, then it is not possible to determine whether \mathcal{A} is properly STLC or STLUC from x_0 using only 1-jets of sections of \mathcal{A} .

Note that there are no obstructions to STLC from x_0 that involve only the 1-jets of sections of \mathcal{A} . This is reflected by the nature of the condition (4.1) being merely a spanning condition. As we shall see, second-order conditions involve notions of convexity rather than simply spanning conditions.

4.2. Statement of results for affine systems. Now we turn to providing second-order conditions on \mathcal{A} for determining controllability (or lack of same) as given in Definition 3.3. First we state a condition for an affine subbundle to be properly STLC from x_0 .

4.3 THEOREM: *Let \mathcal{A} be a real-analytic generalised affine subbundle, let $S_{x_0} = \langle Z_{x_0}(\mathcal{A}), L^{(2)}(\mathcal{A})_{x_0} \rangle$, and suppose that for $x_0 \in M$ the following conditions hold:*

- (i) $0_{x_0} \in \mathcal{A}_{x_0}$;
- (ii) $\langle Z_{x_0}(\mathcal{A}), L^{(3)}(\mathcal{A})_{x_0} \rangle = T_{x_0}M$;
- (iii) $L(S_{x_0}^{\mathcal{A}}) = 0$;
- (iv) $B_{\mathcal{A}}(S_{x_0})$ is indefinite.

Then \mathcal{A} is properly STLC from x_0 .

A condition for an affine subbundle to be STLUC from x_0 can be stated similarly.

4.4 THEOREM: *Let \mathcal{A} be a real-analytic generalised affine subbundle and let $S_{x_0} = \langle Z_{x_0}(\mathcal{A}), L(\mathcal{A}_{x_0}) \rangle + L^{(\infty)}(L(\mathcal{A}))_{x_0}$. First suppose that*

- (i) $0_{x_0} \in \mathcal{A}_{x_0}$;
- (ii) $x_0 \in M$ is a regular point for the distribution $L^{(\infty)}(L(\mathcal{A}))$.

Then $L(S_{x_0}^{\mathcal{A}}) = 0$. If it is additionally the case that

- (iii) $B_{\mathcal{A}}(S_{x_0})$ is definite

then \mathcal{A} is STLUC from x_0 .

4.5 REMARKS: 1. The quadratic forms used in Theorems 4.3 and 4.4 may be verified, using the development of Section 3.5, to be well-defined and symmetric.

2. Note that for Theorem 4.3 the condition that $L(S_{x_0}^A) = 0$ is part of the hypotheses, whereas for Theorem 4.4 it is part of the conclusions. Of course, a legitimate question is then, “When S_{x_0} is defined as in Theorem 4.3, how often is it that $L(S_{x_0}^A) = 0$?” We do not attempt to pose this as a precise question, but merely comment that it seems that for many systems one can expect that $L(S_{x_0}^A) = 0$. Also, Proposition 3.19 ensures that $L(S_{x_0}^A) = 0$ if x_0 is a regular point for $L(A)$. \square

One of the immediate consequences of the preceding geometric results is that they allow one to determine the small-time local controllability from x_0 of an affine system \mathcal{A} by looking at any local control-affine realisation of \mathcal{A} at x_0 . The important upshot is that, provided that the affine subbundle \mathcal{A} has the character of either Theorem 4.3 or 4.4, the small-time local controllability of an affine system in \mathcal{A} can be determined using information at x_0 . To be precise about this, we state two corollaries, the first corresponding to Theorem 4.3.

4.6 COROLLARY: *Suppose that $\mathcal{A} \subset TM$ is a real-analytic generalised affine subbundle satisfying all hypotheses of Theorem 4.3. For an affine system \mathcal{A} in \mathcal{A} , proper at x_0 , the following two statements hold:*

- (i) \mathcal{A} is STLC from x_0 ;
- (ii) any local control-affine realisation $\Sigma = (N, \mathcal{F}, U)$ of \mathcal{A} at x_0 is STLC from x_0 .

4.7 REMARK: If we omit hypothesis (iii) of Theorem 4.3, then Corollary 4.6 is generally not true. We give an instance of this in Example 6.1. \square

Next we state the consequence of Theorem 4.4.

4.8 COROLLARY: *Suppose that $\mathcal{A} \subset TM$ is a real-analytic generalised affine subbundle satisfying all hypotheses of Theorem 4.4. For an affine system \mathcal{A} in \mathcal{A} with $\mathcal{A}(x_0)$ compact, the following two statements hold:*

- (i) \mathcal{A} is not STLC from x_0 ;
- (ii) any local control-affine realisation $\Sigma = (N, \mathcal{F}, U)$ of \mathcal{A} at x_0 is not STLC from x_0 .

5. Statement of results for control-affine systems

In this section we state two results that deal with standard control-affine systems. In order to state the results in this section, we first need some extra notation specific to control-affine systems. Some of the notation we provide is not actually needed for the statement of the results, but will be used in the proofs in Section 7.

5.1. Some control-affine system notation. Let $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ be a control-affine system. As with our specification of the map \mathcal{A} for an affine system, we will allow U to be general, asserting the necessary properties as we go along. We shall frequently ask that U have the property that $0 \in \text{int}(\text{conv}(U))$. In this case we say that U is *proper*. An *admissible control* is a measurable function $u: [0, T] \rightarrow U$, defined for some $T > 0$. If $\xi: [0, T] \rightarrow M$ is a trajectory for the affine system \mathcal{A}_Σ associated with Σ (see Section 3.2), then this implies the existence of an admissible control $u: [0, T] \rightarrow U$ so that

$$\dot{\xi}(t) = f_0(\xi(t)) + \sum_{a=1}^m u_a(t) f_a(\xi(t)).$$

We shall call the pair (ξ, u) a *controlled trajectory* for Σ in this case. We denote by $\text{Traj}(\Sigma, T)$ the collection of controlled trajectories defined on $[0, T]$ and $\text{Traj}(\Sigma) = \bigcup_{T>0} \text{Traj}(\Sigma, T)$. For $x \in M$, $u: [0, T] \rightarrow U$ an admissible control, and $\tau \in [0, T]$, we denote by $[0, T] \ni t \mapsto \xi(x, u, t, \tau) \in M$ the absolutely continuous curve for which $(\xi(x, u, \cdot, \tau), u) \in \text{Traj}(\Sigma, T)$ satisfies $\xi(x, u, \tau, \tau) = x$. For a control $u: [0, T] \rightarrow U$, let $f_u: [0, T] \times M \rightarrow TM$ denote the possibly time-dependent vector field

$$f_u(t, x) = f_0(x) + \sum_{a=1}^m u_a(t) f_a(x).$$

The flow of this time-dependent vector field is then, using our above notation, given by $(t, x) \mapsto \xi(x, u, t, \tau)$. Using our notation from Section 2.1, $T\xi(\cdot, u, t, \tau)$ denotes the derivative of this map with respect to x .

5.1 DEFINITION: Let $\Sigma = (M, \mathcal{F}, U)$ be a control-affine system and let $x_0 \in M$.

- (i) Σ is *small-time locally controllable (STLC)* from x_0 if \mathcal{A}_Σ is STLC from x_0 .
- (ii) A pair $(M, \mathcal{F} = \{f_0, f_1, \dots, f_m\})$ is *properly small-time locally controllable (properly STLC)* from $x_0 \in M$ if $\Sigma_U = (M, \mathcal{F}, U)$ is STLC from x_0 for every proper control set $U \subset \mathbb{R}^m$. \square

We comment that one of the interesting facts that will come up in our development is that it is *not true* that if (M, \mathcal{F}) is properly STLC from x_0 then $\mathcal{A}_\mathcal{F}$ is necessarily properly STLC from x_0 (see Example 6.1).

As in Section 2.1, we denote by $Z_{x_0}(\mathcal{F})$ the vector fields in \mathcal{F} that vanish at x_0 . The remaining vector fields, those that do not vanish at x_0 , we denote by $\tilde{Z}_{x_0}(\mathcal{F})$. Let us also denote $\mathcal{F}_1 = \{f_1, \dots, f_m\}$. Recall from Section 3.2 that we had defined $\mathcal{F}_U = \{f_u \mid u \in U\}$. Consistent with our existing notation, $Z_{x_0}(\mathcal{F}_U)$ denotes those vector fields in \mathcal{F}_U that vanish at x_0 . We shall find it helpful to introduce a little detail concerning the vector fields $Z_{x_0}(\mathcal{F}_U)$. For the moment, suppose that $f_0(x_0) = 0_{x_0}$, as we shall generally make this assumption in the sequel. We let $L_{x_0}^\mathcal{F}: \mathbb{R}^m \rightarrow T_{x_0}M$ be the linear map defined by

$$L_{x_0}^\mathcal{F}(u) = \sum_{a=1}^m u_a f_a(x_0).$$

Note that $Z_{x_0}(\mathcal{F}_U)$ is then comprised of those vector fields f_u for which $u \in \ker(L_{x_0}^\mathcal{F}) \cap U$.

5.2. Statement of results for control-affine systems. In stating the results in this section, we use the vector-valued quadratic form $B_\mathcal{F}(S_{x_0})$ defined corresponding to a family of vector fields $\mathcal{F} = \{f_0, f_1, \dots, f_m\}$ and to a subspace $S_{x_0} \subset T_{x_0}M$. The definition of such forms, and some comments concerning their well-definedness, was provided in Remark 3.17.

The sufficient condition is the following.

5.2 THEOREM: *Let $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ be a real-analytic control-affine system, let $x_0 \in M$, and assume the following:*

- (i) $f_0(x_0) = 0_{x_0}$;
- (ii) U is proper;
- (iii) $\langle Z_{x_0}(\mathcal{F}_{\text{conv}(U)}), L^{(3)}(\mathcal{F})_{x_0} \rangle = T_{x_0}M$;

(iv) $B_{\mathcal{F}}(\langle Z_{x_0}(\mathcal{F}_{\text{conv}(U)}), L^{(2)}(\mathcal{F})_{x_0} \rangle)$ is indefinite.

Then Σ is STLC from x_0 .

- 5.3 REMARKS:**
1. The quadratic form used in Theorem 5.2 is defined as in Remark 3.17, and may be verified to be well-defined and symmetric. However, it may not be the case that $L(S_{x_0}^{\mathcal{F}}) = 0$, but this is not needed here.
 2. This condition was stated by Basto-Gonçalves [1998], except that the quadratic form used in that paper was $B_{\mathcal{F}}(\langle \{f_0\}, L^{(2)}(\mathcal{F})_{x_0} \rangle)$, and the spanning condition (iii) was replaced with $\langle \{f_0\}, L^{(3)}(\mathcal{F})_{x_0} \rangle = T_{x_0}M$. Thus Theorem 5.2 has slightly weaker hypotheses than the result of Basto-Gonçalves. To prove Theorem 5.2, as we shall see in Section 7.1, one needs to slightly extend the notion of the variational cone as defined by Bianchini and Stefani [1993].
 3. In the quotient space defining the quadratic form in condition (iv), note that we use $Z_{x_0}(\mathcal{F}_{\text{conv}(U)})$ and not $Z_{x_0}(\mathcal{F}_U)$. This allows more systems to satisfy the hypotheses of the sufficient condition, since if $B_{\mathcal{F}}(S_{x_0})$ is indefinite and if $\tilde{S}_{x_0} \subset S_{x_0}$ then $B_{\mathcal{F}}(\tilde{S}_{x_0})$ is also indefinite. \square

Let us now state the necessary condition.

5.4 THEOREM: Let $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ be a real-analytic control-affine system, let $x_0 \in M$, and assume the following:

- (i) $f_0(x_0) = 0_{x_0}$;
- (ii) U is compact;
- (iii) $x_0 \in M$ is a regular point for the distribution $L^{(\infty)}(\mathcal{F}_1)$;
- (iv) $B_{\mathcal{F}}(\langle Z_{x_0}(\mathcal{F}_{\text{aff}(U)}), L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0})$ is definite.

Then Σ is not STLC from x_0 .

- 5.5 REMARKS:**
1. As with that in Theorem 5.2, the quadratic form in Theorem 5.4 is well-defined and symmetric. The subspace S_{x_0} also has the property that $L(S_{x_0}^{\mathcal{F}}) = 0$, although that is of no significance, per se, in Theorem 5.4.
 2. A result of a similar character is provided by Stefani [1988] as a sufficient condition for optimality. That result has a slightly different nature than ours in that the smallest eigenvalue of the matrix for some quadratic form is required to satisfy an inequality. Since such eigenvalues are not invariant under changes of basis (although their signs are) it is difficult to ascertain the geometric content of such a result.
 3. The condition of compactness on the control set is the weakest possible to ensure the convergence of the Chen-Fliess-Sussmann series, and this series is the prominent tool used in the proof of Theorem 5.4. \square

6. Examples

In this section we give one example that illustrates how one applies Theorems 4.3 and 4.4. We then provide five examples that illustrate the sharpness of the hypotheses of Theorems 5.2 and 5.4, and thus by implication the hypotheses of Theorems 4.3 and 4.4.

In this way we provide support for speculations that the conditions we give are the best possible ones “of their type.” Apart from their value in relationship to the results stated in Sections 4 and 5, the collection of examples in this section are of independent interest in that they exhibit why it is difficult to obtain general, sharp results for local controllability.

6.1. Discussion. Apart from Example 6.1 which actually does satisfy the hypotheses of either Theorem 4.3 or Theorem 4.4, our examples are specially contrived to violate these hypotheses. Therefore, let us preview our examples together so that, before the details are presented, one can get an idea of what they show. First let us examine the hypotheses of the sufficient condition of Theorem 5.2. The issues behind the conditions (i) and (ii) somewhat peripheral to what we are doing in this paper, so we do not give them detailed consideration. These hypotheses can be improved, but to do so in any generality will involve consideration of issues of a different flavour than are considered in this paper. We refer to a recent paper of the authors [Hirschorn and Lewis 2002] for some techniques for handling such systems. The condition (iii) is a spanning condition. One certainly expects that it cannot be relaxed, and it is easy to see that this is indeed the case. To weaken the quadratic form condition (iv), as suggested in Remark 5.3–3, we might try to enlarge the quotient space of the quadratic form. In Example 6.2 we show that this is not possible by presenting a system whose quotient space is $\langle Z_{x_0}(\mathcal{F}_U), L^{(3)}(\mathcal{F}_1)_{x_0} \rangle$, but which is not controllable, at least not for arbitrarily small controls.

Now let us consider the hypotheses of Theorem 5.4. Again, condition (i) lies somewhat outside the scope of what we are doing in this paper. The condition (ii) on the control set U is the weakest that can be made in order that our use of the Chen-Fliess-Sussmann series be justified. If one allows unbounded controls, then the character of small-time local controllability changes, and we do not consider this. The assumption (iii) of regularity of the involutive closure of the input distribution is essential in our proof. We show in Example 6.6 that if the condition is removed from the hypotheses of Theorem 5.4, it is possible that the system be STLC. Finally, to weaken the hypotheses of the necessary condition, we may consider making the quotient space in the definition of the quadratic form in condition (iv) smaller. We consider doing this in two ways. In Example 6.3 we consider the quotient space $\langle \{f_0\}, L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0}$. That is, we only consider the effects of the vanishing of the drift, and not of the control vector fields at x_0 . Example 6.3 shows that there are controllable systems with a definite quadratic form using this quotient space. In Example 6.4 we consider replacing $L^{(\infty)}(\mathcal{F}_1)_{x_0}$ in the quotient with a subspace that is smaller by one-dimension. The example we construct shows that small-time local controllability is possible in this case.

The next example we consider, Example 6.7, is a parameterised family of “higher-order” examples. For all choices of the parameter, Example 6.7 is one where the hypotheses of neither Theorem 5.2 nor Theorem 5.4 are satisfied. One may readily see that for certain choices of the parameter the system is STLC, and for other choices it is not. Thus we have the unsurprising conclusion that systems falling through the gaps in our results may be either STLC or not.

6.2. The examples. We present the examples as control-affine systems, since it is in this form that they are most easily understood. However, we remark that while we phrase the discussion in terms of illustrating the hypotheses of Theorems 5.2 and 5.4, they apply

equally well to Theorems 4.3 and 4.4 (for Theorem 4.3, we should restrict to those cases where $L(S_{x_0}^A) = 0$). It is only in Example 6.1 that we consider a system where the differences between Theorems 4.3 and 5.2 are present.

Our first example gives the concluding analysis of the system initially presented in Section 1.2. For this system, we justify the claims initially presented for this system.

6.1 EXAMPLE: We take $M = \mathbb{R}^4$ with the control-affine system

$$\begin{aligned}\dot{x}_1 &= u_2 \\ \dot{x}_2 &= u_3 \\ \dot{x}_3 &= x_1 u_3 \\ \dot{x}_4 &= \frac{1}{2} a x_1^2 + b x_1 x_2 + \frac{1}{2} c x_2^2 + d x_1^2 u_1.\end{aligned}$$

We take $x_0 = (0, 0, 0, 0)$, and as a control set we initially take a convex, proper set $U \subset \mathbb{R}^4$. Some useful brackets are

$$\begin{aligned}[f_2, f_3] &= (0, 0, 1, 0), & [f_1, f_2] &= (0, 0, 0, -2d x_1) \\ [f_1, f_3] &= (0, 0, 0, 0), & [f_2, [f_1, f_2]] &= (0, 0, 0, -2d) \\ [f_2, [f_0, f_2]] &= (0, 0, 0, -a) & [f_3, [f_0, f_3]] &= (0, 0, 0, -c) \\ [f_2, [f_0, f_3]] &= [f_3, [f_0, f_2]] = (0, 0, 0, -b) \\ \text{ad}_{f_0}^k f_1 &= \text{ad}_{f_0}^k f_2 = 0, & k \in \mathbb{N} \cup \{0\} & \quad \text{ad}_{f_3}^k f_1 = \text{ad}_{f_3}^k f_2 = 0, \quad k \in \mathbb{N} \cup \{0\}.\end{aligned}$$

Let us first consider the case where $d = 0$, and explain the assertions made in Section 1.2. In this case we see that

$$\begin{aligned}\langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle &= \langle Z_{x_0}(\mathcal{F}_U), L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0} = \\ &= \text{span}_{\mathbb{R}} \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)\}.\end{aligned}$$

Using $\text{span}_{\mathbb{R}} \{(0, 0, 0, 1)\}$ as a model for the quotient spaces of both Theorems 5.2 and 5.4, the quadratic forms of these theorems are represented by the matrix

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

The eigenvalues of this matrix are

$$\frac{1}{2}(a + c \pm \sqrt{4b^2 + (a - c)^2}),$$

and one verifies that these have the same sign exactly when $b^2 - ac < 0$. The numerical value of these eigenvalues have no significance as concerns the quadratic form of Theorems 5.2 and 5.4, but the signs do allow us to deduce that when $b^2 - ac > 0$ the form is indefinite, and when $b^2 - ac < 0$ the form is definite. Thus we do recover, as asserted in Section 1.2, that in the former case the system is properly STLC and in the latter case is not STLC.

Now let us consider the situation when $d \neq 0$. We first note that in this case the quotient space for the quadratic form of Theorem 5.4 is now $T_{x_0} \mathbb{R}^4$, so we may no longer rely on Theorem 5.4 to deduce when the system is not STLC. Also, as we saw in Example 3.16, when $d \neq 0$ the linear map $L(S_{x_0}^A)$ appearing in part (iii) of Theorem 4.3 is nonzero. Thus

we cannot assert, without further justification, that the affine subbundle $\mathcal{A}_{\mathcal{F}}$ is properly STLC when $d \neq 0$. Indeed, let us show that it can happen that for certain choices of a, b, c , and d , it is possible that the affine subbundle $\mathcal{A}_{\mathcal{F}}$ is not properly STLC from x_0 . For the remainder of the example we fix $a = 2, b = 0, c = -2$, and $d = 1$, giving the equations

$$\begin{aligned}\dot{x}_1 &= u_2 \\ \dot{x}_2 &= u_3 \\ \dot{x}_3 &= x_1 u_3 \\ \dot{x}_4 &= x_1^2(1 + u_1) - x_2^2.\end{aligned}$$

We denote by $\mathcal{F} = \{f_0, f_1, f_2, f_3\}$ the family of vector fields defined by these equations. Our bracket computations above allow us to assert that $\Sigma = (M, \mathcal{F}, U)$ is STLC from x_0 for any proper control set U . Indeed, the system is STLC from x_0 without having to use the control vector field f_1 . Now let us take $\tilde{\mathcal{F}} = \{\tilde{f}_0 = f_0 + \alpha f_1, \tilde{f}_1 = f_1, \tilde{f}_2 = f_2, \tilde{f}_3 = f_3\}$. Clearly $\mathcal{A}_{\tilde{\mathcal{F}}} = \mathcal{A}_{\mathcal{F}}$. The differential equations governing this modified system are

$$\begin{aligned}\dot{x}_1 &= u_2 \\ \dot{x}_2 &= u_3 \\ \dot{x}_3 &= x_1 u_3 \\ \dot{x}_4 &= x_1^2(1 + \alpha + u_1) - x_2^2.\end{aligned}$$

It is obvious that for $\alpha < -1$ the system will not be STLC from x_0 if sufficiently small bounds are placed on u_1 . That is to say, for $\alpha < -1$, $(M, \tilde{\mathcal{F}})$ is not properly STLC from x_0 , although (M, \mathcal{F}) is. In particular, if we take $\alpha = -3$ and define $U = [-1, 1]^3$, then we see that $\Sigma = (M, \mathcal{F}, U)$ and $\tilde{\Sigma} = (M, \tilde{\mathcal{F}}, U)$ satisfy $\mathcal{A}_{\Sigma}(x_0) = \mathcal{A}_{\tilde{\Sigma}}(x_0)$, and that Σ is properly STLC from x_0 but that $\tilde{\Sigma}$ is not.

This is explained by the fact that the underlying affine subbundle $\mathcal{A}_{\tilde{\mathcal{F}}}$ does not satisfy the condition (iii) of Theorem 4.3. Indeed, we note that $\ker(L_{x_0}^{\tilde{\mathcal{F}}}) = \text{span}_{\mathbb{R}} \{(1, 0)\}$, and that if we take $S_{x_0} = \langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle$, then $L(S_{x_0}^{\tilde{\mathcal{F}}})(0, 0, \alpha)$ is represented by the matrix

$$\begin{bmatrix} 0 & 0 \\ 0 & 2\alpha \end{bmatrix}.$$

By Proposition 3.14 we conclude that $B_{\tilde{\mathcal{F}}}(S_{x_0})$ is not independent of the choice of local linear generators \mathcal{F} . In this particular example, this is manifested the fact that $(M, \tilde{\mathcal{F}})$, although properly STLC itself, can be transformed to a system not sharing its property of being properly STLC. Generally when it is the case that $B_{\tilde{\mathcal{F}}}(S_{x_0})$ is not independent of \mathcal{F} , the small-time local controllability of the systems that generate the same affine subbundle as $\tilde{\mathcal{F}}$ will be complicated, and will depend on the character of the linear map $L(S_{x_0}^{\tilde{\mathcal{F}}})$. \square

Our next example shows that one cannot weaken hypothesis (iv) of Theorem 5.2 by taking $\langle Z_{x_0}(\mathcal{F}_U), L^{(3)}(\mathcal{F}_1)_{x_0} \rangle$ as the quotient space for the quadratic form. This example is one that is controllable with sufficiently large controls, but which is not generally controllable.

6.2 EXAMPLE: We consider the system on \mathbb{R}^3 defined by

$$\begin{aligned}\dot{x}_1 &= u_1 \\ \dot{x}_2 &= u_2 \\ \dot{x}_3 &= x_1^2(1 + \frac{1}{2}u_2),\end{aligned}$$

and with the control set $U = [-1, 1]^2$. We take as our reference point $x_0 = (0, 0, 0)$. Some relevant Lie brackets for this system are

$$\begin{aligned} [f_0, f_1] &= (0, 0, -2x_1) & \text{ad}_{f_0}^k f_1(x_0) &= \text{ad}_{f_0}^k f_2(x_0) = 0, & k &\geq 1 \\ [f_1, [f_0, f_1]] &= (0, 0, -2) & [f_2, [f_0, f_2]] &= (0, 0, 0) \\ [f_1, [f_0, f_2]] &= (0, 0, 0) & [f_2, [f_0, f_1]] &= (0, 0, 0) \\ [f_1, f_2] &= (0, 0, x_1) & [f_1, [f_1, f_2]] &= (0, 0, 1). \end{aligned}$$

Note that $L^{(3)}(\mathcal{F}_1)_{x_0} = T_{x_0}\mathbb{R}^3$. Therefore, $B_{\mathcal{F}}(L^{(3)}(\mathcal{F}_1)_{x_0})$ takes its image in the trivial vector space, and so is by definition indefinite. However, for the given control set U , the system is obviously not STLC from x_0 .

Let us now show that this system is controllable if the controls are allowed to be sufficiently large. To show this, we use the results from a previous paper of the authors where a class of control variations are developed for general families of vector fields [Hirschorn and Lewis 2002]. For $\alpha > 0$ let us define

$$\begin{aligned} X_\alpha^+ &= f_0 + \alpha f_1, & X_\alpha^- &= f_0 - \alpha f_1, \\ Y_\alpha^+ &= f_0 + \alpha f_2, & Y_\alpha^- &= f_0 - \alpha f_2. \end{aligned}$$

We then clearly have

$$\begin{aligned} \alpha f_1(x_0) &= \left. \frac{d}{ds} \right|_{s=0} e^{sX_\alpha^+}(x_0), & -\alpha f_1(x_0) &= \left. \frac{d}{ds} \right|_{s=0} e^{sX_\alpha^-}(x_0), \\ \alpha f_2(x_0) &= \left. \frac{d}{ds} \right|_{s=0} e^{sY_\alpha^+}(x_0), & -\alpha f_2(x_0) &= \left. \frac{d}{ds} \right|_{s=0} e^{sY_\alpha^-}(x_0). \end{aligned}$$

One also computes

$$\begin{aligned} -\frac{1}{3}\alpha^2[f_1, [f_0, f_1]](x_0) &= \left. \frac{d}{ds} \right|_{s=0} \left(e^{\sqrt[3]{s}X_\alpha^-} \circ e^{\sqrt[3]{s}X_\alpha^+}(x_0) \right) & (6.1) \\ -\frac{17}{6}\alpha^2[f_1, [f_0, f_1]](x_0) - \frac{1}{2}\alpha^3[f_1, [f_1, f_2]](x_0) &= \\ \left. \frac{d}{ds} \right|_{s=0} \left(e^{\sqrt[3]{s}Y_\alpha^-} \circ e^{\sqrt[3]{s}X_\alpha^+} \circ e^{\sqrt[3]{s}Y_\alpha^+} \circ e^{\sqrt[3]{s}X_\alpha^-} \circ e^{\sqrt[3]{s}Y_\alpha^-} \circ e^{\sqrt[3]{s}X_\alpha^+} \circ e^{\sqrt[3]{s}Y_\alpha^+} \circ e^{\sqrt[3]{s}X_\alpha^-} \circ e^{\sqrt[3]{s}Y_\alpha^-} \right. & \\ \left. e^{\sqrt[3]{s}X_\alpha^+} \circ e^{\sqrt[3]{s}Y_\alpha^-} \circ e^{\sqrt[3]{s}X_\alpha^-} \circ e^{\sqrt[3]{s}Y_\alpha^+} \circ e^{\sqrt[3]{s}X_\alpha^+} \circ e^{\sqrt[3]{s}Y_\alpha^-} \circ e^{\sqrt[3]{s}X_\alpha^-} \circ e^{\sqrt[3]{s}Y_\alpha^+} \right) & \end{aligned}$$

Since we shall present similar formulae in subsequent examples, we should indicate where they come from. These are control variations, in fact needle variations, that result from an application of the Campbell-Baker-Hausdorff formula, noting that since the Lie algebra of vector fields is, loosely, the set of *right*-invariant vector fields on the diffeomorphism group, one should apply the Campbell-Baker-Hausdorff formula “backwards.” Thus, for example, to get the formula (6.1), one computes the usual Campbell-Baker-Hausdorff formula for $\log(\exp(\sqrt{s}X_\alpha^+) \exp(\sqrt{s}X_\alpha^-))$. In any event, one can readily see that provided that α is sufficiently large (to be exact, if $\alpha > \frac{34}{3}$), then we have

$$\begin{aligned} 0_{x_0} \in \text{int}(\text{conv}(\alpha f_1(x_0), -\alpha f_1(x_0), \alpha f_2(x_0), -\alpha f_2(x_0), -\frac{1}{3}\alpha^2[f_1, [f_0, f_1]](x_0), \\ -\frac{17}{6}\alpha^2[f_1, [f_0, f_1]](x_0) - \frac{1}{2}\alpha^3[f_1, [f_1, f_2]](x_0))), \end{aligned}$$

provided that $[-\alpha, \alpha]^2 \subset U$. Small-time local controllability of this example for the sufficiently large control set now follows from Theorem 3.7 of [Hirschorn and Lewis 2002]. The lower bound of $\frac{34}{3}$ on the size of the control set to ensure small-time local controllability is undoubtedly not sharp.

We also remark that this example does satisfy condition (iii) of Theorem 4.3. This emphasises the fact that this problem is not properly STLC from x_0 by virtue of a poor choice of local linear generators for the underlying affine subbundle, but due to the fact that the quotient space for $B_{\mathcal{F}}(L^{(3)}(\mathcal{F}_1)_{x_0})$ is too large. We saw in Example 6.1 a situation where dependence on the choice of local linear generators can cause a phenomenon similar on the surface to what we have seen in this example. \square

Our next example shows that one cannot weaken the hypotheses of Theorem 5.4 by only including $\langle \{f_0\}, L^{(1)}(\mathcal{F}_1)_{x_0} \rangle$ in the quotient space defining the quadratic form. One *must* include in the quotient the effects of the associative algebra from *all* vector fields in \mathcal{F}_U that vanish at x_0 .

6.3 EXAMPLE: We consider the system on \mathbb{R}^3 given by

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{x}_2 &= x_2 + x_1 u_2 \\ \dot{x}_3 &= x_1^2 + x_2, \end{aligned}$$

with the control set being any set $U \subset \mathbb{R}^2$ that is convex and proper. (We assume convexity of U for the sake of simplicity, and to ensure that everything we say is correct. One can relax this assumption, since, as Sussmann [1987, Proposition 2.3] shows, if a system is STLC using controls in $\text{conv}(U)$, then it is also STLC using controls in U .) We compute

$$\begin{aligned} [f_0, f_2] &= (0, -x_1, -x_1) & [f_0, f_1] &= (0, 0, -2x_1) \\ [f_1, f_2] &= (0, 1, 0) & [f_1, [f_0, f_1]] &= (0, 0, 2) \\ [f_0, [f_1, f_2]] &= (0, -1, -1) & [f_2, [f_1, f_2]] &= (0, 0, 0) \\ [f_2, [f_0, f_1]] &= (0, 0, 0) & [f_1, [f_1, f_2]] &= (0, 0, 0) \\ [f_1, [f_0, f_2]] &= (0, -1, -1) & \text{ad}_{f_0}^k f_1 &= (0, 0, 0), \quad k \geq 1. \end{aligned}$$

Thus we have

$$\langle \{f_0\}, L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0} = \text{span}_{\mathbb{R}} \{(1, 0, 0), (0, 1, 0)\}.$$

For this system we see that the quadratic form

$$B_{\mathcal{F}}(\langle \{f_0\}, L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0}) \quad (6.2)$$

is definite. Indeed, if we use $\text{span}_{\mathbb{R}} \{(0, 0, 1)\}$ as a model for the quotient then the matrix for the quadratic form in (6.2) is the 1×1 matrix [2].

We claim that this system is STLC from x_0 . We consider the four vector fields

$$X = f_0 + \frac{1}{3}f_1, \quad Y = f_0 - \frac{1}{3}f_1 + f_2.$$

Obviously

$$\frac{1}{3}f_1(x_0) = \left. \frac{d}{ds} \right|_{s=0} e^{sX}(x_0), \quad -\frac{1}{3}f_1(x_0) = \left. \frac{d}{ds} \right|_{s=0} e^{sY}(x_0).$$

Using the machinery in [Hirschorn and Lewis 2002] one then shows that

$$\begin{aligned} \frac{1}{6}[f_1, f_2](x_0) &= \frac{d}{ds} \Big|_{s=0} e^{\sqrt{s}Y} \circ e^{\sqrt{s}X}(x_0) \\ -\frac{1}{6}[f_1, f_2](x_0) &= \frac{d}{ds} \Big|_{s=0} e^{\sqrt{s}X} \circ e^{\sqrt{s}Y}(x_0) \\ -\frac{1}{3}[f_0, [f_1, f_2]](x_0) - \frac{2}{27}[f_1, [f_0, f_1]](x_0) + \frac{4}{9}[f_2, [f_0, f_1]](x_0) &= \\ &= \frac{d}{ds} \Big|_{s=0} e^{\sqrt[3]{s}X} \circ e^{2\sqrt[3]{s}Y} \circ e^{\sqrt[3]{s}X}(x_0) \\ \frac{1}{3}[f_0, [f_1, f_2]](x_0) - \frac{2}{27}[f_1, [f_0, f_1]](x_0) + \frac{1}{9}[f_1, [f_0, f_2]](x_0) &= \\ &= \frac{d}{ds} \Big|_{s=0} e^{\sqrt[3]{s}Y} \circ e^{2\sqrt[3]{s}X} \circ e^{\sqrt[3]{s}Y}(x_0). \end{aligned}$$

One may easily show that

$$\begin{aligned} 0_{x_0} \in \text{int}(\text{conv}(\frac{1}{3}f_1(x_0), -\frac{1}{3}f_1(x_0), \frac{1}{6}[f_1, f_2](x_0), -\frac{1}{6}[f_1, f_2](x_0), \\ -\frac{1}{3}[f_0, [f_1, f_2]](x_0) - \frac{2}{27}[f_1, [f_0, f_1]](x_0) + \frac{4}{9}[f_2, [f_0, f_1]](x_0), \\ \frac{1}{3}[f_0, [f_1, f_2]](x_0) - \frac{2}{27}[f_1, [f_0, f_1]](x_0) + \frac{1}{9}[f_1, [f_0, f_2]](x_0)), \end{aligned}$$

from which it follows that the system is STLC from x_0 by Theorem 3.7 of [Hirschorn and Lewis 2002], at least if $(\pm\frac{1}{3}, 0), (0, 1) \in U$. Furthermore, we have shown that small-time local controllability is possible with piecewise constant controls. Standard homogeneity arguments show that essentially the same computations give small-time local controllability if $(\pm\frac{1}{3}\alpha, 0), (0, \alpha) \in U$ for $\alpha > 0$. If U does not contain these points, then convexity and properness of U ensures that $(\pm\frac{1}{3}\alpha, 0), (0, \alpha) \in U$ for sufficiently small α . Thus this system is properly STLC from x_0 . \square

The next example again deals with possibly weakening the hypotheses of Theorem 5.4. In this example we show that one *must* quotient by the full involutive closure of the input distribution at x_0 .

6.4 EXAMPLE: We consider the system

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{x}_2 &= u_2 \\ \dot{x}_3 &= x_1 u_2 \\ \dot{x}_4 &= x_3 u_2 + \frac{1}{25}(x_1^2 + x_2^2). \end{aligned}$$

The control set we take is again a convex, proper subset $U \subset \mathbb{R}^2$, and we take $x_0 = (0, 0, 0, 0)$. Some brackets for this system are

$$\begin{aligned} [f_1, f_2] &= (0, 0, 1, 0) & [f_2, [f_1, f_2]] &= (0, 0, 0, -1) \\ [f_1, [f_0, f_1]] &= (0, 0, 0, -\frac{2}{25}) & [f_2, [f_0, f_2]] &= (0, 0, 0, -\frac{2}{25}) \\ [f_1, [f_0, f_2]] &= (0, 0, 0, 0) & [f_2, [f_0, f_1]] &= (0, 0, 0, 0). \end{aligned}$$

We also easily see that as a linear map on $T_{x_0}\mathbb{R}^4$ we have $\text{ad}_{f_0} = 0$. Thus for this system we see that the form

$$B_{\mathcal{F}}(\langle Z_{x_0}(\mathcal{F}_U), L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(2)}(\mathcal{F}_1)_{x_0})$$

is definite, but the form

$$B_{\mathcal{F}}(\langle Z_{x_0}(\mathcal{F}_U), L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0})$$

is not, noting that $L^{(\infty)}(\mathcal{F}_1)_{x_0} = L^{(3)}(\mathcal{F}_1)_{x_0}$. Thus this example represents a weakening of the hypotheses we are after.

We claim that this system is STLC from x_0 . To show this, we again use the techniques of [Hirschorn and Lewis 2002]. We define vector fields

$$\begin{aligned} X^+ &= f_0 + f_1, & X^- &= f_0 - f_1, \\ Y^+ &= f_0 + f_2, & Y^- &= f_0 - f_2. \end{aligned}$$

Clearly

$$\begin{aligned} f_1(x_0) &= \frac{d}{ds} \Big|_{s=0} e^{sX^+}(x_0), & -f_1(x_0) &= \frac{d}{ds} \Big|_{s=0} e^{sX^-}(x_0), \\ f_2(x_0) &= \frac{d}{ds} \Big|_{s=0} e^{sY^+}(x_0), & -f_2(x_0) &= \frac{d}{ds} \Big|_{s=0} e^{sY^-}(x_0). \end{aligned}$$

Following [Hirschorn and Lewis 2002] one can show that

$$\begin{aligned} [f_1, f_2](x_0) &= \frac{d}{ds} \Big|_{s=0} \left(e^{\sqrt{s}Y^-} \circ e^{\sqrt{s}X^-} \circ e^{\sqrt{s}Y^+} \circ e^{\sqrt{s}X^+}(x_0) \right) \\ -[f_1, f_2](x_0) &= \frac{d}{ds} \Big|_{s=0} \left(e^{\sqrt{s}X^-} \circ e^{\sqrt{s}Y^-} \circ e^{\sqrt{s}X^+} \circ e^{\sqrt{s}Y^+}(x_0) \right). \end{aligned}$$

In like manner, one shows that

$$\begin{aligned} &4[f_0, [f_1, f_2]](x_0) - \frac{5}{3}[f_1, [f_0, f_1]](x_0) - [f_1, [f_0, f_2]](x_0) - [f_2, [f_0, f_1]](x_0) \\ &\quad - \frac{5}{3}[f_2, [f_0, f_2]](x_0) = \frac{d}{ds} \Big|_{s=0} \left(e^{\sqrt[3]{s}Y^-} \circ e^{\sqrt[3]{s}X^-} \circ e^{\sqrt[3]{s}Y^+} \circ e^{\sqrt[3]{s}X^+} \circ \right. \\ &\quad \left. e^{\sqrt[3]{s}X^-} \circ e^{\sqrt[3]{s}Y^-} \circ e^{\sqrt[3]{s}X^+} \circ e^{\sqrt[3]{s}Y^+}(x_0) \right) \\ -[f_0, [f_1, f_2]](x_0) - \frac{5}{3}[f_1, [f_0, f_1]](x_0) - [f_1, [f_0, f_2]](x_0) - 3[f_2, [f_0, f_1]](x_0) \\ &\quad - 6[f_2, [f_0, f_2]](x_0) + [f_2, [f_1, f_2]](x_0) = \\ &\quad \frac{d}{ds} \Big|_{s=0} \left(e^{\sqrt[3]{s}X^-} \circ e^{\sqrt[3]{s}Y^-} \circ e^{\sqrt[3]{s}X^+} \circ e^{\sqrt[3]{s}Y^+} \circ e^{2\sqrt[3]{s}Y^-} \circ \right. \\ &\quad \left. e^{\sqrt[3]{s}X^-} \circ e^{\sqrt[3]{s}Y^+} \circ e^{\sqrt[3]{s}X^+} \circ e^{\sqrt[3]{s}Y^+}(x_0) \right). \end{aligned}$$

One may readily verify that

$$\begin{aligned} 0_{x_0} \in \text{int}(\text{conv}(f_1(x_0), -f_1(x_0), f_2(x_0), -f_2(x_0), [f_1, f_2](x_0), -[f_1, f_2](x_0), \\ 4[f_0, [f_1, f_2]](x_0) - \frac{5}{3}[f_1, [f_0, f_1]](x_0) - [f_1, [f_0, f_2]](x_0) \\ [f_2, [f_0, f_1]](x_0) - \frac{5}{3}[f_2, [f_0, f_2]](x_0), -[f_0, [f_1, f_2]](x_0) - \frac{5}{3}[f_1, [f_0, f_1]](x_0) \\ - [f_1, [f_0, f_2]](x_0) - 3[f_2, [f_0, f_1]](x_0) - 6[f_2, [f_0, f_2]](x_0) + [f_2, [f_1, f_2]](x_0)), \end{aligned}$$

so it follows from Theorem 3.7 of [Hirschorn and Lewis 2002] that the system is STLC from x_0 , provided that $(\pm 1, 0), (0, \pm 1) \in U$. Arguing as in Example 6.3, we may deduce that the system is STLC for any proper U . \square

6.5 REMARK: Note that while Example 6.2 is, for small control sets, uncontrollable with the involutive closure of the inputs having maximal rank, Example 6.4 is STLC with the same property. This shows that the property of the involutive closure of the input distribution having maximal rank has, *per se*, no bearing on small-time local controllability, at least with bounded controls. If one allows unbounded controls, then Pomet [1999] shows the property of the input distribution having maximal rank at x_0 implies one can follow arbitrarily closely in the C^0 topology any curve near x_0 . Along similar lines, Bianchini and Stefani [1993] show that if $U = \mathbb{R}^m$ then the subspace $L^{(\infty)}(\mathcal{F}_1)_{x_0}$ is comprised of regular variations. Example 6.2 shows that for such systems small-time local controllability can follow not for unbounded controls, but merely for sufficiently large controls. \square

The example we look at now examines what can happen when x_0 is not a regular point for $L^{(\infty)}(\mathcal{F}_1)_{x_0}$.

6.6 EXAMPLE: With $M = \mathbb{R}^3$ and $x_0 = (0, 0, 0)$ we consider the system

$$\begin{aligned}\dot{x}_1 &= \frac{1}{5}u_1 \\ \dot{x}_2 &= x_1 + x_3u_2 \\ \dot{x}_3 &= x_1^2 + x_2u_2.\end{aligned}$$

As U we allow any subset of \mathbb{R}^2 that is convex and proper. For this example we have the following bracket computations:

$$\begin{aligned}[f_0, f_1] &= (0, -\frac{1}{5}, -\frac{2}{5}x_1) & [f_1, f_2] &= (0, 0, 0) \\ [f_1, [f_0, f_1]] &= (0, 0, -\frac{2}{25}) & [f_2, [f_0, f_1]] &= (0, \frac{2}{5}x_1, \frac{1}{5}).\end{aligned}$$

Note that

$$\langle Z_{x_0}(\mathcal{F}_U), L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0} = \text{span}_{\mathbb{R}} \{(1, 0, 0), (0, 1, 0)\},$$

and that the form of part (iv) of Theorem 5.4 is represented by the matrix $[-\frac{2}{25}]$, and so is definite. However, x_0 is not a regular point for $L^{(\infty)}(\mathcal{F}_1)$, so all hypotheses for Theorem 5.4 hold except part (iii).

We claim that this example is STLC from x_0 . To show this define

$$\begin{aligned}X^+ &= f_0 + f_1, & X^- &= f_0 - f_1, \\ Y^+ &= f_0 + f_2, & Y^- &= f_0 - f_2.\end{aligned}$$

We then have

$$f_1(x_0) = \frac{d}{ds}\Big|_{s=0} e^{sX^+}(x_0), \quad -f_1(x_0) = \frac{d}{ds}\Big|_{s=0} e^{sX^-}(x_0).$$

We also compute

$$\begin{aligned}[f_0, f_1](x_0) &= \frac{d}{ds}\Big|_{s=0} \left(e^{\sqrt{s}X^+} \circ e^{\sqrt{s}X^-}(x_0) \right) \\ -[f_0, f_1](x_0) &= \frac{d}{ds}\Big|_{s=0} \left(e^{\sqrt{s}X^-} \circ e^{\sqrt{s}X^+}(x_0) \right) \\ -\frac{2}{3}[f_1, [f_0, f_1]](x_0) &= \frac{d}{ds}\Big|_{s=0} \left(e^{\sqrt[3]{s}X^+} \circ e^{2\sqrt[3]{s}X^-} \circ e^{\sqrt[3]{s}X^+}(x_0) \right) \\ -\frac{2}{3}[f_1, [f_0, f_1]](x_0) - [f_2, [f_0, f_1]](x_0) &= \\ &= \frac{d}{ds}\Big|_{s=0} \left(e^{\sqrt[3]{s}X^+} \circ e^{\sqrt[3]{s}X^-} \circ e^{\sqrt[3]{s}Y^-} \circ e^{\sqrt[3]{s}X^-} \circ e^{\sqrt[3]{s}X^+} \circ e^{\sqrt[3]{s}Y^+} \right).\end{aligned}$$

One now checks that

$$\begin{aligned}0_{x_0} \in \text{int}(\text{conv}(f_1(x_0), -f_1(x_0), [f_0, f_1](x_0), -[f_0, f_1](x_0), -\frac{2}{3}[f_1, [f_0, f_1]](x_0), \\ -\frac{2}{3}[f_1, [f_0, f_1]](x_0) - [f_2, [f_0, f_1]](x_0))),\end{aligned}$$

giving small-time local controllability from x_0 by Theorem 3.7 of [Hirschorn and Lewis 2002]. Again, the usual scaling and convexity arguments give small-time local controllability for arbitrary proper U . \square

Our next example satisfies the hypotheses of neither of Theorems 5.2 and 5.4, and can be made to be STLC or not by the choice of a parameter. Thus this is an example of a system for which small-time local controllability must be determined using higher-order conditions.

6.7 EXAMPLE: We consider the system in \mathbb{R}^3 defined by

$$\begin{aligned}\dot{x}_1 &= u_1 \\ \dot{x}_2 &= u_2 \\ \dot{x}_3 &= x_1^2 + \alpha x_2^4,\end{aligned}$$

with a convex, proper control set $U \subset \mathbb{R}^2$, and with $x_0 = (0, 0, 0)$. Some of the useful Lie brackets are

$$\begin{aligned}[f_1, [f_0, f_1]] &= (0, 0, -2) & [f_2, [f_0, f_2]] &= (0, 0, -12\alpha x_2^2) \\ [f_1, [f_0, f_2]] &= [f_2, [f_0, f_1]] = (0, 0, 0) & [f_0, f_1] &= (0, 0, -2x_1) \\ [f_0, f_2] &= (0, 0, -4\alpha x_2^3), & \text{ad}_{f_0}^k f_1 = \text{ad}_{f_0}^k f_2 &= (0, 0, 0), \quad k \geq 2.\end{aligned}$$

Thus we see that

$$\begin{aligned}\langle Z_{x_0}(\mathcal{F}_U), L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0} = \\ \langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle = \text{span}\{(1, 0, 0), (0, 1, 0)\}.\end{aligned}$$

Therefore, thinking of the quotient $T_{x_0}\mathbb{R}^3 / \text{span}\{(1, 0, 0), (0, 1, 0)\}$ as being \mathbb{R} in the natural manner, the quadratic forms of Theorems 5.2 and 5.4 are the same, and are simply \mathbb{R} -valued quadratic forms. They are represented by the matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus we see that the form is semidefinite, and thus neither Theorem 5.2 nor Theorem 5.4 apply.

However, as is shown in [Hirschorn and Lewis 2002], this system is in fact STLC from x_0 when $\alpha < 0$. Let us record exactly how this may be shown. We define

$$\begin{aligned}X^+ &= f_0 + f_1, & X^- &= f_0 - f_1, \\ Y^+ &= f_0 + f_2, & Y^- &= f_0 - f_2.\end{aligned}$$

We clearly have

$$\begin{aligned} f_1(x_0) &= \frac{d}{ds} \Big|_{s=0} e^{sX^+}(x_0), & -f_1(x_0) &= \frac{d}{ds} \Big|_{s=0} e^{sX^-}(x_0), \\ f_2(x_0) &= \frac{d}{ds} \Big|_{s=0} e^{sY^+}(x_0), & -f_2(x_0) &= \frac{d}{ds} \Big|_{s=0} e^{sY^-}(x_0). \end{aligned}$$

One may also show that

$$\begin{aligned} -\frac{1}{3}[f_1, [f_0, f_1]](x_0) &= \frac{d}{ds} \Big|_{s=0} \left(e^{\mathfrak{V}_{\bar{s}}X^-} \circ e^{\mathfrak{V}_{\bar{s}}X^+}(x_0) \right) \\ -\frac{1}{60}[f_2, [f_2, [f_2, [f_0, f_2]]]](x_0) &= \frac{d}{ds} \Big|_{s=0} \left(e^{\mathfrak{V}_{\bar{s}}Y^-} \circ e^{\mathfrak{V}_{\bar{s}}Y^+}(x_0) \right). \end{aligned}$$

Using the fact that

$$[f_2, [f_2, [f_2, [f_0, f_2]]]](x_0) = (0, 0, -24\alpha),$$

one sees that

$$\begin{aligned} 0_{x_0} \in \text{int}(\text{conv}(f_1(x_0), -f_1(x_0), f_2(x_0), -f_2(x_0), \\ -\frac{1}{3}[f_1, [f_0, f_1]](x_0), -\frac{1}{60}[f_2, [f_2, [f_2, [f_0, f_2]]]](x_0))), \end{aligned}$$

thus giving small-time local controllability from x_0 if $(1, 0), (0, 1) \in U$. Again, scaling arguments give us small-time local controllability for U proper.

For $\alpha \geq 0$, the system is obviously not STLC from x_0 . \square

7. Proofs for the control-affine system results

Let us first prove the results for the usual representation of control-affine systems, as given in Section 3.2. Thus we shall prove Theorems 5.2 and 5.4. This section contains the main technical content of the paper. The techniques used are adaptations of those that are well-established in the existing control literature, with the variational cone playing a significant rôle in the proof of Theorem 5.2, and with the Chen-Fliess-Sussmann series being the key tool in the proof of Theorem 5.4.

7.1. Proof of Theorem 5.2. As we stated in Remark 5.3–2, Theorem 5.2 has almost been stated previously in the literature [Basto-Gonçalves 1998]. However, the manner in which we state the result is not exactly the same as existing statements, and the hypotheses are slightly weaker. In this section we make the appropriate extensions to the existing techniques concerning the variational cone. The main idea in the proof is the demonstration in Proposition 7.6 of the equivalence of the condition (iv) of Theorem 5.2 with a neutralisation condition. We first note that Proposition 2.3 of [Sussmann 1987] allows us to assume that U is closed and convex, and we do so as needed.

The approach we take relies on the notion of a variation. For $x_0 \in M$ and $\alpha > 0$, suppose that $c: [0, T] \rightarrow M$ is a curve satisfying $c(0) = x_0$. The curve c is **tangent to order α** at x_0 if, in a coordinate chart, the limit

$$\lim_{\epsilon \downarrow 0} \frac{c(\epsilon) - x_0}{\epsilon^\alpha} \quad (7.1)$$

exists and is nonzero. One readily verifies that this definition is independent of coordinate chart if M is C^∞ . Clearly c is tangent to order α at x_0 for at most one α , and in this case the limit (7.1) is a tangent vector at x_0 , and is the **α th-order tangent vector to c at x_0** .

We introduce the notions surrounding the variational cone, following Bianchini and Stefani [1993].

7.1 DEFINITION: Let $\Sigma = (M, \mathcal{F}, U)$ be a control-affine system with $(\bar{\xi}, \bar{u}) \in \text{Traj}(\Sigma, T)$ a reference trajectory.

- (i) A **control variation** is a continuous map $\mu: [0, c_0] \times [0, s_0] \rightarrow L_1([0, T], U)$, defined for some $c_0, s_0 > 0$.
- (ii) Let $\alpha > 0$ and $t \in [0, T[$. A vector $v \in T_{\bar{\xi}(t)}M$ is a **variation of order α at time t** if there exists a control variation μ so that cv is the α th-order tangent vector at $\xi(t)$ to the curve

$$s \mapsto \xi(\bar{\xi}(t), \mu(c, s), t, t + s), \bar{u}, t, t - s).$$

Let $\mathcal{V}_{\bar{\xi}}(t) \subset T_{x(t)}M$ denote the set of all variations of any order at time t .

- (iii) The **variational cone** of $(\bar{\xi}, \bar{u})$ at time $t \in [0, T[$ is given by

$$\mathcal{K}_{(\bar{\xi}, \bar{u})}(t) = \begin{cases} \text{conv}(\mathcal{V}_{\bar{\xi}}(0)), & t = 0 \\ \text{conv}(\cup_{\tau \in [0, t]} \{T_{\bar{\xi}(\tau)}\xi(\cdot, \bar{u}, \tau, \tau)(v) \mid v \in \mathcal{V}_{\bar{\xi}}(\tau)\}), & t > 0. \end{cases}$$

\square

Let us recall some basic facts about variations. For this, we suppose that we have a reference trajectory $(\bar{\xi}, \bar{u}) \in \text{Traj}(\Sigma, T)$.

1. For each $t \in [0, T[$, $\mathcal{V}_{\bar{\xi}}(t)$ depends only on $\bar{\xi}$ and not on \bar{u} .
2. Generally, for each $t \in [0, T[$, $\mathcal{K}_{(\bar{\xi}, \bar{u})}(t)$ depends on the reference control \bar{u} .
3. The set of variations is a convex cone [Bianchini and Stefani 1993, Corollary 2.6].
4. If the reference trajectory $(\bar{\xi}, \bar{u})$ is stationary, i.e., $\bar{\xi}(t) = x_0$, then $\mathcal{V}_{\bar{\xi}}(t_1) = \mathcal{V}_{\bar{\xi}}(t_2)$ for $t_1, t_2 \in [0, T[$ [Bianchini and Stefani 1993, bottom page 909].
5. If the reference trajectory is stationary then $\mathcal{K}_{(\bar{\xi}, \bar{u})}(t_1) \subset \mathcal{K}_{(\bar{\xi}, \bar{u})}(t_2)$ for $t_1 < t_2$ (this follows from the preceding statement).
6. If the reference trajectory is stationary and if there exists $T > 0$ so that $\mathcal{K}_{(\bar{\xi}, \bar{u})}(t) = T_{x_0}M$ for all $t \in]0, T[$, then Σ is STLC from x_0 [Bianchini and Stefani 1993, Theorem 1.10].
7. $L^{(2)}(\mathcal{F}_1)_{x_0} \subset \mathcal{K}_{(\bar{\xi}, \bar{u})}(t)$ for each $t \in [0, T[$ [Bianchini and Stefani 1993, Proposition 3.6].

Let us consider more carefully the case where we have a stationary reference trajectory $\bar{\xi}(t) = x_0$. Such a trajectory may be realised with a variety of reference controls, each having to take values in $\ker(L_{x_0}^{\mathcal{F}}) \cap U$. Motivated by this, let $\mathcal{U}(x_0, t)$ be the collection of inputs, defined on an interval $[0, T]$ for $T > t$, taking values in $\ker(L_{x_0}^{\mathcal{F}}) \cap U$. We define

$$\mathcal{K}_{x_0}(t) = \text{conv}(\cup_{\bar{u} \in \mathcal{U}(x_0, t)} \mathcal{K}_{x_0, \bar{u}}(t)).$$

Clearly, if $\mathcal{K}_{x_0}(t) = T_{x_0}M$ for some $t > 0$ we may still assert that Σ is STLC at x_0 .

We now show that fact 7 immediately implies that $\langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F}_1)_{x_0} \rangle \subset \mathcal{K}_{x_0}(t)$ for all positive t . The following lemma is a generalisation of Theorem 2.4 in [Bianchini and Stefani 1993].

7.2 LEMMA: *Let $t_0 > 0$ and let $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ be a control-affine system with $\mathcal{U}(x_0, t_0) \neq \emptyset$ for each $t \in]0, t_0]$. If $W_{x_0} \subset \mathcal{K}_{x_0}(t)$ is a subspace for each $t \in]0, t_0]$, then $\langle Z_{x_0}(\mathcal{F}_U), W_{x_0} \rangle \subset \mathcal{K}_{x_0}(t_0)$.*

Proof: The lemma will follow if we can show that for any $t_0 > 0$, for any $f_{u_1}, \dots, f_{u_k} \in Z_{x_0}(\mathcal{F}_U)$, and for any $w \in W_{x_0}$,

$$\text{ad}_{f_{u_1}} \circ \dots \circ \text{ad}_{f_{u_k}}(w) \in \mathcal{K}_{x_0}(t_0). \quad (7.2)$$

Let $t_1, \dots, t_k \in]0, \frac{t_0}{k}]$. For $j \in \{1, \dots, k\}$ we have

$$f_{u_j} = f_0 + \sum_{a=1}^m u_j^a f_a,$$

defining $u_j \in \ker(L_{x_0}^{\mathcal{F}}) \cap U$, $j \in \{1, \dots, k\}$. For $\mathbf{t} = (t_1, \dots, t_k)$ define

$$\tau_0(\mathbf{t}) = t_0 - \sum_{j=1}^k t_j, \quad \tau_j(\mathbf{t}) = \tau_0(\mathbf{t}) + \sum_{r=1}^j t_r, \quad j \in \{1, \dots, k\}.$$

We define a reference control $u_{\mathbf{t}}$ by

$$u_{\mathbf{t}}(t) = \begin{cases} u_1, & t \in [0, \tau_0(\mathbf{t})] \\ u_j, & t \in]\tau_{j-1}(\mathbf{t}), \tau_j(\mathbf{t})]. \end{cases}$$

By definition of $\mathcal{K}_{x_0, u_{\mathbf{t}}}(t_0)$ we have

$$T_{x_0}\xi(\cdot, u_{\mathbf{t}}, t_0, \tau_0(\mathbf{t}))(w) \in \mathcal{K}_{x_0, u_{\mathbf{t}}}(t_0).$$

This then defines a subspace

$$\text{conv} \{T_{x_0}\xi(\cdot, u_{\mathbf{t}}, t_0, \tau_0(\mathbf{t}))(w) \mid \mathbf{t} \in]0, \frac{t_0}{k}]^k, w \in W_{x_0}\} \subset \mathcal{K}_{x_0}(t_0).$$

We claim that this subspace contains the tangent vector (7.2). Indeed, one may verify from the Campbell-Baker-Hausdorff formula that

$$T_{x_0}\xi(\cdot, u_{\mathbf{t}}, t_0, \tau_0(\mathbf{t}))(w) = \sum_{j_1=0}^{\infty} \dots \sum_{j_k=0}^{\infty} \frac{t_1^{j_1}}{j_1!} \dots \frac{t_k^{j_k}}{j_k!} \text{ad}_{f_{u_1}}^{j_1} \circ \dots \circ \text{ad}_{f_{u_k}}^{j_k}(w).$$

We now see that

$$\left. \frac{d}{dt_1} \right|_{t_1=0} \dots \left. \frac{d}{dt_k} \right|_{t_k=0} T_{x_0}\xi(\cdot, u_{\mathbf{t}}, t_0, \tau_0(\mathbf{t}))(w) = \text{ad}_{f_{u_1}}^{j_1} \circ \dots \circ \text{ad}_{f_{u_k}}^{j_k}(w),$$

giving the lemma. \blacksquare

The lemma tells us that the variational cone contains those tangent vectors at x_0 which form the quotient subspace used in the definition of the quadratic form in condition (iv) of Theorem 5.2. As pointed out by Sussmann [1987], the brackets of the form $[f_a, [f_0, f_a]]$ are obstructions to local controllability, essentially because one cannot move along these brackets in both directions. However, it is possible that these obstructions together may help one another, or at least not cause harm. We make this precise with the following definition.

7.3 DEFINITION: A control-affine system $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ with $f_0(x_0) = 0_{x_0}$ is *second-order neutralisable* at x_0 if

$$\sum_{a=1}^m [f_a, [f_0, f_a]](x_0) \in \langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle. \quad \square$$

The notion of second-order neutralisability is manifestly *not* feedback invariant. We shall now show that condition (iv) of Theorem 5.2 is an intrinsic formulation of second-order neutralisability. First, it is convenient to prove a couple of technical lemmas. The first lemma will be helpful in giving us what we are after here, as well as being useful in the proof of Theorem 5.4 that we give in Section 7.2.

7.4 LEMMA: *Let $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ be a control-affine system with U convex and proper and $f_0(x_0) = 0_{x_0}$. There exists $R \in GL(m; \mathbb{R})$ so that if we define*

- (i) $h_0 = f_0$,
 - (ii) $h_a = \sum_{b=1}^m R_b^a f_b$,
 - (iii) $\mathcal{H} = \{h_0, h_1, \dots, h_m\}$, and
 - (iv) $\tilde{U} = \{R^{-1}u \mid u \in U\}$,
- then, as subsets of $\text{End}(T_{x_0}M)$,

$$\text{span}_{\mathbb{R}}\{Z_{x_0}(\mathcal{H}_{\tilde{U}})\} = \text{span}_{\mathbb{R}}\{Z_{x_0}(\mathcal{H})\}.$$

Proof: In the proof of Proposition 3.8 we constructed $R \in GL(m; \mathbb{R})$ so that if \mathcal{H} is defined as in the statement of the lemma we have

$$Z_{x_0}(\mathcal{H}) = \{h_0, h_1, \dots, h_{\ell}\}, \quad \tilde{Z}_{x_0}(\mathcal{H}) = \{h_{\ell+1}, \dots, h_m\},$$

for some $\ell \in \{0, \dots, m-1\}$. We claim that the conclusions of the lemma hold with R so defined. Since U is convex and proper so is \tilde{U} . Thus for each $a \in \{1, \dots, \ell\}$ the control ρe_a lies in \tilde{U} for some sufficiently small ρ . With ρ so chosen, $\rho h_a(x_0) \in Z_{x_0}(\mathcal{H}_{\tilde{U}})$. Thus we have

$$\text{span}_{\mathbb{R}}\{h_0(x_0), h_1(x_0), \dots, h_{\ell}(x_0)\} \subset \text{span}_{\mathbb{R}}\{Z_{x_0}(\mathcal{H}_{\tilde{U}})\},$$

since $h_0 \in Z_{x_0}(\mathcal{H}_{\tilde{U}})$. Now let $\tilde{u} \in \tilde{U}$ and suppose that

$$h_0 + \sum_{a=1}^m \tilde{u}_a h_a \in Z_{x_0}(\mathcal{H}_{\tilde{U}}).$$

Then

$$h_0(x_0) + \sum_{a=1}^m \tilde{u}_a h_a(x_0) = \sum_{a=\ell+1}^m \tilde{u}_a h_a(x_0) = 0_{x_0}.$$

Since the vectors $\{h_{\ell+1}(x_0), \dots, h_m(x_0)\}$ are linearly independent this implies that $\tilde{u}_{\ell+1} = \dots = \tilde{u}_m = 0$. This shows that

$$\text{span}_{\mathbb{R}}\{Z_{x_0}(\mathcal{H}_{\tilde{V}})\} \subset \text{span}_{\mathbb{R}}\{h_0, h_1, h_1, \dots, h_{\ell}\}.$$

This completes the proof. \blacksquare

The next lemma characterises indefinite \mathbb{R} -valued bilinear maps. We actually prove more in the lemma than is actually needed, as the stronger results are of independent interest. For example, they are used by [Bullo, Lewis, and Lynch \[2002\]](#) to characterise so-called decoupling vector fields that are useful in motion planning for some mechanical control systems.

7.5 LEMMA: *Let V be a finite-dimensional \mathbb{R} -vector space and let $B \in \Sigma_2(V; \mathbb{R})$. For a basis $\mathcal{V} = \{v_1, \dots, v_n\}$ for V , let $B_{\mathcal{V}}$ be the $n \times n$ matrix with components $(B_{\mathcal{V}})_{ij} = B(v_i, v_j)$, $i, j = 1, \dots, n$. The following statements are equivalent:*

- (i) *there exists a basis \mathcal{V} for V for which the sum of the diagonal entries in the matrix $B_{\mathcal{V}}$ is zero;*
- (ii) *there exists a basis \mathcal{V} for V for which the diagonal entries in the matrix $B_{\mathcal{V}}$ are all zero;*
- (iii) *B is indefinite.*

Proof: We will assume that $\text{rank}(B) = n$ since if $\text{rank}(B) < n$ then we will have proved the proposition for the induced nondegenerate form on the quotient $V/\ker(B)$. One may then see that the degeneracy does not alter the result.

(i) \implies (ii) We prove this implication by induction on n . The result is vacuous when $n = 1$. For $n = 2$ we let $\mathcal{V} = \{v_1, v_2\}$ be a basis for which the matrix $B_{\mathcal{V}}$ has a diagonal that sums to zero. By rescaling v_1 and v_2 if necessary we may assume that

$$B_{\mathcal{V}} = \begin{bmatrix} 1 & a \\ a & -1 \end{bmatrix}$$

for some $a \in \mathbb{R}$. If we define a basis $\mathcal{W} = \{w_1, w_2\}$ by

$$\begin{aligned} w_1 &= \frac{-1 - a^2 + a\sqrt{1+a^2}}{2\sqrt[3]{(1+a^2)^2}}v_1 - \frac{1}{2(1+a^2)}v_2 \\ w_2 &= (-a - \sqrt{1+a^2})v_1 + v_2, \end{aligned}$$

then we compute

$$B_{\mathcal{W}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

giving this part of the lemma when $n = 2$. Now suppose this part of the lemma true for $n = k$ and let B be a full rank symmetric bilinear map for which there exists a basis $\mathcal{V} = \{v_1, \dots, v_{k+1}\}$ where the sum of the diagonals of $B_{\mathcal{V}}$ is zero. Let us write

$$B_{\mathcal{V}} = \begin{bmatrix} B_0 & a \\ a^t & \alpha \end{bmatrix}$$

for $a \in \mathbb{R}^k$ and $\alpha \in \mathbb{R}$. If $\alpha = 0$ then this part of the lemma is proved, since by the induction hypotheses we may find a basis $\mathcal{W} = \{w_1, \dots, w_k, w_{k+1} = v_{k+1}\}$ so that all diagonal entries in $B_{\mathcal{W}}$ are zero. So we assume that $\alpha \neq 0$. The sum of the diagonals of B_0 must then be $-\alpha$. Therefore, there must be an $i_0 \in \{1, \dots, k\}$ for which $d_{i_0} \triangleq B(e_{i_0}, e_{i_0})$ is nonzero, and whose sign is the opposite of that of α . Fix any such i_0 . For $i \in \{1, \dots, k\}$ define

$$\beta_i = -\alpha^{-1}(a_i + \sqrt{a_i^2 + \frac{\alpha^2}{k}})$$

and define

$$\beta_{k+1} = -d_{i_0}^{-1}(a_{i_0} + \sqrt{a_{i_0}^2 - \alpha d_{i_0}}).$$

Now define a basis $\tilde{\mathcal{W}} = \{\tilde{w}_1, \dots, \tilde{w}_{k+1}\}$ for V by

$$\tilde{w}_1 = v_1 + \beta_1 v_{k+1}, \dots, \tilde{w}_k = v_k + \beta_k v_{k+1}, \tilde{w}_{k+1} = v_{k+1} + \beta_{k+1} v_{i_0},$$

and compute

$$\begin{aligned} B(\tilde{w}_i, \tilde{w}_i) &= B(v_i, v_i) + \frac{1}{k}\alpha, & i \in \{1, \dots, k\} \\ B(\tilde{w}_{k+1}, \tilde{w}_{k+1}) &= 0. \end{aligned}$$

Now note that the sum of the diagonals of $B_{\tilde{\mathcal{W}}}$ is zero. Therefore, by the induction hypothesis, we may make a change of basis to $\mathcal{W} = \{w_1, \dots, w_k, w_{k+1} = \tilde{w}_{k+1}\}$ so that

$$B_{\mathcal{W}} = \begin{bmatrix} \tilde{B}_0 & \tilde{a} \\ \tilde{a}^t & 0 \end{bmatrix}$$

where the diagonal entries in \tilde{B}_0 are all zero. This proves this part of the lemma.

(ii) \implies (iii) We proceed by induction on n . In the case when $n = 1$ there is nothing to prove. When $n = 2$ let $\mathcal{V} = \{v_1, v_2\}$ be a basis for which $B_{\mathcal{V}}$ has a zero diagonal. Without loss of generality, by rescaling v_1 and v_2 if necessary, we may assume that

$$B_{\mathcal{V}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

If we define a basis $\mathcal{W} = \{w_1 = \frac{1}{\sqrt{2}}(v_1 + v_2), w_2 = \frac{1}{\sqrt{2}}(v_1 - v_2)\}$ for V we see that

$$B_{\mathcal{W}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Thus our assertion holds when $n = 2$. Now suppose it holds for $n = k$ and let B be a full rank quadratic form on a $(k+1)$ -dimensional vector space V with $\mathcal{V} = \{v_1, \dots, v_{k+1}\}$ a basis for V satisfying (ii). Then we write

$$B_{\mathcal{V}} = \begin{bmatrix} B_0 & a \\ a^t & 0 \end{bmatrix},$$

where $B_0 \in \mathbb{R}^{k \times k}$ is symmetric with zero diagonal and $a \in \mathbb{R}^k$. By our induction hypothesis we may choose a basis $\tilde{\mathcal{W}} = \{\tilde{w}_1, \dots, \tilde{w}_k, \tilde{w}_{k+1} = v_{k+1}\}$ with the property that

$$B_{\tilde{\mathcal{W}}} = \begin{bmatrix} \tilde{B}_0 & \tilde{a} \\ \tilde{a}^t & 0 \end{bmatrix},$$

where $\tilde{B}_0 \in \mathbb{R}^{k \times k}$ is diagonal with all diagonal entries being +1 or -1, and with not all diagonal entries having the same sign, and $\tilde{a} \in \mathbb{R}^k$. Next define a basis $\mathcal{W} = \{w_1, \dots, w_{k+1}\}$ for V by

$$w_1 = \tilde{w}_1, \dots, w_k = \tilde{w}_k, w_{k+1} = \tilde{w}_{k+1} - \sum_{i=1}^k s_i \tilde{a}_i \tilde{w}_i,$$

where

$$s_i = \begin{cases} -1, & \text{the } (i, i) \text{ element of } \tilde{B}_0 \text{ is } -1 \\ 1, & \text{the } (i, i) \text{ element of } \tilde{B}_0 \text{ is } 1. \end{cases}$$

One checks by direct calculation that

$$B(w_i, w_{k+1}) = 0, \quad i = 1, \dots, k, \quad B(w_{k+1}, w_{k+1}) = -2 \sum_{i=1}^k s_i \tilde{a}_i^2.$$

Therefore, by scaling w_{k+1} , the matrix $B_{\mathcal{W}}$ will be diagonal with all diagonal entries being +1 or -1, and not all diagonal entries will have the same sign. This is the content of (iii), and so our assertion is proven.

(iii) \implies (i) Condition (iii), with our assumption that B be full rank, is the assertion that we may find a basis $\mathcal{W} = \{w_1, \dots, w_n\}$ for V where the matrix $B_{\mathcal{W}}$ is diagonal with all diagonal entries being either +1 or -1, and not all diagonal entries having the same sign. Let us treat the case where there are more +1's than -1's on the diagonal; the other case follows in exactly the same manner. By permuting the basis elements if necessary, we may then assume that

$$B_{\mathcal{W}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Thus each of the (say) k -1's is paired with a +1 on the diagonal, with the remaining diagonal entries being +1. We define a basis $\mathcal{V} = \{v_1, \dots, v_n\}$ by

$$v_1 = w_1 + w_2, \quad v_2 = w_1 - w_2, \dots, v_{2k-1} = w_{2k-1} + w_{2k}, \quad v_{2k} = w_{2k-1} - w_{2k}, \\ v_{2k+1} = w_{2k+1} + w_{2k}, \dots, v_n = w_n + w_{2k}.$$

One then checks by direct calculation that the diagonal entries in the matrix $B_{\mathcal{V}}$ are zero, which completes the proof. \blacksquare

With these results at hand, we may now show how the condition (iv) of Theorem 5.2 is equivalent to second-order neutralisability.

7.6 PROPOSITION: *Let $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ be a control-affine system with $f_0(x_0) = 0_{x_0}$ and with U convex and proper. The following two statements are equivalent:*

(i) $B_{\mathcal{F}}(\langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle)$ is indefinite;

(ii) there exists $\tilde{m} \geq m$, a full rank matrix $T \in \mathbb{R}^{m \times \tilde{m}}$, and a convex proper subset $\tilde{U} \subset \mathbb{R}^{\tilde{m}}$ so that if $\tilde{\Sigma} = (M, \tilde{\mathcal{F}} = \{\tilde{f}_0 = f_0, \tilde{f}_1, \dots, \tilde{f}_{\tilde{m}}\}, \tilde{U})$ is defined by

$$\tilde{f}_{\alpha} = \sum_{b=1}^m T_{\alpha}^b f_b, \quad \alpha = 1, \dots, \tilde{m},$$

then $\tilde{\Sigma}$ is second-order neutralisable at x_0 .

Proof: Throughout the proof we shall use the fact that if $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ and $\tilde{\Sigma} = (M, \tilde{\mathcal{F}} = \{\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{\tilde{m}}\}, \tilde{U})$ have the property that

$$\tilde{f}_0 = f_0, \quad \tilde{f}_{\alpha} = \sum_{a=1}^m T_{\alpha}^a f_a, \quad \alpha \in \{1, \dots, \tilde{m}\}$$

for a full rank matrix $T \in \mathbb{R}^{m \times \tilde{m}}$, and that both U and \tilde{U} are convex and proper, then the following equalities hold:

$$\langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle = \langle Z_{x_0}(\tilde{\mathcal{F}}_{\tilde{U}}), L^{(2)}(\tilde{\mathcal{F}})_{x_0} \rangle \\ B_{\mathcal{F}}(\langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle) = B_{\tilde{\mathcal{F}}}(\langle Z_{x_0}(\tilde{\mathcal{F}}_{\tilde{U}}), L^{(2)}(\tilde{\mathcal{F}})_{x_0} \rangle).$$

This may be checked directly. The reader may also refer to Propositions 3.7 and 3.8 for the former. The latter is not dealt with directly elsewhere in the paper, but one can easily see how it follows from the discussion surrounding Proposition 3.14.

(i) \implies (ii) The computations of Lemma 7.4 allow us to assume that

$$Z_{x_0}(\mathcal{F}) = \{f_0, f_1, \dots, f_{\ell}\}, \quad \bar{Z}_{x_0}(\mathcal{F}) = \{f_{\ell+1}, \dots, f_m\}, \quad (7.3)$$

with $f_{\ell+1}, \dots, f_m$ linearly independent at x_0 . For brevity denote $S_{x_0} = \langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle$, and write points in $T_{x_0}M/S_{x_0}$ as $v + S_{x_0}$ for $v \in T_{x_0}M$. Also for brevity, denote $B = B_{\mathcal{F}}(S_{x_0})$. Let W_{x_0} be the subspace generated by $\text{image}(Q_B)$. Since B is indefinite, by Lemma 2.2 there exists $k \geq 0$ and $v_1, \dots, v_{m-\ell+k} \in L^{(1)}(\mathcal{F}_1)_{x_0}$ satisfying

$$v_1 = f_{\ell+1}(x_0), \dots, v_{m-\ell} = f_m(x_0),$$

so that

$$0_{x_0} + S_{x_0} \in \text{int}_{W_{x_0}}(\text{conv}(\{Q_B(v_1), \dots, Q_B(v_{m-\ell+k})\})).$$

Thus there exists $\beta_1, \dots, \beta_{m-\ell+k} > 0$ so that

$$\sum_{j=1}^{m-\ell+k} \beta_j = 1, \quad \sum_{j=1}^{m-\ell+k} \beta_j Q_B(v_j) = 0_{x_0} + S_{x_0}.$$

Let us take g_1, \dots, g_k to be real-analytic vector fields that arbitrarily extend $v_{m-\ell+1}, \dots, v_{m-\ell+k}$. Now define $\tilde{m} = m + k$ and define $\tilde{\mathcal{F}} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{\tilde{m}})$ by

$$\tilde{f}_a = f_a, \quad a \in \{0, \dots, \ell\} \\ \tilde{f}_{\ell+a} = \sqrt{\beta_a} f_{\ell+a}, \quad a \in \{1, \dots, m-\ell\} \\ \tilde{f}_{m+a} = \sqrt{\beta_{m-\ell+a}} g_a, \quad a \in \{1, \dots, k\}.$$

Now define

$$D = \text{diag}(\underbrace{1, \dots, 1}_{\ell \text{ of these}}, \sqrt{\beta_1}, \dots, \sqrt{\beta_{m-\ell}})$$

and let $M \in \mathbb{R}^{m \times k}$ be any matrix satisfying

$$\tilde{f}_{m+r}(x_0) = \sum_{a=1}^m M_r^a f_a(x_0), \quad r \in \{1, \dots, k\}.$$

We then see that

$$\tilde{f}_\alpha = \sum_{b=1}^m T_\alpha^b f_b$$

where

$$T = [D \mid M].$$

Thus T has full rank. We also take $\tilde{U} \subset \mathbb{R}^{\tilde{m}}$ to be any convex, proper set. For $\alpha \in \{1, \dots, \ell\}$ we have $[\tilde{f}_\alpha, [\tilde{f}_0, \tilde{f}_\alpha]](x_0) = 0_{x_0}$. Therefore, by the definition of $\tilde{\mathcal{F}}$ we have

$$\sum_{\alpha=1}^{\tilde{m}} Q_B(\tilde{f}_\alpha(x_0)) = 0_{x_0} + S_{x_0},$$

showing that $\tilde{\Sigma} = (M, \tilde{\mathcal{F}}, \tilde{U})$ is second-order neutralisable.

(ii) \implies (i) Let $\tilde{m} \geq m$, $T \in \mathbb{R}^{\tilde{m} \times m}$, and \tilde{U} have the properties of part (ii). We continue using our above abbreviations $S_{x_0} = \langle Z_{x_0}(\mathcal{F}_U), L^{(2)}(\mathcal{F})_{x_0} \rangle$ and $\tilde{B} = B_{\tilde{\mathcal{F}}}(S_{x_0})$. Let us assume as in Lemma 7.4 that

$$Z_{x_0}(\tilde{\mathcal{F}}) = \{\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{\tilde{\ell}}\}, \quad \tilde{Z}_{x_0}(\tilde{\mathcal{F}}) = \{\tilde{f}_{\tilde{\ell}+1}, \dots, \tilde{f}_{\tilde{m}}\},$$

and that $\{\tilde{f}_{\tilde{\ell}+1}(x_0), \dots, \tilde{f}_{\tilde{m}}(x_0)\}$ form a basis for $L^{(1)}(\tilde{\mathcal{F}})_{x_0}$. We then directly have, by definition of second-order neutralisability and since $\{\tilde{f}_1, \dots, \tilde{f}_{\tilde{\ell}}\}$ vanish at x_0 ,

$$\sum_{\alpha=1}^{\tilde{m}} [\tilde{f}_\alpha, [\tilde{f}_0, \tilde{f}_\alpha]](x_0) = \sum_{\alpha=\tilde{\ell}+1}^{\tilde{m}} [\tilde{f}_\alpha, [\tilde{f}_0, \tilde{f}_\alpha]](x_0) \in S_{x_0}.$$

This implies that

$$\sum_{\alpha=\tilde{\ell}+1}^{\tilde{m}} Q_{\tilde{B}}(\tilde{f}_\alpha(x_0)) = 0_{x_0} + S_{x_0},$$

giving

$$\sum_{\alpha=\tilde{\ell}+1}^{\tilde{m}} \lambda Q_{\tilde{B}}(\tilde{f}_\alpha(x_0)) = 0$$

for each $\lambda \in (T_{x_0}M/S_{x_0})^*$, since $(T_{x_0}M/S_{x_0})^* \simeq \text{ann}(S_{x_0})$. Lemma 7.5 then implies that $\lambda \tilde{B}$ is indefinite for each $\lambda \in (T_{x_0}M/S_{x_0})^* \setminus \text{ann}(W_{x_0})$, giving the lemma since $\{\tilde{f}_{\tilde{\ell}+1}(x_0), \dots, \tilde{f}_{\tilde{m}}(x_0)\}$ form a basis for $L^{(1)}(\tilde{\mathcal{F}})_{x_0}$. \blacksquare

7.7 REMARK: That the matrix $T \in \mathbb{R}^{\tilde{m} \times m}$ is surjective allows us to assert that the interiors of the reachable sets for Σ and $\tilde{\Sigma}$ will have nonempty intersection. That is to say, we have not lost any information in changing the system vector fields from \mathcal{F} to $\tilde{\mathcal{F}}$. \square

This preceding result tells us that the hypothesis (iv) of Theorem 5.2 enables us to neutralise the second-order obstructions to local controllability. It now follows from Theorem 3.5 of [Bianchini and Stefani 1993] that the tangent vectors from $L^{(3)}(\mathcal{F})_{x_0}$ comprise a subspace of variations, and so, by Lemma 7.2, the tangent vectors from $\langle Z_{x_0}(\mathcal{F}_{\text{conv}(U)}), L^{(3)}(\mathcal{F})_{x_0} \rangle$ also comprise a subspace of variations. Theorem 5.2 now follows from the assumption (iii) in its list of hypotheses.

7.2. Proof of Theorem 5.4. We first recall the Chen-Fliess-Sussmann series for a real-analytic function along a trajectory of a control-affine system. This series has been presented in the papers [Chen 1957, Fliess 1981, Sussmann 1983]. We let $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ be a control-affine system with U compact. The series involves a sum over all multi-indices in $\{0, 1, \dots, m\}$. If $I = (a_1, \dots, a_k)$ is such a multi-index, then denote $|I| = k$. Given a single multi-index $I = (a_1, \dots, a_k)$ with $a_j \in \{0, 1, \dots, m\}$, $j = 1, \dots, k$, one defines

$$\int_0^t u_I = \int_0^t \int_0^{\tau_k} \int_0^{\tau_{k-1}} \dots \int_0^{\tau_2} u_{a_k}(\tau_k) u_{a_{k-1}}(\tau_{k-1}) \dots u_{a_2}(\tau_2) u_{a_1}(\tau_1) d\tau_1 d\tau_2 \dots d\tau_k.$$

In such expressions, u_0 is always taken to be 1. Also, given the multi-index $I = (a_1, \dots, a_k)$ one defines a differential operator f_I on $\mathcal{F}(M)$ by

$$f_I \phi = f_{a_1} f_{a_2} \dots f_{a_k} \phi$$

(by $f_a \phi$ we mean the Lie derivative of the function ϕ with respect the vector field f_a). In the general development expounded on by Sussmann [1983], one studies formal series of differential operators

$$\sum_I \left(\int_0^T u_I \right) f_I.$$

Sussmann shows that for a given real-analytic function ϕ on M , the series

$$\sum_{|I| \leq N} \left(\int_0^t u_I \right) f_I \phi(\xi(0))$$

converges to $\phi(x(t))$ as $N \rightarrow \infty$ for $t \in [0, T]$ provided T is sufficiently small. Furthermore, the convergence is uniform in the control and in the initial condition $\xi(0)$.

Now we proceed with the proof proper. Note that we can always find a compact, convex, and proper set \tilde{U} containing U (for example, by taking \tilde{U} to be a sufficiently large closed ball). Clearly if the system is not STLC with the control set \tilde{U} it will not be STLC with the control set U . Thus we make, without loss of generality, the assumption that U is compact, convex, and proper.

We now prove a lemma that defines the key element in our proof: a real-analytic function having certain properties.

7.8 LEMMA: *If the hypotheses of Theorem 5.4 hold, with the possible exception that the control set U is convex and proper, then there exists a neighbourhood \mathcal{N} of x_0 , a real-analytic function $\phi: \mathcal{N} \rightarrow \mathbb{R}$, and $Q \in GL(m; \mathbb{R})$ with the following properties:*

- (i) $X\phi(x) = 0$ for all $X \in \mathcal{L}^{(\infty)}(\tilde{\mathcal{F}}_1)$ and for all $x \in \mathcal{N}$;
- (ii) $X\phi(x_0) = 0$ if $X(x_0) \in \langle Z_{x_0}(\tilde{\mathcal{F}}), L^{(1)}(\tilde{\mathcal{F}}_1)_{x_0} \rangle$;
- (iii) $[\tilde{f}_a, [\tilde{f}_0, \tilde{f}_a]]\phi(x_0) = -1$ for $\tilde{f}_a \in \bar{Z}_{x_0}(\tilde{\mathcal{F}})$;
- (iv) $[\tilde{f}_a, [\tilde{f}_0, \tilde{f}_a]]\phi(x_0) = 0$ for $\tilde{f}_a \in Z_{x_0}(\tilde{\mathcal{F}})$;
- (v) $[\tilde{f}_a, [\tilde{f}_0, \tilde{f}_b]]\phi(x_0) = 0$ for all $a, b \in \{1, \dots, m\}$ with $a \neq b$;
- (vi) in any neighbourhood $\tilde{\mathcal{N}} \subset \mathcal{N}$ of x_0 , ϕ can take on both positive and negative values, where
- (vii) $\tilde{f}_0 = f_0$;
- (viii) $\tilde{f}_a = \sum_{b=1}^m Q_a^b f_b$, $a \in \{1, \dots, m\}$;
- (ix) $\tilde{\mathcal{F}} = \{\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_m\}$.

Proof: Let $R \in GL(m; \mathbb{R})$, $\ell \in \{0, \dots, m-1\}$, and \mathcal{H} be defined as in the proof of Lemma 7.4. For brevity let us denote

$$\begin{aligned} S_{x_0} &= \langle Z_{x_0}(\mathcal{H}), L^{(1)}(\mathcal{H})_{x_0} \rangle + L^{(\infty)}(\mathcal{H})_{x_0} \\ &= \langle Z_{x_0}(\mathcal{F}), L^{(1)}(\mathcal{F}_1)_{x_0} \rangle + L^{(\infty)}(\mathcal{F}_1)_{x_0} \end{aligned}$$

and denote $B = B_{\mathcal{F}}(S_{x_0})$. Since B is definite, by Lemma 2.2 there exists $\lambda \in (T_{x_0}M/S_{x_0})^*$ so that λB is positive-definite. Therefore, there exists a matrix $\tilde{S} \in GL(m-\ell; \mathbb{R})$ so that if

$$\tilde{f}_{\ell+a} = \sum_{b=1}^r \tilde{S}_a^b h_{\ell+b}, \quad a \in \{1, \dots, m-\ell\},$$

then

$$\lambda B(\tilde{f}_{\ell+a}, \tilde{f}_{\ell+b}) = \begin{cases} 1, & a = b \\ 0, & a \neq b. \end{cases}$$

Now we take

$$S = \begin{bmatrix} I_\ell & 0_{\ell \times (m-\ell)} \\ 0_{(m-\ell) \times \ell} & \tilde{S} \end{bmatrix}$$

where I_k is the $k \times k$ identity matrix and $0_{k \times r}$ is the $k \times r$ matrix of zeros. Defining $Q = RS$ gives us our desired collection of vector fields $\tilde{\mathcal{F}} = \{\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_m\}$ where

$$\tilde{f}_0 = f_0, \quad \tilde{f}_a = \sum_{b=1}^m Q_a^b f_b, \quad a = 1, \dots, m.$$

Note that

$$\mathcal{L}^{(\infty)}(\tilde{\mathcal{F}}_1) = \mathcal{L}^{(\infty)}(\mathcal{F}_1) \quad \text{and} \quad L^{(1)}(\tilde{\mathcal{F}}_1)_{x_0} = L^{(1)}(\mathcal{F}_1)_{x_0}.$$

Now recall that $(T_{x_0}M/S_{x_0})^* \simeq \text{ann}(S_{x_0})$. Thus we can think of $\lambda \in T_{x_0}^*M$. What's more, with this identification we have

$$\lambda B(v_1, v_2) = \langle \lambda; [V_1, [f_0, V_2]](x_0) \rangle, \quad (7.4)$$

where, as usual, $V_1, V_2 \in \Gamma(TM)$ extend $v_1, v_2 \in L^{(1)}(\tilde{\mathcal{F}}_1)_{x_0}$. Now choose a set of coordinates (x^1, \dots, x^n) for M about x_0 so that

1. x_0 is mapped to $(0, \dots, 0) \in \mathbb{R}^n$ by the coordinate chart,
2. $L^{(\infty)}(\tilde{\mathcal{F}}_1)_x = \text{span}_{\mathbb{R}} \left\{ \frac{\partial}{\partial x^1} \Big|_x, \dots, \frac{\partial}{\partial x^\ell} \Big|_x \right\}$ for all $x \in \mathcal{N}$,
3. $\langle Z_{x_0}(\tilde{\mathcal{F}}), L^{(1)}(\tilde{\mathcal{F}}_1)_{x_0} \rangle + L^{(\infty)}(\tilde{\mathcal{F}}_1)_{x_0} = \text{span}_{\mathbb{R}} \left\{ \frac{\partial}{\partial x^1} \Big|_{x_0}, \dots, \frac{\partial}{\partial x^\ell} \Big|_{x_0} \right\}$, and
4. $\text{d}x^{\ell+1} \Big|_{x_0} = -\lambda$ where λ is as chosen above.

We claim that with these coordinates, the function $\phi(x^1, \dots, x^n) = x^{\ell+1}$ satisfies the conditions (i)–(vi) of Lemma 7.8. The condition 2 for the coordinates (x^1, \dots, x^n) ensures that $\text{d}\phi$ annihilates $L^{(\infty)}(\tilde{\mathcal{F}}_1)$ in a neighbourhood of x_0 . Thus ϕ satisfies condition (i) of Lemma 7.8. The condition 3 for the coordinates (x^1, \dots, x^n) ensures that $\text{d}\phi$ annihilates $\langle Z_{x_0}(\tilde{\mathcal{F}}), L^{(1)}(\tilde{\mathcal{F}}_1)_{x_0} \rangle$. Thus ϕ satisfies condition (ii) of Lemma 7.8. The definition of $Q \in GL(m; \mathbb{R})$ to initiate the proof, the property 4 of the coordinates (x^1, \dots, x^n) , and the equation (7.4) ensures that the condition (iii) of Lemma 7.8 is satisfied. Note that if $\tilde{f}_a \in Z_{x_0}(\tilde{\mathcal{F}})$ then we have $[\tilde{f}_a, [\tilde{f}_0, \tilde{f}_a]](x_0) = 0_{x_0}$. Thus condition (iv) of Lemma 7.8 holds. By our definition of Q above, by the property 4 of the coordinates (x^1, \dots, x^n) , and by the equation (7.4), the condition (v) of Lemma 7.8 holds for ϕ . That condition (vi) of Lemma 7.8 holds follows from property 1 of the coordinates (x^1, \dots, x^n) . \blacksquare

Note that $\text{Traj}(\Sigma) = \text{Traj}(\tilde{\Sigma})$ if we take

$$\tilde{U} = \{Q^{-1}u \mid u \in U\}.$$

Thus, we shall replace $\Sigma = \tilde{\Sigma}$ for the remainder of the proof. We shall also rearrange the vector fields in \mathcal{F} so that $Z_{x_0}(\mathcal{F}) = \{f_0, f_1, \dots, f_\ell\}$.

Our strategy in the remainder of the proof is to consider the Chen-Fliess-Sussmann series for ϕ along an arbitrary $(\xi, u) \in \text{Traj}(\Sigma)$ (with $\xi(0) = x_0$), and show that ϕ assumes only positive values for all sufficiently small times. We suppose that ξ and u are defined on $[0, T]$. The computations here are reminiscent of those in the proof of Proposition 6.3 in [Sussmann 1983], although the multi-input setting we consider here does necessitate careful consideration when adapting the proof in [Sussmann 1983].

In writing the Chen-Fliess-Sussmann series, one must take a sum over all multi-indices in $\{0, 1, \dots, m\}$. The next lemma breaks these multi-indices into groups that we shall separately analyse.

7.9 LEMMA: *Any multi-index I in $\{0, 1, \dots, m\}$ fits into at least one of the following groups:*

Type 1a. $I = (a_1, \dots, a_{r-1}, a)$ where $f_a \in \bar{Z}_{x_0}(\mathcal{F})$ and for any $r \geq 1$;

Type 1b. $I = (a, a_2, \dots, a_r)$ where $f_a \in Z_{x_0}(\mathcal{F})$ and for any $r \geq 1$;

Type 2. *multi-indices of length greater than 1 and of the form $I = (a, a_2, \dots, a_r)$ where $f_a \in Z_{x_0}(\mathcal{F})$ and where $f_{a_j} \in Z_{x_0}(\mathcal{F})$, $j \in \{2, \dots, r\}$;*

Type 3a. $I = (a, a, 0)$ where $f_a \in \bar{Z}_{x_0}(\mathcal{F})$;

Type 3b. $I = (a, a, 0)$ where $f_a \in Z_{x_0}(\mathcal{F})$;

Type 3c. $I = (a, b, 0)$ for any distinct $a, b \neq 0$;

Type 3d. $I = (a, b, c)$ with $f_a, f_b \in \bar{Z}_{x_0}(\mathcal{F})$ and $f_c \in Z_{x_0}(\mathcal{F}) \setminus \{f_0\}$;

Type 4. I satisfies

- (i) I contains at least two indices $a, b \in \{1, \dots, m\}$ for which $f_a, f_b \in \bar{Z}_{x_0}(\mathcal{F})$,
- (ii) $I = (a, a_2, \dots, a_{r-1}, b)$ with $f_a \in \bar{Z}_{x_0}(\mathcal{F})$ and $f_b \in Z_{x_0}(\mathcal{F})$, and
- (iii) $|I| \geq 4$.

Proof: We consider multi-indices of increasing length, categorising all for a given length.

If $|I| = 1$ then clearly I is either Type 1a or Type 1b.

If $|I| = 2$ then $I = (a, b)$ with $f_a, f_b \in Z_{x_0}(\mathcal{F})$ (Type 1b or Type 2), $f_a \in Z_{x_0}(\mathcal{F})$ and $f_b \in \bar{Z}_{x_0}(\mathcal{F})$ (Type 1a or Type 1b), $f_a \in \bar{Z}_{x_0}(\mathcal{F})$ and $f_b \in Z_{x_0}(\mathcal{F})$ (Type 2), or $f_a, f_b \in \bar{Z}_{x_0}(\mathcal{F})$ (Type 1a).

If $|I| = 3$ then $I = (a, b, c)$ with $f_a \in Z_{x_0}(\mathcal{F})$ (Type 1b), $f_a \in \bar{Z}_{x_0}(\mathcal{F})$ and $f_b, f_c \in Z_{x_0}(\mathcal{F})$ (Type 2), $f_a, f_b \in \bar{Z}_{x_0}(\mathcal{F})$ and $f_c \in Z_{x_0}(\mathcal{F})$ (Type 3), $f_a, f_c \in \bar{Z}_{x_0}(\mathcal{F})$ and $f_b \in Z_{x_0}(\mathcal{F})$ (Type 1a), or $f_a, f_b, f_c \in \bar{Z}_{x_0}(\mathcal{F})$ (Type 1a).

If $|I| \geq 4$ then $I = (a, a_2, \dots, a_r)$ with $f_a \in Z_{x_0}(\mathcal{F})$ (Type 1b), $I = (a, a_2, \dots, a_{r-1}, b)$ with $f_a, f_b \in \bar{Z}_{x_0}(\mathcal{F})$ (Type 1a), or $I = (a, a_2, \dots, a_{r-1}, b)$ with $f_a \in \bar{Z}_{x_0}(\mathcal{F})$ and $f_b \in Z_{x_0}(\mathcal{F})$ (Type 4). ■

Next we compute the contribution of each of the multi-indices of Lemma 7.9 to the Chen-Fliess-Sussmann series.

7.10 LEMMA: *If I is a multi-index of Type 1a or Type 1b then $f_I\phi(x_0) = 0$.*

Proof: If $I = (a_1, \dots, a_{r-1}, a)$ with $f_a \in \bar{Z}_{x_0}(\mathcal{F})$ then $f_a\phi$ is identically zero in a neighbourhood of x_0 by property (i) of Lemma 7.8. If $I = (a, a_2, \dots, a_r)$ with $f_a \in Z_{x_0}(\mathcal{F})$ then

$$f_I\phi(x_0) = f_a f_{a_2} \cdots f_{a_r} \phi(x_0).$$

Since $f_a(x_0) = 0_{x_0}$, this gives $f_I\phi(x_0) = 0$. ■

7.11 LEMMA: *If I is a multi-index of Type 2 then $f_I\phi(x_0) = 0$.*

Proof: Suppose that $I = (a, a_1, \dots, a_k)$ with $f_{a_1}, \dots, f_{a_k} \in Z_{x_0}(\mathcal{F})$. We claim that

$$f_a f_{a_1} \cdots f_{a_k} \phi(x_0) = (-1)^k \text{ad}_{f_{a_k}} \cdots \text{ad}_{f_{a_1}} f_a \phi(x_0). \quad (7.5)$$

Indeed, by the definition of the Lie bracket we have

$$f_a f_{a_1} = -[f_{a_1}, f_a] + f_{a_1} f_a.$$

Since $f_{a_1}(x_0) = 0_{x_0}$ we have $f_{a_1} f_a f_{a_2} \cdots f_{a_k} \phi(x_0) = 0$. Thus

$$f_a f_{a_1} \cdots f_{a_k} \phi(x_0) = -[f_{a_1}, f_a] f_{a_2} \cdots f_{a_k} \phi(x_0).$$

Applying the same argument again

$$\begin{aligned} -[f_{a_1}, f_a] f_{a_2} \cdots f_{a_k} &= [f_{a_2}, [f_{a_1}, f_a]] f_{a_3} \cdots f_{a_k} - f_{a_2} [f_{a_1}, f_a] f_{a_3} \cdots f_{a_k} \\ \implies f_a f_{a_1} \cdots f_{a_k} \phi(x_0) &= [f_{a_2}, [f_{a_1}, f_a]] f_{a_3} \cdots f_{a_k} \phi(x_0). \end{aligned}$$

This may be iterated $k - 2$ more times to finally get (7.5). The lemma now follows by property (ii) of Lemma 7.8. ■

7.12 LEMMA: *If I is a multi-index of Type 3a then $f_I\phi(x_0) = 1$.*

Proof: In this proof we perform the basic computation needed for the next two lemmas as well. For $f_a, f_b \in \bar{Z}_{x_0}(\mathcal{F})$ and $f_c \in Z_{x_0}(\mathcal{F})$ we have

$$\begin{aligned} f_a f_b f_c &= f_a [f_b, f_c] + f_a f_c f_b \\ &= [f_a, [f_b, f_c]] + [f_b, f_c] f_a + [f_a, f_c] f_b + f_c f_a f_b. \end{aligned}$$

Since $f_a\phi$ and $f_b\phi$ are identically zero in a neighbourhood of x_0 by property (i) of Lemma 7.8, it now follows that

$$f_a f_b f_c \phi(x_0) = -[f_a, [f_c, f_b]] \phi(x_0). \quad (7.6)$$

Now we take the special case where $f_a = f_b$ and $f_c = f_0$ with $f_a \in \bar{Z}_{x_0}(\mathcal{F})$. In this case, the condition (iii) of Lemma 7.8 ensures that $f_a f_b f_0 \phi(x_0) = 1$. ■

7.13 LEMMA: *If I is a multi-index of Type 3b then $f_I\phi(x_0) = 0$.*

Proof: We utilise equation (7.6). In this case the condition (iv) of Lemma 7.8 ensures that $f_a f_a f_0 \phi(x_0) = 0$. ■

7.14 LEMMA: *If I is a multi-index of Type 3c then $f_I\phi(x_0) = 0$.*

Proof: Again, using (7.6), this time along with condition (v) of Lemma 7.8, we see that $f_a f_b f_0 \phi(x_0) = 0$. ■

7.15 LEMMA: *If I is a multi-index of Type 3d then $f_I\phi(x_0) = 0$.*

Proof: Here we note that for $f_c \in Z_{x_0}(\mathcal{F})$, $f_c\phi$ is identically zero in a neighbourhood of x_0 by property (i) of Lemma 7.8. Thus $f_a f_b f_c \phi$ also identically vanishes. ■

Of the multi-indices of the five preceding lemmas we see that the only one that contributes to the Chen-Fliess-Sussmann series for ϕ is the Type 3a multi-indices. We continue to fix an arbitrary $(\xi, u) \in \text{Traj}(\Sigma)$, asking only that $\xi(0) = x_0$. The following lemma gives the coefficient of the Type 3a terms as they appear in the series.

7.16 LEMMA: *If $I = (a, a, 0)$ is a multi-index of Type 3a then*

$$\int_0^t u_I = \frac{1}{2} \int_0^t (v_a(s))^2 ds,$$

where

$$v_a(t) = \int_0^t u_a(s) ds. \quad (7.7)$$

Proof: By definition we have

$$\begin{aligned} \int_0^t u_I &= \int_0^t \int_0^{\tau_3} \int_0^{\tau_2} u_0(\tau_3) u_a(\tau_2) u_a(\tau_1) d\tau_1 d\tau_2 d\tau_3 \\ &= \int_0^t \int_0^{\tau_3} u_a(\tau_2) v_a(\tau_2) d\tau_2 d\tau_3 \\ &= \int_0^t \int_0^{\tau_3} \dot{v}_a(\tau_2) v_a(\tau_2) d\tau_2 d\tau_3 \\ &= \int_0^t \int_0^{\tau_3} \frac{d}{d\tau_2} \left(\frac{1}{2} v_a^2(\tau_2) \right) d\tau_2 d\tau_3 \\ &= \int_0^t \frac{1}{2} v_a^2(\tau_3) d\tau_3, \end{aligned}$$

and this is the stated result. \blacksquare

We introduce the following notation for measurable functions $f, g: [0, T] \rightarrow \mathbb{R}$ and $t \in]0, T[$:

$$\langle f, g \rangle_{2,t} = \int_0^t f(s)g(s) \, ds, \quad \|f\|_{2,t} = (\langle f, f \rangle_{2,t})^{1/2}.$$

Also, for our given $(\xi, u) \in \text{Traj}(\Sigma)$, denote by

$$\|u_\emptyset\|_{\max,t} = \max_{a \in \{\ell+1, \dots, m\}} \|v_a\|_{2,t},$$

where we recall that ℓ is defined so that $Z_{x_0}(\mathcal{F}) = \{f_0, f_1, \dots, f_\ell\}$, and where v_a is defined as in (7.7). The following result summarises the preceding six lemmas.

7.17 LEMMA: For $(\xi, u) \in \text{Traj}(\Sigma)$ the contribution, denoted $C_3(t)$, to the Chen-Fließ-Sussmann series of the Type 1, Type 2, and Type 3 terms for $\phi(\xi(t))$ is bounded below by $\frac{1}{2}\|u_\emptyset\|_{\max,t}$.

Next we need to estimate the contribution from the Type 4 terms. The proof of the following lemma contains the majority of the necessary computations.

7.18 LEMMA: Let I be a multi-index of Type 4, and write $I = (a, Z, b, J)$ where $f_a, f_b \in \bar{Z}_{x_0}(\mathcal{F})$, $Z = (z_1, \dots, z_k)$ has the property that $f_{z_j} \in Z_{x_0}(\mathcal{F})$, $j \in \{1, \dots, k\}$, and $|J| = r \geq 1$. Since the control set U is compact, choose $R > 0$ so that

$$U \subset [-R, R] \times \dots \times [-R, R].$$

Let $\bar{I} = (b, Z, a, J)$. Then there exists a constant $D > 0$ so that

(i) if $k = 0$ we have

$$\left| \int_0^t u_I + \int_0^t u_{\bar{I}} \right| \leq \frac{DR^r t^{r-1}}{(r-1)!} \|u_\emptyset\|_{\max,t}^2,$$

(ii) if $k = 1$ we have

$$\left| \int_0^t u_I \right| \leq \frac{DR^{r+1} t^r}{(r-1)!} \|u_\emptyset\|_{\max,t}^2,$$

and

(iii) if $k > 1$ we have

$$\left| \int_0^t u_I \right| \leq \frac{DR^{r+k} t^{r+k-1}}{(r-1)!(k-2)!} \|u_\emptyset\|_{\max,t}^2.$$

Proof: Let us write $J = (a_1, \dots, a_r)$. We then have

$$\int_0^t u_I = \int_0^t \int_0^{s_r} \dots \int_0^{s_2} u_{a_r}(s_r) \dots u_{a_1}(s_1) \psi_k(a, b, s_1) \, ds_1 \dots ds_r, \quad (7.8)$$

where

$$\psi_k(a, b, t) = \int_0^t \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_1} u_b(\tau_b) u_{z_k}(\tau_k) \dots u_{z_1}(\tau_1) u_a(\tau_a) \, d\tau_a d\tau_1 \dots d\tau_k d\tau_b.$$

Each of the controls u_{z_1}, \dots, u_{z_k} is bounded in magnitude by R . Thus we have

$$\begin{aligned} |\psi_k(a, b, t)| &= \left| \int_0^t \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_1} u_b(\tau_b) u_{z_k}(\tau_k) \dots u_{z_1}(\tau_1) u_a(\tau_a) \, d\tau_a d\tau_1 \dots d\tau_k d\tau_b \right| \\ &\leq R^k \left| \int_0^t \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_1} u_b(\tau_b) u_a(\tau_a) \, d\tau_a d\tau_1 \dots d\tau_k d\tau_b \right|. \end{aligned}$$

With this as motivation, define

$$\bar{\psi}_k(a, b, t) = \int_0^t \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_1} u_b(\tau_b) u_a(\tau_a) \, d\tau_a d\tau_1 \dots d\tau_k d\tau_b,$$

and compute

$$\begin{aligned} \bar{\psi}_k(a, b, t) &= \int_0^t u_b(\tau_b) \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_2} \int_0^{\tau_1} u_a(\tau_a) \, d\tau_a d\tau_1 d\tau_2 \dots d\tau_k d\tau_b \\ &= \int_0^t u_b(\tau_b) \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_1} \int_{\tau_a}^{\tau_2} u_a(\tau_a) \, d\tau_1 d\tau_a d\tau_2 \dots d\tau_k d\tau_b \\ &= \int_0^t u_b(\tau_b) \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_3} \int_0^{\tau_1} (\tau_2 - \tau_a) u_a(\tau_a) \, d\tau_a d\tau_2 d\tau_3 \dots d\tau_k d\tau_b \\ &= \int_0^t u_b(\tau_b) \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_1} \int_{\tau_a}^{\tau_3} (\tau_2 - \tau_a) u_a(\tau_a) \, d\tau_2 d\tau_a d\tau_3 \dots d\tau_k d\tau_b \\ &= \int_0^t u_b(\tau_b) \int_0^{\tau_b} \int_0^{\tau_k} \dots \int_0^{\tau_4} \int_0^{\tau_1} \frac{1}{2} (\tau_3 - \tau_a)^2 u_a(\tau_a) \, d\tau_a d\tau_3 d\tau_4 \dots d\tau_k d\tau_b \\ &\vdots \\ &= \frac{1}{k!} \int_0^t u_b(s) \int_0^s (s - \tau)^k u_a(\tau) \, d\tau \, ds. \end{aligned}$$

An application of (7.8) gives

$$\begin{aligned} \left| \int_0^t u_I \right| &\leq R^r \int_0^t \int_0^{s_r} \dots \int_0^{s_3} \int_0^{s_2} |\psi_k(a, b, s_1)| \, ds_1 ds_2 \dots ds_r \\ &= R^r \int_0^t \int_0^{s_r} \dots \int_0^{s_2} \int_{s_1}^{s_3} |\psi_k(a, b, s_1)| \, ds_2 ds_1 \dots ds_r \\ &= R^r \int_0^t \int_0^{s_r} \dots \int_0^{s_2} (s_3 - s_1) |\psi_k(a, b, s_1)| \, ds_1 \dots ds_r \\ &\vdots \\ &= \frac{R^r}{(r-1)!} \int_0^t (t - \tau)^{r-1} |\psi_k(a, b, \tau)| \, d\tau \\ &\leq \frac{R^r t^{r-1}}{(r-1)!} \int_0^t |\psi_k(a, b, \tau)| \, d\tau \\ &\leq \frac{R^{r+k} t^{r-1}}{(r-1)!} \int_0^t |\bar{\psi}_k(a, b, \tau)| \, d\tau. \end{aligned} \quad (7.9)$$

Thus we need to estimate $|\tilde{\psi}_k(a, b, t)|$.

For $k = 0$ we have

$$\begin{aligned}\tilde{\psi}_0(a, b, t) &= \int_0^t u_b(s) v_a(s) \, ds \\ &= \int_0^t \frac{dv_b(s)}{ds} v_a(s) \, ds \\ &= v_b(t) v_a(t) - \int_0^t \frac{dv_a(s)}{ds} v_b(s) \, ds.\end{aligned}$$

Thus we have

$$\begin{aligned}\tilde{\psi}_0(a, b, t) + \tilde{\psi}_0(b, a, t) &= 2v_a(t)v_b(t) - \int_0^t \left(\frac{dv_a(s)}{ds} v_b(s) + \frac{dv_b(s)}{ds} v_a(s) \right) ds \\ &= 2v_a(t)v_b(t) - \int_0^t \frac{d}{ds} (v_a(s)v_b(s)) \, ds \\ &= v_a(t)v_b(t).\end{aligned}$$

Proceeding as in the derivation of (7.9) gives

$$\begin{aligned}\left| \int_0^t u_I + \int_0^t u_{\bar{I}} \right| &= \left| \int_0^t \int_0^{s_r} \cdots \int_0^{s_2} u_{a_r}(s_r) \cdots u_{a_1}(s_1) (\tilde{\psi}_0(a, b, s_1) + \tilde{\psi}_0(b, a, s_1)) \, ds_1 \cdots ds_r \right| \\ &\leq \frac{R^r}{(r-1)!} \int_0^t (t-\tau)^{r-1} |\tilde{\psi}_0(a, b, \tau) + \tilde{\psi}_0(b, a, \tau)| \, d\tau \\ &\leq \frac{R^r t^{r-1}}{(r-1)!} \int_0^t |v_a(\tau)v_b(\tau)| \, d\tau \\ &\leq \frac{R^r t^{r-1}}{(r-1)!} \|v_a\|_{2,t} \|v_b\|_{2,t} \\ &\leq \frac{R^r t^{r-1}}{(r-1)!} \|u_\emptyset\|_{\max,t}^2\end{aligned}\tag{7.10}$$

when $k = 0$, using the Cauchy-Bunyakovsky-Schwarz inequality in the penultimate step.

For $k \geq 1$ let us apply an integration by parts to $\tilde{\psi}_k(a, b, t)$ to obtain

$$\begin{aligned}\tilde{\psi}_k(a, b, t) &= \frac{1}{k!} \int_0^t u_b(s) \int_0^s (s-\tau)^k u_a(\tau) \, d\tau \, ds \\ &= \frac{1}{k!} \int_0^t \frac{dv_b(s)}{ds} \int_0^s (s-\tau)^k u_a(\tau) \, d\tau \, ds \\ &= \frac{v_b(t)}{k!} \int_0^t (t-\tau)^k u_a(\tau) \, d\tau - \frac{1}{(k-1)!} \int_0^t v_b(s) \int_0^s (s-\tau)^{k-1} u_a(\tau) \, d\tau \, ds.\end{aligned}\tag{7.11}$$

Again using integration by parts we have

$$\begin{aligned}\int_0^t (t-\tau)^k u_a(\tau) \, d\tau &= \int_0^t (t-\tau)^k \frac{dv_a(\tau)}{d\tau} \, d\tau \\ &= -k \int_0^t (t-\tau)^{k-1} v_a(\tau) \, d\tau.\end{aligned}\tag{7.12}$$

For $k = 1$, a substitution of (7.12) into (7.11) gives

$$\tilde{\psi}_1(a, b, t) = -v_b(t) \int_0^t v_a(\tau) \, d\tau - \int_0^t v_b(s) v_a(s) \, ds.$$

Thus in this case we have

$$\begin{aligned}\left| \int_0^t \tilde{\psi}_1(a, b, s) \, ds \right| &\leq \left| \int_0^t v_b(s) \int_0^s v_a(\tau) \, d\tau \, ds \right| + \left| \int_0^t \int_0^s v_b(\tau) v_a(\tau) \, d\tau \, ds \right| \\ &\leq \int_0^t |v_b(s)| \left(\int_0^s d\tau \right)^{1/2} \left(\int_0^s v_a^2(\tau) \, d\tau \right)^{1/2} \, ds + \int_0^t \|v_a\|_{2,s} \|v_b\|_{2,s} \, ds \\ &\leq \int_0^t |v_b(s)| s^{1/2} \|v_a\|_{2,s} \, ds + \|v_a\|_{2,t} \|v_b\|_{2,t} \int_0^t ds \\ &\leq \|v_a\|_{2,t} \left(\int_0^t s \, ds \right)^{1/2} \left(\int_0^t v_b^2(s) \, ds \right)^{1/2} + \|v_a\|_{2,t} \|v_b\|_{2,t} t \\ &= \frac{1}{\sqrt{2}} \|v_a\|_{2,t} \|v_b\|_{2,t} t + \|v_a\|_{2,t} \|v_b\|_{2,t} t,\end{aligned}$$

where we have twice used the Cauchy-Bunyakovsky-Schwarz inequality in the second step and once in the fourth step. Thus we have

$$\left| \int_0^t u_I \right| \leq \left(1 + \frac{1}{\sqrt{2}} \right) \frac{R^{r+1} t^r}{(r-1)!} \|v_a\|_{2,t} \|v_b\|_{2,t},\tag{7.13}$$

valid when $k = 1$.

For $k > 1$ we twice substitute (7.12) into (7.11) to get

$$\tilde{\psi}_k(a, b, t) = -\frac{v_b(t)}{(k-1)!} \int_0^t (t-\tau)^{k-1} v_a(\tau) \, d\tau - \frac{1}{(k-2)!} \int_0^t v_b(s) \int_0^s (s-\tau)^{k-2} u_a(\tau) \, d\tau \, ds.$$

Now we compute, using the Cauchy-Bunyakovsky-Schwarz inequality,

$$\begin{aligned}\left| \int_0^t (t-\tau)^k v_a(\tau) \, d\tau \right| &\leq \left(\int_0^t (t-\tau)^{2k} \, d\tau \right)^{1/2} \left(\int_0^t v_a^2(\tau) \, d\tau \right)^{1/2} \\ &\leq \left(t^{2k} \int_0^t d\tau \right)^{1/2} \|v_a\|_{2,t} \\ &= t^{k+1/2} \|v_a\|_{2,t}.\end{aligned}\tag{7.14}$$

Twice using (7.14) we obtain the estimate

$$\begin{aligned}|\tilde{\psi}_k(a, b, t)| &\leq \frac{|v_b(t)| \|v_a\|_{2,t} t^{k-1/2}}{(k-1)!} + \frac{1}{(k-2)!} \int_0^t v_b(s) s^{k-3/2} \|v_a\|_{2,s} \, ds \\ &\leq \frac{|v_b(t)| \|v_a\|_{2,t} t^{k-1/2}}{(k-1)!} + \frac{\|v_a\|_{2,t}}{(k-2)!} \left(\int_0^t s^{2k-3} \, ds \right)^{1/2} \|v_b\|_{2,t} \\ &= \frac{|v_b(t)| \|v_a\|_{2,t} t^{k-1/2}}{(k-1)!} + \frac{\|v_a\|_{2,t} \|v_b\|_{2,t} t^{k-1}}{(k-2)!(2k-2)^{1/2}},\end{aligned}$$

where we have used the Cauchy-Bunyakovsky-Schwarz inequality in the second step. Therefore, we obtain the estimate, valid for $k > 1$,

$$\begin{aligned} \int_0^t |\tilde{\psi}_k(a, b, s)| \, ds &\leq \int_0^t \frac{\|v_b(s)\| \|v_a\|_{2,s} s^{k-1/2}}{(k-1)!} \, ds + \int_0^t \frac{\|v_a\|_{2,s} \|v_b\|_{2,s} s^{k-1}}{(k-2)!(2k-2)^{1/2}} \, ds \\ &= \frac{\|v_a\|_{2,t}}{(k-1)!} \int_0^t s^{k-1/2} |v_b(s)| \, ds + \frac{\|v_a\|_{2,t} \|v_b\|_{2,t}}{(k-2)!(2k-2)^{1/2}} \int_0^t s^{k-1} \, ds \\ &\leq \frac{\|v_a\|_{2,t}}{(k-1)!} \left(\int_0^t s^{2k-1} \, ds \right)^{1/2} \|v_b\|_{2,t} + \frac{\|v_a\|_{2,t} \|v_b\|_{2,t} t^k}{k(k-2)!(2k-2)^{1/2}} \\ &= \frac{\|v_a\|_{2,t} \|v_b\|_{2,t} t^k}{(2k)^{1/2}(k-1)!} + \frac{\|v_a\|_{2,t} \|v_b\|_{2,t} t^k}{k(k-2)!(2k-2)^{1/2}} \\ &\leq \frac{\|v_a\|_{2,t} \|v_b\|_{2,t} t^k}{(k-1)!} + \frac{\|v_a\|_{2,t} \|v_b\|_{2,t} t^k}{(k-2)!} \\ &\leq \frac{2 \|v_a\|_{2,t} \|v_b\|_{2,t} t^k}{(k-2)!}, \end{aligned}$$

where we again use the Cauchy-Bunyakovsky-Schwarz inequality, this time in the third step. Thus we have the estimate for $k > 1$ given by

$$\left| \int_0^t u_I \right| \leq 2 \frac{R^{r+k} t^{r+k-1}}{(r-1)!(k-2)!} \|v_a\|_{2,t} \|v_b\|_{2,t}. \quad (7.15)$$

Referring to equations (7.10), (7.13), and (7.15) we see that the lemma follows provided we choose $D = 2$. \blacksquare

This then gives us the coefficient of the Type 4 terms. To analyse the contribution of these terms, we also need the following estimate which is Lemma 4.2 of [Sussmann 1983].

7.19 LEMMA: *For $\{g_1, \dots, g_k\} \in \Gamma(TM)$, $\psi \in \mathcal{F}(M)$, $K \subset M$ compact, and $J = (\alpha_1, \dots, \alpha_r)$ a multi-index in $\{1, \dots, k\}$, there exists a constant σ depending only on K so that*

$$|g_J \psi(x)| \leq r! \sigma^r$$

for every $x \in K$.

Let us collect together the estimates for the Type 4 terms in the Chen-Fliess-Sussmann series, these corresponding to a multi-index of the form $I = (a, Z, b, J)$ with $|Z| = k$ and $|J| = r \geq 1$. The case $k = 0$ demands that $r \geq 2$. Since the estimate in part (i) of Lemma 7.18 involves I and \bar{I} , for each fixed J there will be $\binom{m}{2} = \frac{1}{2}m(m-1)$ such terms for $m > 1$ and 1 such term if $m = 1$. Let us write

$$P_m = \begin{cases} \frac{1}{2}m(m-1), & m > 1 \\ 1, & m = 1. \end{cases}$$

If $k \geq 1$ then for each fixed Z and J there are m^2 contributions as described by parts (ii) and (iii) of Lemma 7.18. Thus, if $C_4(t)$ denotes the contribution by all Type 4 terms to the

Chen-Fliess-Sussmann series for $\phi(\xi(t))$ we have

$$\begin{aligned} |C_4(t)| &\leq \left(\sum_{r=2}^{\infty} \frac{DR^r \sigma^{r+2} P_m m^r t^{r-1}}{(r-1)!} + \sum_{r=1}^{\infty} \frac{DR^{r+1} \sigma^{r+3} m^{r+3} t^r}{(r-1)!} \right. \\ &\quad \left. + \sum_{r=1}^{\infty} \sum_{k=2}^{\infty} \frac{DR^{r+k} \sigma^{r+k+2} m^{r+k+2} t^{r+k-1}}{(r-1)!(k-2)!} \right) \|u_{\theta}\|_{\max,t}^2 \\ &= t \left(\sum_{r=2}^{\infty} \frac{DR^r \sigma^{r+2} P_m m^r t^{r-2}}{(r-1)!} + \sum_{r=1}^{\infty} \frac{DR^{r+1} \sigma^{r+3} m^{r+3} t^{r-1}}{(r-1)!} \right. \\ &\quad \left. + \sum_{r=1}^{\infty} \sum_{k=2}^{\infty} \frac{DR^{r+k} \sigma^{r+k+2} m^{r+k+2} t^{r+k-2}}{(r-1)!(k-2)!} \right) \|u_{\theta}\|_{\max,t}^2. \end{aligned}$$

Now we note that for $T > 0$ the three series

$$\begin{aligned} &\sum_{r=2}^{\infty} \frac{DR^r \sigma^{r+2} P_m m^r T^{r-2}}{(r-1)!} \\ &\sum_{r=1}^{\infty} \frac{DR^{r+1} \sigma^{r+3} m^{r+3} T^{r-1}}{(r-1)!} \\ &\sum_{r=1}^{\infty} \sum_{k=2}^{\infty} \frac{DR^{r+k} \sigma^{r+k+2} m^{r+k+2} T^{r+k-2}}{(r-1)!(k-2)!} \end{aligned} \quad (7.16)$$

are convergent by, for example, the ratio test. Let us denote the limits of the series in (7.16) by $C_{4,1}$, $C_{4,2}$, and $C_{4,3}$, respectively. We note that these constants are independent of t and of $(\xi, u) \in \text{Traj}(\Sigma)$. Then we have shown that

$$|C_4(t)| \leq t(C_{4,1} + C_{4,2} + C_{4,3}) \|u_{\theta}\|_{\max,t}^2.$$

By making T sufficiently small we guarantee that $|C_4(t)| < \frac{1}{2} \|u_{\theta}\|_{\max,t}^2$ for all $t \in [0, T]$. Now, with T so defined, for each $t \in [0, T]$ we have

$$\begin{aligned} \phi(\xi(t)) &= C_3(t) + C_4(t) \\ &\geq \frac{1}{2} \|u_{\theta}\|_{\max,t}^2 - |C_4(t)| \\ &\geq 0, \end{aligned}$$

using Lemma 7.17. Theorem 5.4 now follows from property (vi) of Lemma 7.8 of ϕ .

8. Proofs of geometric results for affine systems

Next we prove Theorems 4.3 and 4.4 which concern the properties of a real-analytic generalised affine subbundle of TM . The idea behind our proof is that we use Theorems 5.2 and 5.4 as jumping off points. An important rôle is also played by the results of Section 3 that relate objects (subspaces, linear maps, vector-valued quadratic forms) constructed from an affine subbundle to similar objects constructed from a given set of local linear generators of that affine subbundle. In fact, there is one small generalisation that we need to make

to Proposition 3.8 to allow for control sets more general than all of \mathbb{R}^m . We shall say that U is **proper at x_0** if $\text{aff}(\ker(L_{x_0}^{\mathcal{F}}) \cap U) = \ker(L_{x_0}^{\mathcal{F}})$. If U is proper then both $\text{aff}(U)$ and $\text{conv}(U)$ are proper at x_0 . With this notation, we have the following generalisation of Proposition 3.8.

8.1 LEMMA: *Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle with the property that for $x_0 \in M$, $0_{x_0} \in \mathcal{A}_{x_0}$. Also let $\Sigma = (M, \mathcal{F} = \{f_0, f_1, \dots, f_m\}, U)$ be a control-affine system satisfying*

- (i) $f_0(x_0) = 0_{x_0}$,
- (ii) $\mathcal{A}_{\mathcal{F}} = \mathcal{A}$, and
- (iii) U is proper at x_0 .

If $S_{x_0} \subset T_{x_0}M$ is a subspace containing $L(\mathcal{A}_{x_0})$ then $\langle Z_{x_0}(\mathcal{A}), S_{x_0} \rangle = \langle Z_{x_0}(\mathcal{F}_U), S_{x_0} \rangle$.

Proof: The result will follow from Proposition 3.8 if we can show that $Z_{x_0}(\mathcal{F}_U) = Z_{x_0}(\mathcal{F}_{\mathbb{R}^m})$. Clearly $Z_{x_0}(\mathcal{F}_U) \subset Z_{x_0}(\mathcal{F}_{\mathbb{R}^m})$, so it is the opposite inclusion that needs proof. Let $u \in \mathbb{R}^m$. Since U is proper at x_0 there exists $\beta_1, \dots, \beta_k \in \mathbb{R}$ and $u_1, \dots, u_k \in \ker(L_{x_0}^{\mathcal{F}}) \cap U$ so that

$$\sum_{j=1}^k \beta_j = 1, \quad u = \sum_{j=1}^k \beta_j u_j.$$

Therefore

$$f_u = f_0 + \sum_{a=1}^m \sum_{j=1}^k \beta_j u_{j,a} f_a = \sum_{j=1}^k \beta_j \left(f_0 + \sum_{a=1}^m u_{j,a} f_a \right) = \sum_{j=1}^k \beta_j f_{u_j}.$$

Since $f_{u_j} \in Z_{x_0}(\mathcal{F}_U)$, $j \in \{1, \dots, k\}$, the lemma follows. \blacksquare

8.1. Proof of Theorem 4.3. Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle and let \mathcal{A} be an affine system in \mathcal{A} that is proper at x_0 . Let S_{x_0} be defined as in Theorem 4.3, noting by Propositions 3.7 and 3.8 that

$$S_{x_0} = \langle Z_{x_0}(\mathcal{A}_{\mathbb{R}^k}), L^{(2)}(\mathcal{A})_{x_0} \rangle$$

for any collection $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ of local linear generators for \mathcal{A} defined about x_0 . Property (iv) of Definition 3.1 ensures that there is a control-affine system $(\tilde{\mathcal{N}}, \tilde{\mathcal{F}} = \{\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_m\}, \tilde{U})$ with $\tilde{\mathcal{N}}$ a neighbourhood of x_0 , $\tilde{f}_0(x_0) = 0_{x_0}$, and \tilde{U} convex and proper, and so that $\tilde{\mathcal{F}}_{\tilde{U}}(x) \subset \text{conv}(\mathcal{A}(x))$ for all x in some neighbourhood of x_0 . Following Proposition 3.8 and Remark 3.17, the hypotheses of Theorem 4.3 ensure that there exists local linear generators $\mathcal{F} = \{f_0, f_1, \dots, f_m\}$ of \mathcal{A} defined on a neighbourhood \mathcal{N} of x_0 and satisfying

1. $f_0(x_0) = 0_{x_0}$,
2. $\langle Z_{x_0}(\mathcal{F}_{\mathbb{R}^m}), L^{(3)}(\mathcal{F})_{x_0} \rangle = T_{x_0}M$;
3. $B_{\mathcal{F}}(\langle Z_{x_0}(\mathcal{F}_{\mathbb{R}^m}), L^{(2)}(\mathcal{F})_{x_0} \rangle)$ is indefinite.

By Lemma 8.1 and Proposition 3.8, since \tilde{U} is convex and proper, we have

$$\langle Z_{x_0}(\mathcal{F}_{\mathbb{R}^m}), L^{(3)}(\mathcal{F})_{x_0} \rangle = \langle Z_{x_0}(\tilde{\mathcal{F}}_{\tilde{U}}), L^{(3)}(\tilde{\mathcal{F}})_{x_0} \rangle.$$

By Proposition 3.14 and Remark 3.17, since S_{x_0} is assumed to be second-order invariant with respect to \mathcal{A} at x_0 , we know that $B_{\tilde{\mathcal{F}}}(S_{x_0}) = B_{\tilde{\mathcal{F}}}(S_{x_0})$. This means that $\tilde{\Sigma}$ satisfies the hypotheses of Theorem 5.2, and so is STLC from x_0 . Let $\text{conv}(\mathcal{A})$ be the affine system in \mathcal{A} defined by $\text{conv}(\mathcal{A})(x) = \text{conv}(\mathcal{A}(x))$. Filippov [1988, Theorem 2.4] asserts that for any trajectory $\xi: [0, T] \rightarrow M$ of $\text{conv}(\mathcal{A})$, there is a sequence $\{\xi_j: [0, T] \rightarrow M\}$ of trajectories of \mathcal{A} with the property that

$$\lim_{j \rightarrow \infty} \xi_j(t) = \xi(t), \quad t \in [0, T],$$

with this convergence uniform in t . Since $\text{Traj}(\mathcal{A}_{\tilde{\Sigma}}, T) \subset \text{Traj}(\text{conv}(\mathcal{A}), T)$ for $T > 0$ sufficiently small, and since $\tilde{\Sigma}$ is STLC from x_0 , it now follows that \mathcal{A} is STLC from x_0 .

8.2. Proof of Theorem 4.4. Let $\mathcal{A} \subset TM$ be a real-analytic generalised affine subbundle and let \mathcal{A} be an affine system in \mathcal{A} for which $\mathcal{A}(x_0)$ is convex. Let S_{x_0} be defined as in Theorem 4.4, noting by Propositions 3.7 and 3.8 that

$$S_{x_0} = \langle Z_{x_0}(\mathcal{A}_{\mathbb{R}^k}), L^{(1)}(\mathcal{A})_{x_0} + L^{(\infty)}(\mathcal{A})_{x_0} \rangle$$

for any collection $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ of local linear generators for \mathcal{A} .

We first prove that $L(S_{x_0}^{\mathcal{A}}) = 0$. Following Proposition 3.15, it suffices to show that $L(S_{x_0}^{\mathcal{X}}) = 0$ where $S_{x_0}^{\mathcal{X}}$ is defined as in (3.2) for a collection $\mathcal{X} = \{X_0, X_1, \dots, X_k\}$ of local linear generators for \mathcal{A} . Let $\alpha, \beta \in \ker(L_{x_0}^{\mathcal{X}})$ and let $V_1, V_2 \in \Gamma(TM)$ satisfy $V_1(x_0), V_2(x_0) \in L^{(1)}(\mathcal{A})_{x_0}$. We then compute

$$\begin{aligned} \pi_{S_{x_0}}([V_1, [X_\alpha, V_2]](x_0)) - \pi_{S_{x_0}}([V_1, [X_\beta, V_2]](x_0)) &= \\ &= \sum_{j=1}^k (\alpha_j - \beta_j) \pi_{S_{x_0}}([V_1, [X_j, V_2]](x_0)). \end{aligned}$$

Since $[V_1, [X_j, V_2]](x_0) \in L^{(\infty)}(\mathcal{A})_{x_0}$, it follows that $S_{x_0}^{\mathcal{X}}$ is a constant map, or equivalently that $L(S_{x_0}^{\mathcal{X}}) = 0$.

Now we prove that any affine system \mathcal{A} in \mathcal{A} with $\mathcal{A}(x_0)$ compact is not STLC from x_0 . Continuity of \mathcal{A} ensures that $\mathcal{A}(x)$ is compact for x in a neighbourhood $\tilde{\mathcal{N}}$ of x_0 . If $\text{cl}(\tilde{\mathcal{N}})$ is compact, then, possibly shrinking $\tilde{\mathcal{N}}$, we can choose a control-affine system $\tilde{\Sigma} = (\tilde{\mathcal{N}}, \tilde{\mathcal{F}} = \{\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_m\}, \tilde{U})$ with $\tilde{\mathcal{N}}$ a neighbourhood of x_0 , with \tilde{U} proper, compact, and convex, and so that $\mathcal{A}(x) \subset \tilde{\mathcal{F}}_{\tilde{U}}(x)$ for each $x \in \tilde{\mathcal{N}}$. The hypotheses of Theorem 4.4 ensure the existence of local linear generators $\mathcal{F} = \{f_0, f_1, \dots, f_m\}$ for \mathcal{A} defined on a neighbourhood \mathcal{N} of x_0 and satisfying

1. $f_0(x_0) = 0_{x_0}$,
2. x_0 is a regular point for the distribution $L^{(\infty)}(\mathcal{F})$, and
3. $B_{\mathcal{F}}(S_{x_0})$ is definite.

Since \tilde{U} is proper and convex, the fact that $L(S_{x_0}^A) = 0$ implies that $B_{\mathcal{F}}(S_{x_0}) = B_{\tilde{\mathcal{F}}}(S_{x_0})$. Also, clearly x_0 is a regular point for $L^{(\infty)}(\tilde{\mathcal{F}}_1)$, implying that $\tilde{\Sigma}$ satisfies the hypotheses of Theorem 5.4, and so is not STLC from x_0 . Since $\text{Traj}(\mathcal{A}, T) \subset \text{Traj}(\mathcal{A}_{\tilde{\Sigma}}, T)$ for $T > 0$ sufficiently small, it follows that \mathcal{A} is also not STLC from x_0 .

References

- Agrachev, A. A. [1990] *Quadratic mappings in geometric control theory*, Journal of Soviet Mathematics, **51**(6), 2667–2734.
- [1999] *Is it possible to recognize local controllability in a finite number of differentiations?*, in *Open Problems in Mathematical Systems and Control Theory*, pages 15–18, Communications and Control Engineering Series, Springer-Verlag, New York-Heidelberg-Berlin, ISBN 1-85233-044-9.
- Basto-Gonçalves, J. [1985] *Local controllability of nonlinear systems*, Systems & Control Letters, **6**(3), 213–217.
- [1998] *Second-order conditions for local controllability*, Systems & Control Letters, **35**(5), 287–290.
- Bianchini, R. M. and Stefani, G. [1984] *Normal local controllability of order one*, International Journal of Control, **39**(4), 701–714.
- [1993] *Controllability along a trajectory: a variational approach*, SIAM Journal on Control and Optimization, **31**(4), 900–927.
- Bullo, F., Cortés, J., Lewis, A. D., and Martínez, S. [2004] *Vector-valued quadratic forms in control theory*, in *Sixty Open Problems in Mathematical Systems and Control Theory*, V. Blondel and A. Megretski, editors, pages 315–320, Princeton University Press, Princeton, New Jersey, ISBN 0-691-11748-9.
- Bullo, F., Lewis, A. D., and Lynch, K. M. [2002] *Controllable kinematic reductions for mechanical systems: Concepts, computational tools, and examples*, in *Proceedings of MTNS 2002*, South Bend, IN.
- Chen, K.-T. [1957] *Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula*, Annals of Mathematics. Second Series, **65**, 163–178.
- Filippov, A. F. [1988] *Differential Equations with Discontinuous Righthand Sides*, number 18 in Mathematics and its Applications (Soviet Series), Kluwer Academic Publishers, Dordrecht, ISBN 90-277-2699-X.
- Fliess, M. [1981] *Fonctionnelles causales non linéaires et indéterminées non commutatives*, Bulletin de la Société Mathématique de France, **109**, 3–40.
- Hermes, H. [1982] *On local controllability*, SIAM Journal on Control and Optimization, **20**(2), 211–220.
- [1991] *Nilpotent and high-order approximations of vector field systems*, SIAM Review. A Publication of the Society for Industrial and Applied Mathematics, **33**(2), 238–264.
- Hermes, H. and Kawski, M. [1987] *Local controllability of a single input, affine system*, in *Nonlinear Analysis and Applications*, volume 109 of *Lecture Notes in Pure and Applied Mathematics*, pages 235–248, Dekker Marcel Dekker, New York.
- Hirschorn, R. M. and Lewis, A. D. [2001] *Geometric first-order controllability conditions for affine connection control systems*, in *Proceedings of the 40th IEEE Conference on Decision and Control*, pages 4216–4221, Institute of Electrical and Electronics Engineers, Orlando, FL.
- [2002] *Local controllability of families of vector fields*, SIAM Journal on Control and Optimization, to appear.
- Kawski, M. [1987] *A necessary condition for local controllability*, in *Differential Geometry: the Interface Between Pure and Applied Mathematics*, volume 68 of *Contemporary Mathematics*, pages 143–155, American Mathematical Society, Providence, Rhode Island.
- [1990] *The complexity of deciding controllability*, Systems & Control Letters, **15**(1), 9–14.
- [1992] *High-order conditions for local controllability in practice*, in *Recent Advances in Mathematical Theory of Systems, Control, Networks and Signal Processing, II*, pages 271–276, Mita, Tokyo.
- Knobloch, H.-W. and Wagner, K. [1984] *On local controllability of nonlinear systems*, in *Dynamical Systems and Microphysics*, pages 243–286, Academic Press, New York.
- Petrov, N. N. [1977] *Local controllability*, Differential Equations, **12**(12), 1445–1550.
- Pomet, J.-B. [1999] *On the curves that may be approached by trajectories of a smooth control affine system*, Systems & Control Letters, **36**(2), 143–149.
- Sontag, E. D. [1988] *Controllability is harder to decide than accessibility*, SIAM Journal on Control and Optimization, **26**(5), 1106–1118.
- Stefani, G. [1986] *On the local controllability of a scalar-input control system*, in *Theory and Applications of Nonlinear Control Systems*, pages 167–179, North-Holland, Amsterdam/New York.
- [1988] *A sufficient condition for extremality*, in *Analysis and Optimization of Systems*, pages 270–281, number 111 in *Lecture Notes in Control and Information Sciences*, Springer-Verlag, New York-Heidelberg-Berlin.
- Sussmann, H. J. [1978] *A sufficient condition for local controllability*, SIAM Journal on Control and Optimization, **16**(5), 790–802.
- [1983] *Lie brackets and local controllability: a sufficient condition for scalar-input systems*, SIAM Journal on Control and Optimization, **21**(5), 686–713.
- [1987] *A general theorem on local controllability*, SIAM Journal on Control and Optimization, **25**(1), 158–194.

Sussmann, H. J. and Jurdjevic, V. [1972] *Controllability of nonlinear systems*, Journal of Differential Equations, **12**, 95-116.