# CONFIGURATION CONTROLLABILITY OF MECHANICAL SYSTEMS UNDERACTUATED BY ONE CONTROL*

JORGE CORTÉS†§ AND SONIA MARTÍNEZ‡§

**Abstract.** We investigate local configuration controllability for mechanical control systems within the affine connection formalism. We rely on previous results on controllability and series expansions for the evolution of mechanical systems starting from rest. Extending the work by Lewis for the single-input case, we are able to characterize local configuration controllability for systems with $n$ degrees of freedom and $n-1$ input forces.

**Key words.** nonlinear control, configuration controllability, symmetric product

**AMS subject classifications.** 53B05, 70Q05, 93B03, 93B05, 93B29

## 1. Introduction.

Mechanical control systems belong to a class of nonlinear systems whose controllability properties have not been fully characterized yet. Much work has been devoted to the study of their rich geometrical structure, both in the Hamiltonian framework (see [30] and references therein) and in the Lagrangian one, which is receiving increasing attention in the last years [5, 8, 17, 21, 23, 24, 25, 31]. This research is providing new insights and a bigger understanding of the accessibility and controllability aspects associated with them. In particular, the affine connection formalism has revealed to be very useful modeling different types of mechanical systems, such as natural ones (Lagrangian equal to kinetic energy minus potential energy) [24, 25], with symmetries [5, 9], with nonholonomic constraints [6, 23], etc. and, on the other hand, it has led to the development of some new techniques and control algorithms for approximate trajectory generation in controller design [4, 37]. Certainly, we shall see further progress in these directions in the next years.

Underactuated mechanical control systems are interesting to study both from a theoretical and a practical point of view. From a theoretical perspective, they offer a control challenge as they have non-zero drift, their linearization at zero velocity is not controllable, they are not static feedback linearizable and it is not known if they are dynamic feedback linearizable. That is, they are not amenable to standard techniques in control theory [13, 30]. From the practical point of view, they appear in numerous applications as a result of design choices motivated by the search of less costly devices, or as a result of a failure regime in fully actuated mechanical systems.

The work by Lewis and Murray [24, 25] on simple mechanical control systems has rendered strong conditions for configuration accessibility and sufficient conditions for configuration controllability. The conditions for the latter are based on the sufficient conditions that Sussmann obtained for general affine control systems [35].

*This research was partially supported by FPU and FPI grants from the Spanish Ministerio de Educación y Cultura and Ministerio de Ciencia y Tecnología, respectively, and grant DGICYT PGC2000-2191-E.

†Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1308 W. Main St., Urbana, IL 61801, USA, jcortes@uiuc.edu, Ph. +1-217-244-8734, Fax. +1-217-244-1653

‡Escola Universitària Politèctnica de Vilanova i la Geltrú, Universidad Politécnica de Cataluña, Av. V. Balaguer s/n, Vilanova i la Geltrú 08800, Spain, soniam@mat.upc.es, Ph. +34 938967713, Fax. +34 938967700

§*Former address*: Laboratory of Dynamical Systems, Mechanics and Control, Instituto de Matemáticas y Física Fundamental, Consejo Superior de Investigaciones Científicas, Serrano 123, 28006 Madrid, Spain

It is worth noting that these conditions are not invariant under input transformations. As controllability is the more interesting property in practice, more research is needed in order to sharpen the configuration controllability conditions. Whatever these conditions might be, they will result harder to check than the ones for accessibility, since controllability is inherently a more difficult property to establish [14, 33]. Lewis [21] investigated and fully solved the single-input case, building on previous results by Sussmann for general scalar-input systems [34]. The recent work by Bullo [3] on series expansions for the evolution of a mechanical control system starting from rest has given the necessary tools to tackle this problem in the much more involved multi-input case. In this paper, we characterize local configuration controllability for systems whose number of inputs and degrees of freedom differs by one. Examples include autonomous vehicles (like aircraft takeoff and landing models [11, 28], underwater vehicles [32]), robotic manipulators with a passive joint [26] and locomotion devices (such as the robotic leg [23] or the quadrotor [29]). In addition, fully actuated mechanical systems may temporarily suffer from an actuator failure turning them into underactuated systems by one control, in which case, the knowledge of their controllability properties becomes relevant within a robust design perspective. Interestingly, the differential flatness properties of this type of underactuated mechanical control systems have also been characterized in intrinsic geometric terms [32].

Both results, Lewis' and ours, can be seen as particular cases of the following conjecture, which remains open: *The system is locally configuration controllable at a point if and only if there exists a basis of inputs satisfying the sufficient conditions for local configuration controllability at that point.* The conjecture relies on the fact we have mentioned before: the lack of invariance of the sufficient conditions under input transformations. It is remarkable to note that local controllability has not been characterized yet for general control systems, even for the single input case (in this regard see [12, 34, 35]).

The paper is organized as follows. In Section 2, we describe the affine connection framework for mechanical control systems and recall the controllability notions we shall consider on them. In Section 3 we review the existing results concerning configuration controllability [24, 25] and the series expansion for the evolution of a mechanical control system starting from rest developed by Bullo in [3]. In Section 4 we briefly recall the single-input case solved by Lewis and properly state his conjecture. Section 5 contains the main contributions of this paper. In Section 6 we treat two examples to illustrate the results. Finally, we present our conclusions in Section 7.

**2. Simple mechanical control systems.** Let $Q$ be a $n$-dimensional manifold. We will denote by $TQ$ the tangent bundle of $Q$, by $\mathfrak{X}(Q)$ the set of vector fields on $Q$ and by $C^\infty(Q)$ the set of smooth functions on $Q$. Throughout the paper, the manifold $Q$ and the mathematical objects defined on it will be assumed analytic.

A *simple mechanical control system* is defined by a triple $(Q, g, \mathcal{F})$, where $Q$ is the manifold of configurations of the system, $g$ is a Riemannian metric on $Q$ and $\mathcal{F} = \{F^1, \ldots, F^m\}$ is a set of $m$ linearly independent 1-forms on $Q$, which physically correspond to forces or torques.

Associated with the metric $g$ there is a natural affine connection, called the *Levi-Civita* connection. An *affine connection* [1, 18] is defined as an assignment

$$\begin{array}{rccc} \nabla : & \mathfrak{X}(Q) \times \mathfrak{X}(Q) & \longrightarrow & \mathfrak{X}(Q) \\ & (X, Y) & \longmapsto & \nabla_X Y \end{array}$$

which is $\mathbb{R}$-bilinear and satisfies $\nabla_{fX} Y = f \nabla_X Y$ and $\nabla_X (fY) = f \nabla_X Y + X(f) Y$, for any $X,\, Y \in \mathfrak{X}(Q)$, $f \in C^\infty(Q)$. A curve $c : [a,b] \longrightarrow Q$ is a *geodesic* for $\nabla$ if $\nabla_{\dot c(t)}\dot c(t) = 0$. Locally, the condition for a curve $t \longmapsto (q^1(t), \ldots, q^n(t))$ to be a geodesic can be expressed as

$$\ddot q^a + \Gamma^a_{bc}\dot q^b \dot q^c = 0\,, \quad 1 \le a \le n\,, \tag{2.1}$$

where the $\Gamma^a_{bc}(q)$ are the Christoffel symbols of the affine connection, that is, they are given by $\nabla_{\frac{\partial}{\partial q^b}} \dfrac{\partial}{\partial q^c} = \Gamma^a_{bc} \dfrac{\partial}{\partial q^a}$. The geodesic equation (2.1) is a first-order differential equation on $TQ$. The vector field corresponding to this first-order equation is given in coordinates by

$$S = v^a \frac{\partial}{\partial q^a} - \Gamma^a_{bc} v^b v^c \frac{\partial}{\partial v^a}\,,$$

and is called the *geodesic spray* of the affine connection $\nabla$. Hence, the integral curves of the geodesic spray $S$, $(q^a, \dot q^a)$ are the solutions of the geodesic equation.

The Levi-Civita connection $\nabla^g$ is determined by the formula

$$\begin{aligned} 2\, g(\nabla^g_X Y, Z) = {} & (X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) \\ & + g(Y,[Z,X]) - g(X,[Y,Z]) + g(Z,[X,Y]))\,, \quad X, Y, Z \in \mathfrak{X}(Q)\,. \end{aligned}$$

One can compute the Christoffel symbols of $\nabla^g$ to be

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad} \left( \frac{\partial g_{db}}{\partial q^c} + \frac{\partial g_{dc}}{\partial q^b} - \frac{\partial g_{bc}}{\partial q^d} \right)\,,$$

where $(g^{ad})$ denotes the inverse of the inertia matrix $(g_{da}) = (g(\frac{\partial}{\partial q^d}, \frac{\partial}{\partial q^a}))$.

The metric tensor $g$ induces a bundle isomorphism $\flat_g : TQ \longrightarrow T^*Q$ given by $\flat_g(X)(Y) = g(X,Y)$. Instead of the input forces $F^1, \ldots, F^m$, we shall make use of the vector fields $Y_1, \ldots, Y_m$, defined as $Y_i = \flat_g^{-1}(F^i)$. Roughly speaking, this corresponds to consider "accelerations" rather than forces. If $Y_i = Y_i^a(q)\frac{\partial}{\partial q^a}$, the control equations for the simple mechanical control system read in coordinates

$$\begin{aligned} \dot q^a &= v^a \\ \dot v^a &= -\Gamma^a_{bc}\dot q^b \dot q^c + \textstyle\sum_{i=1}^m u_i(t) Y_i^a(q)\,, \quad 1 \le a \le n\,. \end{aligned}$$

These equations can be written in a coordinate-free way as

$$\nabla^g_{\dot c(t)} \dot c(t) = \sum_{i=1}^m u^i(t) Y_i(c(t))\,. \tag{2.2}$$

The inputs we will consider come from the set $\mathcal{U} = \{u : [0,T] \to \mathbb{R}^m \,|\, T > 0$, $u$ is measurable and $\|u\| \le 1\}$, where

$$\|u\| = \sup_{t\in[0,T]} \|u(t)\|_\infty = \sup_{t\in[0,T]} \max_{l=1,\ldots,m} |u_l(t)|\,.$$

We can use a general affine connection in (2.2) instead of the Levi-Civita connection without changing the structure of the equation. This is particularly interesting, since nonholonomic mechanical control systems give also rise to equations of the form (2.2) by means of the so-called nonholonomic affine connection (see [23]). Therefore, the discussion throughout the paper is carried out for a general affine connection $\nabla$.

We can turn (2.2) into a general affine control system with drift

$$\dot{x}(t) = f(x(t)) + \sum u^i(t) g_i(x(t)). \tag{2.3}$$

To do this we need another bit of notation. The vertical lift of a vector field $X$ on $Q$ is the vector field $X^v$ on $TQ$ defined as

$$X^v(v_q) = \frac{d}{dt}\Big|_{t=0} (v_q + tX(q)) .$$

In coordinates, if $X = X^a \frac{\partial}{\partial q^a}$, one can check that $X^v = X^a \frac{\partial}{\partial v^a}$. Then, the second-order equation (2.2) on $Q$ can be written as the first-order system on $TQ$

$$\dot{v} = S(v) + \sum_{i=1}^{m} u^i(t) Y_i^v(v) , \tag{2.4}$$

where $S$ is the geodesic spray associated with the affine connection $\nabla$.

**2.1. Controllability notions.** The control equations for the mechanical system (2.4) are nonlinear. The standard techniques in control theory [30], as for example the linearization around an equilibrium point or linearization by feedback, do not yield satisfactory results in the analysis of its controllability properties, in the sense that they do not provide necessary and sufficient conditions characterizing them.

The point in the approach of Lewis and Murray to simple mechanical control systems is precisely to focus on what is happening to configurations, rather than to states, since in many of these systems, configurations may be controlled, but not configurations and velocities at the same time. The basic question they pose is "what is the set of configurations which are attainable from a given configuration starting from rest?" Moreover, since we deal with objects defined on the configuration manifold $Q$, we expect to find answers on $Q$, although the control system (2.4) lives in $TQ$.

DEFINITION 2.1. *A* solution *of (2.2) is a pair $(c,u)$, where $c : [0,T] \longrightarrow Q$ is a piecewise smooth curve and $u \in \mathcal{U}$ such that $(\dot{c},u)$ satisfies the first order control system (2.4).*

Consider $q_0 \in Q$, $(q_0, 0_{q_0}) \in T_{q_0}Q$ and let $U \subset Q$, $\overline{U} \subset TQ$ be neighborhoods of $q_0$ and $(q_0, 0_{q_0})$, respectively. Define

$$\mathcal{R}_Q^U(q_0,T) = \left\{ q \in Q \;\middle|\; \begin{array}{c} \text{there exists a solution } (c,u) \text{ of (2.2) such that} \\ \dot{c}(0) = 0_{q_0}, c(t) \in U \text{ for } t \in [0,T] \text{ and } \dot{c}(T) \in T_qQ \end{array} \right\}$$

$$\mathcal{R}_{TQ}^{\overline{U}}(q_0,T) = \left\{ (q,v) \in TQ \;\middle|\; \begin{array}{c} \text{there exists a solution } (c,u) \text{ of (2.2) such that } \dot{c}(0) = \\ 0_{q_0}, (c(t),\dot{c}(t)) \in \overline{U} \text{ for } t \in [0,T] \text{ and } \dot{c}(T) = v \in T_qQ \end{array} \right\}$$

and denote

$$\mathcal{R}_Q^U(q_0, \leq T) = \cup_{0 \leq t \leq T} \mathcal{R}_Q^U(q_0,t) \;, \quad \mathcal{R}_{TQ}^{\overline{U}}(q_0, \leq T) = \cup_{0 \leq t \leq T} \mathcal{R}_{TQ}^{\overline{U}}(q_0,t) \,.$$

Now, we recall the notions of accessibility considered in [24].

DEFINITION 2.2. *The system (2.2) is* locally configuration accessible (LCA) *at $q_0 \in Q$ if there exists $T > 0$ such that $\mathcal{R}_Q^U(q_0, \leq t)$ contains a non-empty open set of $Q$, for all neighborhoods $U$ of $q_0$ and all $0 \leq t \leq T$. If this holds for any $q_0 \in Q$ then the system is called locally configuration accessible.*

DEFINITION 2.3. *The system (2.2) is* locally accessible (LA) at $q_0 \in Q$ and zero velocity *if there exists $T > 0$ such that $\mathcal{R}^{\overline{U}}_{TQ}(q_0, \leq t)$ contains a non-empty open set of $TQ$, for all neighborhoods $\overline{U}$ of $(q_0, 0_{q_0})$ and all $0 \leq t \leq T$. If this holds for any $q_0 \in Q$ then the system is called locally accessible at zero velocity.*

We shall focus our attention on the following concepts of controllability [24].

DEFINITION 2.4. *The system (2.2) is* small-time locally configuration controllable (STLCC) at $q_0 \in Q$ *if there exists $T > 0$ such that $\mathcal{R}^U_Q(q_0, \leq t)$ contains a non-empty open set of $Q$ to which $q_0$ belongs, for all neighborhoods $U$ of $q_0$ and all $0 \leq t \leq T$. If this holds for any $q_0 \in Q$ then the system is called small-time locally configuration controllable.*

DEFINITION 2.5. *The system (2.2) is* small-time locally controllable (STLC) at $q_0 \in Q$ and zero velocity *if there exists $T > 0$ such that $\mathcal{R}^{\overline{U}}_{TQ}(q_0, \leq t)$ contains a non-empty open set of $TQ$ to which $(q_0, 0_{q_0})$ belongs, for all neighborhoods $\overline{U}$ of $(q_0, 0_{q_0})$ and all $0 \leq t \leq T$. If this holds for any $q_0 \in Q$ then the system is called small-time locally controllable at zero velocity.*

**3. Existing results.** Here we review some accessibility and controllability results obtained in [24, 25] and expose the work by Bullo [3] in describing the evolution of mechanical control systems via a series expansion.

**3.1. On controllability.** Given an affine connection $\nabla$ on $Q$, the *symmetric product* of two vector fields $X, Y \in \mathfrak{X}(Q)$ is defined by

$$\langle X : Y \rangle = \nabla_X Y + \nabla_Y X \,.$$

The geometric meaning of the symmetric product is the following [22]: a *geodesically invariant* distribution $\mathcal{D}$ is a distribution such that for every geodesic $c(t)$ of $\nabla$ starting from a point in $\mathcal{D}$, $\dot{c}(0) \in \mathcal{D}_{c(0)}$, we have that $\dot{c}(t) \in \mathcal{D}_{c(t)}$. Then, one can prove that $\mathcal{D}$ is geodesically invariant if and only if $\langle X : Y \rangle \in \mathcal{D}$, $\forall X, Y \in \mathcal{D}$.

Given the input vector fields $\mathcal{Y} = \{Y_1, \ldots, Y_m\}$, let us denote by $\overline{Sym}(\mathcal{Y})$ the distribution obtained by closing the set $\mathcal{Y}$ under the symmetric product and by $\overline{Lie}(\mathcal{Y})$ the involutive closure of $\mathcal{Y}$. With these ingredients, one can prove

THEOREM 3.1. *([24]) The control system (2.2) is locally configuration accessible at $q$ (respectively locally accessible at $q$ and zero velocity) if $\overline{Lie}(\overline{Sym}(\mathcal{Y}))_q = T_qQ$ (respec. $\overline{Sym}(\mathcal{Y})_q = T_qQ$).*

If $P$ is a symmetric product of vector fields in $\mathcal{Y}$, we let $\gamma_i(P)$ denote the number of occurrences of $Y_i$ in $P$. The *degree* of $P$ will be $\gamma_1(P) + \cdots + \gamma_m(P)$. We shall say that $P$ is *bad* if $\gamma_i(P)$ is even for each $1 \leq i \leq m$. We say that $P$ is *good* if it is not bad. The following theorem gives sufficient conditions for STLCC.

THEOREM 3.2. *Suppose that the system is LCA at $q$ (respectively, LA at $q$ and zero velocity) and that $\mathcal{Y}$ is such that every bad symmetric product $P$ at $q$ in $\mathcal{Y}$ can be written as a linear combination of good symmetric products at $q$ of lower degree than $P$. Then (2.2) is STLCC at $q$ (respec. STLC at $q$ and zero velocity).*

This theorem was proved in [24], adapting previous work by Sussmann [35] on general control systems of the form (2.3). Throughout the paper, we will refer to the conditions of every bad symmetric product at $q$ being a linear combination of good symmetric products at $q$ of lower degree as the *sufficient conditions for STLCC*.

**3.2. Series expansion.** Within the realm of geometric control theory, series expansions play a key role in the study of nonlinear controllability [2, 15, 34, 35], trajectory generation and motion planning problems [4, 19, 20, 29], etc. In [27], Magnus describes the evolution of systems on a Lie group. In [7, 10, 16, 36] a general framework is developed to describe the evolution of a nonlinear system via the so-called Chen-Fliess series and its factorization.

In the context of mechanical control systems, the work by Bullo in [3] describes the evolution of the trajectories with zero initial velocity via a series expansion on the configuration manifold $Q$. In this section we describe the series expansion, which will be key in the subsequent discussion. Before doing so, however, we need to introduce some notation on analyticity over complex neighborhoods.

Let $q_0 \in Q$. By selecting a coordinate chart around $q_0$, we locally identify $Q \equiv \mathbb{R}^n$. In this way, we write $q_0 \in \mathbb{R}^n$. Let $\sigma$ be a positive scalar, and define the complex $\sigma$-neighborhood of $q_0$ in $\mathbb{C}^n$ as $B_\sigma(q_0) = \{z \in \mathbb{C}^n \mid \|z - q_0\| < \sigma\}$. Let $f$ be a real analytic function on $\mathbb{R}^n$ that admits a bounded analytic continuation over $B_\sigma(q_0)$. The norm of $f$ is defined as

$$\|f\|_\sigma \triangleq \max_{z \in B_\sigma(q_0)} |f(z)|,$$

where $f$ denotes both the function over $\mathbb{R}^n$ and its analytic continuation. Given a time-varying vector field $(q,t) \mapsto Z(q,t) = Z_t(q)$, let $Z_t^i$ be its $i$th component with respect to the usual basis on $\mathbb{R}^n$. Assuming $t \in [0,T]$, and assuming that every component function $Z_t^i$ is analytic over $B_\sigma(q_0)$, we define the norm of $Z$ as

$$\|Z\|_{\sigma,T} \triangleq \max_{t \in [0,T]} \max_{i \in \{1,\dots,n\}} \|Z_t^i\|_\sigma.$$

In what follows, we will often simplify notation by neglecting the subscript $T$ in the norm of a time-varying vector field. Finally, given an affine connection $\nabla$ with Christoffel symbols $\{\Gamma^i_{jk} \mid i,j,k \in \{1,\dots,n\}\}$, introduce the notation:

$$\|\Gamma\|_\sigma \triangleq \max_{ijk} \left\|\Gamma^i_{jk}\right\|_\sigma.$$

In the sequel, we let

$$Z(q,t) = \sum_{i=1}^m u_i(t) Y_i(q).$$

THEOREM 3.3. *([3]) Let $c(t)$ be the solution of equation (2.2) with input given by $Z(q,t)$ and with initial conditions $c(0) = q_0$, $\dot c(0) = 0$. Let the Christoffel symbols $\Gamma^i_{jk}(q)$ and the vector field $Z(q,t)$ be uniformly integrable and bounded analytic in $Q$. Define recursively the time varying vector fields*

$$V_1(q,t) = \int_0^t Z(q,s)ds,$$
$$V_k(q,t) = -\frac{1}{2}\sum_{j=1}^{k-1} \int_0^t \langle V_j(q,s) : V_{k-j}(q,s)\rangle\, ds,\ k \geq 2,$$

*where $q$ is maintained fixed at each integral. Select a coordinate chart around the point $q_0 \in Q$, let $\sigma > \sigma'$ be two positive constants, and assume that*

$$\|Z\|_\sigma T^2 < L \triangleq \min\left\{\frac{\sigma - \sigma'}{2^4 n^2(n+1)}, \frac{1}{2^4 n(n+1)\|\Gamma\|_\sigma}, \frac{\eta^2(\sigma' n^2 \|\Gamma\|_{\sigma'})}{n^2\|\Gamma\|_{\sigma'}}\right\}. \quad (3.1)$$

*Then the series* $(q,t) \longmapsto \sum_{k=1}^{\infty} V_k(q,t)$ *converges absolutely and uniformly in* $t$ *and* $q$*, for all* $t \in [0,T]$ *and for all* $q \in B_{\sigma'}(q_0)$*, with the* $V_k$ *satisfying the bound*

$$\|V_k\|_{\sigma'} \leq L^{1-k}\,\|Z\|_{\sigma}^{k}\,t^{2k-1}, \tag{3.2}$$

*Over the same interval, the solution* $c(t)$ *satisfies*

$$\dot c(t) = \sum_{k=1}^{\infty} V_k(c(t),t)\,. \tag{3.3}$$

This theorem generalizes previous results obtained in [4] under the assumption of small amplitude forcing. The first few terms of the series (3.3) can be computed to obtain

$$\begin{aligned}\dot c(t)=&\overline{Z}(c(t),t) - \frac{1}{2}\overline{\langle \overline{Z} : \overline{Z}\rangle}(c(t),t) + \frac{1}{2}\overline{\left\langle \overline{\langle \overline{Z} : \overline{Z}\rangle} : \overline{Z}\right\rangle}(c(t),t) \\ &-\frac{1}{2}\overline{\left\langle \overline{\left\langle \overline{\langle \overline{Z} : \overline{Z}\rangle} : \overline{Z}\right\rangle} : \overline{Z}\right\rangle}(c(t),t) - \frac{1}{8}\overline{\left\langle \overline{\langle \overline{Z} : \overline{Z}\rangle} : \overline{\langle \overline{Z} : \overline{Z}\rangle}\right\rangle}(c(t),t) + O(\|Z\|_{\sigma}^{5}t^{9}),\end{aligned} \tag{3.4}$$

where $\overline{Z}(q,t) \equiv \int_0^t Z(q,s)ds$ and so on.

**4. The single-input case.** Theorem 3.2 gives us sufficient conditions for small-time local configuration controllability. A natural concern both from the theoretical and the practical point of view is to try to sharpen this controllability test. Lewis [21] investigated the single-input case and proved the next result.

THEOREM 4.1. *Let* $(Q,g)$ *be an analytic manifold with an affine connection* $\nabla$*. Let* $Y$ *be an analytic vector field on* $Q$ *and* $q_0 \in Q$*. Then the system*

$$\nabla_{\dot c(t)}\dot c(t) = u(t)Y(c(t))$$

*is locally configuration controllable at* $q_0 \in Q$ *if and only if* $\dim Q = 1$*.*

The fact of being able to completely characterize STLCC in the single-input case (something which has not been accomplished yet for general control systems of the form (2.3)) suggests that understanding local configuration controllability for mechanical systems may be possible. More precisely, examining the single-input case, one can deduce that if (2.2) is STLCC at $q_0$ then $\dim Q = 1$, which implies $\langle Y : Y\rangle(q_0) \in \operatorname{span}\{Y(q_0)\}$, i.e. sufficient conditions for STLCC are also necessary. Can this be extrapolated to the multi-input case? The following conjecture was posed by Lewis:

> *Let a mechanical control system* (2.2) *be locally configuration accessible at* $q_0 \in Q$*. Then it is STLCC at* $q_0$ *if and only if there exists a basis of input vector fields which satisfies the sufficient conditions for STLCC at* $q_0$*.*

Theorem 4.1 implies that the conjecture is true for $m = 1$. In the following section we prove that this conjecture is also valid for $m = n-1$.

**5. Mechanical systems underactuated by one control.** Here we focus our attention on mechanical control systems of the form (2.2) which has $n$ degrees of freedom and $m = n-1$ control input vector fields. The following lemma, taken from [34], will be helpful in the proof of the theorem of this section.

LEMMA 5.1. *Let* $Q$ *be a* $n$*-dimensional analytic manifold. Given* $q_0 \in Q$ *and* $X_1,\ldots,X_p \in \mathfrak{X}(Q)$*,* $p \leq n$*, linearly independent vector fields, there exists a function* $\phi : Q \longrightarrow \mathbb{R}$ *satisfying the properties*

1. *$\phi$ is analytic*
2. $\phi(q_0) = 0$
3. $X_1(\phi) = \cdots = X_{p-1}(\phi) = 0$ *on a neighborhood $V$ of $q_0$*
4. $X_p(\phi)(q_0) = -1$
5. *Within any neighborhood of $q_0$ there exists points $q$ where $\phi(q) < 0$ and $\phi(q) > 0$.*

*Proof.* Let $Z_1, \ldots, Z_n$ be vector fields defined in a neighborhood of $q_0$ such that $\{Z_1(q_0), \ldots, Z_n(q_0)\}$ forms a basis for $T_{q_0}Q$ and $Z_i = X_i$, $1 \leq i \leq p-1$, $Z_p = -X_p$. Let $t_i \longmapsto \Psi_i(t)$ be the flow of $Z_i$, $1 \leq i \leq n$. In a sufficiently small neighborhood $V$ of $q_0$, any point $q$ may be expressed as $q = \Psi_1(t_1) \circ \cdots \circ \Psi_n(t_n)(q_0)$ for some unique $n$-tuple $(t_1, \ldots, t_n) \in \mathbb{R}^n$. Define $\phi(q) = t_p$. It is a simple exercise to verify that $\phi$ satisfies the required properties. □

Next, we state and prove the main result of the paper.

THEOREM 5.2. *Let $Q$ be a $n$-dimensional analytic manifold and let $Y_1, \ldots, Y_{n-1}$ be analytic vector fields on $Q$. Consider the control system*

$$\nabla_{\dot{c}(t)} \dot{c}(t) = \sum_{i=1}^{n-1} u_i(t) Y_i(c(t)) \,, \tag{5.1}$$

*and assume that it is locally configuration accessible at $q_0 \in Q$. Then the system is locally configuration controllable at $q_0$ if and only if there exists a basis of input vector fields satisfying the sufficient conditions for STLCC at $q_0$.*

A rough sketch of the proof is the following: because of the hypotheses of the theorem, we only need to check that the symmetric products of degree two of a given basis of the input distribution, when evaluated at $q_0$, are linear combinations of good products of degree one. To verify this, we associate with the given basis a symmetric matrix $A$, in such a way that this basis satisfies the sufficient conditions for STLCC if and only if the diagonal elements of $A$ are all zero. If this is not the case, we search for a change of basis $B$ such that the new basis has an associated matrix $A$ with zeroes in its diagonal. This is equivalent to solving a quadratic equation in $B$. In order to ensure that a solution to this equation exists, we have to explore the different possibilities that may occur regarding the various radicands involved. Finally, we discard the situations in which the equation is not solvable by a contradiction argument with the controllability assumption (see Figure 5.1).

*Proof.* We only need to prove one implication (the other one is Theorem 3.2). Let us suppose that the system is locally configuration controllable at $q_0$. Let $\mathcal{D}$ denote the input distribution. Either one of the following is true,

1. $\forall Y_1, Y_2 \in \mathcal{D}$, $\langle Y_1 : Y_2 \rangle (q_0) \in \mathcal{D}_{q_0}$.
2. There exist $Y_1$, $Y_2 \in \mathcal{D}$ such that $\langle Y_1 : Y_2 \rangle (q_0) \notin \mathcal{D}_{q_0}$.

In case (i), there is nothing to prove since any basis of input vector fields satisfies the sufficient conditions for STLCC at $q_0$. In case (ii), it is clear that one can choose $Y_1$, $Y_2 \in \mathcal{D}$, linearly independent at $q_0$ and such that $\langle Y_1 : Y_2 \rangle (q_0) \notin \mathcal{D}_{q_0}$ (if $Y_1$, $Y_2$ in (ii) are linearly dependent, then $\langle Y_1 : Y_1 \rangle (q_0) \notin \mathcal{D}_{q_0}$. Take any $Y_2$ linearly independent with $Y_1$. If $\langle Y_1 : Y_2 \rangle (q_0) \in \mathcal{D}_{q_0}$, and define a new $Y_2'$ by $Y_1 + Y_2$). Then, we can complete the set $\{Y_1(q_0), Y_2(q_0)\}$ to a basis of $\mathcal{D}_{q_0}$,

$$\{Y_1(q_0), Y_2(q_0), \ldots, Y_m(q_0)\}$$

such that $\operatorname{span}\{Y_1(q_0), Y_2(q_0), \ldots, Y_m(q_0), \langle Y_1 : Y_2 \rangle (q_0)\} = T_{q_0}Q$. In this basis, the symmetric products of degree two of the vector fields $\{Y_1,\dots,Y_m\}$ at $q_0$ are expressed,

$$\begin{aligned}
\langle Y_1 : Y_1\rangle(q_0) &= lc(Y_1(q_0),\dots,Y_m(q_0)) + a_{11}\langle Y_1 : Y_2\rangle(q_0)\\
&\vdots\\
\langle Y_m : Y_m\rangle(q_0) &= lc(Y_1(q_0),\dots,Y_m(q_0)) + a_{mm}\langle Y_1 : Y_2\rangle(q_0)\\
\langle Y_1 : Y_2\rangle(q_0) &= \qquad\qquad\qquad\qquad\qquad a_{12}\langle Y_1 : Y_2\rangle(q_0)\\
\langle Y_1 : Y_3\rangle(q_0) &= lc(Y_1(q_0),\dots,Y_m(q_0)) + a_{13}\langle Y_1 : Y_2\rangle(q_0)\\
&\vdots\\
\langle Y_{m-1} : Y_m\rangle(q_0) &= lc(Y_1(q_0),\dots,Y_m(q_0)) + a_{m-1m}\langle Y_1 : Y_2\rangle(q_0)\,,
\end{aligned}$$

where $lc(Y_1(q_0),\dots,Y_m(q_0))$ means a linear combination of $Y_1(q_0),\dots,Y_m(q_0)$. The coefficients $a_{ij}$ define a symmetric matrix $A=(a_{ij})\in\mathbb{R}^{m\times m}$. Observe that if $a_{11}=\cdots=a_{mm}=0$, then the bad symmetric products $\langle Y_i : Y_i\rangle(q_0)$ are in $\mathcal{D}_{q_0}$ and we have finished. Suppose then that the opposite situation is true, that is, there exists $s=s_1$ such that $a_{s_1s_1}\neq 0$.

What we are going to prove now is that, under the hypothesis of STLCC at $q_0$, there exists a change of basis $B=(b_{jk})$, $\det B\neq 0$, providing new vector fields in $\mathcal{D}$,

$$Y'_j=\sum_{k=1}^{m} b_{jk}Y_k\,,\quad 1\le j\le m\,,$$

which satisfy the sufficient conditions for STLCC at $q_0$. Since

$$\begin{aligned}
\langle Y'_j : Y'_j\rangle(q_0) &= \sum_{k,l=1}^{m} b_{jk}b_{jl}\langle Y_k : Y_l\rangle(q_0)\\
&= \sum_{k=1}^{m} b_{jk}^2\langle Y_k : Y_k\rangle(q_0) + 2\sum_{1\le k<l\le m} b_{jk}b_{jl}\langle Y_k : Y_l\rangle(q_0) \qquad (5.2)\\
&= lc(Y'_1(q_0),\dots,Y'_m(q_0)) + \left(\sum_{k=1}^{m} b_{jk}^2a_{kk} + 2\sum_{1\le k<l\le m} b_{jk}b_{jl}a_{kl}\right)\langle Y_1 : Y_2\rangle(q_0)\,,
\end{aligned}$$

the matrix $B$ we are looking for must fulfill

$$\sum_{k=1}^{m} b_{jk}^2a_{kk} + 2\sum_{1\le k<l\le m} b_{jk}b_{jl}a_{kl}=0\,,\quad 1\le j\le m\,, \qquad (5.3)$$

or, equivalently,

$$(BAB^T)_{jj}=0\,,\quad 1\le j\le m\,.$$

Note that, since $a_{s_1s_1}\neq 0$, this is equivalent to

$$b_{js_1}=\frac{-\sum_{k\neq s_1} b_{jk}a_{ks_1}}{a_{s_1s_1}} \pm\frac{\sqrt{(\sum_{k\neq s_1} b_{jk}a_{ks_1})^2 - a_{s_1s_1}(\sum_{k\neq s_1} b_{jk}^2a_{kk} + 2\sum_{k<l,k,l\neq s_1} b_{jk}b_{jl}a_{kl})}}{a_{s_1s_1}}\,,$$

for each $1 \leq j \leq m$. After some computations, the radicand of this expression becomes

$$\sum_{k \neq s_1} b_{jk}^2 (a_{ks_1}^2 - a_{s_1 s_1} a_{kk}) + 2 \sum_{k<l, k,l \neq s_1} b_{jk} b_{jl} (a_{ks_1} a_{ls_1} - a_{s_1 s_1} a_{kl}) .$$

If this radicand is zero, it would imply that the matrix $B$ should be singular in order to satisfy (5.3). We must ensure then that it is possible to select $B$ such that the radicand is different from zero. We do this in the following, studying several cases that can occur. Denoting by

$$a_{kl}^{(2)} = a_{ks_1} a_{ls_1} - a_{s_1 s_1} a_{kl} , \quad k,l \in \{1, \ldots, m\} \setminus \{s_1\} ,$$

we have that the radicand would vanish if

$$\sum_{k \neq s_1} b_{jk}^2 a_{kk}^{(2)} + 2 \sum_{k<l, k,l \neq s_1} b_{jk} b_{jl} a_{kl}^{(2)} = 0 . \tag{5.4}$$

Note the similarity between (5.3) and (5.4). Define recursively

$$\begin{aligned} a_{kl}^{(1)} &= a_{kl} , \\ a_{kl}^{(i)} &= a_{ks_{i-1}}^{(i-1)} a_{ls_{i-1}}^{(i-1)} - a_{s_{i-1}s_{i-1}}^{(i-1)} a_{kl}^{(i-1)} , \ i \geq 2 , \quad k,l \in \{1, \ldots, m\} \setminus \{s_1, \ldots, s_{i-1}\} . \end{aligned} \tag{5.5}$$

*Case A: Here we treat the case when for each $i$ there exists $s_i$ such that $a_{s_i s_i}^{(i)} \neq 0$. Several subcases are discussed.*

Reasoning as before, (5.4) would imply that for $1 \leq j \leq m$

$$\begin{aligned} b_{js_2} = lc(b_{j1}, \ldots, \hat{b}_{js_1}, \ldots, \hat{b}_{js_2}, \ldots, b_{jm}) \\ \pm \frac{1}{a_{s_2 s_2}^{(2)}} \sqrt{\sum_{k \neq s_1, s_2} b_{jk}^2 a_{kk}^{(3)} + 2 \sum_{k<l, k,l \neq s_1, s_2} b_{jk} b_{jl} a_{kl}^{(3)}} , \end{aligned}$$

where the symbol $\hat{b}$ means that the term $b$ has been removed. Iterating this procedure, we finally obtain the following equations for the $b_{js_{m-1}}$,

$$b_{js_{m-1}} = b_{js_m} \frac{-a_{s_{m-1}s_m}^{(m-1)} \pm \sqrt{(a_{s_{m-1}s_m}^{(m-1)})^2 - a_{s_{m-1}s_{m-1}}^{(m-1)} a_{s_m s_m}^{(m-1)}}}{a_{s_{m-1}s_{m-1}}^{(m-1)}} , \quad 1 \leq j \leq m .$$

Let $(b_{js_m})_{1 \leq j \leq m}$ be a non-zero vector in $\mathbb{R}^m$. Now, we distinguish three possibilities.

*Case A1: We show that if the radicand $(a_{s_{m-1}s_m}^{(m-1)})^2 - a_{s_{m-1}s_{m-1}}^{(m-1)} a_{s_m s_m}^{(m-1)}$ is positive, then it is possible to obtain the desired change of basis.*

If $(a_{s_{m-1}s_m}^{(m-1)})^2 - a_{s_{m-1}s_{m-1}}^{(m-1)} a_{s_m s_m}^{(m-1)} > 0$, then the quadratic polynomial in $b_{js_{m-1}}$

$$a_{s_{m-1}s_{m-1}}^{(m-1)} b_{js_{m-1}}^2 + 2 a_{s_{m-1}s_m}^{(m-1)} b_{js_{m-1}} b_{js_m} + a_{s_m s_m}^{(m-1)} b_{js_m}^2 , \tag{5.6}$$

has two real roots and we can choose $(b_{js_{m-1}})_{1 \leq j \leq m} \in \mathbb{R}^m$ linearly independent with $(b_{js_m})_{1 \leq j \leq m}$ and such that (5.6) be positive for all $1 \leq j \leq m$. As this polynomial is the radicand of the preceding one,

$$\sum_{k \neq s_1, \ldots, s_{m-3}} b_{jk}^2 a_{kk}^{(m-2)} + 2 \sum_{k<l, k,l \neq s_1, \ldots, s_{m-3}} b_{jk} b_{jl} a_{kl}^{(m-2)} , \tag{5.7}$$

our choice of $(b_{js_{m-1}})_{1\leq j\leq m}$ ensures that we can again take $(b_{js_{m-2}})_{1\leq j\leq m}\in\mathbb{R}^m$, linearly independent with $(b_{js_{m-1}})_{1\leq j\leq m}$ and $(b_{js_m})_{1\leq j\leq m}$ such that (5.7) is positive for all $1\leq j\leq m$. This is propagated step by step through the iteration process and we are able to choose a non-singular matrix $(b_{jk})$ satisfying (5.3).

*Case A2: We show that when the radicand $(a^{(m-1)}_{s_{m-1}s_m})^2-a^{(m-1)}_{s_{m-1}s_{m-1}}a^{(m-1)}_{s_m s_m}$ is negative, then either it is possible to find the change of basis or the system is not STLCC at $q_0$.*

If $(a^{(m-1)}_{s_{m-1}s_m})^2-a^{(m-1)}_{s_{m-1}s_{m-1}}a^{(m-1)}_{s_m s_m}<0$, then (5.6) does not change its sign for all $b_{js_{m-1}}$, $b_{js_m}$. If this sign is positive, the same argument as in case *A1* ensures us the choice of the desired matrix. If negative, it implies that (5.7) does not change its sign for all $b_{js_{m-2}}$, $b_{js_{m-1}}$, $b_{js_m}$. Then, the unique problem we must face is when, through the iteration process, all the radicands are negative. In the following, we discard this latter case by contradiction with the hypothesis of controllability. Apply Lemma 5.1 to the vector fields $\{Y_1,\ldots,Y_m,\langle Y_1:Y_2\rangle\}$ to find a function $\phi$ satisfying the properties 1.-5. By (3.4), we have that

$$\begin{aligned}\dot c(t)&=\sum_{i=1}^m \bar u_i Y_i-\frac{1}{2}\overline{\langle\sum_{j=1}^m \bar u_j Y_j:\sum_{k=1}^m \bar u_k Y_k\rangle}+O(\|Z\|^3_\sigma t^5)\\&=\sum_{i=1}^m \bar u_i Y_i-\frac{1}{2}\overline{\sum_{j=1}^m \bar u_j^2\langle Y_j:Y_j\rangle+2\sum_{j<k}\bar u_j\bar u_k\langle Y_j:Y_k\rangle}+O(\|Z\|^3_\sigma t^5)\,,\end{aligned}$$

where $Z=\sum_{i=1}^m u_iY_i$. Now, observe that $\dfrac{d}{dt}(\phi(c(t)))=\dot c(t)(\phi)$. Then, using properties *(iii)* and *(iv)* of $\phi$, we get

$$\frac{d}{dt}(\phi(c(t)))=\frac{1}{2}\overline{\sum_{j=1}^m a_{jj}\bar u_j^2+2\sum_{j<k}a_{jk}\bar u_j\bar u_k}+O(\|Z\|^3_\sigma t^5)\,.$$

The expression $\sum_{j=1}^m a_{jj}\bar u_j^2+2\sum_{j<k}a_{jk}\bar u_j\bar u_k$ does not change its sign, whatever the functions $u_1(t),\ldots,u_m(t)$ might be, because as a quadratic polynomial in $\bar u_{s_1}$ its radicand is always negative. Therefore, $\dfrac{d}{dt}(\phi(c(t)))$ has constant sign for sufficiently small $t$, since $\overline{\sum_{j=1}^m a_{jj}\bar u_j^2+2\sum_{j<k}a_{jk}\bar u_j\bar u_k}=O(\|u\|^2t^3)$ and dominates $O(\|Z\|^3_\sigma t^5)=O(\|u\|^3t^5)$ when $t\to 0$. Finally,

$$\phi(c(t))=\phi(q_0)+\int_0^t\frac{d}{ds}(\phi(c(s)))=\int_0^t\frac{d}{ds}(\phi(c(s)))$$

will have constant sign for $t$ small enough. As a consequence, all the points in a neighborhood of $q_0$ where $\phi$ has the opposite sign (property *(v)*) are unreachable in small time, which contradicts the hypothesis of controllability.

*Case A3: We show that if the radicand $(a^{(m-1)}_{s_{m-1}s_m})^2-a^{(m-1)}_{s_{m-1}s_{m-1}}a^{(m-1)}_{s_m s_m}$ vanishes, then an intermediate change of basis reduces the problem to considering $m-1$ input vector fields. The preceding discussion can be then reproduced.*

The situation now is similar to that of case *A2*. However, the argument employed above to discard the possibility of all the radicands being negative does not apply, since in this case there *do* exist controls such that $\sum_{j=1}^m a_{jj}\bar u_j^2+2\sum_{j<k}a_{jk}\bar u_j\bar u_k$ is zero and hence we should really investigate the sign of $O(\|Z\|^3_\sigma t^5)$ to reach a contradiction.

Instead, what we are going to do is to get a new basis $\{Y'_j\}$ such that $\langle Y'_1 : Y'_j\rangle(q_0) \in \mathcal{D}_{q_0}$, $1 \leq j \leq m$, and thus remove one vector field ($Y'_1$) from the discussion. By repeating this procedure, we finally come to consider a limit case, which we will discard by contradiction with the controllability hypothesis.

For $j = 1$, we choose $b_{1s_m} \neq 0$ and

$$\begin{aligned} b_{1s_{m-1}} &= -b_{1s_m}\frac{a^{(m-1)}_{s_{m-1}s_m}}{a^{(m-1)}_{s_{m-1}s_{m-1}}} = C_{s_{m-1}}b_{1s_m} \\ b_{1s_{m-2}} &= -\frac{a^{(m-2)}_{s_{m-2}s_{m-1}}b_{1s_{m-1}} + a^{(m-2)}_{s_{m-2}s_m}b_{1s_m}}{a^{(m-2)}_{s_{m-2}s_{m-2}}} = C_{s_{m-2}}b_{1s_m} \\ &\vdots \\ b_{1s_1} &= -\frac{\sum_{k\neq s_1} b_{1k}a_{ks_1}}{a_{s_1s_1}} = C_{s_1}b_{1s_m}\,. \end{aligned} \tag{5.8}$$

We denote $C_{s_m} = 1$. For $j > 1$, we select the $(b_{jk})_{1\leq k\leq m}$ such that the matrix $B$ be non-singular. Consequently, we change our original basis $\{Y_1, \ldots, Y_m\}$ to a new one $\{Y'_1, \ldots, Y'_m\}$. In this basis, following (5.2), one has

$$\begin{aligned} &\langle Y'_1 : Y'_1\rangle(q_0) = lc(Y'_1(q_0), \ldots, Y'_m(q_0)) \\ &\langle Y'_j : Y'_j\rangle(q_0) = lc(Y'_1(q_0), \ldots, Y'_m(q_0)) + a'_{jj}\,\langle Y_1 : Y_2\rangle\,(q_0)\,, \quad 2 \leq j \leq m\,. \end{aligned}$$

In addition, one can check that for each $2 \leq j \leq m$,

$$\begin{aligned} \langle Y'_1 : Y'_j\rangle(q_0) &= lc(Y'_1(q_0), \ldots, Y'_m(q_0)) + \left(\sum_{k,l} a_{kl}b_{1k}b_{jl}\right)\langle Y_1 : Y_2\rangle(q_0) \\ &= lc(Y'_1(q_0), \ldots, Y'_m(q_0)) + b_{1s_m}\left(\sum_l b_{jl}\left(\sum_k a_{kl}C_k\right)\right)\langle Y_1 : Y_2\rangle(q_0)\,. \end{aligned}$$

Now, when the $C_k$ are given by (5.8), we have

$$\sum_k a_{kl}C_k = 0\,, \quad 1 \leq l \leq m\,,$$

(see Lemma A.1 in the Appendix) and this guarantees that

$$\langle Y'_1 : Y'_j\rangle(q_0) = lc(Y'_1(q_0), \ldots, Y'_m(q_0))\,, \quad 2 \leq j \leq m\,.$$

If the $a'_{jj} = 0$, $2 \leq j \leq m$, we are done. Assume then that $a'_{33} \neq 0$, reordering the input vector fields if necessary. Assume further that $\langle Y'_2 : Y'_3\rangle(q_0)$ is not a linear combination of $\{Y'_1, \ldots, Y'_m\}$ (otherwise, redefine a new $Y''_2$ as $Y'_2 + Y'_3$). Then,

$$\begin{aligned} \langle Y'_2 : Y'_2\rangle\,(q_0) &= lc(Y'_1(q_0), \ldots, Y'_m(q_0)) + a'_{22}\,\langle Y'_2 : Y'_3\rangle\,(q_0) \\ &\vdots \\ \langle Y'_m : Y'_m\rangle\,(q_0) &= lc(Y'_1(q_0), \ldots, Y'_m(q_0)) + a'_{mm}\,\langle Y'_2 : Y'_3\rangle\,(q_0) \\ \langle Y'_2 : Y'_3\rangle\,(q_0) &= \qquad\qquad\qquad\qquad\qquad\; a'_{23}\,\langle Y'_2 : Y'_3\rangle\,(q_0) \\ \langle Y'_2 : Y'_4\rangle\,(q_0) &= lc(Y'_1(q_0), \ldots, Y'_m(q_0)) + a'_{24}\,\langle Y'_2 : Y'_3\rangle\,(q_0) \\ &\vdots \\ \left\langle Y'_{m-1} : Y'_m\right\rangle(q_0) &= lc(Y'_1(q_0), \ldots, Y'_m(q_0)) + a'_{m-1m}\,\langle Y'_2 : Y'_3\rangle\,(q_0)\,, \end{aligned}$$

where we have denoted with a slight abuse of notation by $a'_{jk}$ the new coefficients corresponding to $\langle Y'_2 : Y'_3 \rangle$. Consequently, we can now reproduce the preceding discussion, but with the $m-1$ vector fields $\{Y'_2, \dots, Y'_m\}$. That is, we look for one change of basis $B'$ in the vector fields $\{Y'_2, \dots, Y'_m\}$ such that the new ones $\{Y''_2, \dots, Y''_m\}$ together with $Y'_1$ verify the sufficient conditions for STLCC at $q_0$. Accordingly, we must consider the vanishing of the new polynomials

$$\sum_{k=2}^{m} b_{jk}^{2}{}' a'_{kk} + 2 \sum_{2 \le k < l \le m} b'_{jk} b'_{jl} a'_{kl} = 0 \,, \quad 2 \le j \le m \,.$$

The cases in which the last radicand $(a^{(m-1)}_{s_{m-1}s_m}{}')^2 - a^{(m-1)}_{s_{m-1}s_{m-1}}{}' a^{(m-1)}_{s_m s_m}{}'$ does not vanish are treated as before (cases *A1* and *A2*). When it vanishes, we obtain a new basis $\{Y''_1 = Y'_1, Y''_2, \dots, Y''_m\}$ such that

$$\begin{aligned} &\langle Y''_1 : Y''_1 \rangle (q_0) \,, \ \langle Y''_2 : Y''_2 \rangle (q_0) \in \mathcal{D}_{q_0} \\ &\langle Y''_j : Y''_j \rangle (q_0) = lc(Y''_1(q_0), \dots, Y''_m(q_0)) + c'_{jj} \, \langle Y'_2 : Y'_3 \rangle \, (q_0) \,, \ \ 3 \le j \le m \\ &\quad \langle Y''_1 : Y''_j \rangle \,, \ \langle Y''_2 : Y''_{j+1} \rangle \in \mathcal{D}_{q_0} \,, \ \ 2 \le j \le m \,, \end{aligned}$$

where there could exit some $3 \le j \le m$ such that $c'_{jj} \neq 0$. By an induction procedure, we finally come to consider discarding the case of a certain basis $\{Z_1 = Y'_1, Z_2 = Y''_2, \dots, Z_m\}$ of $\mathcal{D}$ satisfying $\langle Z_i : Z_j \rangle (q_0) \in \operatorname{span}\{Z_1(q_0), \dots, Z_m(q_0)\}$, $1 \le i < j \le m$, and the sufficient conditions for STLCC at $q_0$ for $Z_1, \dots, Z_{m-1}$, but such that $\langle Z_m : Z_m \rangle (q_0) \notin \operatorname{span}\{Z_1(q_0), \dots, Z_m(q_0)\}$. Similarly as we have done above, the application of Lemma 5.1 with the vector fields $\{Z_1, \dots, Z_m, \langle Z_m : Z_m \rangle\}$ implies that the system is not controllable at $q_0$, yielding a contradiction.

*Case B: Finally, we prove that if there exists an* $i \ge 2$ *such that* $a^{(i)}_{kk} = 0$*, for all* $k \in \{1, \dots, m\} \setminus \{s_1, \dots, s_{i-1}\}$*, then either the desired change of basis is straightforward or an intermediate step can be done that reduces the problem to considering* $i-1$ *input vector fields.*

In this case, the polynomial

$$\sum_{k \neq s_1, \dots, s_{i-1}} b_{jk}^2 a^{(i)}_{kk} + 2 \sum_{k<l, k,l \neq s_1, \dots, s_{i-1}} b_{jk} b_{jl} a^{(i)}_{kl}$$

takes the form

$$2 \sum_{k<l, k,l \neq s_1, \dots, s_{i-1}} b_{jk} b_{jl} a^{(i)}_{kl} \,. \tag{5.9}$$

If any of the $a^{(i)}_{kl}$ is different from zero, then it is clear that we can choose the $b_{jk}$, $k \notin \{s_1, \dots, s_{i-1}\}$, such that (5.9) be positive. Then, reasoning as before, we find a regular matrix $B$ yielding the desired change of basis. If this is not the case, i.e. $a^{(i)}_{kl} = 0$, for all $k < l, k, l \notin \{s_1, \dots, s_{i-1}\}$, we can do the following. Choose $\{(b_{jk})_{1 \le j \le m}\}$, with $k \notin \{s_1, \dots, s_{i-1}\}$, $m - i + 1$ linearly independent vectors in $\mathbb{R}^m$ such that the minor $\{b_{jk}\}^{k \neq s_1, \dots, s_{i-1}}_{1 \le j \le m-i+1}$ is regular. Now, let $j$ in equation (5.8) vary between 1 and $m - i + 1$; that is, take

$$b_{js_{i-1}} = - \frac{\sum^{m}_{k \neq s_1, \dots, s_{i-1}} b_{jk} a^{(i-1)}_{s_{i-1}k}}{a^{(i-1)}_{s_{i-1}s_{i-1}}}$$

$$b_{js_{i-2}} = - \frac{\sum^{m}_{k \neq s_1, \dots, s_{i-2}} b_{jk} a^{(i-2)}_{s_{i-2}k}}{a^{(i-2)}_{s_{i-2}s_{i-2}}} \,, \quad \dots \,, \quad b_{js_1} = - \frac{\sum_{k \neq s_1} b_{jk} a_{ks_1}}{a_{s_1 s_1}} \,, \tag{5.10}$$

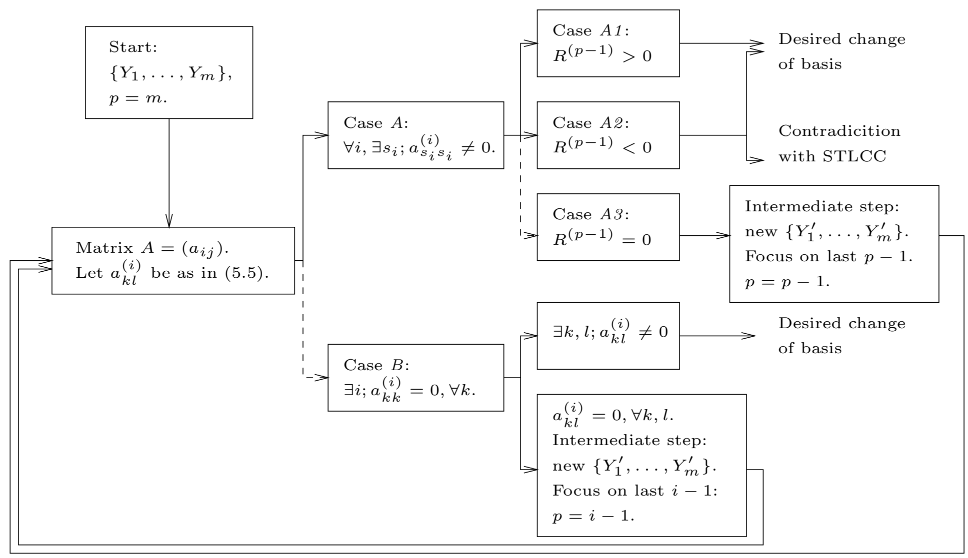


FIG. 5.1. *Illustration of the proof of Theorem 5.2.* $R^{(p-1)}$ *denotes* $(a_{s_{p-1}s_p}^{(p-1)})^2 - a_{s_{p-1}s_{p-1}}^{(p-1)} a_{s_p s_p}^{(p-1)}$. *The dashed lines mean that one cannot fall repeatedly in cases A3 or B without contradicting STLCC.*

for $1 \leq j \leq m-i+1$. Finally, for $j > m-i+1$, we select the $b_{jk}$ such that the matrix $B$ is non-singular. In this manner, in a unique step, we would change to a new basis $\{Y_1', \dots, Y_m'\}$ verifying

$$\langle Y_1' : Y_1' \rangle(q_0)\,, \ \dots, \langle Y_{m-i+1}' : Y_{m-i+1}' \rangle(q_0) \in \mathcal{D}_{q_0}$$
$$\langle Y_j' : Y_j' \rangle(q_0) = lc(Y_1'(q_0), \dots, Y_m'(q_0)) + a_{jj}' \langle Y_1 : Y_2 \rangle (q_0)\,, \quad m-i+2 \leq j \leq m$$
$$\langle Y_k' : Y_l' \rangle(q_0) \in \mathcal{D}_{q_0}\,, \quad k < l, 1 \leq k \leq m-i+1\,,$$

with possibly some of the $(a_{jj}')_{m-i+1 \leq j \leq m}$ being different from zero. Now, the above discussion can be redone in this context to assert the validity of the theorem. That is, we have to look for a change of basis $B'$ in the vector fields $\{Y_{m-i+2}', \dots, Y_m'\}$ such that the new ones, $\{Y_{m-i+2}'', \dots, Y_m''\}$ together with $\{Y_1', \dots, Y_{m-i+1}'\}$ verify the sufficient conditions for STLCC at $q_0$. To find the change of basis for $\{Y_{m-i+2}', \dots, Y_m'\}$, we have to consider the corresponding versions of cases $A$ and $B$. If we repeatedly fall in case $B$, then we come to discard the same possibility that we encountered in the treatment of case *A3*, which can be done again by means of Lemma 5.1. □

To recap, the steps of the proof can be summarized as follows (see Figure 5.1): first, we have considered the case when there exists for all $i$ a $s_i$ such that $a_{s_i s_i}^{(i)} \neq 0$. We have seen that this case can be subdivided into three: one (case *A1*) ensuring the desired change of basis, another one (case *A2*) in which either one obtains the basis or one contradicts the hypothesis of small-time local configuration controllability and a third one (case *A3*) where an intermediate change of basis is performed that allows us to focus on the search of a change of basis for $m-1$ of the new vector fields. Then, under the same assumption on the new coefficients, $a_{jk}'$ (i.e. for all $i$, there exists a $s_i$ such that ${a_{s_i s_i}^{(i)}}' \neq 0$), we can reproduce the former discussion. We cannot repeatedly fall into case *A3*, since we would contradict the controllability assumption. Finally, we have treated the case when this type of "circular" process is broken (case $B$): that is, when there exists an $i$ such that $a_{kk}^{(i)} = 0$, for all $k \neq s_1, \dots, s_{i-1}$. What we have shown then is that this leads to either a new basis of input vector fields satisfying the sufficient conditions for STLCC or a reduced situation where we can "get rid" at the same time of the problems associated with $m-i+1$ vector fields.

REMARK 5.3. Notice that the proof of this result can be reproduced for the corresponding notions of accessibility and controllability at zero velocity. Indeed, a mechanical control system of the form (2.2) with $m=n-1$, which is STLC at $q_0$ and zero velocity is in particular STLCC at $q_0$. Then, Theorem 5.2 implies that there exists a basis of input vector fields $\mathcal{Y}$ satisfying the sufficient conditions of Theorem 3.2, so the same result is also valid for local controllability at zero velocity.

COROLLARY 5.4. *Let $Q$ be a 3-dimensional analytic manifold and let $Y_1, Y_2$ be analytic vector fields on $Q$. Consider the control system* (5.1) *and assume that it is locally configuration accessible at $q_0 \in Q$. Let $A$ be the $2\times 2$ symmetric matrix whose elements are given by*

$$\begin{aligned}
\langle Y_1 : Y_1\rangle (q_0) &= lc(Y_1(q_0), Y_2(q_0)) + a_{11}\langle Y_1 : Y_2\rangle (q_0)\\
\langle Y_2 : Y_2\rangle (q_0) &= lc(Y_1(q_0), Y_2(q_0)) + a_{22}\langle Y_1 : Y_2\rangle (q_0)\\
\langle Y_1 : Y_2\rangle (q_0) &= a_{12}\langle Y_1 : Y_2\rangle (q_0)\,.
\end{aligned}$$

*Then the system is locally configuration controllable at $q_0$ if and only if* $\det A < 0$.

*Proof.* The results follows from the proof of Theorem 5.2 by noting that $\det A < 0$ corresponds to case *A1*, $\det A > 0$ to case *A2* and $\det A = 0$ to case *A3*. □

REMARK 5.5. Note that Corollary 5.4 together with Theorem 4.1 completely characterize the configuration controllability properties of mechanical control systems with 3 degrees of freedom, since fully actuated systems are obviously STLCC.

## 6. Examples.

**6.1. The planar rigid body.** Consider a planar rigid body [24]. Fix a point $P \in \mathbb{R}^2$ and let $\{e_1, e_2\}$ be the standard orthonormal frame at that point. Let $\{d_1, d_2\}$ be an orthonormal frame attached to the body at its center of mass. The configuration manifold is then $SE(2)$, with coordinates $(x, y, \theta)$, where $(x, y)$ describe the position of the center of mass and $\theta$ the orientation of the frame $\{d_1, d_2\}$ with respect to $\{e_1, e_2\}$.

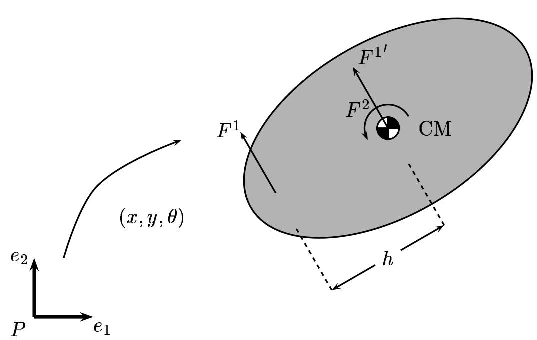


FIG. 6.1. *The planar rigid body.*

The inputs of the system consist of a force $F^1$ applied at a distance $h$ from the center of mass $CM$ and a torque, $F^2$, about $CM$ (see Figure 6.1). In coordinates, the input forces are given by

$$F^1 = -\sin\theta dx + \cos\theta dy - hd\theta\ , \quad F^2 = d\theta\,.$$

The Riemannian metric is

$$g = m dx \otimes dx + m dy \otimes dy + J d\theta \otimes d\theta ,$$

where $m$ is the mass of the body and $J$ its moment of inertia.

The input vector fields can be computed via $\flat_g^{-1}$ as

$$Y_1 = -\frac{\sin\theta}{m}\frac{\partial}{\partial x} + \frac{\cos\theta}{m}\frac{\partial}{\partial y} - \frac{h}{J}\frac{\partial}{\partial\theta} d\theta , \quad Y_2 = \frac{1}{J}\frac{\partial}{\partial\theta} .$$

One can easily show that the planar body is locally configuration accessible [24]. However, the inputs $Y_1$, $Y_2$ fail to satisfy the sufficient conditions for STLCC. In fact,

$$\begin{aligned}
\langle Y_1 : Y_1\rangle &= \frac{2h\cos\theta}{mJ}\frac{\partial}{\partial x} + \frac{2h\sin\theta}{mJ}\frac{\partial}{\partial y} , \\
\langle Y_1 : Y_2\rangle &= -\frac{\cos\theta}{mJ}\frac{\partial}{\partial x} - \frac{\sin\theta}{mJ}\frac{\partial}{\partial y} , \\
\langle Y_2 : Y_2\rangle &= 0 .
\end{aligned}$$

Therefore, $\{Y_1, Y_2, \langle Y_1 : Y_2\rangle\}$ are linearly independent and $\langle Y_1 : Y_1\rangle = -2h \langle Y_1 : Y_2\rangle$. Theorem 5.2 ensures us STLCC if and only if there exist a basis of input vector fields satisfying the sufficient conditions. We have that

$$\det A = \det \begin{pmatrix} -2h & 1 \\ 1 & 0 \end{pmatrix} = -1 < 0 ,$$

and consequently, by Corollary 5.4, the system is locally configuration controllable. Indeed, this example falls into case *A1* of the proof of Theorem 5.2. Accordingly, we obtain the change of basis: $Y_1' = Y_1 + hY_2$, $Y_2' = Y_2$. This yields

$$\langle Y_1' : Y_1'\rangle = \langle Y_2' : Y_2'\rangle = 0 , \quad \langle Y_1' : Y_2'\rangle = \langle Y_1 : Y_2\rangle ,$$

which satisfies the sufficient conditions for STLCC. The new input vector field precisely corresponds to the force $F^{1'}$ in Figure 6.1.

**6.2. A simple example.** The following example does not necessarily correspond to a physical example, but illustrates the proof of Theorem 5.2. Consider a mechanical control system on $\mathbb{R}^3$, with coordinates $(x, y, z)$. The Riemannian metric is given by

$$g = dx \otimes dx + dy \otimes dy + dz \otimes dz ,$$

and the input vector fields

$$Y_1 = z\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{1}{4}\frac{\partial}{\partial z} , \quad Y_2 = y\frac{\partial}{\partial x} + \frac{1}{4}\frac{\partial}{\partial y} - \frac{1}{2}\frac{\partial}{\partial z} .$$

In coordinates, we have the following control equations

$$\ddot{x} = u_1 z + u_2 y , \quad \ddot{y} = u_1 + \frac{u_2}{4} , \quad \ddot{z} = \frac{u_1}{4} - \frac{u_2}{2} . \qquad (6.1)$$

Since

$$\langle Y_1 : Y_1\rangle = \langle Y_1 : Y_2\rangle = \langle Y_2 : Y_2\rangle = \frac{1}{2}\frac{\partial}{\partial x} ,$$

we deduce that $\mathrm{span}\{Y_1(q), Y_2(q), \langle Y_1 : Y_2\rangle(q)\} = T_qQ$ for all $q \in Q$ and the system (6.1) is locally configuration accessible. However, Corollary 5.4 implies that it is not STLCC, since $\det A = 0$. Going through the proof of Theorem 5.2, we see that this example falls into case *A3*. Choosing the change of basis

$$B = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix},$$

we get the new input vector fields $Y_1' = -Y_1 + Y_2$ and $Y_2' = Y_1 + Y_2$. Now, we have

$$\langle Y_1' : Y_1'\rangle = 0\,, \quad \langle Y_1' : Y_2'\rangle = 0\,, \quad \langle Y_2' : Y_2'\rangle = 2\frac{\partial}{\partial x}\,.$$

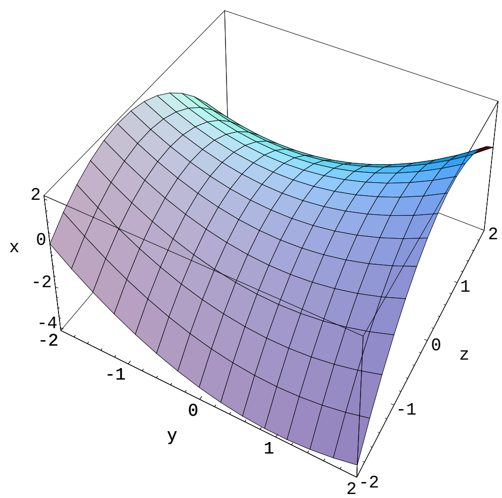


FIG. 6.2. *The level surface* $\phi(x, y, z) = 0$.

We can compute explicitely the function $\phi$ of Lemma 5.1 for this example. The flows of $Z_1 = Y_1'$, $Z_2 = Y_2'$, $Z_3 = -\langle Y_2' : Y_2'\rangle$ are given by

$$\begin{aligned}
\Psi_1(t)(x, y, z) &= (x + (y - z)t, y - 3t/4, z - 3t/4) \\
\Psi_2(t)(x, y, z) &= (x + (y + z)t + t^2/2, y + 5t/4, z - t/4) \\
\Psi_3(t)(x, y, z) &= (x - 2t, y, z)
\end{aligned}$$

Letting $(x_0, y_0, z_0)$ be an arbitrary point, one verifies

$$\Psi_1(t_1) \circ \Psi_2(t_2) \circ \Psi_3(t_3)(x_0, y_0, z_0) =$$
$$\left(x_0 - 2t_3 + (y_0 + z_0 + \frac{1}{2}t_2)t_2 + t_1(y_0 - z_0 - \frac{3}{2}t_2), y_0 - \frac{3}{4}t_1 + \frac{5}{4}t_2, z_0 - \frac{3}{4}t_1 - \frac{1}{4}t_2\right).$$

We may solve for $\phi(x, y, z) = t_3$ as

$$\phi(x, y, z) =$$
$$\frac{1}{18}\left(-9(x - x_0) + 4(y^2 - yy_0 + yz - 5y_0z - 2z^2 + yz_0 + 3y_0z_0 + 5zz_0 - 3z_0^2)\right)\,.$$

In Figure 6.2, we show the level set $\phi(x, y, z) = 0$ for $(x_0, y_0, z_0) = (0, 0, 0)$. The locally accessible configurations from $(0, 0, 0)$ are contained below the surface, where $\phi(x, y, z) \geq 0$.

**7. Conclusions.** In this paper, we have built on previous results on controllability and series expansions for the evolution of mechanical control systems within the affine connection formalism to demonstrate that the sufficient conditions encountered in [24] for STLCC are also necessary when the configuration manifold is $n$-dimensional and the system is actuated by $n-1$ inputs, in the sense that there exists some basis of input vector fields that verifies them.

However $n-1$ controls is a special case and is the simplest case next to fully actuated systems, which are always STLCC. For an arbitrary number of inputs, higher-order controllability will necessarily play a key role. Future research will be devoted to investigate the validity of the controllability conjecture in the full general case.

**Acknowledgments.** We wish to thank F. Cantrijn for several helpful suggestions and the Department of Mathematical Physics and Astronomy of the University of Ghent for its kind hospitality. We would also like to thank the anonymous reviewers for their useful comments which helped us improve the presentation of the manuscript.

REFERENCES

[1] R. Abraham, J.E. Marsden: *Foundations of Mechanics*. 2nd ed., Benjamin-Cummings, Reading (Ma), 1978.

[2] A.A. Agrachev, R.V Gamkrelidze: Local controllability and semigroups of diffeomorphisms. *Acta Appl. Math.* **32** (1993), 1-57.

[3] F. Bullo: Series expansions for the evolution of mechanical control systems. *SIAM J. Control Optim.* **40** (1) (2001), 166-190.

[4] F. Bullo, N.E. Leonard, A.D. Lewis: Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups. *IEEE Trans. Automat. Control* **45** (8) (2000), 1437-1454.

[5] F. Bullo, A.D. Lewis: Configuration controllability of mechanical systems on Lie groups. *Proceedings of MTNS'96*, St. Louis, Missouri, 1996.

[6] F. Bullo, M. Zefran: On mechanical control systems with nonholonomic constraints and symmetries. *Systems Control Lett.* **45** (2) (2002), 133-143.

[7] K.T. Chen: Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Ann. of Math.* **67** (1957), 164-178.

[8] J. Cortés, S. Martínez, F. Bullo: On nonlinear controllability and series expansions for Lagrangian systems with dissipative forces. To appear in *IEEE Trans. Automat. Control.*

[9] J. Cortés, S. Martínez, J.P. Ostrowski, H. Zhang: Simple mechanical control systems with constraints and symmetry. To appear in *SIAM J. Control Optim.*

[10] M. Fliess: Fonctionelles causales non linéares et indeterminées non commutatives. *Bull. Soc. Math.* **109** (1981), 3-40.

[11] J.E. Hauser, S.S. Sastry, G. Meyer: Nonlinear control design for slightly nonminimum phase systems: application to V/STOL aircraft. *Automatica* **28** (4) (1992), 665-679.

[12] H. Hermes: Control systems which generate decomposable Lie algebras. *J. Differential Equations* **44** (1982), 166-187.

[13] A. Isidori: *Nonlinear Control Systems*. 3nd ed., Communications and Control Engineering Series, Springer-Verlag, Berlin, 1995.

[14] M. Kawski: The complexity of deciding controllability. *Systems Control Lett.* **15** (1) (1990), 9-14.

[15] M. Kawski: High-order small-time local controllability. In *Nonlinear Controllability and Optimal Control*, ed. H.J. Sussmann, Dekker, 1990, pp. 441-477.

[16] M. Kawski, H.J. Sussmann: Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory. In *Operators, Systems and Linear Algebra*, eds. U. Helmke, D. Pratzel-Wolters, E. Zerz, Teubner, 1997, pp. 111-128.

[17] S.D. Kelly, R.M. Murray: Geometric phases and robotic locomotion. *J. Robotic Systems* **12** (6) (1995), 417-431.

[18] S. Kobayashi, K. Nomizu: *Foundations of Differential Geometry*. Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers, Wiley, New-York, 1963.

[19] G. Lafferriere, H.J. Sussmann: A differential geometric approach to motion planning. In *Nonholonomic Motion Planning*, eds. Z.X. Li, J.F. Canny, Kluwer, 1997, pp. 235-270.

[20] N.E. Leonard, P.S. Krishnaprasad: Motion control of drift-free, left-invariant systems on Lie groups. *IEEE Trans. Automat. Control* **40** (9) (1995), 1539-1554.
[21] A.D. Lewis: Local configuration controllability for a class of mechanical systems with a single input, *Proceedings of the 1997 European Control Conference*, Brussels, Belgium, 1997.
[22] A.D. Lewis: Affine connections and distributions with applications to nonholonomic mechanics. *Rep. Math. Phys.* **42** (1/2) (1998), 135-164.
[23] A.D. Lewis: Simple mechanical control systems with constraints. *IEEE Trans. Automat. Control* **45** (8) (2000), 1420-1436.
[24] A.D. Lewis, R.M. Murray: Configuration controllability of simple mechanical control systems. *SIAM J. Control Optim.* **35** (3) (1997), 766-790.
[25] A.D. Lewis, R.M. Murray: Configuration controllability of simple mechanical control systems. *SIAM Rev.* **41** (3) (1999), 555-574.
[26] K.M. Lynch, N. Shiroma, H. Arai, K. Tanie: Collision-free trajectory planning for a 3-DOF robot with a passive joint. *Int. J. Robotics Research* **19** (12) (2000), 1171-1184.
[27] W. Magnus: On the exponential solution of differential equations for a linear operator. *Comm. Pure and Appl. Math.* **VII** (1954), 649-673.
[28] P. Martin, S. Devasia, B. Paden: A different look at output tracking: Control of a VTOL aircraft. *Automatica* **32** (1) (1996), 101-107.
[29] S. Martínez, J. Cortés: Motion control algorithms for mechanical systems with symmetries. To appear in *Acta Appl. Math.*
[30] H. Nijmeijer, A.J. van der Schaft: *Nonlinear Dynamical Control Systems*. Springer-Verlag, New York, 1990.
[31] J.P. Ostrowski, J.W. Burdick: Controllability tests for mechanical systems with symmetries and constraints. *J. Appl. Math. Comp. Sci.* **7** (2) (1997), 101-127.
[32] M. Rathinam, R.M. Murray: Configuration flatness for Lagrangian systems underactuated by one control. *SIAM J. Control Optim.* **36** (1) (1998), 164-179.
[33] E.D. Sontag: Controllability is harder to decide than accessibility. *SIAM J. Control Optim.* **26** (5) (1988), 1106-1118.
[34] H.J. Sussmann: Lie brackets and local controllability: a sufficient condition for scalar-input systems. *SIAM J. Control Optim.* **21** (5) (1983), 686-713.
[35] H.J. Sussmann: A general theorem on local controllability. *SIAM J. Control Optim.* **25** (1) (1987), 158-194.
[36] H.J. Sussmann: A product expansion of the Chen series. *Theory and Applications of Nonlinear Control Systems*, eds. C.I. Byrnes, A. Lindquist, Elsevier, 1986, pp. 323-335.
[37] H. Zhang, J.P. Ostrowski: Control algorithms using affine connections on principal fiber bundles. IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control, Princeton, USA, 2000.

## Appendix A. A simple lemma.

LEMMA A.1. *With the notation of Theorem 5.2, assume that* $(a^{(m-1)}_{s_{m-1}s_m})^2 - a^{(m-1)}_{s_{m-1}s_{m-1}} a^{(m-1)}_{s_m s_m} = 0$. *Then the coefficients* $C_k$ *given by (5.8) verify*

$$\sum_{k=1}^{m} a_{kl} C_k = 0 \,, \quad 1 \leq l \leq m \,.$$

*Proof.* From (5.8), one can obtain the following recurrence formula for the coefficients $C_k$,

$$C_{s_m} = 1 \,, \quad C_{s_j} = -\frac{1}{a^{(j)}_{s_j s_j}} \left( \sum_{i=j+1}^{m} a^{(j)}_{s_i s_j} C_{s_i} \right) \,, \quad 1 \leq j \leq m-1 \,. \tag{A.1}$$

Let us denote

$$\Sigma(l) = \sum_{k=1}^{m} a_{kl} C_k \,.$$

It is easy to see that $\Sigma(s_1) = 0$. Indeed, using (A.1), we have that

$$\Sigma(s_1) = a_{s_1 s_1} C_{s_1} + \sum_{i=2}^{m} a_{s_i s_1} C_{s_i} = -\sum_{i=2}^{m} a_{s_i s_1} C_{s_i} + \sum_{i=2}^{m} a_{s_i s_1} C_{s_i} = 0\,.$$

To prove the result for the remaining indices we can do the following. First, note that

$$a_{s_1 s_j} C_{s_1} = -\frac{a_{s_1 s_j}}{a_{s_1 s_1}} \left( \sum_{i=2}^{m} a_{s_i s_j} C_{s_i} \right) = -\sum_{i=2}^{m} \left( \frac{a_{s_1 s_j} a_{s_i s_j}}{a_{s_1 s_1}} \right) C_{s_i}$$

Then, substituting in $\Sigma(s_j)$, we get

$$\begin{aligned} \Sigma(s_j) &= -\sum_{i=2}^{m} \left( \frac{a_{s_1 s_j} a_{s_i s_j}}{a_{s_1 s_1}} \right) C_{s_i} + \sum_{i=2}^{m} a_{s_i s_j} C_{s_i} \\ &= \sum_{i=2}^{m} \left( \frac{a_{s_i s_j} a_{s_1 s_1} - a_{s_1 s_j} a_{s_i s_j}}{a_{s_1 s_1}} \right) C_{s_i} = -\frac{1}{a_{s_1 s_1}} \left( \sum_{i=2}^{m} a^{(2)}_{s_i s_j} C_{s_i} \right), \end{aligned}$$

where we have used the definition (5.5) for the coefficients $a^{(j)}_{kl}$. This procedure can be iterated to obtain the general expression

$$\Sigma(s_j) = \frac{(-1)^k}{a_{s_1 s_1} a^{(2)}_{s_2 s_2} \ldots a^{(k)}_{s_k s_k}} \left( \sum_{i=k+1}^{m} a^{(k+1)}_{s_i s_j} C_{s_i} \right), \tag{A.2}$$

which is valid for any $1 \leq k \leq m-2$.

Now, consider the cases $2 \leq j \leq m-1$. Take $k = j-1$. Then, using (A.2),

$$\begin{aligned} \Sigma(s_j) &= \frac{(-1)^{j-1}}{a_{s_1 s_1} a^{(2)}_{s_2 s_2} \ldots a^{(j-1)}_{s_{j-1} s_{j-1}}} \left( \sum_{i=j}^{m} a^{(j)}_{s_i s_j} C_{s_i} \right) \\ &= \frac{(-1)^{j-1}}{a_{s_1 s_1} a^{(2)}_{s_2 s_2} \ldots a^{(j-1)}_{s_{j-1} s_{j-1}}} \left( a^{(j)}_{s_j s_j} C_{s_j} + \sum_{i=j+1}^{m} a^{(j)}_{s_i s_j} C_{s_i} \right) = 0\,, \end{aligned}$$

where in the last equality we have used (A.1). Finally, if $j = m$, we have that

$$\begin{aligned} \Sigma(s_m) &= \frac{(-1)^{m-2}}{a_{s_1 s_1} a^{(2)}_{s_2 s_2} \ldots a^{(m-2)}_{s_{m-2} s_{m-2}}} \left( a^{(m-1)}_{s_{m-1} s_m} C_{s_{m-1}} + a^{(m-1)}_{s_m s_m} C_{s_m} \right) \\ &= \frac{(-1)^{m-2}}{a_{s_1 s_1} a^{(2)}_{s_2 s_2} \ldots a^{(m-2)}_{s_{m-2} s_{m-2}}} \left( -\frac{(a^{(m-1)}_{s_{m-1} s_m})^2}{a^{(m-1)}_{s_{m-1} s_{m-1}}} + a^{(m-1)}_{s_m s_m} \right). \end{aligned}$$

From the hypothesis $(a^{(m-1)}_{s_{m-1} s_m})^2 - a^{(m-1)}_{s_{m-1} s_{m-1}} a^{(m-1)}_{s_m s_m} = 0$, we conclude that $\Sigma(s_m) = 0$, and this completes the proof. □