

Configuration controllability of simple mechanical control systems

Andrew D. Lewis* Richard M. Murray†

21/11/1995

Last updated: 17/12/2001

Abstract

In this paper we present a definition of “configuration controllability” for mechanical systems whose Lagrangian is kinetic energy with respect to a Riemannian metric minus potential energy. A computable test for this new version of controllability is derived. This condition involves an object which we call the *symmetric product*. Of particular interest is a definition of “equilibrium controllability” for which we are able to derive computable sufficient conditions. Examples illustrate the theory.

Keywords. mechanics, Riemannian geometry, controllability, symmetric product

AMS Subject Classifications. 53B20, 70H03, 70Q05, 93B03, 93B05, 93B29

1. Introduction

Mechanical systems form a large subset of control systems which have many diverse applications. These systems are characterised by a rich structure which has been under-exploited in the current controls literature. In this paper we utilise the structure of a specific class of mechanical systems to obtain controllability results which are meaningful for these systems. These results are important in two respects. First, they identify the structure of mechanical systems which lends to controllability of these systems. Second, the results provide computable checks for useful notions of controllability. One important aspect of our work is that the computations for checking controllability are performed on the configuration space, and not on the phase space. This is important since the phase space has twice the dimension of the configuration space for mechanical systems.

Much of the previous work in the area of mechanical control systems has relied on specific structure of these systems. In an early paper, Brockett [1977] outlined the use of special structures in analytical mechanics as they relate to control theory. Bloch and Crouch [1992] study mechanical systems on Riemannian manifolds. Under suitable hypotheses on the inputs, and assuming some group symmetries for the systems under investigation, the authors

*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 in part supported by a grant from the Powell Foundation

Work performed while a Graduate Research Assistant at the California Institute of Technology.

†Associate Professor, CONTROL AND DYNAMICAL SYSTEMS, MAIL CODE 104–44 CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91125, U.S.A.

Email: murray@indra.caltech.edu, URL: <http://avalon.caltech.edu/~murray/>

are able to use a result of San Martin and Crouch [1984] to arrive at a controllability result. Mechanical systems with nonholonomic constraints are studied by Bloch, Reyhanoglu, and McClamroch [1992]. In this paper the authors are able to show that the systems considered are controllable if the inputs span a complement to the constraint forces. In both of the above papers, the results are limited by the hypotheses placed on the system: symmetries in the first case, and constraints in the second. In this paper we attempt to develop control theoretic tools for *mechanical* control systems. We emphasise mechanical because it is our intent to use the mechanical structure to advantage in the control problem rather than any additional structure imposed on the system.

In Section 2 we motivate the development of the paper by posing various controllability questions for a simple example. In this section we also preview the results of the paper by stating a simplified form of the most general results. In Section 3 we present enough background from the theory of free Lie algebras and Riemannian geometry that we can use these ideas in Section 5. In Section 3.2 we introduce the notion of a symmetric algebra which is new and will be particularly interesting to us. The *symmetric product* is defined in Section 3.3. This is an interesting object whose geometric meaning is not fully utilised in this paper. However, it proves to be a useful computational tool for expressing our controllability results. In Section 4 we state a result of Sussmann [1987] which we shall use to prove some controllability results in Section 5. The main results of the paper are stated in Section 5. Illustrative examples are given in Section 6.

2. Preliminary statement of results

It is possible to state a subset of the results of the paper without going through all of the formality needed to state the most general results. In this section we give some idea of the questions that we answer in the paper as well as state the results in the case when no potential energy is present.

Consider the planar rigid body system of Figure 1. On this body we consider two possible sets of forces. In one case we are able to apply a force in any direction to the body at a point away from the centre of mass (case (a) in the figure). In the other case, we can only apply a force which is in a direction perpendicular to the line joining the point of application of the force with the centre of mass (case (b) in the figure). The reader may wish to consider the former case as corresponding to having a thruster on the body whose direction may be varied, while in the second case the thruster can only provide thrust in one direction. In each of these cases one may ask certain questions about the controllability of this system. We list some of these questions below and in parentheses give the name of the general notion corresponding to this question.

1. Starting from rest at a given configuration, is it possible to reach an open set of configurations? (local configuration accessibility)
2. Starting from rest in a given configuration, is it possible to reach a neighbourhood of the initial configuration? (local configuration controllability)
3. Is it possible to get to these configurations with zero velocity? (equilibrium controllability)

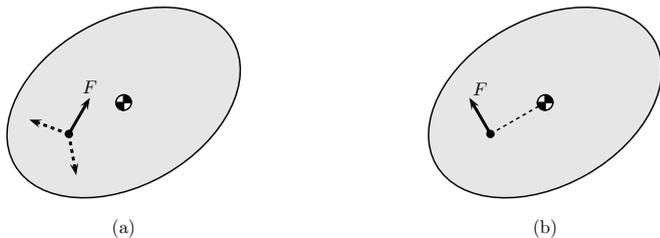


FIGURE 1. A planar rigid body with a variable direction thruster (a) and a fixed directional thruster (b)

It is exactly these questions which we address in this paper. Observe that the above controllability questions have the feature that the initial velocity is assumed to be zero. This turns out to greatly simplify the controllability computations. We observe that for this example the linearisation is not controllable so, if the system is controllable, nonlinear tools must be employed.

Although we delay answering the above questions for the planar rigid body until Section 6.2, we may state general results for a class of systems smaller than the general class we consider in the sequel. Let us consider, for the moment, mechanical systems whose Lagrangian is kinetic energy with respect to a Riemannian metric g on the configuration manifold Q . Suppose that the inputs are modelled by vector fields $\mathcal{V} = \{Y_1, \dots, Y_m\}$. We may define the *symmetric product* between two vector fields on Q by

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

where $\nabla_X Y$ is the *covariant derivative* of Y with respect to X . If $\mathcal{T}(Q)$ denotes the set of vector fields on Q , and if $\mathcal{V} \subset \mathcal{T}(Q)$, we denote by $\overline{\text{Sym}}(\mathcal{V})$ the set of vector fields on Q obtained by taking iterated symmetric products of vector fields from \mathcal{V} . The usual involutive closure of \mathcal{V} will be denoted $\overline{\text{Lie}}(\mathcal{V})$. We shall say that a symmetric product from $\overline{\text{Sym}}(\mathcal{V})$ is *bad* if it contains an even number of each of the vector fields in \mathcal{V} . Otherwise we shall call a symmetric product from $\overline{\text{Sym}}(\mathcal{V})$ *good*. The *degree* of an iterated symmetric product of factors from \mathcal{V} will denote the total number of factors.

Notice that with the Lagrangian given by just kinetic energy, all configurations with zero velocity are equilibrium points for the unforced mechanical system. We shall say the system is *locally configuration accessible* at $q \in Q$ if the set of points reachable starting from q at zero velocity is open in Q . We shall say the system is *equilibrium controllable* if, starting from a given configuration at zero velocity, we can reach an open set of final configurations at zero velocity. Now we may state two results.

THEOREM: *Consider the mechanical control system on the configuration manifold Q whose Lagrangian is the kinetic energy with respect to a Riemannian metric g and whose input vector fields are $\mathcal{V} = \{Y_1, \dots, Y_m\}$. Then*

- (i) *the system is locally configuration accessible at q if the distribution defined by $\overline{\text{Lie}}(\overline{\text{Sym}}(\mathcal{V}))$ has maximal rank at q , and*
- (ii) *the system is equilibrium controllable if it is locally configuration accessible and if every bad symmetric product is a linear combination of good symmetric products of lower degree.*

To prove this result, one basically proceeds as follows. Compute the accessibility distribution on TQ for the mechanical control system and evaluate at zero velocity. This will describe the set of *states* accessible from points of zero velocity. However, since we are interested in controllability of the *configurations*, we can project the accessibility distribution to Q with $T\tau_Q$, the derivative of the tangent bundle projection. It turns out that this is exactly the distribution $\overline{\text{Lie}}(\overline{\text{Sym}}(\mathcal{V}))$. In this way we see that the conditions in (i) give local configuration accessibility. To prove (ii), we appeal to the controllability results of Sussmann [1987] on local controllability. An application of Sussmann's results to the systems we are considering yields (ii).

The sections which follow formalise the above definitions and results and also generalise them to the case where the system has potential energy.

3. Mathematical preliminaries

In this section we present the necessary mathematical ideas we shall need for our exposition of Section 5.

3.1. Free lie algebras and families of vector fields. In this section we recall some ideas for Lie algebras as presented in [Serre 1992]. These ideas will be important in our adaptation of the conditions for small-time local controllability of Sussmann [1987] as well as for some bracket calculations in Section 5.1.

Let X be a set and let $A(X)$ be the free algebra of associative but not necessarily commutative products of elements in X . Let I be the two-sided ideal of $A(X)$ generated by elements of the form $a \cdot a$ and $a \cdot (b \cdot c) + c \cdot (a \cdot b) + b \cdot (c \cdot a)$. The algebra $L(X) = A(X)/I$ is called the *free Lie algebra* generated by X . The inherited product on this algebra satisfies the usual Lie bracket properties of a Lie algebra. We denote by $\text{Br}(X)$ the subset of $L(X)$ consisting of brackets whose elements are in X . This subset generates $L(X)$ as a real vector space. In fact, the following proposition, whose proof may be found in [Jacobson 1962], gives a subset of $\text{Br}(X)$ which generates $L(X)$.

3.1 PROPOSITION: *Every element of $L(X)$ is a linear combination of repeated brackets of the form*

$$[X_k, [X_{k-1}, [\dots, [X_2, X_1] \dots]]] \quad (3.1)$$

where $X_i \in X$, $i = 1, \dots, k$.

We will need the notion of what we shall call the components of an element $u \in L(X)$. Every such element u has a unique decomposition as $u = [u_1, u_2]$. In turn, each of u_1 and u_2 may be uniquely expressed as $u_1 = [u_{11}, u_{12}]$ and $u_2 = [u_{21}, u_{22}]$. This process may be continued until we end up with elements which are not decomposable. All such elements $u_{i_1 \dots i_m}$, $i_a \in \{1, 2\}$, shall be called *components* of u .

If $\mathbf{X} = \{X_0, \dots, X_l\}$, for $B \in \text{Br}(\mathbf{X})$ we define $\delta_a(B)$, $a = 0, \dots, l$, to be the number of times that X_a occurs in B . The sum of the δ_a 's we shall call the *degree* of B .

Given a family of vector fields on a manifold M , $\mathcal{V} \subset \mathcal{T}(M)$, we may define a distribution on M by

$$D_{\mathcal{V}}(x) = \text{span}_{\mathbb{R}} \{X(x) \mid X \in \mathcal{V}\}.$$

Since $\mathcal{T}(M)$ is a Lie algebra, we may ask for the smallest Lie subalgebra of $\mathcal{T}(M)$ which contains a family of vector fields \mathcal{V} . It is convenient to describe this subalgebra using the ideas from free Lie algebras presented above.

Let X be a set which is bijective to \mathcal{V} with bijection ϕ . Thus, to each element of X we associate a vector field in \mathcal{V} . We establish a Lie algebra homomorphism, $\text{Ev}(\phi): L(X) \rightarrow \mathcal{T}(M)$, in a natural manner. Thus we define $\text{Ev}(\phi)$ so that $[\text{Ev}(\phi)(B_1), \text{Ev}(\phi)(B_2)] = \text{Ev}(\phi)([B_1, B_2])$ for $B_1, B_2 \in \text{Br}(X)$ and then extend this to $L(X)$ by \mathbb{R} -linearity. The smallest Lie subalgebra of $\mathcal{T}(M)$ which contains \mathcal{V} may now be stated in a simple manner. It is simply the image of $L(X)$ under the homomorphism $\text{Ev}(\phi)$. We shall denote this subalgebra by $\overline{\text{Lie}}(\mathcal{V})$ and call it the *involutive closure* of \mathcal{V} .

For $x \in M$ we define the map $\text{Ev}_x(\phi): L(X) \rightarrow T_x M$ by

$$\text{Ev}_x(\phi)(u) = (\text{Ev}(\phi)(u))(x).$$

We shall say that \mathcal{V} satisfies the *Lie algebra rank condition (LARC)* at x if $\text{Ev}_x(\phi)(L(X)) = T_x M$.

It is possible to talk about the involutive closure and the LARC without using free Lie algebras. However, since we will have to use free Lie algebras later in the paper, using them here provides us an opportunity to introduce the ideas in a more straightforward setting.

3.2. Symmetric algebras. As far as we know, the idea of a symmetric algebra does not appear in the literature. However, the concept is a natural one and shall be useful to us. A *symmetric algebra* is an algebra, A , where the multiplication (which we shall denote by $\langle u, v \rangle$) satisfies $\langle u : v \rangle = \langle v : u \rangle$ for $u, v \in A$. A map, $\sigma: A \rightarrow A'$, between symmetric algebras is called a *symmetric algebra homomorphism* if $\langle \sigma(u) : \sigma(v) \rangle = \langle u : v \rangle$ for each $u, v \in A$.

We now construct a symmetric algebra which is generated by a given set X . To construct this algebra, let X be a set and recall that $A(X)$ is the free algebra on X . The *free symmetric algebra* on X , denoted $S(X)$, is the quotient algebra obtained by taking the quotient of $A(X)$ by the two-sided ideal generated by all elements of the form $a \cdot b - b \cdot a$ where $a, b \in A(X)$. We shall denote the product in $S(X)$ by $\langle u : v \rangle$. Note that, by construction, $\langle u : v \rangle = \langle v : u \rangle$ for every $u, v \in S(X)$. We denote by $\text{Pr}(X)$ the subset of $S(X)$ consisting of the symmetric products whose elements are in X .

As with free Lie algebras, the finitely generated case is the most interesting to us. Let $\mathbf{Y} = \{X_1, \dots, X_{l+1}\}$ (the reason for the slightly unusual enumeration will become clear in Section 5.5). For $P \in \text{Pr}(\mathbf{Y})$ define $\gamma_a(P)$ to be the number of times the element X_a occurs in $P \in \text{Pr}(\mathbf{Y})$ for $a = 1, \dots, l+1$. We shall call the sum of the γ_a 's the *degree* of P .

3.3. Some Riemannian geometry. The kinetic energy of a mechanical system may be regarded as being determined by a *Riemannian metric* on the configuration space. A Riemannian metric, g , on a manifold, M , is simply a smooth assignment of an inner

product for each tangent space of the manifold. In a set of coordinates (x^1, \dots, x^n) for M , the components of the metric are given by $g_{ij} = g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$. For each $x \in M$, we may define isomorphisms $\sharp: T_x^* M \rightarrow T_x M$ and $\flat: T_x M \rightarrow T_x^* M$ in the usual manner (see [Klingenberg 1982]). These maps naturally extend to isomorphisms from $\mathcal{T}(M)$, the set of vector fields on M , to $\mathcal{T}^*(M)$, the set of one-forms on M . In this case, given a function $f \in C^\infty(M)$, we define $\text{grad } f = (\mathbf{d}f)^\sharp$.

A Riemannian manifold is endowed with a unique *affine connection* (called the Levi-Civita connection) which is characterised by being torsion free and by its parallel transportation being metric preserving (see [Klingenberg 1982]). This affine connection defines $\nabla_X Y$, which is called the *covariant derivative* of Y with respect to X . In coordinates we have

$$\nabla_X Y = \left(\frac{\partial Y^i}{\partial x^j} X^j + \Gamma_{jk}^i X^j Y^k \right) \frac{\partial}{\partial x^i}.$$

The Γ_{jk}^i are the *Christoffel symbols* and are given by

$$\Gamma_{jk}^i = \frac{1}{2} g^{il} \left(\frac{\partial g_{lj}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l} \right).$$

Here g^{ij} is the inverse of the matrix g_{ij} . On TM we may define a second-order vector field called the *geodesic spray* which we denote by Z_g . This vector field is characterised by the fact that the projection to M of the integral curves of Z_g by the tangent bundle projection are *geodesics*. In coordinates we have

$$Z_g = v^i \frac{\partial}{\partial x^i} - \Gamma_{jk}^i v^j v^k \frac{\partial}{\partial v^i}.$$

Here we are denoting by $(x^1, \dots, x^n, v^1, \dots, v^n)$ the natural coordinates for TM corresponding to coordinates (x^1, \dots, x^n) for M .

We shall need the concept of a "symmetric subalgebra" of $\mathcal{T}(M)$ which is generated by a family of vector fields $\mathcal{V} \subset \mathcal{T}(M)$. This construction relies on the covariant derivative discussed above. We may make $\mathcal{T}(M)$ into a symmetric algebra by defining the symmetric product

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

We remark that this product first appeared in the work of Crouch [1981] on gradient dynamical systems. Let \mathcal{V} be a family of vector fields on M and let X be a set which is bijective to \mathcal{V} with bijection $\psi: X \rightarrow \mathcal{V}$. As in Section 3.2, let $S(X)$ be the free symmetric algebra on X and let $\text{Pr}(X)$ be the symmetric products with elements in X . We may define a symmetric algebra homomorphism from $S(X)$ to $\mathcal{T}(M)$ by extending ψ in the natural way much as we did for Lie brackets in Section 3.1. We denote the resulting map from $S(X)$ to $\mathcal{T}(M)$ by $\text{Ev}(\psi)$. We also define $\text{Ev}_x(\psi)(P) = (\text{Ev}(\psi)(P))(x)$ for $x \in M$. We denote by $\overline{\text{Sym}}(\mathcal{V})$ the image of $S(X)$ under this homomorphism and call this the *symmetric closure* of \mathcal{V} .

4. Sufficient conditions for small-time local controllability

Sussmann [1987] gives a general result concerning so-called small-time local controllability. We are interested in a version of Sussmann's result and so will present only as much

background as is necessary to state this result. We consider control systems of the form

$$\dot{x} = X(x) + u^a Y_a(x) \quad (4.1)$$

on a manifold M where X, Y_1, \dots, Y_m are analytic. (Here and in the sequel, when we write $u^a Y_a$ there will be an implied sum over a from 1 to m .) We shall consider inputs from the set \mathcal{U} of piecewise constant inputs. Let $x_0 \in M$, let V be a neighbourhood of x_0 , and let $T > 0$. We denote by $\mathcal{R}^V(x_0, T)$ the set of points which can be reached from x_0 in time T while remaining in V using inputs from \mathcal{U} . We also denote $\mathcal{R}^V(x_0, \leq T) = \bigcup_{t=0}^T \mathcal{R}^V(x_0, t)$. We say that the system (4.1) is **locally accessible** at x_0 if $\mathcal{R}^V(x_0, \leq T)$ contains an open subset of M for each V and for each T sufficiently small. Furthermore, we say that (4.1) is **small-time locally controllable** (STLC) if it is locally accessible and if x_0 is in the interior of $\mathcal{R}^V(x_0, \leq T)$ for each V and for each T sufficiently small.

Let $\mathbf{X} = \{X_0, \dots, X_m\}$. An element $B \in \text{Br}(\mathbf{X})$ is said to be **bad** if $\delta_0(B)$ is odd and $\delta_a(B)$ is even for each $a = 1, \dots, m$. A bracket is **good** if it is not bad. Let S_m denote the permutation group on m symbols. For $\pi \in S_m$ and $B \in \text{Br}(\mathbf{X})$, define $\bar{\pi}(B)$ to be the bracket obtained by fixing X_0 and sending X_a to $X_{\pi(a)}$ for $a = 1, \dots, m$. Now define

$$\beta(B) = \sum_{\pi \in S_m} \bar{\pi}(B).$$

We may state sufficient conditions for STLC.

4.1 THEOREM: ([Sussmann 1987]) *Consider the bijection $\phi: \mathbf{X} \rightarrow \{X, Y_1, \dots, Y_m\}$ which sends X_0 to X and X_a to Y_a for $a = 1, \dots, m$. Suppose that (4.1) is such that every bad bracket $B \in \text{Br}(\mathbf{X})$ has the property that*

$$\text{Ev}_x(\phi)(\beta(B)) = \sum_{a=1}^m \xi^a \text{Ev}_x(\phi)(C_a)$$

where C_a are good brackets in $\text{Br}(\mathbf{X})$ of lower degree than B and $\xi^a \in \mathbb{R}$ for $a = 1, \dots, m$. Also suppose that (4.1) satisfies the LARC at x . Then (4.1) is STLC at x .

Sussmann [1987] gives this result as a corollary of a special case originally conjectured by Hermes [1982] and proven by Sussmann [1983].

5. Controllability of simple mechanical control systems

In this section we present the main results of the paper. First we make explicit the class of control systems we are considering. All problem data will be assumed to be analytic so that we may use piecewise constant inputs. The data for the systems we consider is: an n -dimensional configuration manifold Q , a Riemannian metric, g , on Q which represents the kinetic energy, a \mathbb{R} -valued function, V , on Q which represents the potential energy, and m linearly independent one-forms, F^1, \dots, F^m , on Q which represent the input forces for the system. A system described by this data we call a **simple mechanical control system**. Although the one-forms F^1, \dots, F^m describe the forces for the problem, it is the vector fields $Y_a = (F^a)^\sharp$, $a = 1, \dots, m$, which will appear in the computations. Nevertheless, it is the one-forms which are basic in the problem description.

Given a vector field X on Q , we define the **vertical lift** (see [Abraham and Marsden 1978]) of X as the vector field on TQ defined by

$$X^{\text{lift}}(v) = \frac{d}{dt} \Big|_{t=0} (v + tX(\tau_Q(v)))$$

for $v \in TQ$ and where $\tau_Q: TQ \rightarrow Q$. If (q^1, \dots, q^n) are coordinates for Q , we shall denote the corresponding natural coordinates for TQ by $(q^1, \dots, q^n, v^1, \dots, v^n)$. In coordinates we have

$$X^{\text{lift}}(v_q) = X^i(q) \frac{\partial}{\partial v^i}$$

for $v_q \in T_q Q$. We may now define the vector field X_L on TQ by $X_L = Z_g - \text{grad } V^{\text{lift}}$ where we recall that Z_g is the geodesic spray introduced in Section 3.3. With this notation, the Euler-Lagrange equations for the forced system may be shown to be equivalent to the first order system

$$\dot{v} = X_L + u^a Y_a^{\text{lift}} \quad (5.1)$$

on TQ . Thus the **drift vector field** for the system is X_L and the control vector fields are $Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}$. It is this first order affine control system which we study in this section. We are particularly interested in the following problem:

PROBLEM STATEMENT: Describe the set of configurations reachable from a given configuration when starting at rest.

Observe that we place no restriction on the final velocities of the system. The reader will further observe that this problem statement involves only *configurations* and not velocities. It would be desirable, therefore, to derive an answer to this problem in terms of quantities on the *configuration space*. As we shall see, this can in fact be done and is one of the more compelling aspects of this approach.

Since the computations in this section are quite involved, let us outline them here before we begin. The main goal of the computations is to describe the accessibility distribution for (5.1) at points of zero velocity in TQ . Thus we need to compute the involutive closure of the family of vector fields $\mathcal{Y}' = \{X_L, Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}\}$. Observe that since $X_L = Z_g - \text{grad } V^{\text{lift}}$, we may write vector fields in $\overline{\text{Lie}}(\mathcal{Y}')$ as \mathbb{R} -linear combinations of vector fields in $\mathcal{Y} = \{Z_g, Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}, \text{grad } V^{\text{lift}}\}$. This is made precise using free Lie algebras in Section 5.1. When we evaluate the brackets which are used in the computation of the accessibility distribution at zero velocity, only a small number of them make a contribution, and the rest vanish. The brackets which vanish do so in one of two ways. Either they are identically zero, or they are polynomial in the velocity coordinates and so go to zero when the velocity goes to zero. Therefore, we have three possible classes of brackets: one class which is non-zero when the velocity is zero, one-class which is identically zero, and one class which is not identically zero, but is zero when the velocity is zero. In Section 5.1 we categorise these three types of brackets. There we shall see that the brackets which make a contribution to the accessibility distribution at zero velocity may be written as linear combinations of special brackets which we call **primitive** brackets. The computations in Section 5.1 are done at the level of free Lie algebras since this provides a rigorous way to perform the necessary computations. In Section 5.2 we give expressions for primitive brackets in terms of the geometry of the problem. It is here that the symmetric product

introduced in Section 3.3 makes its appearance. In Section 5.3 we assemble the results of Sections 5.1 and 5.2 to arrive at the form of the accessibility distribution for (5.1) at points of zero velocity. In Section 5.4 we provide a precise statement of the types of controllability we consider, and in Section 5.5 we provide computable conditions for these versions of controllability.

We remark that most of the complexity of this section is a consequence of including potential energy in the formulation. In [Lewis and Murray 1995] the authors provide sufficient conditions for controllability when there is no potential energy function. Due to space considerations, some of the free Lie algebra proofs from Section 5.1 are omitted. We refer the reader to the dissertation of Lewis [1995] for these proofs.

5.1. Computations with free Lie algebras. In this section we perform some calculations with a pair of free Lie algebras which are suited to our purposes. Rather than just using a generating set which is in 1–1 correspondence with the set $\mathcal{Y}' = \{X_L, Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}\}$ of control vector fields and the drift vector field, we also use a generating set which is in 1–1 correspondence with the set $\mathcal{Y} = \{Z_g, Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}, \text{grad } V^{\text{lift}}\}$. The reason for this is that vector fields in \mathcal{Y}' are \mathbb{R} -linear combinations of vector fields in \mathcal{Y} , and, as we shall see in Section 5.3, it is comparatively easy to describe the involutive closure of \mathcal{Y} .

Let $\mathbf{X} = \{X_0, \dots, X_{m+1}\}$ and let $L(\mathbf{X})$ be the free Lie algebra generated by the set \mathbf{X} . We can simplify many of our computations for the controllability analysis of (5.1) by making simplifications to a set of generators for $L(\mathbf{X})$. We first need some notation. Let

$$\begin{aligned} \text{Br}^k(\mathbf{X}) &= \{B \in \text{Br}(\mathbf{X}) \mid \text{the degree of } B \text{ is } k\}, \\ \text{Br}_k(\mathbf{X}) &= \left\{ B \in \text{Br}(\mathbf{X}) \mid \delta_0(B) - \sum_{a=1}^{m+1} \delta_a(B) = k \right\}. \end{aligned}$$

We shall see in Section 5.2 that, when we restrict to zero velocities, only a small subset of $\text{Br}(\mathbf{X})$ will evaluate to something non-zero. In turn, these brackets will be seen to be linear combinations of a special class of brackets which we shall call **primitive** brackets. Recall from Section 3.1 the notion of components in $L(\mathbf{X})$.

5.1 DEFINITION: Let $B \in \text{Br}_0(\mathbf{X}) \cup \text{Br}_{-1}(\mathbf{X})$ and let $B_1, B_2, B_{11}, B_{12}, B_{21}, B_{22}, \dots$ be the decomposition of B into its components. We shall say that B is **primitive** if each of its components is in $\text{Br}_{-1}(\mathbf{X}) \cup \text{Br}_0(\mathbf{X}) \cup \{X_0\}$. \square

The relevant observations that need to be made regarding primitive brackets are:

Prim1. If $B \in \text{Br}_{-1}(\mathbf{X})$ is primitive then, up to sign, we may write $B = [B_1, B_2]$ with $B_1 \in \text{Br}_{-1}(\mathbf{X})$ and $B_2 \in \text{Br}_0(\mathbf{X})$ both primitive.

Prim2. If $B \in \text{Br}_0(\mathbf{X})$ is primitive then, up to sign, B may have one of two forms. Either $B = [X_0, B_1]$ with $B_1 \in \text{Br}_{-1}(\mathbf{X})$ primitive, or $B = [B_1, B_2]$ with $B_1, B_2 \in \text{Br}_0(\mathbf{X})$ both primitive.

Using these two rules, it is possible to construct primitive brackets of any degree. For

example, the primitive brackets of degrees one through four are, up to sign

$$\begin{aligned} \text{Degree 1: } & \{X_a \mid a = 1, \dots, m\} \\ \text{Degree 2: } & \{[X_0, X_a] \mid a = 1, \dots, m\} \\ \text{Degree 3: } & \{[X_a, [X_0, X_b]] \mid a, b = 1, \dots, m\} \\ \text{Degree 4: } & \{[X_0, [X_a, [X_0, X_b]]] \mid a, b = 1, \dots, m\} \cup \\ & \{[[X_0, X_a], [X_0, X_b]] \mid a, b = 1, \dots, m\}. \end{aligned}$$

From Proposition 3.1 we know that to generate $L(\mathbf{X})$ we need only look at brackets of the form

$$[X_{a_k}, [X_{a_{k-1}}, \dots, [X_{a_2}, X_{a_1}]]] \quad (5.2)$$

where $a_i \in \{0, \dots, m+1\}$ for $i = 1, \dots, k$. We shall see in Section 5.2 that brackets from $\text{Br}_j(\mathbf{X})$, where $j \geq 1$ or $j \leq -2$, will not be of interest to us. In particular, we shall see that when $j \leq -2$ these brackets evaluate identically to zero. Therefore, in this section we concentrate our attention on brackets in $\text{Br}_0(\mathbf{X}) \cup \text{Br}_{-1}(\mathbf{X})$ which satisfy certain requirements. We state the form of these brackets in the following lemma.

5.2 LEMMA: *Let us impose the condition on elements of $\text{Br}(\mathbf{X})$ that we shall consider a bracket to be zero if any of its components is in $\text{Br}_{-j}(\mathbf{X})$ for $j \geq 2$. Let $B \in \text{Br}_0(\mathbf{X}) \cup \text{Br}_{-1}(\mathbf{X})$. Then we may write B as a finite sum of primitive brackets.*

The inductive proof is straightforward and we refer the interested reader to [Lewis 1995]. However, in lieu of a proof an example is illustrative.

5.3 EXAMPLE: Consider the bracket $B = [X_{m+1}, [X_0, [X_0, X_a]]] \in \text{Br}_0(\mathbf{X})$. This bracket is in $\text{Br}_0(\mathbf{X})$ but is not primitive. However, by Lemma 5.2, we may write B as a finite sum of primitive brackets. Indeed, by Jacobi's identity we have

$$\begin{aligned} B &= [X_{m+1}, [X_0, [X_0, X_a]]] = -[[X_0, X_a], [X_{m+1}, X_0]] - [X_0, [[X_0, X_a], X_{m+1}]] \\ &= [[X_0, X_a], [X_0, X_{m+1}]] + [X_0, [X_{m+1}, [X_0, X_a]]]. \end{aligned}$$

The proof of Lemma 5.2 is essentially a generalisation of this example.

Now we relate the free Lie algebra $L(\mathbf{X})$ with a free Lie algebra which corresponds to the set $\mathcal{Y}' = \{X_L, Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}\}$. As we mentioned above, the reason we wish to do this is because the vector fields in \mathcal{Y}' are \mathbb{R} -linear combinations of vector fields in $\mathcal{Y} = \{Z_g, Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}, \text{grad } V^{\text{lift}}\}$, the latter family of vector fields being bijective with the set \mathbf{X} . Let $\mathbf{X}' = \{X'_0, \dots, X'_m\}$. We formally set $X'_0 = X_0 - X_{m+1}$ and $X'_a = X_a$ for $a = 1, \dots, m$. We may now write brackets in $\text{Br}(\mathbf{X}')$ as linear combinations of brackets in $\text{Br}(\mathbf{X})$ by \mathbb{R} -linearity of the bracket. We may, in fact, be even more precise about this.

Let $B' \in \text{Br}(\mathbf{X}')$. We define a subset, $\mathbf{S}(B')$, of $\text{Br}(\mathbf{X})$ by saying that $B \in \mathbf{S}(B')$ if each occurrence of X'_a in B' is replaced with X_a for $a = 1, \dots, m$, and if each occurrence of X'_0 in B' is replaced with *either* X_0 *or* X_{m+1} . An example is illustrative. Suppose that

$$B' = [[X'_0, X'_1], [X'_2, [X'_0, X'_3]]].$$

Then

$$\begin{aligned} \mathbf{S}(B') &= \{[[X_0, X_1], [X_2, [X_0, X_3]]], [[X_0, X_1], [X_2, [X_{m+1}, X_3]]], \\ & \quad [[X_{m+1}, X_1], [X_2, [X_0, X_3]]], [[X_{m+1}, X_1], [X_2, [X_{m+1}, X_3]]]\}. \end{aligned}$$

Now we may precisely state how we write brackets in $\text{Br}(\mathbf{X}')$.

5.4 LEMMA: *Let $B' \in \text{Br}(\mathbf{X}')$. Then*

$$B' = \sum_{B \in \mathcal{S}(B')} (-1)^{\delta_{m+1}(B)} B.$$

The proof is by induction and may be found in [Lewis 1995].

We shall only be interested in terms in the above decomposition of B' which are in $\text{Br}_0(\mathbf{X}) \cup \text{Br}_{-1}(\mathbf{X})$ since, as we shall see in Section 5.2, these are the only ones which will contribute to $\text{Ev}_{0_q}(\phi')(B')$. Here 0_q is the zero vector in $T_q Q$.

5.2. Distribution computations for simple mechanical control systems. In this section we use the simplifications of Section 5.1 to get a complete description of the brackets which contribute to the accessibility distribution for (5.1) restricted to $Z(TQ)$, the zero section of TQ . Note that we restrict to $Z(TQ)$ because we are interested in determining the reachable points starting with zero initial velocity. To make the correspondence between the free Lie algebra $L(\mathbf{X})$ used in Section 5.1 and the accessibility algebra for (5.1), we use the family of vector fields $\mathcal{V} = \{Z_g, Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}, \text{grad } V^{\text{lift}}\}$ and establish a bijection, ϕ , from \mathbf{X} to \mathcal{V} by mapping X_0 to X_L , X_a to Y_a^{lift} for $a = 1, \dots, m$, and X_{m+1} to $\text{grad } V^{\text{lift}}$. Please note that \mathcal{V} is *not* the family of vector fields which generates the accessibility algebra. The accessibility algebra is generated by the family $\mathcal{V}' = \{X_L, Y_1^{\text{lift}}, \dots, Y_m^{\text{lift}}\}$. We establish a bijection, ϕ' , from \mathbf{X}' to \mathcal{V}' by mapping X'_0 to X_L and X'_a to Y_a^{lift} for $a = 1, \dots, m$. By Lemma 5.4, each vector field in $\text{Lie}(\mathcal{V}')$ is a \mathbb{R} -linear sum of vector fields in $\text{Lie}(\mathcal{V})$. That lemma also completely describes the sum.

Now we shall show that it is possible to compute the brackets from $\text{Br}(\mathbf{X})$ in terms of the problem data. We first present a lemma which gives the basic structure of primitive brackets. In this lemma we see that a large number of brackets are computable in terms of quantities defined on Q . This is worth noting since the vector fields themselves are defined on TQ . Of particular interest in the lemma is the appearance of the symmetric product which was introduced in Section 3.3.

We need to say a few words about the structure of TQ . We denote by $Z(TQ)$ the zero section of TQ . Since Q is naturally diffeomorphic to $Z(TQ)$, there is a natural inclusion of $T_q Q$ into $T_{0_q} TQ$ for each $q \in Q$. We shall call the image of this inclusion in $T_{0_q} TQ$ the **horizontal subspace**. We shall call the subspace of $T_{0_q} TQ$ which is tangent to the fibre of TQ at q the **vertical subspace** and denote it by $V_{0_q} TQ$. We have $T_{0_q} TQ = T_q Q \oplus V_{0_q} TQ$ for each $q \in Q$. We mention that this notion of vertical is valid at any point in TQ . However, the definition of horizontal is only valid on $Z(TQ)$.

5.5 LEMMA: *Suppose that $B \in \text{Br}^k(\mathbf{X})$ is primitive.*

- (i) *If $B \in \text{Br}_{-1}(\mathbf{X})$ then $\text{Ev}(\phi)(B)$ is the vertical lift of a vector field on Q .*
- (ii) *If $B \in \text{Br}_0(\mathbf{X})$ then $U = \text{Ev}(\phi)(B)$ has the property that, when expressed in a local chart, the vertical components of U are linear in the fibre coordinates v and the horizontal components are independent of v . In particular, we may define a vector field on Q by $U_Q: q \mapsto U(0_q) \in T_q Q \subset T_{0_q} TQ$. There are two cases to consider.*

- (a) *$B = [X_0, B_1]$ with $B_1 \in \text{Br}_{-1}(\mathbf{X})$: Define U_1 to be the vector field on Q such that $\text{Ev}(\phi)(B_1) = U_1^{\text{lift}}$. Then $U(0_q) = \text{Ev}(\phi)(B)(0_q) = -U_1(q)$. Let $U_2 \in \mathcal{F}(Q)$. Then $[U_2^{\text{lift}}, U] = (\nabla_{U_1} U_2 + \nabla_{U_2} U_1)^{\text{lift}}$.*

- (b) *$B = [B_1, B_2]$ with $B_1, B_2 \in \text{Br}_0(\mathbf{X})$: Define $U_{1,Q}, U_{2,Q}$ to be the vector fields on Q corresponding to $\text{Ev}(\phi)(B_1), \text{Ev}(\phi)(B_2)$, respectively. Then $\text{Ev}(\phi)(B)(0_q) = [U_{1,Q}, U_{2,Q}](q)$.*

Proof: The proof is by induction on k . The result is true for $k = 1$ trivially. If X and Y are vector fields on Q it is a straightforward coordinate computation to show that

$$[X^{\text{lift}}, Y^{\text{lift}}] = 0.$$

If X is a vector field on Q we compute

$$[Z_g, X^{\text{lift}}] = -Y^i \frac{\partial}{\partial q^i} + \left(\frac{\partial Y^i}{\partial q^j} v^j + \Gamma_{jk}^i Y^j v^k + \Gamma_{kj}^i v^k Y^j \right) \frac{\partial}{\partial v^i}. \quad (5.3)$$

Inspecting (5.3) shows that $[Z_g, X^{\text{lift}}](0_q) = -X(q)$. Now let $Y \in \mathcal{F}(Q)$. We compute

$$[Y^{\text{lift}}, [Z_g, X^{\text{lift}}]] = \left(\frac{\partial Y^i}{\partial q^j} X^j + \frac{\partial X^i}{\partial q^j} Y^j + 2\Gamma_{jk}^i X^j Y^k \right) \frac{\partial}{\partial v^i} \quad (5.4)$$

which is the coordinate representation of $(\nabla_X Y + \nabla_Y X)^{\text{lift}}$. This shows that the lemma is true for $k = 2$.

Now suppose the lemma true for $k = 1, \dots, l$ for $l \geq 2$ and let $B \in \text{Br}^{l+1}(\mathbf{X})$ be primitive.

(i) Suppose that $B \in \text{Br}_{-1}(\mathbf{X})$. Without loss of generality (by Prim1 we may suppose that $B = [B_1, B_2]$ with $B_1 \in \text{Br}_{-1}(\mathbf{X})$ and $B_2 \in \text{Br}_0(\mathbf{X})$). Then, by the induction hypotheses, we have

$$\text{Ev}(\phi)(B_1) = \alpha^i(q) \frac{\partial}{\partial v^i}, \quad \text{Ev}(\phi)(B_2) = \lambda^i(q) \frac{\partial}{\partial q^i} + \mu_j^i(q) v^j \frac{\partial}{\partial v^i}.$$

Now we compute

$$\text{Ev}(\phi)([B_1, B_2]) = \left(\mu_j^i \alpha^j - \frac{\partial \alpha^i}{\partial q^j} \lambda^j \right) \frac{\partial}{\partial v^i}.$$

Note that the components in the q -direction are zero and the components in the v -direction are only functions of q . This means that this vector field is the vertical lift of a vector field on Q . This proves (i).

(ii) Suppose that $B \in \text{Br}_0(\mathbf{X})$. Without loss of generality (by Prim2) we may suppose that either (ii a) $B = [X_0, B_1]$ with $B_1 \in \text{Br}_{-1}(\mathbf{X})$ or that (ii b) $B = [B_1, B_2]$ with $B_1, B_2 \in \text{Br}_0(\mathbf{X})$. Let us deal with the first case. Equation (5.3) gives $\text{Ev}(B)(\phi)(0_q) = -U_1(q)$ where U_1 is the vector field on Q so that $\text{Ev}(\phi)(B_1) = U_1^{\text{lift}}$ (such a vector field exists by (i)). For every vector field U_2 on Q we have $[U_2^{\text{lift}}, [Z_g, U_1^{\text{lift}}]] = (\nabla_{U_1} U_2 + \nabla_{U_2} U_1)^{\text{lift}}$ by (5.4). This proves (ii a).

Now suppose that we have $B_1, B_2 \in \text{Br}_0(\mathbf{X})$. Then, by the induction hypotheses, we have

$$\text{Ev}(\phi)(B_1) = \alpha^i(q) \frac{\partial}{\partial q^i} + \beta_j^i(q) v^j \frac{\partial}{\partial v^i}, \quad \text{Ev}(\phi)(B_2) = \lambda^i(q) \frac{\partial}{\partial q^i} + \mu_j^i(q) v^j \frac{\partial}{\partial v^i}.$$

We compute

$$\begin{aligned} \text{Ev}(\phi)([B_1, B_2]) &= \left(\frac{\partial \lambda_i}{\partial q^j} \alpha^j - \frac{\partial \alpha^i}{\partial q^j} \lambda^j \right) \frac{\partial}{\partial q^i} + \\ &\quad \left(\frac{\partial \mu_k^i}{\partial q^j} \alpha^j v^k + \mu_j^i \beta_k^j v^k - \frac{\partial \beta_k^i}{\partial q^j} \lambda^j v^k - \beta_j^i \mu_k^j v^k \right) \frac{\partial}{\partial v^i}. \end{aligned}$$

The components have the order in v specified by the lemma. Also, it is clear that the vector fields on Q defined by B_1 and B_2 are

$$U_{1,Q} = \alpha^i(q) \frac{\partial}{\partial q^i}, \quad \text{and} \quad U_{2,Q} = \lambda^i(q) \frac{\partial}{\partial q^i},$$

respectively. It is easy to see that $\text{Ev}(\phi)(B)(0_q) = [U_{1,Q}, U_{2,Q}](q)$. This completes the proof of the lemma. \blacksquare

This lemma provides us with a positive step towards computing the value of all primitive brackets when evaluated using $\text{Ev}(\phi)$. The following lemma shows that these are the *only* brackets we need to consider.

5.6 LEMMA: (i) *Let $l \geq 1$ be an integer and let $B \in \text{Br}_l(\mathbf{X})$. Then $\text{Ev}(\phi)(B)(0_q) = 0$ for each $q \in Q$.*

(ii) *Let $l \geq 2$ be an integer and let $B \in \text{Br}^k(\mathbf{X}) \cap \text{Br}_{-l}(\mathbf{X})$ for $k \geq 2$. Then $\text{Ev}(\phi)(B) = 0$.*

The proof of this lemma may be found in [Lewis 1995]. It goes very much like the proof of Lemma 5.5.

Let us summarise what we have done in this section. First we obtained a characterisation of primitive brackets in \mathbf{X} when we evaluate them in \mathcal{V} via $\text{Ev}(\phi)$. This characterisation involved Lie brackets and covariant derivatives of the vector fields $Y_1, \dots, Y_m, \text{grad } V$. Then we showed in Lemma 5.6 that primitive brackets are the only ones we need be concerned with if we are evaluating the vector fields on the zero section of TQ .

5.3. The form of the accessibility distribution restricted to $Z(TQ)$ for simple mechanical control systems.

In this section we compute the accessibility distribution for (5.1) when restricted to the zero section of TQ . By Lemma 5.4 we know that we may write the vector fields in the accessibility distribution in terms of vector fields in $\overline{\text{Lie}}(\mathcal{V})$. In Section 5.2 we saw some hints that we might be able to write vector fields in $\overline{\text{Lie}}(\mathcal{V})$ in terms of covariant derivatives and Lie brackets of the input vector fields and $\text{grad } V$. First we resolve this issue by saying exactly what the vector fields in $\overline{\text{Lie}}(\mathcal{V})$ look like when we restrict them to $Z(TQ)$. We denote by $D_{\overline{\text{Lie}}(\mathcal{V})}$ the distribution defined by

$$D_{\overline{\text{Lie}}(\mathcal{V})}(v) = \text{span}_{\mathbb{R}} \{U(v) \mid U \in \overline{\text{Lie}}(\mathcal{V})\}.$$

The reader will also wish to recall the ideas from symmetric algebras presented in Section 3.3. We denote $\mathcal{S} = \{Y_1, \dots, Y_m\}$.

The following lemma describes the horizontal and vertical parts of the involutive closure of \mathcal{V} restricted to $Z(TQ)$. The reader may wish to recall our remarks about the structure of the tangent bundle preceding Lemma 5.5.

5.7 LEMMA: *Let $q \in Q$. Then*

$$D_{\overline{\text{Lie}}(\mathcal{V})}(0_q) \cap V_{0_q} TQ = (D_{\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\})}(q))^{\text{lift}}$$

and

$$D_{\overline{\text{Lie}}(\mathcal{V})}(0_q) \cap T_q Q = D_{\overline{\text{Lie}}(\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\}))}(q).$$

Proof: From Lemma 5.6 we know that the only brackets from $\text{Br}(\mathbf{X})$ which we need to consider are the primitive brackets. From Lemma 5.5 we know that the brackets which are in $\text{Br}_{-1}(\mathbf{X})$ will generate the vertical directions, and the brackets which are in $\text{Br}_0(\mathbf{X})$ will generate the horizontal directions.

First we show that $(D_{\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\})}(q))^{\text{lift}} \subset D_{\overline{\text{Lie}}(\mathcal{V})}(0_q)$. This may be done inductively. Define $\text{Sym}^{(1)}(\mathcal{S} \cup \{\text{grad } V\}) = \mathcal{S} \cup \{\text{grad } V\}$ and inductively define

$$\text{Sym}^{(k)}(\mathcal{S} \cup \{\text{grad } V\}) = \{\langle U_1 : U_2 \rangle \mid U_i \in \text{Sym}^{(k_i)}(\mathcal{S} \cup \{\text{grad } V\}), k_1 + k_2 = k\}.$$

Clearly

$$\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\}) = \bigcup_{k \in \mathbb{Z}^+} \text{Sym}^{(k)}(\mathcal{S} \cup \{\text{grad } V\}).$$

It is trivially true that $(\text{Sym}^{(1)}(\mathcal{S} \cup \{\text{grad } V\}))^{\text{lift}} \subset \overline{\text{Lie}}(\mathcal{V})$. Now suppose that $(\text{Sym}^{(k)}(\mathcal{S} \cup \{\text{grad } V\}))^{\text{lift}} \subset \overline{\text{Lie}}(\mathcal{V})$ for $k = 1, \dots, l$ for $l \geq 1$. We see that $(\text{Sym}^{(l+1)}(\mathcal{S} \cup \{\text{grad } V\}))^{\text{lift}} \subset \overline{\text{Lie}}(\mathcal{V})$ since we may generate all elements of $(\text{Sym}^{(l+1)}(\mathcal{S} \cup \{\text{grad } V\}))^{\text{lift}}$ by considering brackets of the form $[U_1^{\text{lift}}, [Z_g, U_2^{\text{lift}}]]$ where $U_i \in \text{Sym}^{(i)}(\mathcal{S}, V)$ and $l_1 + l_2 = l + 1$. This follows from (5.4). This shows that $(D_{\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\})}(q))^{\text{lift}} \subset D_{\overline{\text{Lie}}(\mathcal{V})}(0_q)$.

Now we show that $D_{\overline{\text{Lie}}(\mathcal{V})}(0_q) \subset (D_{\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\})}(q))^{\text{lift}}$. To do this we must show that the image under $\text{Ev}(\phi)$ of all primitive brackets in $\text{Br}_{-1}(\mathbf{X})$ may be written as a linear combination of vector fields in $\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\})$. A primitive bracket in $\text{Br}_{-1}(\mathbf{X})$ may be written as $B = [B_1, B_2]$ with $B_1 \in \text{Br}_{-1}(\mathbf{X})$ and $B_2 \in \text{Br}_0(\mathbf{X})$ both being primitive. Therefore, either $B_2 = [X_0, B_2']$ with B_2' primitive and in $\text{Br}_{-1}(\mathbf{X})$ or $B_2 = [B_2', B_2'']$ with $B_2', B_2'' \in \text{Br}_0(\mathbf{X})$ both primitive. In the first case $\text{Ev}(\phi)(B) \in \text{Sym}^{(k)}(\mathcal{S} \cup \{\text{grad } V\})$ for some k by (5.4). In the second case we may use Jacobi's identity to obtain

$$B = -[B_2'', [B_1, B_2']] + [B_2', [B_1, B_2'']].$$

We may apply the above argument to the terms $[B_1, B_2']$ and $[B_1, B_2'']$ repeatedly using (5.4) until they are expressed in terms of covariant derivatives. When this is done, $\text{Ev}(\phi)(B)$ will then be a \mathbb{R} -linear combination of elements in $\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\})$. This shows that $D_{\overline{\text{Lie}}(\mathcal{V})}(0_q) \subset (D_{\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\})}(q))^{\text{lift}}$.

To demonstrate the proposed form of $D_{\overline{\text{Lie}}(\mathcal{V})} \cap T_q Q$, by Lemma 5.5(ii b) we need only show that $\overline{\text{Sym}}(\mathcal{S} \cup \{\text{grad } V\})(q) \subset D_{\overline{\text{Lie}}(\mathcal{V})}(0_q)$. But this is clear from Lemma 5.5(ii a). This completes the proof of the lemma. \blacksquare

5.8 REMARK: Notice that the constructions in the above lemma depend only upon $\{Y_1, \dots, Y_m, \text{grad } V\}$. The effects of the geodesic spray do not appear explicitly. However, its contribution is obviously important in the computations performed in Section 5.2.

From Lemmas 5.4 and 5.7 we know that the vector fields which contribute to $\overline{\text{Lie}}(\mathcal{V}')$ when we evaluate on $Z(TQ)$ will be \mathbb{R} -linear combinations of vector fields from $\overline{\text{Lie}}(\overline{\text{Sym}}(\mathcal{V} \cup \{\text{grad } V\}))$. Thus, to compute these vector fields, we need to figure out which vector fields need to be “removed” from $\overline{\text{Lie}}(\overline{\text{Sym}}(\mathcal{V} \cup \{\text{grad } V\}))$. We present an algorithm which we shall prove determines exactly which \mathbb{R} -linear combinations from $\overline{\text{Lie}}(\overline{\text{Sym}}(\mathcal{V} \cup \{\text{grad } V\}))$ we need to compute. We define two sequences of families of vector fields on Q which we shall denote by $\mathcal{C}_{\text{ver}}^{(k)}(\mathcal{V}, V)$ and $\mathcal{C}_{\text{hor}}^{(k)}(\mathcal{V}, V)$ where $k \in \mathbb{Z}^+$. In Figure 2 the algorithm is presented for computing these families. When we have computed these sequences we define

$$\mathcal{C}_{\text{ver}}(\mathcal{V}, V) = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_{\text{ver}}^{(k)}(\mathcal{V}, V), \quad \mathcal{C}_{\text{hor}}(\mathcal{V}, V) = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_{\text{hor}}^{(k)}(\mathcal{V}, V).$$

The distributions defined by these families of vector fields shall be denoted $C_{\text{ver}}(\mathcal{V}, V)$ and $C_{\text{hor}}(\mathcal{V}, V)$, respectively.

We may now state the form of the accessibility distribution $\overline{\text{Lie}}(\mathcal{V}')$ for (5.1) when restricted to the zero section of TQ .

5.9 PROPOSITION: *Let $q \in Q$. Then*

$$D_{\overline{\text{Lie}}(\mathcal{V}')} (0_q) \cap V_{0_q} TQ = (C_{\text{ver}}(\mathcal{V}, V)(q))^{\text{lift}}$$

and

$$D_{\overline{\text{Lie}}(\mathcal{V}')} (0_q) \cap T_q Q = C_{\text{hor}}(\mathcal{V}, V)(q).$$

Proof: Studying the algorithm that we have used to compute $\mathcal{C}_{\text{ver}}(\mathcal{V}, V)$ and $\mathcal{C}_{\text{hor}}(\mathcal{V}, V)$, the reader will notice that we have exactly taken each primitive bracket $B \in \text{Br}(\mathbf{X})$ and computed which \mathbb{R} -linear combinations from $\text{Br}(\mathbf{X})$ appear along with B in the decomposition of some $B' \in \text{Br}(\mathbf{X}')$ given by Lemma 5.4. Since it is only these primitive brackets which appear in $\overline{\text{Lie}}(\mathcal{V}') \mid Z(TQ)$, this will, by construction, generate $D_{\overline{\text{Lie}}(\mathcal{V}')} \mid Z(TQ)$.

We need to prove that, as stated in the first step of the algorithm, if $\delta_{m+1}(B) = 0$, then $\text{Ev}_{0_q}(\phi)(B) \in D_{\overline{\text{Lie}}(\mathcal{V}')} (0_q)$. To show that this is in fact the case, let $B' \in \text{Br}(\mathbf{X}')$ be the bracket obtained by replacing X_a with X'_a for $a = 0, \dots, m$. We claim that the only bracket in $\mathbf{S}(B')$ which contributes to $\text{Ev}(\phi')(B')$ is B . This is true since any other brackets in $\mathbf{S}(B')$ are obtained by replacing X_0 in B with X_{m+1} . Such a replacement will result in a bracket which has at least one component which is in $\text{Br}_{-l}(\mathbf{X})$ for $l \geq 2$. These brackets evaluate to zero by Lemma 5.6(ii).

We also need to show that if B has components of the form $[X_0, X_{m+1}]$, then it will not contribute to $\overline{\text{Lie}}(\mathcal{V}') \mid Z(TQ)$. This is clear since, when constructing B' in the algorithm, the component $[X_0, X_{m+1}]$ will become $[X'_0, X'_0]$ which means that B' will be identically zero. \blacksquare

It is perhaps useful to construct a few of the families $\mathcal{C}_{\text{ver}}^{(k)}(\mathcal{V}, V)$ and $\mathcal{C}_{\text{hor}}^{(k)}(\mathcal{V}, V)$ to show how the algorithm works. We shall do this for $k = 1, 2$. Our notation in these calculations follows that in the algorithm.

Let $i = 1$. The only primitive brackets in $\text{Br}^{(1)}(\mathbf{X})$ are X_1, \dots, X_{m+1} . For the brackets $B = X_a$, $a = 1, \dots, m$, $\delta_{m+1}(B) = 0$. Note that $\text{Ev}(\phi)(B) = Y_a^{\text{lift}}$ so $Y_a \in \mathcal{C}_{\text{ver}}^{(1)}(\mathcal{V}, V)$ for $a = 1, \dots, m$. The bracket X_{m+1} has no components of the form $[X_0, X_{m+1}]$ so it is a candidate for providing an element of $\mathcal{C}_{\text{ver}}^{(1)}(\mathcal{V}, V)$. If $B = X_{m+1}$ we compute $B' = X'_0$.

5.1 ALGORITHM:

```

For  $i \in \mathbb{Z}^+$  do
  For  $B \in \text{Br}^{(i)}(\mathbf{X})$  primitive do
    If  $\delta_{m+1}(B) = 0$  then
      If  $B \in \text{Br}_{-1}(\mathbf{X})$  then
         $U \in \mathcal{C}_{\text{ver}}^{\frac{1}{2}(i+1)}(\mathcal{V}, V)$  where  $\text{Ev}(\phi)(B) = U^{\text{lift}}$ 
      else
         $U \in \mathcal{C}_{\text{hor}}^{(i/2)}(\mathcal{V}, V)$  where  $U(q) = \text{Ev}_{0_q}(\phi)(B)$ 
      end
    end
  end
  If  $B$  has no components of the form  $[X_0, X_{m+1}]$  then
    Compute  $B' \in \text{Br}(\mathbf{X})$  by replacing every occurrence of  $X_0$  and
     $X_{m+1}$  in  $B$  with  $X'_0$  and by replacing every occurrence of  $X_a$  in  $B$ 
    with  $X'_a$  for  $a = 1, \dots, m$ .
    Let  $B'' = 0$ .
    For  $\tilde{B} \in \mathbf{S}(B') \cap (\text{Br}_{-1}(\mathbf{X}) \cup \text{Br}_0(\mathbf{X}))$  do
      Write  $\tilde{B}$  as a finite sum of primitive brackets in  $\text{Br}(\mathbf{X})$  by
      Lemma 5.2.
       $B'' = B'' + (-1)^{\delta_{m+1}(\tilde{B})} \tilde{B}$ 
    end
    If  $B \in \text{Br}_{-1}(\mathbf{X})$  then
       $U \in \mathcal{C}_{\text{ver}}^{\frac{1}{2}(i+1)}(\mathcal{V}, V)$  where  $\text{Ev}(\phi)(B'') = U^{\text{lift}}$ 
    else
       $U \in \mathcal{C}_{\text{hor}}^{(i/2)}(\mathcal{V}, V)$  where  $U(q) = \text{Ev}_{0_q}(\phi)(B'')$ 
    end
  end
end
end
end
END

```

FIGURE 2. Algorithm for computing $\overline{\text{Lie}}(\mathcal{V}') \mid Z(TQ)$

Therefore, $\mathbf{S}(B') = \{X_0, X_{m+1}\}$. The only element in $\mathbf{S}(B')$ which is in $\text{Br}_{-1}(\mathbf{X}) \cup \text{Br}_0(\mathbf{X})$ is X_{m+1} . Therefore, $B'' = -X_{m+1}$. We then see that $\text{Ev}(\phi)(B'') = -\text{grad } V^{\text{lift}}$ from which we conclude that $\text{grad } V \in \mathcal{C}_{\text{ver}}^{(1)}(\mathcal{V}, V)$. In summary,

$$\mathcal{C}_{\text{ver}}^{(1)}(\mathcal{V}, V) = \{Y_1, \dots, Y_m, \text{grad } V\}.$$

Now we look at the case when $i = 2$. The primitive brackets in $\text{Br}^{(2)}(\mathbf{X})$ are $\{[X_0, X_1], \dots, [X_0, X_{m+1}]\}$. The brackets $B = [X_0, X_a]$, $a = 1, \dots, m$ have the property that $\delta_{m+1}(B) = 0$. We compute $\text{Ev}_{0_q}(\phi)(B) = -Y_a(q)$ and so conclude that $Y_a \in \mathcal{C}_{hor}^{(1)}(\mathcal{Y}, V)$. The bracket $[X_0, X_{m+1}]$ is not a candidate for providing an element of $\mathcal{C}_{hor}^{(1)}(\mathcal{Y}, V)$ so we have

$$\mathcal{C}_{hor}^{(1)}(\mathcal{Y}, V) = \{Y_1, \dots, Y_m\}.$$

In a similar manner we may compute

$$\mathcal{C}_{ver}^{(2)}(\mathcal{Y}, V) = \{Y_a : Y_b \mid a, b = 1, \dots, m\} \cup \{Y_a : \text{grad } V \mid a = 1, \dots, m\}$$

and

$$\begin{aligned} \mathcal{C}_{hor}^{(2)}(\mathcal{Y}, V) = & \{Y_a : Y_b \mid a, b = 1, \dots, m\} \cup \{Y_a, Y_b \mid a, b = 1, \dots, m\} \cup \\ & \{2\langle Y_a : \text{grad } V \rangle + [Y_a, \text{grad } V] \mid a = 1, \dots, m\}. \end{aligned}$$

To compute the terms $2\langle Y_a : \text{grad } V \rangle + [Y_a, \text{grad } V]$ in $\mathcal{C}_{hor}^{(2)}(\mathcal{Y}, V)$, we have used the computations of Example 5.3.

It would be interesting to be able to derive an inductive formula for computing the families $\mathcal{C}_{ver}^{(k)}(\mathcal{Y}, V)$ and $\mathcal{C}_{hor}^{(k)}(\mathcal{Y}, V)$. However, such an inductive formula appears to be quite complex.

There are some important statements which can easily be made regarding the distributions $C_{hor}(\mathcal{Y}, V)$ and $C_{ver}(\mathcal{Y}, V)$.

5.10 REMARKS: 1. The generators that we have written for $\mathcal{C}_{ver}^{(k)}(\mathcal{Y}, V)$ and $\mathcal{C}_{hor}^{(k)}(\mathcal{Y}, V)$ are not linearly independent. Thus one should be able to generate these families with fewer calculations than are necessary to compute the generators we give. One way to do this is to choose a Philip Hall basis for $L(\mathbf{X}')$ and compute the image of these brackets under $\text{Ev}(\phi')$. This will work for any given example. However, we are unable to give the general form for the image of a Philip Hall basis under $\text{Ev}(\phi')$.

2. We claim that $C_{hor}(\mathcal{Y}, V)$ is involutive. Let $B'_1, B'_2 \in \text{Br}(\mathbf{X}')$ be brackets which, when evaluated under $\text{Ev}_{0_q}(\phi')$, give vector fields $U_1, U_2 \in \mathcal{C}_{hor}(\mathcal{Y}, V)$. Then the decomposition of B_i given by Lemma 5.4 has the form $B'_i = B_i + \tilde{B}_i$ where $B_i \in \text{Br}_0(\mathbf{X})$ and \tilde{B}_i is a sum of brackets in $\text{Br}_j(\mathbf{X})$ for $j \geq 2$. Therefore, $[B'_1, B'_2] = [B_1, B_2] + B''$ where B'' is a sum of brackets in $\text{Br}_j(\mathbf{X})$ for $j \geq 2$. This shows that $[U_1, U_2] \in \mathcal{C}_{hor}(\mathcal{Y}, V)$. Here we have imposed the condition that brackets in $\text{Br}_{-j}(\mathbf{X})$ are taken to be zero for $j \geq 2$ (see Lemma 5.2).

3. An interesting special case, and one that we shall see in the examples in Section 6, is that when $V = 0$. In this case we have $\mathcal{C}_{ver}(\mathcal{Y}, V) = \overline{\text{Sym}}(\mathcal{Y})$ and $\mathcal{C}_{hor}(\mathcal{Y}, V) = \overline{\text{Lie}}(\overline{\text{Sym}}(\mathcal{Y}))$. This is easily seen in the algorithm by following the path when $\delta_{m+1}(B) = 0$.

4. The calculations of this section and Section 5.2 remain valid if we replace $\text{grad } V$ with an arbitrary vector field on Q .

5.4. Controllability definitions for simple mechanical control systems. It is possible to simply adopt the controllability definitions from nonlinear control theory since our system may be written as a standard control system on TQ . However, since we are dealing with simple control mechanical systems, it is of more interest to us to know what is happening to the *configurations*. A good example of a question of interest in mechanics is ‘‘What is the set of configurations which are reachable from a given configuration if we start at rest?’’ This is in fact exactly the question we pose.

5.11 DEFINITION: A *solution* of (5.1) is a pair, (c, u) , where $c: [0, T] \rightarrow Q$ is a piecewise smooth curve and $u \in \mathcal{U}$ such that (c', u) satisfies the first order control system (5.1). \square

Note that since X_L is a second-order vector field on TQ , every solution of the control system (5.1) will be of the form (c', u) for some curve c on Q . We refer the reader to [Abraham and Marsden 1978] for a discussion of second-order, and particularly Lagrangian, vector fields.

Let $q_0 \in Q$ and let U be a neighbourhood of q_0 . We define

$$\begin{aligned} \mathcal{R}_Q^U(q_0, T) = & \{q \in Q \mid \text{there exists a solution } (c, u) \text{ of (5.1)} \\ & \text{such that } c'(0) = 0_{q_0}, c(t) \in U \text{ for } t \in [0, T], \text{ and } c'(T) \in T_q Q\} \end{aligned}$$

and denote $\mathcal{R}_Q^U(q_0, \leq T) = \bigcup_{0 \leq t \leq T} \mathcal{R}_Q^U(q_0, t)$. Notice that our definitions for reachable configurations do not require us to get to a point in the reachable set at *zero* velocity. They merely ask that we be able to reach that point at *some* velocity. It is, however, required that the initial velocity be zero.

We shall say that $q \in Q$ is an *equilibrium point* for L if $X_L(0_q) = 0$. Let $\mathfrak{E}(L)$ denote the set of equilibrium points for L .

We now introduce our notions of controllability.

5.12 DEFINITION: We shall say that (5.1) is *locally configuration accessible* 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 neighbourhoods U of q_0 and all $0 < t \leq T$. If this holds for any $q_0 \in Q$ then the system is called *locally configuration accessible*.

We say that (5.1) is *small-time locally configuration controllable* (STLCC) at q_0 if it is locally configuration accessible at q_0 and if there exists $T > 0$ such that q_0 is in the interior of $\mathcal{R}_Q^U(q_0, \leq t)$ for every neighbourhood U of q_0 and $0 < t \leq T$. If this holds for any $q_0 \in Q$ then the system is called *small-time locally configuration controllable*.

We shall say that (5.1) is *equilibrium controllable* if, for $q_1, q_2 \in \mathfrak{E}(L)$, there exists a solution (c, u) of (5.1) where $c: [0, T] \rightarrow Q$ is such that $c(0) = q_1$, $c(T) = q_2$ and both $c'(0)$ and $c'(T)$ are zero. \square

Note that these definitions may be made to apply to any control system which evolves on TQ .

5.5. Conditions for controllability of simple mechanical control systems. In [Lewis and Murray 1995] the authors present sufficient conditions for local configuration accessibility in the absence of potential energy. Here, since we have a complete description of $\overline{\text{Lie}}(\mathcal{Y}') \mid Z(TQ)$, we can give stronger results.

5.13 THEOREM: *The control system (5.1) is locally configuration accessible at q if $C_{hor}(\mathcal{U}, V)(q) = T_q Q$.*

Proof: Let C denote the accessibility distribution. Since $C_{hor}(\mathcal{U}, V)(q) \subset C(0_q)$ by Proposition 5.9, and $C_{hor}(\mathcal{U}, V)(q) = T_q Q$ by hypothesis, $Z(TQ)$ must be an integral manifold of C . Let Λ be the maximal integral manifold which contains $Z(TQ)$. Since C is the accessibility distribution, Λ must be invariant under the system (5.1) and the system must be locally accessible when restricted to Λ . Thus the set $\mathcal{R}^{\tilde{U}}(0_q, \leq T)$ is open in Λ for every neighbourhood $\tilde{U} \subset \Lambda$ of 0_q and for every T sufficiently small. Now let U be a neighbourhood of q and define a neighbourhood of 0_q in Λ by $\tilde{U} = \tau_Q^{-1}(U) \cap \Lambda$. The set $\tau_Q(\mathcal{R}^{\tilde{U}}(0_q, \leq T))$ is open in Q for T sufficiently small since τ_Q is an open mapping. This proves the theorem. ■

We also have a partial converse to Theorem 5.13 in the case when there is no potential energy.

5.14 THEOREM: *Suppose that $V = 0$ and (5.1) is locally configuration accessible. Then $C_{hor}(\mathcal{U}, V)(q) = T_q Q$ for q in an open dense subset of Q .*

Proof: First note that if $C_{hor}(\mathcal{U}, V)(q_0) = T_{q_0} Q$ then $C_{hor}(\mathcal{U}, V)(q) = T_q Q$ in a neighbourhood of q_0 . This proves that the set of points q where $C_{hor}(\mathcal{U}, V)(q) = T_q Q$ is open. Now suppose that $C_{hor}(\mathcal{U}, V)(q) \subsetneq T_q Q$ in an open subset U of Q . Then there exists an open subset $\tilde{U} \subset U$ so that $\text{rank}(C_{hor}(\mathcal{U}, V)(q)) = k < n$ for all $q \in \tilde{U}$. However, this contradicts local configuration accessibility. Therefore, there can be no open subset of Q on which $C_{hor}(\mathcal{U}, V)(q) \subsetneq T_q Q$. Thus the set of points q where $C_{hor}(\mathcal{U}, V)(q) = T_q Q$ is dense. This completes the proof. ■

We may also prove an easy statement about STLCC. We need to say a few things about “good” and “bad” symmetric products. Let $\mathbf{Y} = \{X_1, \dots, X_{m+1}\}$ and establish a bijection $\psi: \mathbf{Y} \rightarrow \mathcal{U} \cup \{\text{grad } V\}$ by asking that $\psi(X_a) = Y_a$ for $a = 1, \dots, m$ and $\psi(X_{m+1}) = \text{grad } V$. If $P \in \text{Pr}(\mathbf{Y})$ we shall say that P is **bad** if $\gamma_a(P)$ is even for each $a = 1, \dots, m$. We say that P is **good** if it is not bad. Let S_m denote the permutation group on m symbols. For $\pi \in S_m$ and $P \in \text{Pr}(\mathbf{Y})$ define $\tilde{\pi}(P)$ to be the bracket obtained by fixing X_{m+1} and sending X_a to $X_{\pi(a)}$ for $a = 1, \dots, m$. Now define

$$\rho(P) = \sum_{\pi \in S_m} \tilde{\pi}(P).$$

We may now state the sufficient conditions for STLCC.

5.15 THEOREM: *Suppose that $\mathcal{U} \cup \{\text{grad } V\}$ is such that every bad symmetric product in $\text{Pr}(\mathbf{Y})$ has the property that*

$$\text{Ev}_{0_q}(\psi)(\rho(P)) = \sum_{a=1}^m \xi_a \text{Ev}_{0_q}(\psi)(C_a)$$

where C_a are good symmetric products in $\text{Pr}(\mathbf{Y})$ of lower degree than P and $\xi_a \in \mathbb{R}$ for $a = 1, \dots, m$. Also, suppose that (5.1) is locally configuration accessible at q . Then (5.1) is STLCC at q .

Proof: First recall from the proof of Theorem 5.13 that if (5.1) is locally configuration accessible at q , then $Z(TQ)$ is an integral manifold for the accessibility distribution. We let Λ be the maximal integral manifold for the accessibility distribution which contains $Z(TQ)$. Restricted to Λ , (5.1) is locally accessible. To show that (5.1) is STLCC at q , it clearly suffices to show that (5.1) is STLC at 0_q when restricted to Λ . We do this by showing that (5.1) satisfies the hypotheses of Theorem 4.1 if it satisfies the stated hypotheses on the symmetric products. To do this we shall show that there is a 1–1 correspondence between bad brackets in $\text{Br}(\mathbf{X}')$ and bad symmetric products in $\text{Pr}(\mathbf{Y})$ and good brackets in $\text{Br}(\mathbf{X}')$ and good symmetric products in $\text{Pr}(\mathbf{Y})$.

Suppose that $B' \in \text{Br}(\mathbf{X}')$ is bad. Thus $\delta_a(B')$ is even for $a = 1, \dots, m$ and $\delta_0(B')$ is odd. When we evaluate $\text{Ev}_{0_q}(\phi')(B')$, the only terms that will remain in the decomposition of $\text{Ev}(\phi')(B')$ given by Lemma 5.4 are the terms obtained from brackets in $\mathbf{S}(B')$ which are in $\text{Br}_0(\mathbf{X}) \cup \text{Br}_{-1}(\mathbf{X})$. Since B' is bad, we must have $\delta_a(B)$ even and $\delta_0(B) + \delta_{m+1}(B)$ odd for each $B \in \mathbf{S}(B')$. If $\delta_0(B)$ is odd then $\delta_{m+1}(B)$ must be even. In this case we get $\sum_{a=1}^{m+1} \delta_a(B)$ as even and $\delta_0(B)$ as odd. Thus the only brackets in $\mathbf{S}(B')$ which contribute to $\text{Ev}(\phi')(B')$ must be in $\text{Br}_{-1}(\mathbf{X})$. This will give us a vector in $V_{0_q} TQ$ which comes from a symmetric product which is bad. Now suppose that $\delta_0(B)$ is even for $B \in \mathbf{S}(B')$. Then $\delta_{m+1}(B)$ must be odd. In this case $\sum_{a=1}^{m+1} \delta_a(B)$ is odd and $\delta_0(B)$ is even and again, the only brackets in $\mathbf{S}(B')$ which contribute to $\text{Ev}(\phi')(B')$ must be in $\text{Br}_{-1}(\mathbf{X})$. We then conclude that $\text{Ev}_{0_q}(\phi')(B')$ must be of the form $(\text{Ev}_q(\psi)(P))^{\text{lift}}$ where $P \in \text{Pr}(\mathbf{Y})$ is bad.

Now suppose that $B' \in \text{Br}(\mathbf{X}')$ is good. It is clear that if $\delta_a(B')$ is odd for any $a = 1, \dots, m$ then B' cannot give rise to a bad symmetric product. Thus we may suppose that $\delta_a(B')$ is even for each $a = 0, \dots, m$. Now let's look at what the brackets look like from $\mathbf{S}(B')$ which contribute to $\text{Ev}(\phi')(B')$. Let B be such a bracket. We must have $\delta_a(B)$ even for $a = 1, \dots, m$ and $\delta_0(B) + \delta_{m+1}(B)$ even. If $\delta_0(B)$ is odd then $\delta_{m+1}(B)$ must be odd. Since B is primitive this means that $\sum_{a=1}^{m+1} \delta_a(B)$ and $\delta_0(B)$ are odd. Therefore, B must be in $\text{Br}_0(\mathbf{X})$. Now suppose that $\delta_0(B)$ is even. Then $\delta_{m+1}(B)$ must also be even. Thus $\sum_{a=1}^{m+1} \delta_a(B)$ and $\delta_0(B)$ are even and so $B \in \text{Br}_0(\mathbf{X})$. Therefore, good brackets from $\text{Br}(\mathbf{X}')$ do not generate any bad symmetric products. ■

Since the system restricted to the integral manifold Λ in the proof of the above theorem is STLC, the hypotheses of the theorem imply more than STLCC. In fact, the following corollary is easily seen to be true.

5.16 COROLLARY: *Suppose that the hypotheses of Theorem 5.15 hold for each $q \in Q$. Then the system (5.1) is equilibrium controllable.*

5.17 REMARKS: 1. We have shown that it is not necessary to be able to generate all directions on TQ to obtain controllability in the configuration variables. Indeed, the only vertical directions we generate are $C_{ver}(\mathcal{U}, V)$ which need not span $V_{0_q} TQ$. This means that the notion of configuration controllability is genuinely weaker than are the standard notions of controllability if we are to simply regard the system (5.1) as a typical nonlinear control system.

2. The result 5.16 may be made even stronger if we allow a point $q \in Q$ to be an equilibrium point if $\text{grad } V(q)$ is in the span of the inputs at q .

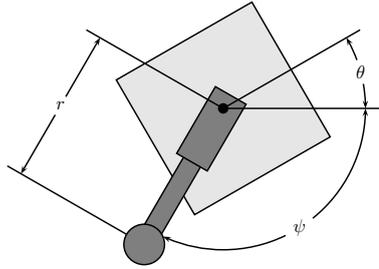


FIGURE 3. The robotic leg

6. Examples of mechanical control systems

In this section we present some examples. The examples are rather simple and are intended to illustrate the concepts put forward by the theory. One of the advantages of the conditions for local configuration accessibility given in Theorem 5.13 is that it lends itself to symbolic computation. Indeed, a *Mathematica* package was written to facilitate the computations in this section.

6.1. The robotic leg. This example, although simple, exhibits much of the subtle behaviour that makes the study of mechanical systems interesting. The example is a rigid body with inertia J which is pinned to ground at its centre of mass. The body has attached to it an extensible massless leg and the leg has a point mass with mass m at its tip. The coordinate θ will describe the angle of the body, and ψ will describe the angle of the leg from an inertial reference frame. The coordinate r will describe the extension of the leg. Thus the configuration space for this problem is $Q = \mathbb{T}^2 \times \mathbb{R}^+$. See Figure 3. In the coordinates (θ, ψ, r) the Riemannian metric for the robotic leg is

$$g = Jd\theta \otimes d\theta + mr^2d\psi \otimes d\psi + mdr \otimes dr,$$

the input one-forms are $F^1 = d\theta - d\psi$ and $F^2 = dr$, and the potential energy function is zero. We may compute the input vector fields to be

$$Y_1 = \frac{1}{J} \frac{\partial}{\partial \theta} - \frac{1}{mr^2} \frac{\partial}{\partial \psi}, \quad Y_2 = \frac{1}{m} \frac{\partial}{\partial r}.$$

Since there is no potential energy present, the distribution $C_{hor}(\mathcal{D}, V)$ is simply generated by the vector fields $\text{Lie}(\text{Sym}(\mathcal{D}))$.

We will find the following computations to be sufficient:

$$\begin{aligned} \langle Y_1 : Y_1 \rangle &= -\frac{2}{m^2 r^3} \frac{\partial}{\partial r}, & \langle Y_1 : Y_2 \rangle &= 0, & \langle Y_2 : Y_2 \rangle &= 0, \\ [Y_1, Y_2] &= -\frac{2}{m^2 r^3} \frac{\partial}{\partial \psi}, & [Y_1, \langle Y_1 : Y_1 \rangle] &= \frac{4}{m^3 r^6} \frac{\partial}{\partial \psi}. \end{aligned}$$

Inputs	Locally configuration accessible?	Satisfies sufficient conditions for STLCC?	STLCC?
Y_1 (torque)	yes	no	no
Y_2 (extension)	no	no	no
Y_1 and Y_2	yes	yes	yes

TABLE 1. Controllability results for the robotic leg. The first column displays which inputs are present, the second column indicates whether the system is locally configuration accessible with these inputs, the third column indicates whether the system with these inputs satisfies the sufficient conditions of Theorem 5.15 for STLCC, and the last column indicates whether the system with these inputs is actually STLCC.

The controllability results for the robotic leg are displayed in Table 1.

6.1 REMARKS: 1. The linearisation of this system at points of zero velocity is not controllable with any combination of inputs, so the controllability does not follow from linear results.

2. When only the input Y_2 is present the equations are

$$\begin{aligned} \ddot{r} - r\dot{\psi}^2 &= \frac{1}{m} u_1 \\ \ddot{\theta} &= 0 \\ \ddot{\psi} + \frac{2}{r} \dot{r}\dot{\psi} &= 0. \end{aligned}$$

Note that when the initial velocity is zero, the top equation decouples from the bottom two equations. Physically this means that we are simply moving the leg back and forth with no effect on the configuration of the body since the initial velocity is zero.

3. Although the system only violates the *sufficient* conditions for STLCC with the input Y_1 , one may easily see by looking at the r -component of the equations of motion that the system is, in fact, not STLCC. The reason for this is that, since $\ddot{r} \geq 0$, r will always increase no matter what happens to the other variables. Thus our initial configuration will never be in the interior of the set of reachable configurations.

6.2. The forced planar rigid body. In this section we study the planar rigid body discussed in the introduction with various combinations of forces and torques. The configuration space for the system is the Lie group $SE(2)$. To establish the correspondence between the configuration of the body and $SE(2)$, fix a point $O \in \mathbb{R}^2$ and let $\{e_1 = \frac{\partial}{\partial x}, e_2 = \frac{\partial}{\partial y}\}$ be the standard orthonormal frame at that point. Let $\{f_1, f_2\}$ be an orthonormal frame attached to the body at its centre of mass. The configuration of the body is determined by the element $g \in SE(2)$ which maps the point O with its frame $\{e_1, e_2\}$ to the position,

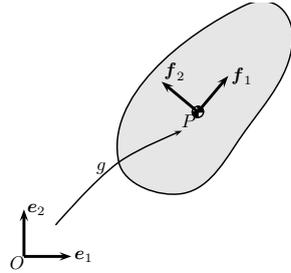


FIGURE 4. The configuration of a planar rigid body as an element of $SE(2)$

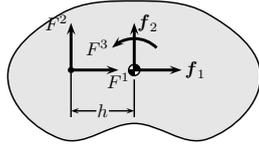


FIGURE 5. Positions for application of forces on a planar rigid body after simplifying assumptions

P , of the centre of mass of the body with its frame $\{\mathbf{f}_1, \mathbf{f}_2\}$. See Figure 4. The inputs for this problem consist of forces applied at an arbitrary point and a torque about the centre of mass. Without loss of generality (by redefining our body reference frame $\{\mathbf{f}_1, \mathbf{f}_2\}$) we may suppose that the point of application of the force is a distance h along the \mathbf{f}_1 body-axis from the centre of mass. The situation is illustrated in Figure 5.

With this convention fixed, we shall use coordinates (x, y, θ) for the planar rigid body where (x, y) describe the position of the centre of mass and θ describes the orientation of the frame $\{\mathbf{f}_1, \mathbf{f}_2\}$ with respect to the frame $\{\mathbf{e}_1, \mathbf{e}_2\}$. In these coordinates, the Riemannian metric for the system is

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

Here m is the mass of the body and J is its moment of inertia about the centre of mass. The inputs are described by the one-forms

$$F^1 = \cos \theta dx + \sin \theta dy, \quad F^2 = -\sin \theta dx + \cos \theta dy - h d\theta, \quad F^3 = d\theta$$

from which we compute the input vector fields as

$$Y_1 = \frac{\cos \theta}{m} \frac{\partial}{\partial x} + \frac{\sin \theta}{m} \frac{\partial}{\partial y},$$

$$Y_2 = -\frac{\sin \theta}{m} \frac{\partial}{\partial x} + \frac{\cos \theta}{m} \frac{\partial}{\partial y} - \frac{h}{J} \frac{\partial}{\partial \theta}, \quad Y_3 = \frac{1}{J} \frac{\partial}{\partial \theta}.$$

Again, as with the robotic leg, there is no potential energy so the distribution $C_{hor}(\mathcal{D}, V)$ may be computed by calculating $\overline{\text{Lie}}(\overline{\text{Sym}}(\mathcal{D}))$.

Inputs	Locally configuration accessible?	Satisfies sufficient conditions for STLCC?	STLCC?
Y_1 (force at CM)	no	no	no
Y_2 (force \perp to CM)	yes	no	no
Y_3 (torque)	no	no	no
Y_1 and Y_2	yes	yes	yes
Y_1 and Y_3	yes	yes	yes
Y_2 and Y_3	yes	no	yes

TABLE 2. Controllability results for the planar rigid body. The first column displays which inputs are present, the second column indicates whether the system is locally configuration accessible with these inputs, the third column indicates whether the system with these inputs satisfies the sufficient conditions of Theorem 5.15 for STLCC, and the last column indicates whether the system with these inputs is actually STLCC.

The following computations are sufficient to obtain the results we desire:

$$\langle Y_1 : Y_1 \rangle = 0, \quad \langle Y_1 : Y_2 \rangle = \frac{h \sin \theta}{mJ} \frac{\partial}{\partial x} - \frac{h \cos \theta}{mJ} \frac{\partial}{\partial y},$$

$$\langle Y_1 : Y_3 \rangle = -\frac{\sin \theta}{mJ} \frac{\partial}{\partial x} + \frac{\cos \theta}{mJ} \frac{\partial}{\partial y}, \quad \langle Y_2 : Y_2 \rangle = \frac{2h \cos \theta}{mJ} \frac{\partial}{\partial x} + \frac{2h \sin \theta}{mJ} \frac{\partial}{\partial y},$$

$$\langle Y_2 : Y_3 \rangle = -\frac{\cos \theta}{mJ} \frac{\partial}{\partial x} - \frac{\sin \theta}{mJ} \frac{\partial}{\partial y}, \quad \langle Y_3 : Y_3 \rangle = 0,$$

$$\langle Y_1, Y_2 \rangle = -\frac{h \sin \theta}{mJ} \frac{\partial}{\partial x} + \frac{h \cos \theta}{mJ} \frac{\partial}{\partial y}, \quad \langle Y_1, Y_3 \rangle = \frac{\sin \theta}{mJ} \frac{\partial}{\partial x} - \frac{\cos \theta}{mJ} \frac{\partial}{\partial y},$$

$$\langle Y_2, Y_3 \rangle = \frac{\cos \theta}{mJ} \frac{\partial}{\partial x} + \frac{\sin \theta}{mJ} \frac{\partial}{\partial y}, \quad \langle Y_2, \langle Y_2 : Y_2 \rangle \rangle = \frac{2h^2 \sin \theta}{mJ^2} \frac{\partial}{\partial x} - \frac{2h^2 \cos \theta}{mJ^2} \frac{\partial}{\partial y}.$$

With the computations done, we may proceed to determine configuration controllability for the planar rigid body with various combinations of inputs. The results are displayed in Table 2.

6.2 REMARKS: 1. The linearisation of this system around points of zero velocity is not controllable so the cases where the system is STLCC do not follow from the linear calculations.

2. In this example, in the cases when the system fails to satisfy the sufficient conditions for STLCC of Theorem 5.15, we are not able to say whether the system is, in fact, not STLCC. In fact, when the inputs Y_2 and Y_3 are present, even though the system does not satisfy the sufficient conditions of Theorem 5.15, it is easy to see that it is STLCC. Recent work, beyond the scope of this paper, shows that when only the input Y_2 is present, the system is not STLCC (see [Lewis 1997]).

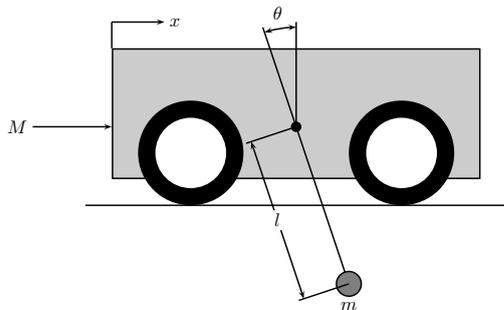


FIGURE 6. Pendulum suspended from a cart

3. In the case when only the input Y_1 is present, it is illustrative to represent the equations in the coordinates $(\xi, \eta, \psi) = (x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, \theta)$. In these coordinates the equations have the form

$$\begin{aligned}\ddot{\xi} - 2\dot{\eta}\dot{\psi} - \xi\dot{\psi}^2 &= \frac{1}{m}u \\ \dot{\eta} + 2\dot{\xi}\dot{\psi} - \eta\dot{\psi}^2 &= 0 \\ \ddot{\psi} &= 0.\end{aligned}$$

Notice that the top equation decouples from the last two equations when the initial velocity is zero. Since Y_1 is directed towards the centre of mass, applying this input will cause the body to move in this direction and none of the other degrees of freedom are affected.

4. In the case when the input Y_3 is present, the equations have the form

$$\begin{aligned}\ddot{\theta} &= \frac{1}{J}u_3 \\ \ddot{x} &= 0 \\ \ddot{y} &= 0.\end{aligned}$$

Again, the top equation decouples from the bottom two equations. This time the coupling is true for all initial velocities, and not just zero initial velocity. In this case we see that the input simply causes a rotation of the body about its centre of mass. The position of the centre of mass is not affected if the initial velocity is zero.

6.3. The pendulum on a cart. To illustrate the effects of potential energy, consider the problem of a pendulum suspended from a cart. The configuration manifold for the system

is $Q = \mathbb{R} \times \mathbb{S}^1$. As coordinates we shall use (x, θ) as shown in Figure 6. In this case the Riemannian metric for the system is

$$g = (M + m)dx \otimes dx + ml \cos \theta dx \otimes d\theta + ml \cos \theta d\theta \otimes dx + ml^2 d\theta \otimes d\theta.$$

Here M is the mass of the cart and m is the mass of the pendulum. The potential energy is

$$V = ma_g l(1 - \cos \theta)$$

where a_g is the acceleration due to gravity. The input is given by the one-form

$$F^1 = dx.$$

The input vector field is then readily computed to be

$$Y_1 = \frac{ml^2}{m^2l^2 + Mml^2 - m^2l^2 \cos^2 \theta} \frac{\partial}{\partial x} + \frac{ml \cos \theta}{m^2l^2 + Mml^2 - m^2l^2 \cos^2 \theta} \frac{\partial}{\partial \theta}.$$

To compute $C_{hor}(\mathcal{Y}, V)$ we need the following computations:

$$\begin{aligned}\langle Y_1 : Y_1 \rangle &= \frac{16m \cos^2 \theta \sin \theta}{l(m + 2M - m \cos 2\theta)^3} \frac{\partial}{\partial x} + \frac{8(M + m) \sin \theta}{l^2(m \cos 2\theta - m - 2M)^3} \frac{\partial}{\partial \theta}, \\ \langle Y_1 : \text{grad } V \rangle &= \frac{4a_g m \cos \theta (m - m \cos 2\theta - 2M \cos 2\theta)}{l(m \cos 2\theta - m - 2M)^3} \frac{\partial}{\partial x} + \\ &\quad \frac{4a_g (2M^2 \cos 2\theta + 3Mm \cos 2\theta + m^2 \cos 2\theta - Mm - m^2)}{l^2(m \cos 2\theta - m - 2M)^3} \frac{\partial}{\partial \theta}.\end{aligned}$$

Note that at all points $q \in Q$ except those where $\theta \in \{0, \pi\}$, the vector fields $\{Y_1, \langle Y_1 : Y_1 \rangle\}$ generate the tangent space at q . This means that the system is locally configuration accessible at these points. Also, at these points the bad symmetric product $\langle Y_1 : Y_1 \rangle$ is not a multiple of Y_1 so the system may not be STLCC at these points. At points where $\theta \in \{0, \pi\}$, the vector fields $\{Y_1, \langle Y_1 : \text{grad } V \rangle\}$ span $T_q Q$ and so the system is also locally configuration accessible at these points. Most importantly, however, the bad symmetric product vanishes at these two points so the system is STLCC at these equilibria. This must be so as, at these two points, the linearised system is controllable.

7. Conclusions and future work

In this paper we have outlined what we regard as a *beginning* of a thorough program for analysis and synthesis for simple mechanical control systems. The first part of such a program is to determine the pertinent versions of controllability (local configuration accessibility and STLCC) and determine algebraic tests for these notions of controllability. The conditions we present for checking our versions of controllability involve only computations on the configuration space. In determining these conditions, the symmetric product proved to play an important rôle. As we have presented it, the symmetric product is a useful computational tool. Our recent work provides a fairly complete description of the geometric rôle of the symmetric product in the control of mechanical systems. This will be the subject of an upcoming paper.

In the examples in Section 6 some interesting circumstances may be observed. The most interesting of these is a comparison of the robotic leg with input Y_2 and the planar rigid body with the inputs Y_2 and Y_3 . In the former case the system does not satisfy the sufficient conditions for STLCC and may be shown to indeed not be STLCC. However, in the latter case, even though the sufficient conditions for STLCC are not met, the system *is* STLCC. It would be interesting to better understand why this happens, and perhaps arrive at a stronger condition for STLCC.

Finally we mention that, from a practical point of view, perhaps the most useful contribution is that of the notion, mentioned in Section 5.4, of equilibrium controllability. If a system satisfies the hypotheses of Theorem 5.15 at each configuration, it would be interesting to determine a means of generating paths which connect points in the configuration manifold at zero velocity. Such an algorithm may involve a deeper understanding of the symmetric product.

In summary, we feel that this paper provides an effective initial understanding of mechanical control systems, and we hope that it will prove to be a useful foundation for further work in the area of mechanical control theory.

Acknowledgements

We would like to thank Jerry Marsden and Jim Ostrowski for helpful conversations. The anonymous reviewers were also very helpful in improving the presentation of the paper during the review process.

References

- Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics*, second edition, Addison Wesley, Reading, MA, ISBN 0-8053-0102-X.
- Bloch, A. M. and Crouch, P. E. [1992] *Kinematics and dynamics of nonholonomic control systems on Riemannian manifolds*, in *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 1–5, Institute of Electrical and Electronics Engineers, Tucson, AZ.
- Bloch, A. M., Reyhanoglu, M., and McClamroch, N. H. [1992] *Control and stabilization of nonholonomic dynamic systems*, Institute of Electrical and Electronics Engineers. Transactions on Automatic Control, **37**(11), 1746–1757.
- Brockett, R. W. [1977] *Control theory and analytical mechanics*, in *Geometric Control Theory*, C. Martin and R. Hermann, editors, pages 1–48, Math Sci Press, 53 Jordan Road, Brookline, Massachusetts.
- Crouch, P. E. [1981] *Geometric structures in systems theory*, Institution of Electrical Engineers. Proceedings. D. Control Theory and Applications, **128**(5), 242–252.
- Hermes, H. [1982] *On local controllability*, SIAM Journal on Control and Optimization, **20**(2), 211–220.
- Jacobson, N. [1962] *Lie Algebras*, number 10 in Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers, New York.

- Klingenberg, W. [1982] *Riemannian Geometry*, number 1 in Studies in Mathematics, Walter de Gruyter, Berlin-New York, ISBN 3-11-008673-5.
- Lewis, A. D. [1995] *Aspects of Geometric Mechanics and Control of Mechanical Systems*, Ph.D. thesis, California Institute of Technology, Pasadena, California, USA. URL: <http://www.cds.caltech.edu/>
- [1997] *Local configuration controllability for a class of mechanical systems with a single input*, in *Proceedings of the European Control Conference*, Brussels, Belgium.
- Lewis, A. D. and Murray, R. M. [1995] *Configuration controllability of a class of mechanical systems*, in *Proceedings of the 34th IEEE Conference on Decision and Control*, pages 1–5, Institute of Electrical and Electronics Engineers, New Orleans, LA.
- San Martin, L. and Crouch, P. E. [1984] *Controllability on principal fibre bundles with compact structure group*, Systems & Control Letters, **5**(1), 35–40.
- Serre, J.-P. [1992] *Lie Algebras and Lie Groups*, number 1500 in Lecture Notes in Mathematics, Springer-Verlag, New York-Heidelberg-Berlin.
- Sussmann, H. J. [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.