CDS Technical Report 95-015
California Institute of Technology

Configuration Controllability of Simple Mechanical Control Systems*

Andrew D. Lewis and Richard M. Murray

Division of Engineering and Applied Science 104–44,
California Institute of Technology,
Pasadena, CA 91125

1 June, 1995

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 also derived. This condition involves a new 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.

1. Introduction

The class of mechanical control systems is a large and interesting subset of all control systems. In this paper we present some basic notions for studying a subset of mechanical control systems which we call "simple mechanical control systems." These control systems are characterised by their Lagrangian being "kinetic energy minus potential energy." The main point of interest is that the definitions of controllability we propose involve only the configuration variables as it is these which are often interesting in mechanical systems. We are then able to derive computable conditions for our versions of controllability which involve a new object, the *symmetric product*, which may be defined on a Riemannian manifold. One of the versions of controllability is what we call "equilibrium controllability" which involves being able to steer between any two equilibrium points for the system. Using sufficient conditions for small-time local controllability by Sussmann [1987] we are able to derive sufficient conditions for this version of controllability.

Much of the previous work in the area of mechanical control systems has relied on specific structure of these systems. 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 are able to use the result of San Martin and Crouch [1984] to arrive at a controllability result. Mechanical systems

*Submitted to *Siam Journal on Control and Optimization.*

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 set of constraint forces. Lewis [1995] proves a result of this type for a general class of constraints and Lagrangians. 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 a control theoretic tool bag 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.

We present a simple example in Section 2 which is meant to motivate the need to investigate mechanical control systems in some detail. The example also serves as a guide for some of the calculations which will be done in Section 5.

Since a precise statement of our results requires some background, mathematical preliminaries are presented in Section 3. In this section, the most important and new concept is that of a *symmetric product*. This is presented in algebraic form in Section 3.3 and in geometric form in Section 3.4.

In Section 5 we present the main results of Lewis [1995] and in Section 6 some illustrative examples are given.

2. A Motivating Example

In this section we describe in some detail a "simple mechanical control system" which illustrates the need to refine the treatment of mechanical systems in nonlinear control theory. In particular, this example demonstrates that the nonlinear control calculations which one often performs do not provide a satisfactory resolution to the controllability problem for all mechanical systems. We propose that a weaker notion of controllability may be useful. We also do some computations with this example which hint at how the general calculations will proceed in Section 5.

A Description of the System

The example we consider is a rigid body with inertia J which is pinned to ground at its centre of mass. This example was first presented by Li, Montgomery, and Raibert [1989].[1] 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 1. The Lagrangian is

$$L = \frac{1}{2}J\dot{\theta}^2 + \frac{1}{2}m(\dot{r}^2 + r^2\dot{\psi}^2).$$

[1] In their paper the example considered is actually in free flight. We present the robotic leg fixed to a point as this simplifies the analysis, but removes none of the essential structure.

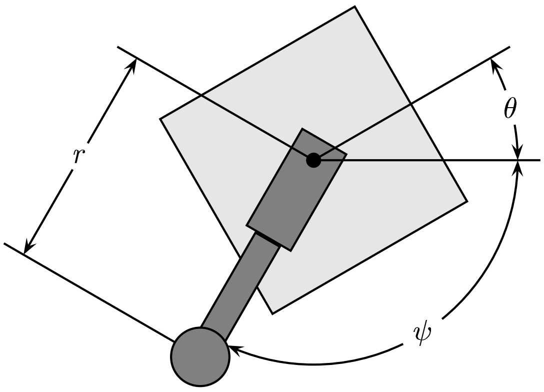

FIGURE 1. The robotic leg

If we consider forces applied in the $(\theta - \psi)$ and r-directions, Lagrange's equations are

$$J\ddot{\theta} = u_1 \tag{2.1a}$$
$$mr^2\ddot{\psi} + 2mr\dot{r}\dot{\psi} = -u_1 \tag{2.1b}$$
$$m\ddot{r} - mr\dot{\psi}^2 = u_2. \tag{2.1c}$$

Contradictory Controllability Results

We may rewrite Lagrange's equations as a vector field on TQ in the form

$$X_L = Z_g = v_\theta \frac{\partial}{\partial \theta} + v_\psi \frac{\partial}{\partial \psi} + v_r \frac{\partial}{\partial r} - \frac{2v_r v_\psi}{r}\frac{\partial}{\partial v_\psi} + r v_\psi^2 \frac{\partial}{\partial v_r}.$$

The control vector fields on TQ may be computed as the *vertical lifts* (see Section 5) of the vector fields

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

on Q. The distribution calculations may be performed to obtain the accessibility distribution as

$$C(\theta, \psi, r, v_\theta, v_\psi, v_r) = \left\langle 2mrv_\psi \frac{\partial}{\partial v_\theta} - J\frac{\partial}{\partial r}, mr^2\frac{\partial}{\partial v_\theta} - J\frac{\partial}{\partial v_\psi}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \psi}, \frac{\partial}{\partial v_r} \right\rangle_{\mathbb{R}}.$$

Since this distribution does not span TQ, we conclude that the system is not locally accessible. Nevertheless, it is possible to steer the system from one configuration to another. Indeed we have the following result, some of which was proven by Murray and Sastry [1993].

CLAIM: *Select two configurations, $q_1 = (\theta_1, \psi_1, r_1)$ and $q_2 = (\theta_2, \psi_2, r_2)$. Suppose that the system starts at rest in configuration q_1. Then there exists inputs u_1, u_2 which steer the system to rest at q_2.*

Proof: We first note that the inputs leave the total angular momentum,

$$\mu = J\dot{\theta} + mr^2\dot{\psi},$$

of the system conserved. Thus, when we start at rest at q_1, all consequent motions of the system will have zero angular momentum. This may be thought of as imposing a constraint given by

$$J\dot{\theta} + mr^2\dot{\psi} = 0. \tag{2.2}$$

Let us first answer the question: How many configurations are accessible from q_1 along paths which preserve zero angular momentum? Let D be the distribution defined by (2.2). This distribution has dimension two and the Lie bracket between any two basis vector fields for $\mathcal{D}$ will not lie in $\mathcal{D}$. This shows that D is controllable. Therefore, from q_1 it is possible to reach any other configuration while maintaining the constraint of zero angular momentum. To prove the claim, we need to show that all motions of the system which preserve zero angular momentum are realisable using suitable inputs, u_1, u_2. Let c be a path in Q which satisfies the constraint (2.2) and which connects q_1 with q_2. We may suppose that c is parameterised so that we start at rest at q_1 and end at rest at q_2. From (2.1c) and (2.1a) we immediately have $u_2 = m\ddot{r} - mr\dot{\psi}^2$ and $u_1 = J\ddot{\theta}$. We need only show that, so defined, u_1 satisfies (2.1b). From (2.2) we have

$$J\ddot{\theta} = -mr^2\ddot{\psi} - 2mr\dot{r}\dot{\psi}.$$

Therefore,

$$mr^2\ddot{\psi} + 2mr\dot{r}\dot{\psi} = -u_1$$

which is simply (2.1b). This completes the proof. ■

A Closer Look at the Distribution Calculations

The above claim indicates that we would like to be able to consider this problem controllable in some sense. Let us try to understand how we might do this by taking a closer look at the distribution computations which yield the accessibility distribution. Since we are interested in describing the set of points reachable from initial conditions with zero velocity, we will evaluate all brackets on the zero section of TQ which we shall denote by $Z(TQ)$. We denote the zero tangent vector at $q \in Q$ by 0_q. We may compute

$$\begin{aligned}
[Y_1^{lift}, Y_2^{lift}] &= 0 \\
[Z_g, Y_1^{lift}](0_q) &= -Y_1(q) \\
[Z_g, Y_2^{lift}](0_q) &= -Y_2(q) \\
[Y_1^{lift}, [Z_g, Y_1^{lift}]] &= -\frac{2}{m^2 r^3} \frac{\partial}{\partial v_r} \\
[Y_1^{lift}, [Z_g, Y_2^{lift}]] &= 0 \\
[Y_2^{lift}, [Z_g, Y_2^{lift}]] &= 0 \\
[Z_g, [Z_g, Y_1^{lift}]](0_q) &= 0 \\
[Z_g, [Z_g, Y_2^{lift}]](0_q) &= 0 \\
[Z_g, [Y_1^{lift}, [Z_g, Y_1^{lift}]]](0_q) &= -\frac{2}{m^2 r^3} \frac{\partial}{\partial r} \\
[[Z_g, Y_1^{lift}], [Z_g, Y_2^{lift}]](0_q) &= [Y_1, Y_2](q).
\end{aligned}$$

(See (5.3) for a definition of X^{lift}.) These turn out to be the only interesting brackets for the robotic leg. If we examine these bracket calculations, we make the following informal observations.

1. The brackets between the input vector fields are zero.
2. The brackets in which the drift vector field appears the same number of times as the control vector fields give brackets in the "q-direction" when we evaluate them at zero velocity.
3. The brackets which contain the control vector fields one more time than the drift vector field are vertical lifts of vector fields on Q.
4. The brackets which contain the drift vector field more often than the control vector fields are zero when evaluated at points of zero velocity.

These observations suggest what may happen with general systems of the form (5.4). In Section 5 we formally go through the calculations needed to prove the form of the accessibility distribution for these systems when restricted to the zero section of TQ. The reader may wish to refer back to the above bracket calculations at various times during the general exposition.

General Considerations

With the information given in this example, we are in a position to give some preliminary general results. 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 modeled by vector fields $\mathcal{Y} = \{Y_1, \dots, Y_m\}$. 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 $\mathfrak{X}(Q)$ denote the set of vector fields on Q, and if $\mathcal{V} \subset \mathfrak{X}(Q)$, we denote by $\overline{\mathrm{Sym}}(\mathcal{V})$ the set of vector fields on Q obtain by taking iterated symmetric products of vector fields from $\mathcal{V}$. The usual involutive closure of $\mathcal{V}$ will be denoted $\overline{\mathrm{Lie}}(\mathcal{V})$. We shall say that s symmetric product from $\overline{\mathrm{Sym}}(\mathcal{Y})$ is *bad* if it contains an even number of each of the vector fields in $\mathcal{Y}$. Otherwise we shall call a symmetric product from $\overline{\mathrm{Sym}}(\mathcal{Y})$ *good*.

Notice that with the Lagrangian given by just kinetic energy, all configurations with zero velocity are equilibrium point 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 $\mathcal{Y} = \{Y_1, \ldots, Y_m\}$. Then*

i) the system is locally configuration equilibrium accessible at q if the distribution defined by $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))$ 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 order.

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

3. Lie Algebras and Symmetric Algebras

When studying control systems it is useful to have in hand some basic notions of Lie algebras. In Section 5 we will need the notion of what we shall call a *symmetric algebra*. In order to be precise about how we define certain types of brackets and symmetric products, we need to introduce free Lie algebras and symmetric products. These also turn out to be convenient for describing the *involutive closure* and the *symmetric closure*.

3.1. Free Lie Algebras Our discussion of free Lie algebras is an abbreviated version of that found in Serre [1992]. We shall not be fully precise here. See Lewis [1995] for details.

We denote by $A(X)$ the algebra of associative but not necessarily commutative products of indeterminants from the set X. We will suppose the coefficients to be in $\mathbb{R}$ although arbitrary definitions are possible over a commutative ring with unit. To construct the free Lie algebra generated by X, let I be the two-sided ideal of $A(X)$ generated by elements of the form $a \cdot a$ and $a \cdot (b \cdot c) + c \cdot (a \cdot b) + b \cdot (c \cdot a)$ for $a, b, c \in A(X)$. The *free Lie algebra* generated by X is the quotient algebra, $L(X) = A(X)/I$. The inherited product on $L(X)$ is typically denoted by $[\cdot, \cdot]$. We denote by $\mathrm{Br}(X)$ the subset of $L(X)$ containing products of elements in X. This subset generates $L(X)$ as a $\mathbb{R}$-vector space. However, it is not a linearly independent subset since, for example, $[u, v] = -[u, v]$ for each $u, v \in L(X)$. Below we construct a set of generators which is contained in $\mathrm{Br}(X)$.

It may be shown that there is an algebra homomorphism from $L(X)$ to $T(\mathbb{R}^X)$, the tensor algebra of the free vector space generated by X. Serre [1992] shows that the image of $L(X)$ under this homomorphism is a subalgebra of the tensor algebra. We shall use this fact when we discuss representing free Lie algebras in the Lie algebra of vector fields in Section 3.2.

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 whose lengths are one. All such elements $u_{i_1 \cdots i_m}$, $i_a \in \{1, 2\}$, shall be called *components* of u.

Of special interest to us is the case where the set X is finite. We shall denote $\boldsymbol{X} = \{X_0, \ldots, X_l\}$ as a finite set with $l+1$ elements. In this case we develop some extra notation. Let $B \in \mathrm{Br}(\boldsymbol{X})$. We define $\delta_a(B)$ to be the number of times the element X_a occurs in B for $a = 0, \ldots, l$. The *degree* of B is the sum the δ_a's.

We will find it helpful to write down a generating set for $L(X)$. It is possible to determine linearly independent generating sets, called *Philip Hall bases* in the literature (see Serre [1992]). However, we shall not need such sophisticated techniques and it is good enough to just determine a generating set without the condition that it be linearly independent.

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

$$[X_k, [X_{k-1}, [\cdots, [X_2, X_1] \cdots]]] \tag{3.1}$$

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

Proof: Denote by $\bar{L}(X)$ the subspace of $L(X)$ generated by brackets of the form (3.1). It is clear that $\bar{L}(X) \subset L(X)$ by definition. Also, $X \subset \bar{L}(X)$. Thus, to show that $\bar{L}(X) = L(X)$, we need only show that $\bar{L}(X)$ is a subalgebra of $L(X)$ since $L(X)$ is the smallest subalgebra containing X. Note that k in (3.1) is the *degree* of the expression. Now consider two such expressions of degree j and l,

$$U = [U_j, [U_{j-1}, [\cdots, [U_2, U_1] \cdots]]] \tag{3.2a}$$

$$V = [V_l, [V_{l-1}, [\cdots, [V_2, V_1] \cdots]]]. \tag{3.2b}$$

We shall prove by induction that $[U, V] \in \bar{L}(X)$ for any j and l. Note that $[U, V] \in \bar{L}(X)$ for all V and l, and for $j = 1$. Now suppose this is true for $j = 1, \ldots, k$. Then, taking $j = k+1$ in (3.2a), we have

$$[U, V] = [[U_{k+1}, U^1], V]$$

where $U^1 = [U_{j-1}, [\cdots, [U_2, U_1] \cdots]]$. By the Jacobi identity we have

$$[[U_{k+1}, U^1], V] + [[V, U_{k+1}], U^1] + [[U^1, V], U_{k+1}] = 0.$$

This gives

$$[U, V] = [U^1, [V, U_{k+1}]] + [U_{k+1}, [U^1, V]].$$

By the induction hypothesis, $[U^1, [U_{k+1}, V]] \in \bar{L}(X)$ since the degree of U^1 is k. Also $[U^1, V] \in \bar{L}(X)$ so the second term is in $\bar{L}(X)$. Thus $\bar{L}(X)$ is a subalgebra and hence $\bar{L}(X) = L(X)$. ∎

3.2. Distributions Generated by a Family of Vector Fields A *family of vector fields* on a differentiable manifold M is simply a subset $\mathcal{V} \subset \mathfrak{X}(M)$. Given a family of vector fields $\mathcal{V}$, we may define a distribution on M by

$$D_{\mathcal{V}}(x) = \langle X(x) \mid \quad X \in \mathcal{V} \rangle_{\mathbb{R}} .$$

Since $\mathfrak{X}(M)$ is a Lie algebra, we may ask for the smallest Lie subalgebra of $\mathfrak{X}(M)$ which contains a family of vector fields $\mathcal{V}$. This will be the set of vector fields on M generated by repeated Lie brackets of elements in $\mathcal{V}$. It is most convenient to describe this subalgebra using the ideas from free Lie algebras presented in Section 3.1.

Let X be a set which is bijective to $\mathcal{V}$. Thus each element of X is in 1–1 correspondence with a vector field in $\mathcal{V}$. Recall that $T(\mathbb{R}^X)$ is the tensor algebra of the free vector space on X. Thus each element of $T(\mathbb{R}^X)$ is an associative, but not necessarily commutative, product of finite linear combinations of elements from X. Given a bijection $\phi\colon X \to \mathcal{V}$, we may define a $\mathbb{R}$-algebra homomorphism from $T(\mathbb{R}^X)$ to $\mathfrak{X}(M)$ by "plugging in" the vector field $\phi(u)$ for the element $u \in X$ in expressions in $T(\mathbb{R}^X)$. The map is explicitly given by

$$\begin{aligned} \mathrm{Ev}(\phi)\colon\ & T(\mathbb{R}^X) \to \mathfrak{X}(M) \\ & u_1 \otimes \cdots \otimes u_k \mapsto \phi(u_1) \circ \cdots \circ \phi(u_k). \end{aligned}$$

Here we are using the algebra structure on $\mathfrak{X}(M)$ given by its being the set of derivations on $C^\infty(M)$, the ring of smooth functions on M. Since elements of $L(X)$ may be regarded naturally as elements of $T(\mathbb{R}^X)$, the map $\mathrm{Ev}(\phi)$ restricts to $L(X)$ and so defines a Lie algebra homomorphism from $L(X)$ to $\mathfrak{X}(M)$.

The smallest Lie subalgebra of $\mathfrak{X}(M)$ which contains $\mathcal{V}$ may now be stated in a simple manner. It is simply the image of $L(X)$ under the homomorphism $\mathrm{Ev}(\phi)$. We shall denote this subalgebra by $\overline{\mathrm{Lie}}(\mathcal{V})$ and call it the *involutive closure* of $\mathcal{V}$.

For $x \in M$ we define the map $\mathrm{Ev}_x(\phi)\colon T(\mathbb{R}^X) \to T_xM$ by

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

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

3.3. Free Symmetric Algebras As far as we know, the idea of a symmetric algebra does not appear in the literature. However, the concept is a very natural one and shall be useful to us.

A *symmetric algebra* is an algebra, A, where the multiplication (which we shall denote by $(u,v) \mapsto \langle u : v\rangle$) is symmetric. Thus $\langle u : v\rangle = \langle v : u\rangle$ for $u, v \in A$. A map, $\phi\colon A \to A'$, between symmetric algebras is called a *symmetric algebra homomorphism* if $\phi(\langle u : v\rangle) = \langle \phi(u) : \phi(v)\rangle$ 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 $\mathrm{Pr}(X)$ 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 $\boldsymbol{Y} = \{X_1, \ldots, X_{l+1}\}$ (the reason for the slightly unusual enumeration will become clear in Section 5.6). For $P \in \mathrm{Pr}(\boldsymbol{Y})$ define $\gamma_a(P)$ to be the number of times the element X_a occurs in $P \in \mathrm{Pr}(\boldsymbol{Y})$ for $a = 1, \ldots, l+1$. We shall call the sum of the γ_a's the *degree* of P.

3.4. The Symmetric Algebra Generated by a Family of Vector Fields It turns out that we may define a special product on a Riemannian manifold which we shall call the *symmetric product.* First we give some basic notation from Riemannian geometry.

Recall that a *Riemannian manifold* is a is a pair, (M, g), where M is a differentiable manifold and g is a Riemannian metric on M. Thus g is a symmetric positive-definite tensor field of type $(0,2)$ on M. Given a Riemannian metric, we may define two isomorphisms of $C^\infty(M)$ modules; $\sharp\colon \Omega^1(M) \to \mathfrak{X}(M)$ and $\flat\colon \mathfrak{X}(M) \to \Omega^1(M)$ in the usual manner. In particular, if f is a function on Q, we define its *gradient* by $\operatorname{grad} f = (\boldsymbol{d}f)^\sharp$.

A Riemannian manifold is endowed with an *affine connection* which defines the operation, $\nabla_X Y$, called the *covariant derivative* of Y with respect to X. Given an affine connection and a set of coordinates $(x^1, \ldots, x^n)$ for M, we define the *Christoffel symbols* for the affine connection in these coordinates by

$$\nabla_{\partial/\partial x^j}\left(\frac{\partial}{\partial x^k}\right) = \Gamma^i_{jk}\frac{\partial}{\partial x^i}.$$

Given the properties of an affine connection, it may be easily verified that

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

If (M, g) is a Riemannian manifold, there exists a unique affine connection on M with the properties that $\nabla_X Y - \nabla_Y X = [X, Y]$ and that parallel translation with respect to this affine connection is an isometry. This affine connection is often called the *Levi-Civita* connection. It may be verified that the Christoffel symbols of the Levi-Civita connection are given by

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

Here g^{ij} is the inverse of the matrix g_{ij}. A curve $c\colon [0, T] \to M$ on a Riemannian manifold is said to be a *geodesic* if $\nabla_{c'(t)}c'(t) = 0$. In local coordinates, a geodesic is given by the solution of the following second-order differential equation:

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = 0.$$

This differential equation is, of course, the local representative of a vector field on TM. This vector field is called the *geodesic spray* or simply the *spray.* We shall denote it by Z_g. In local coordinates

$$Z_g = v^i\frac{\partial}{\partial q^i} - \Gamma^i_{jk}v^j v^k\frac{\partial}{\partial v^i}.$$

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

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

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 $\psi\colon X \to \mathcal{V}$. As in Section 3.3, let $S(X)$ be the free symmetric algebra on X and let $\mathrm{Pr}(X)$ be the symmetric products with elements in X. We may define a define a symmetric algebra homomorphism from $S(X)$ to $\mathfrak{X}(M)$ by extending ψ in the natural way (i.e., $\psi(\langle P_1 : P_1 \rangle) \mapsto \langle \psi(P_1) : \psi(P_2) \rangle$) to yield a map from $\mathrm{Pr}(X)$ to $\mathfrak{X}(M)$. This map may then be extended by $\mathbb{R}$-linearity to take values from $S(X)$. We denote the resulting map from $S(X)$ to $\mathfrak{X}(M)$ by $\mathrm{Ev}(\psi)$. We also define $\mathrm{Ev}_x(\psi)(P) = (\mathrm{Ev}(\psi)(P))(x)$ for $x \in M$. We denote by $\overline{\mathrm{Sym}}(\mathcal{V})$ 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

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

on a manifold M. We shall consider inputs from the set

$$\mathcal{U} = \{u\colon \mathbb{R} \to \mathbb{R}^m \mid \ u \text{ is piecewise constant}\}.$$

A *solution* of (4.1) is a pair, (c, u), where $c\colon [0, T] \to M$ is a piecewise smooth curve on M and $u \in \mathcal{U}$ such that

$$c'(t) = X(c(t)) + u^a(t) Y_a(c(t))$$

for each $t \in [0, T]$. For $x_0 \in M$, a neighborhood V of x_0, and $T > 0$ denote

$$\begin{aligned}\mathcal{R}^V(x_0, T) = \{x \in M \mid \ & \text{there exists a solution } (c, u) \text{ of } (4.1) \\ & \text{such that } c(0) = x_0,\ c(t) \in V \text{ for } t \in [0, T], \text{ and } c(T) = x\}\end{aligned}$$

and denote

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

Now we can define the versions of controllability.

4.1 Definition: The system (4.1) is *locally accessible* from x_0 if there exists $T > 0$ so that $\mathcal{R}^V(x_0, \leq t)$ contains a non-empty open set of M for all neighborhoods V of x_0 and all $0 < t \leq T$. If this holds for any $x_0 \in M$ then the system is called *locally accessible.*

The system (4.1) is *small-time locally controllable* (STLC) from $x_0 \in M$ if it is locally accessible from x_0 and if there exists $T > 0$ so that x_0 is in the interior of $\mathcal{R}^V(x_0, \leq t)$ for each $0 < t \leq T$ and each neighborhood V of x_0. If this holds for any $x_0 \in M$ then the system is called STLC. □

Let $\boldsymbol{X} = \{X_0, \ldots, X_m\}$. We will need some of the notation from Section 3.1 regarding free Lie algebras. In particular, $\mathrm{Br}(\boldsymbol{X})$ is the set of "brackets" of elements from $\boldsymbol{X}$ and $\delta_a(B)$ is the number of occurrences of X_a in $B \in \mathrm{Br}(\boldsymbol{X})$. The reader should also recall the Lie algebra rank condition (LARC) and that this is a sufficient condition for local accessibility. With further conditions on the types of brackets that a control system possesses, it may also be STLC.

An element $B \in \mathrm{Br}(\boldsymbol{X})$ is said to be *bad* if $\delta_0(B)$ is odd and $\delta_a(B)$ is even for each $a = 1, \ldots, m$. 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 $B \in \mathrm{Br}(\boldsymbol{X})$, define $\bar{\pi}(B)$ to be the bracket obtained by fixing X_0 and sending X_a to $X_{\pi(a)}$ for $a = 1, \ldots, m$. Now define

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

We may state sufficient conditions for STLC.

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

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

where C_a are good brackets in $\mathrm{Br}(\boldsymbol{X})$ of lower degree than B and $\xi_a \in \mathbb{R}$ for $a = 1, \ldots, m$. 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 [1978] and proven by Sussmann [1983].

5. Lagrangian Control Theory for Simple Mechanical Control Systems

In this section we study a specific, but large, class of mechanical control systems. Our presentation is from a Lagrangian point of view since this framework seems best adapted to the computations we do.

The systems studied are the so-called *simple mechanical control systems.* Such systems are characterised by the following data:

1. a Riemannian metric g on the n-dimensional configuration manifold Q which defines the kinetic energy of the system,

2. a function V on the configuration manifold which is the potential energy function, and

3. m linearly independent one-forms, $F^1, \ldots, F^m$, on Q which define the input forces.

The Lagrangian for the control system we consider is defined by

$$L(v) = \frac{1}{2} g(v,v) - V \circ \tau_Q(v) \tag{5.1}$$

where $\tau_Q \colon TQ \to Q$ is the tangent bundle projection. Thus we consider the Lagrangian to be "kinetic energy minus potential energy." The control torques take their values in the subset of T^*Q defined by

$$\Lambda_q = \left\langle F^1(q), \ldots, F^m(q) \right\rangle_{\mathbb{R}}.$$

This means that we will allow the possible directions for application of force to be functions of position only. More generally, one may want these directions to be functions of time and velocity as well.

With this data, the Lagrangian control system in local coordinates has the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = u_a F_i^a. \tag{5.2}$$

For the given Lagrangian, these equations may be expressed in a convenient invariant form. To express this we need the notion of the vertical lift of a vector field. Let X be a vector field on Q. Its *vertical lift* is the vector field on TQ defined by

$$X^{\mathit{lift}}(v) = \frac{d}{dt}\left(X(\tau_Q(v)) + tv\right)|_{t=0}. \tag{5.3}$$

In local coordinates, if

$$X(q) = X^i(q)\frac{\partial}{\partial q^i}$$

then we have

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

The reader may also wish to recall the definition of the geodesic spray, Z_g, from Section 3.4. We shall define

$$X_L = Z_g - \operatorname{grad} V^{\mathit{lift}}.$$

5.1 LEMMA: *Let L be the Lagrangian defined by (5.1). Then the equations (5.2) are equivalent to the equations*

$$\dot{v}(t) = X_L(v(t)) + u_a(t) Y_a(\tau_Q(v(t))) \tag{5.4}$$

where $Y_a = (F^a)^\sharp$ for $a = 1, \ldots, m$.

Proof: Let $c\colon [0,T] \to Q$ be an integral curve of X_L. Thus, in local coordinates,

$$\ddot{q}^i + \Gamma^i_{jk}\dot{q}^j\dot{q}^k = -g^{ij}\frac{\partial V}{\partial q^j} + u_a g^{ij} F^a_j$$

where

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

Note that

$$\frac{\partial L}{\partial q^k} = \frac{1}{2}\frac{\partial g_{ij}}{\partial q^k} v^i v^j - \frac{\partial V}{\partial q^k}, \qquad \frac{\partial L}{\partial v^k} = g_{kj} v^j.$$

Therefore,

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} &= g_{ij}\ddot{q}^j + \frac{\partial g_{ij}}{\partial q^k}\dot{q}^j\dot{q}^k - \frac{1}{2}\frac{\partial g_{jk}}{\partial q^i}\dot{q}^j\dot{q}^k + \frac{\partial V}{\partial q^i} \\ &= g_{ij}\ddot{q}^j + \left(\frac{\partial g_{ij}}{\partial q^k} - \frac{1}{2}\frac{\partial g_{jk}}{\partial q^i}\right)\dot{q}^j\dot{q}^k + \frac{\partial V}{\partial q^i}. \end{aligned}$$

Now note that

$$\begin{aligned} \Gamma^l_{jk}\dot{q}^j\dot{q}^k &= \frac{1}{2} g^{li}\left(\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{ij}}{\partial q^k} - \frac{\partial g_{jk}}{\partial q^i}\right)\dot{q}^j\dot{q}^k \\ &= g^{li}\left(\frac{\partial g_{ij}}{\partial q^k} - \frac{1}{2}\frac{\partial g_{jk}}{\partial q^i}\right)\dot{q}^j\dot{q}^k. \end{aligned}$$

The lemma now follows by multiplying Lagrange's equations by the "inverse" of g. ■

Note that we may also write (5.4) as

$$\nabla_{c'(t)} c'(t) = \operatorname{grad} V(c(t)) + u^a(t) Y_a(c(t)).$$

We shall use this form of the equations when we define a solution for a simple mechanical control system in Section 5.5.

With systems of this type there are some things that are worth noticing before proceeding to the calculations. In particular, note that all of the data for the problem is defined by quantities on the configuration manifold. Therefore, we would like to be able to compute the answers to interesting questions in terms of these quantities. An example of such an interesting question is the following:

Problem Statement: Describe the set of configurations which are reachable from a given configuration when starting at rest. □

It is exactly this question which we are interested in and which we shall answer. Furthermore, as we shall see, our answer is obtainable in terms of quantities defined on Q.

Since some rather detailed calculations are required in this section, let us outline what we plan to do. The reader may wish to refer to Section 2 where we presented an example which illustrated what we wish to do and why it is interesting. This example shows that the conventional definitions of controllability in the nonlinear control literature are not so well adapted to the mechanical systems we are considering. We also performed a few calculations for this example which foreshadow the general results developed in the succeeding sections. In Section 5.1 we do some computations with free Lie algebras. The reader should be warned that the presentation in this section may be difficult to follow, but is very important in understanding the basic premise of the sections which follow. We will also find it useful to know some tangent bundle structure. This is presented in Section 5.2. This structure becomes of consequence when we restrict the accessibility distribution to $Z(TQ)$. The distribution computations are performed in Section 5.3. With these computations, in Section 5.4 we are able to state the form of the accessibility distribution restricted to the zero section of TQ. In Section 5.5 we present controllability definitions for systems of the form (5.4). These formalise the problem statement given above. Using the computations from Section 5.3, we may obtain conditions for our notions of controllability. These are presented in Section 5.6. Finally, in Section 5.7 some decomposition results are presented which are analogous to the accessibility decompositions which can be made for nonlinear control systems.

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. The reader should be warned that they may not see what they expect here. Rather than just using a generating set which is in 1–1 correspondence with the set $\{X_L, Y_1^{lift}, \dots, Y_m^{lift}\}$ of control vector fields and the drift vector field, we also use a generating set which is in 1–1 correspondence with the set $\{Z_g, Y_1^{lift}, \dots, Y_m^{lift}, \operatorname{grad} V^{lift}\}$. The reason for this will become clear when we perform the distribution calculations in Section 5.3.

Let $\boldsymbol{X} = \{X_0, \dots, X_{m+1}\}$ and let $L(\boldsymbol{X})$ be the free Lie algebra generated by the set $\boldsymbol{X}$. We can simplify many of our computations for the controllability analysis of (5.4) by making simplifications to a set of generators for $L(\boldsymbol{X})$.

We first need some notation. Let

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

We will also need the concept of a *primitive* bracket.

5.2 Definition: Let $B \in \mathrm{Br}_0(\boldsymbol{X}) \cup \mathrm{Br}_{-1}(\boldsymbol{X})$ 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 $\mathrm{Br}_{-1}(\boldsymbol{X}) \cup \mathrm{Br}_0(\boldsymbol{X}) \cup \{X_0\}$. □

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

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

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

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

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

$$[X_{a_k}, [X_{a_{k-1}}, \dots, [X_{a_2}, X_{a_1}]]] \tag{5.5}$$

where $a_i \in \{0, \dots, m+1\}$ for $i = 1, \dots, k$. We shall see in Section 5.3 that brackets from $\mathrm{Br}_j(\boldsymbol{X})$, where $j \geq 1$ or $j \leq -2$, will not be of interest to us. In particular, we shall see that when $j \leq -2$ the brackets evaluate identically to zero. Therefore, in this section we concentrate our attention on brackets in $\mathrm{Br}_0(\boldsymbol{X}) \cup \mathrm{Br}_{-1}(\boldsymbol{X})$ which satisfy certain requirements. We state this in the following lemma.

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

Proof: It is sufficient to prove the lemma for brackets of the form (5.5). We proceed by induction on k in (5.5). The lemma is true for $k = 1, 2$ by inspection. Now suppose the lemma true for $k = 1, \dots, l$ and let B be of the form (5.5) for $k = l + 1$. Then we have two cases. Either $B \in \mathrm{Br}_{-1}(\boldsymbol{X})$ or $B \in \mathrm{Br}_0(\boldsymbol{X})$.

We look first at the case where $B \in \mathrm{Br}_{-1}(\boldsymbol{X})$. Since we are considering brackets in $\mathrm{Br}_{-2}(\boldsymbol{X})$ to be zero, we may write $B = [X_a, B']$ with $B' \in \mathrm{Br}_0(\boldsymbol{X})$ of the form (5.5) and $a \in \{1, \dots, m+1\}$. By the induction hypothesis, B' is a finite sum of primitive brackets and the lemma is proved in this case since B will then also be a finite sum of primitive brackets.

Now we look at the case where $B \in \mathrm{Br}_0(\boldsymbol{X})$. There are two possibilities in this case. The first possibility is that $B = [X_0, B']$ with $B' \in \mathrm{Br}_{-1}(\boldsymbol{X})$. In this case B' is a finite sum of primitive brackets by the induction hypothesis and, therefore, B is also a finite sum of primitive brackets.

The final case is when $B = [X_{a_1}, B']$ with $B' \in \mathrm{Br}_{+1}(\boldsymbol{X})$ of the form (5.5). If $B' = [X_0, B'']$ with $B'' \in \mathrm{Br}_0(\boldsymbol{X})$ then, by Jacobi's identity, we have

$$B = [X_{a_1}, [X_0, B'']] = -[B'', [X_{a_1}, X_0]] - [X_0, [B'', X_{a_1}]].$$

Since $B'' \in \mathrm{Br}_0(\boldsymbol{X})$, by the induction hypotheses it may be written as a finite sum of primitive brackets in $\mathrm{Br}_0(\boldsymbol{X})$. Clearly $[X_{a_1}, X_0]$ is primitive which proves that $[B'', [X_{a_1}, X_0]]$ is a finite sum of primitive brackets. The bracket $[B'', X_{a_1}]$ is in $\mathrm{Br}_{-1}(\boldsymbol{X})$. Therefore, by the induction hypotheses it may be written as a finite sum of primitive brackets. Thus the term $[X_0, [B'', X_{a_1}]]$, and hence B, may be written as a finite sum of primitive brackets.

Now suppose that $B' = [X_{a_2}, B'']$ with $B'' \in \mathrm{Br}_{+2}(\boldsymbol{X})$. First look at the case where $B'' = [X_0, B''']$ with $B''' \in \mathrm{Br}_{+1}(\boldsymbol{X})$. In this case we have

$$\begin{aligned}B =&[X_{a_1}, [X_{a_2}, [X_0, B''']]] = -[X_{a_1}, [B''', [X_{a_2}, X_0]]] - [X_{a_1}, [X_0, [B''', X_{a_2}]]] \\ =&[[X_{a_2}, X_0], [X_{a_1}, B''']] + [B''', [[X_{a_2}, X_0], X_{a_1}]] + \\ &[[B''', X_{a_2}], [X_{a_1}, X_0]] + [X_0, [[B''', X_{a_2}], X_{a_1}]].\end{aligned}$$

The first, third and fourth terms can be written as finite sums of primitive brackets by the induction hypothesis, and the second term is zero by our condition that brackets in $\mathrm{Br}_{-2}(\boldsymbol{X})$ are taken to be zero.

If $B'' = [X_{a_3}, B''']$ then we keep stripping factors off of B''' until we encounter an X_0. When we do, we repeatedly apply the above procedure. This proves the lemma. ■

An example is useful in illustrating what is behind the lemma.

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

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

Now we relate the free Lie algebra $L(\boldsymbol{X})$ with a free Lie algebra which corresponds to the set $\{X_L, Y_1^{lift}, \dots, Y_a^{lift}\}$. Let $\boldsymbol{X}' = \{X_0', \dots, X_m'\}$. 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 $\mathrm{Br}(\boldsymbol{X}')$ as linear combinations of brackets in $\mathrm{Br}(\boldsymbol{X})$ by $\mathbb{R}$-linearity of the bracket. We may, in fact, be even more precise about this.

Let $B' \in \mathrm{Br}(\boldsymbol{X}')$. We define a subset, $\boldsymbol{S}(B')$, of $\mathrm{Br}(\boldsymbol{X})$ by saying that $B \in \boldsymbol{S}(B')$ 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

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

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

5.5 Lemma: *Let $B' \in \mathrm{Br}(\boldsymbol{X}')$. Then*

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

Proof: It suffices to prove the lemma for the case when B' is of the form

$$B' = [X'_{a_k}, [X'_{a_{k-1}}, [\cdots, [X'_{a_2}, X'_{a_1}]]]] \tag{5.6}$$

since these brackets generate $L(\boldsymbol{X}')$ by Proposition 3.1. We proceed by induction on k. The lemma is true for $k = 1$. Now suppose the lemma true for $k = 1, \ldots, l$ where $l \geq 1$ and let B' be of the form (5.6) with $k = l + 1$. Then either $B' = [X'_a, B'']$, $a = 1, \ldots, m$ or $B' = [X_0, B'']$ with B'' of the form (5.6) with $k = l$. In the first case, by the induction hypotheses, we have

$$\begin{aligned} B' &= \sum_{B \in \boldsymbol{S}(B'')} [X_a, (-1)^{\delta_{m+1}(B)} B] \\ &= \sum_{B \in \boldsymbol{S}(B')} (-1)^{\delta_{m+1}(B)} B. \end{aligned}$$

In the second case we have

$$\begin{aligned} B' &= \sum_{B \in \boldsymbol{S}(B'')} [X_0 - X_{m+1}, (-1)^{\delta_{m+1}(B)} B] \\ &= \sum_{B \in \boldsymbol{S}(B')} (-1)^{\delta_{m+1}(B)} B. \end{aligned}$$

This proves the lemma. ■

We shall only be interested in terms in the above decomposition of B' which are in $\mathrm{Br}_0(\boldsymbol{X}) \cup \mathrm{Br}_{-1}(\boldsymbol{X})$ since, as we shall see in Section 5.3, these are the only ones which will contribute to $\mathrm{Ev}_{0_q}(\phi')(B')$.

A good understanding of this section is important in any effort to understand the proofs of Proposition 5.11 and Theorem 5.17 which follow. The reader should come back to this section if they are having difficulty with these proofs.

5.2. Some Useful Tangent Bundle Structure Since we are interested in restricting the accessibility distribution to the zero section of TQ, there are some useful properties of the tangent bundle which we shall need.

Since $Z(TQ)$, the zero section of the tangent bundle, is a submanifold of TQ which is canonically diffeomorphic to Q, it is possible to realise T_qQ as a subspace of $T_{0_q}TQ$. At each point $0_q \in Z(TQ)$ we shall call this subspace *horizontal*. Note that this version of horizontal is valid *only* at those points in TQ which are on the zero section. Present as a subspace of $T_{v_q}TQ$ for *any* $v_q \in TQ$ is the *vertical* subspace. Recall that this subspace is the kernel of the map $T_{v_q}\tau_Q$. Also note that at points $0_q \in Z(TQ)$, $T_{0_q}TQ = T_qQ \oplus V_{0_q}Q$. By T_qQ in this decomposition we mean the horizontal subspace of $T_{0_q}TQ$ which is canonically isomorphic to T_qQ. The reader should be aware that this identification will be implicitly made in the sequel.

5.3. 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 (5.4) restricted to $Z(TQ)$. To make the correspondence between the free Lie algebra $L(\boldsymbol{X})$ used in Section 5.1 and the accessibility algebra for (5.4), we define a family of vector fields

$$\mathcal{V} = \{Z_g, Y_1^{lift}, \ldots, Y_m^{lift}, \operatorname{grad} V^{lift}\}$$

and establish a bijection, ϕ, from $\boldsymbol{X}$ to $\mathcal{V}$ by mapping X_0 to X_L, X_a to Y_a^{lift} for $a = 1, \ldots, m$, and X_{m+1} to $\operatorname{grad} V^{lift}$. 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 $\mathcal{V}' = \{X_L, Y_1^{lift}, \ldots, Y_m^{lift}\}$. We establish a bijection, ϕ', from $\boldsymbol{X}'$ to $\mathcal{V}'$ by mapping X_0' to X_L and X_a' to Y_a^{lift} for $a = 1, \ldots, m$. By Lemma 5.5, each vector field in $\overline{\mathrm{Lie}}(\mathcal{V}')$ is a $\mathbb{R}$-linear sum of vector fields in $\overline{\mathrm{Lie}}(\mathcal{V})$.

Now we shall show that it is possible to compute the brackets from $\mathrm{Br}(\boldsymbol{X})$ 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 covariant derivative which was introduced in Section 3.4.

5.6 LEMMA: *Suppose that $B \in \mathrm{Br}^k(\boldsymbol{X})$ is primitive.*

i) If $B \in \mathrm{Br}_{-1}(\boldsymbol{X})$ then $\mathrm{Ev}(\phi)(B)$ is the vertical lift of a vector field on Q.

ii) If $B \in \mathrm{Br}_0(\boldsymbol{X})$ then $U = \mathrm{Ev}(\phi)(B)$ 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 $U_Q \colon q \mapsto U(0_q) \in T_qQ \subset T_{0_q}TQ$. There are two cases to consider.

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

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

Proof: The proof is by induction on k. The result is true for $k = 1$ trivially. To prove the result for $k = 2$ we introduce some notation which we will find handy for doing the bracket calculations in coordinates. If we have two *general* vector fields

$$X_1 = X_{1,h}^i(q,v)\frac{\partial}{\partial q^i} + X_{1,v}^i(q,v)\frac{\partial}{\partial v^i}, \qquad X_2 = X_{2,h}^i(q,v)\frac{\partial}{\partial q^i} + X_{2,v}^i(q,v)\frac{\partial}{\partial v^i},$$

their Lie bracket will be represented by

$$[X_1, X_2] \sim \begin{bmatrix} \frac{\partial X_{2,h}^i}{\partial q^j} & \frac{\partial X_{2,h}^i}{\partial v^j} \\ \frac{\partial X_{2,v}^i}{\partial q^j} & \frac{\partial X_{2,v}^i}{\partial v^j} \end{bmatrix} \begin{pmatrix} X_{1,h}^j \\ X_{1,v}^j \end{pmatrix} - \begin{bmatrix} \frac{\partial X_{1,h}^i}{\partial q^j} & \frac{\partial X_{1,h}^i}{\partial v^j} \\ \frac{\partial X_{1,v}^i}{\partial q^j} & \frac{\partial X_{1,v}^i}{\partial v^j} \end{bmatrix} \begin{pmatrix} X_{2,h}^j \\ X_{2,v}^j \end{pmatrix}.$$

This is somewhat imprecise, but is convenient notationally.

If X, Y are vector fields on Q we may compute

$$[X^{lift}, Y^{lift}] \sim \begin{bmatrix} 0 & 0 \\ \frac{\partial Y^i}{\partial q^j} & 0 \end{bmatrix} \begin{pmatrix} 0 \\ X^j \end{pmatrix} - \begin{bmatrix} 0 & 0 \\ \frac{\partial X^i}{\partial q^j} & 0 \end{bmatrix} \begin{pmatrix} 0 \\ Y^j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{5.7}$$

If X is a vector field on Q we compute

$$[Z_g, X^{lift}] \sim \begin{bmatrix} 0 & 0 \\ \frac{\partial X^i}{\partial q^j} & 0 \end{bmatrix} \begin{pmatrix} v^j \\ -\Gamma^j_{kl} v^k v^l \end{pmatrix} - \begin{bmatrix} 0 & \delta^i_j \\ -\frac{\partial \Gamma^i_{kl}}{\partial q^j} & -2\Gamma^i_{jk} v^k \end{bmatrix} \begin{pmatrix} 0 \\ X^j \end{pmatrix}. \tag{5.8}$$

Inspecting (5.8) shows that $[Z_g, X^{lift}](0_q) = -X(q)$. Now let $Y \in \mathfrak{X}(Q)$. We compute

$$\begin{aligned}[Y^{lift}, [Z_g, X^{lift}]] \sim & \begin{bmatrix} -\frac{\partial X^i}{\partial q^j} & 0 \\ \frac{\partial^2 X^i}{q^j q^k} v^k + 2\frac{\partial \Gamma^i_{kl}}{\partial q^j} X^k v^l + 2\Gamma^i_{kl} \frac{\partial X^k}{\partial q^j} v^l & \frac{\partial X^i}{\partial q^j} + 2\Gamma^i_{kj} X^k \end{bmatrix} \begin{pmatrix} 0 \\ Y^j \end{pmatrix} - \\ & \begin{bmatrix} 0 & 0 \\ \frac{\partial Y^i}{\partial q^j} & 0 \end{bmatrix} \begin{pmatrix} -X^j \\ \frac{\partial X^j}{\partial q^k} v^k + 2\Gamma^j_{kl} X^k v^l \end{pmatrix}.\end{aligned}$$

Reading the coefficients gives

$$[Y^{lift}, [Z_g, X^{lift}]] = \left(\frac{\partial Y^i}{\partial q^j} X^j + \frac{\partial X^i}{\partial q^j} Y^j + 2\Gamma^i_{jk} X^j Y^k \right) \frac{\partial}{\partial v^i} \tag{5.9}$$

which is the coordinate representation of $(\nabla_X Y + \nabla_Y X)^{lift}$. 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 $B \in \mathrm{Br}^{l+1}(\boldsymbol{X})$ be primitive.

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

$$\mathrm{Ev}(\phi)(B_1) = \alpha^i(q) \frac{\partial}{\partial v^i}, \qquad \mathrm{Ev}(\phi)(B_2) = \lambda^i(q) \frac{\partial}{\partial q^i} + \mu^i_j(q) v^j \frac{\partial}{\partial v^i}.$$

Now we compute

$$\mathrm{Ev}(\phi)([B_1, B_2]) \sim \begin{bmatrix} \frac{\partial \lambda^i}{\partial q^j} & 0 \\ \frac{\partial \mu^i_k}{\partial q^j} v^k & \mu^i_j \end{bmatrix} \begin{pmatrix} 0 \\ \alpha^j \end{pmatrix} - \begin{bmatrix} 0 & 0 \\ \frac{\partial \alpha^i}{\partial q^j} & 0 \end{bmatrix} \begin{pmatrix} \lambda^j \\ \mu^j_k v^k \end{pmatrix}.$$

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 $B \in \mathrm{Br}_0(\boldsymbol{X})$. Without loss of generality (by Prim2) we may suppose that either (a) $B = [X_0, B_1]$ with $B_1 \in \mathrm{Br}_{-1}(\boldsymbol{X})$ or that (b) $B = [B_1, B_2]$ with $B_1, B_2 \in \mathrm{Br}_0(\boldsymbol{X})$. Let us deal with the first case. Equation (5.8) gives $\mathrm{Ev}(B)(\phi)(0_q) = -U_1(q)$ where U_1 is the vector field on Q so that $\mathrm{Ev}(\phi)(B_1) = U_1^{lift}$ (such a vector field exists by i). For every vector field U_2 on Q we have $[U_2^{lift}, [Z_g, U_1^{lift}]] = (\nabla_{U_1} U_2 + \nabla_{U_2} U_1)^{lift}$ by (5.9). This proves ii(a).

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

$$\mathrm{Ev}(\phi)(B_1) = \alpha^i(q)\frac{\partial}{\partial q^i} + \beta^i_j(q)v^j\frac{\partial}{\partial v^i}, \qquad \mathrm{Ev}(\phi)(B_2) = \lambda^i(q)\frac{\partial}{\partial q^i} + \mu^i_j(q)v^j\frac{\partial}{\partial v^i}.$$

We compute

$$\mathrm{Ev}(\phi)([B_1, B_2]) \sim \begin{bmatrix} \frac{\partial \lambda^i}{\partial q^j} & 0 \\ \frac{\partial \mu^i_k}{\partial q^j} v^k & \mu^i_j \end{bmatrix} \begin{pmatrix} \alpha^j \\ \beta^j_k v^k \end{pmatrix} - \begin{bmatrix} \frac{\partial \alpha^i}{\partial q^j} & 0 \\ \frac{\partial \beta^i_k}{\partial q^j} v^k & \beta^i_j \end{bmatrix} \begin{pmatrix} \lambda^j \\ \mu^j_k v^k \end{pmatrix}.$$

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

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

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

This lemma provides us with a strong step towards computing the value of all primitive brackets when evaluated using $\mathrm{Ev}(\phi)$. Next we show that these are the *only* types of brackets we need to consider. First we look at brackets in $\mathrm{Br}_l(\boldsymbol{X})$ for $l \geq 1$.

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

Proof: The lemma may be proved by showing that, in a coordinate chart for TQ, the horizontal components of $U = \mathrm{Ev}(\phi)(B)$ are polynomial in the fibre coordinates of degree l, and the vertical components of U are polynomial of degree $l+1$ in the fibre coordinates. This will follow if we can show that bracketing by X_a, $a = 1, \ldots, m$ reduces the polynomial order of the components by one and bracketing by X_0 increases the polynomial order of the components by one. This is a simple calculation which follows along the same lines as the calculations done for Lemma 5.6. ■

Now we look at the remaining brackets, those in $\mathrm{Br}_{-l}(\boldsymbol{X})$ for $l \geq 2$.

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

Proof: We prove the lemma by induction on k for brackets of the form (5.5). The result makes no sense for $k = 1$ and is true for $k = 2$ by (5.7). Now suppose the lemma true for $k = 2, \ldots, j$ and let $B \in \mathrm{Br}^{j+1}(\boldsymbol{X}) \cap \mathrm{Br}_{-l}(\boldsymbol{X})$ for $l \geq 2$ be of the form (5.5). Then either $B = [X_0, B']$ with $B' \in \mathrm{Br}_{-l-1}(\boldsymbol{X})$ or $B = [X_a, B']$ with $B' \in \mathrm{Br}_{-l+1}(\boldsymbol{X})$ and $a = 1, \ldots, m+1$. In either case the result follows immediately from the induction hypotheses and (5.7). ■

Let us summarise what we have done in this section. First we obtained a characterisation of primitive brackets in $\boldsymbol{X}$ when we evaluate them in $\mathcal{V}$ via $\mathrm{Ev}(\phi)$. This characterisation involved Lie brackets and covariant derivatives of the vector fields $Y_1, \ldots, Y_m, \operatorname{grad} V$. Then we showed in Lemmas 5.7 and 5.8 that the primitive brackets are the only ones we need be concerned with if we are evaluating the vector fields on the zero section of TQ.

5.4. The Form of the Accessibility Distribution Restricted to $Z(TQ)$ for Simple Mechanical Control Systems In this section we compute the accessibility distribution for (5.4) when restricted to the zero section of TQ. By Lemma 5.5 we know that we may write the vector fields in the accessibility algebra in terms of vector fields in $\overline{\mathrm{Lie}}(\mathcal{V})$. In Section 5.3 we saw some hints that we might be able to write vector fields in $\overline{\mathrm{Lie}}(\mathcal{V})$ in terms of covariant derivatives and Lie brackets of the input vector fields and $\operatorname{grad} V$. First we resolve this issue by saying exactly what the vector fields in $\overline{\mathrm{Lie}}(\mathcal{V})$ look like when we restrict them to $Z(TQ)$. We denote by $D_{\overline{\mathrm{Lie}}(\mathcal{V})}$ the distribution defined by

$$D_{\overline{\mathrm{Lie}}(\mathcal{V})}(v) = \big\langle U(v) \mid \ U \in \overline{\mathrm{Lie}}(\mathcal{V}) \big\rangle_{\mathbb{R}}.$$

The reader will also wish to recall the ideas from symmetric algebras presented in Section 3.4. We denote $\mathcal{Y} = \{Y_1, \ldots, Y_m\}$. Recall from Section 5.2 that T_qQ may be canonically included in $T_{0_q}TQ$. Also recall from that section that VTQ is the bundle of vertical vectors on TQ.

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

$$D_{\overline{\mathrm{Lie}}(\mathcal{V})}(0_q) \cap V_{0_q}TQ = (D_{\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\operatorname{grad} V\})}(q))^{\mathit{lift}}$$

and

$$D_{\overline{\mathrm{Lie}}(\mathcal{V})}(0_q) \cap T_qQ = D_{\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\operatorname{grad} V\}))}(q).$$

Proof: From Lemmas 5.7 and 5.8 we know that the only brackets from $\mathrm{Br}(\boldsymbol{X})$ which we need to consider are the primitive brackets. From Lemma 5.6 we know that the brackets which are in $\mathrm{Br}_{-1}(\boldsymbol{X})$ will generate the vertical directions, and the brackets which are in $\mathrm{Br}_0(\boldsymbol{X})$ will generate the horizontal directions.

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

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

Clearly

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

It is trivially true that $(\mathrm{Sym}^{(1)}(\mathcal{Y}\cup\{\operatorname{grad} V\}))^{\mathit{lift}} \subset \overline{\mathrm{Lie}}(\mathcal{V})$. Now suppose that $(\mathrm{Sym}^{(k)}(\mathcal{Y}\cup \{\operatorname{grad} V\}))^{\mathit{lift}} \subset \overline{\mathrm{Lie}}(\mathcal{V})$ for $k = 1, \ldots, l$ for $l \geq 1$. We see that $(\mathrm{Sym}^{(l+1)}(\mathcal{Y}\cup\{\operatorname{grad} V\}))^{\mathit{lift}} \subset$

$\overline{\mathrm{Lie}}(\mathcal{V})$ since we may generate all elements of $(\mathrm{Sym}^{(l+1)}(\mathcal{Y}\cup\{\mathrm{grad}\,V\}))^{lift}$ by considering brackets of the form $[U_1^{lift},[Z_g,U_2^{lift}]]$ where $U_i\in\mathrm{Sym}^{(l_i)}(\mathcal{Y},V)$ and $l_1+l_2=l+1$. This follows from (5.9). This shows that $(D_{\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\})}(q))^{lift}\subset D_{\overline{\mathrm{Lie}}(\mathcal{V})}(0_q)$.

Now we show that $D_{\overline{\mathrm{Lie}}(\mathcal{V})}(0_q)\subset(D_{\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\})}(q))^{lift}$. To do this we must show that the image under $\mathrm{Ev}(\phi)$ of all primitive brackets in $\mathrm{Br}_{-1}(\boldsymbol{X})$ may be written as a linear combination of vector fields in $\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\})$. A primitive bracket in $\mathrm{Br}_{-1}(\boldsymbol{X})$ may be written as $B=[B_1,B_2]$ with $B_1\in\mathrm{Br}_{-1}(\boldsymbol{X})$ and $B_2\in\mathrm{Br}_0(\boldsymbol{X})$ both being primitive. Therefore, either $B_2=[X_0,B_2']$ with B_2' primitive and in $\mathrm{Br}_{-1}(\boldsymbol{X})$ or $B_2=[B_2',B_2'']$ with $B_2',B_2''\in\mathrm{Br}_0(\boldsymbol{X})$ both primitive. In the first case $\mathrm{Ev}(\phi)(B)\in\mathrm{Sym}^{(k)}(\mathcal{Y}\cup\{\mathrm{grad}\,V\})$ for some k by (5.9). 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 (5.9) until they are expressed in terms of covariant derivatives. When this is done, $\mathrm{Ev}(\phi)(B)$ will then be a $\mathbb{R}$-linear combination of elements in $\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\})$. This shows that $D_{\overline{\mathrm{Lie}}(\mathcal{V})}(0_q)\subset(D_{\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\})}(q))^{lift}$.

To demonstrate the proposed form of $D_{\overline{\mathrm{Lie}}(\mathcal{V})}\cap T_qQ$, by Lemma 5.6 ii(b) we need only show that $\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\})(q)\subset D_{\overline{\mathrm{Lie}}(\mathcal{V})}(0_q)$. But this is clear from Lemma 5.6 ii(a). This completes the proof of the lemma. ■

5.10 Remark: Notice that the constructions in the above lemma only depend upon $\{Y_1,\dots,Y_m,\mathrm{grad}\,V\}$. The effects of the geodesic spray do not appear explicitly. However, its contribution is obviously important in the essential computations performed in Section 5.3. □

From Lemma 5.5 we know that the vector fields which contribute to $\overline{\mathrm{Lie}}(\mathcal{V}')$ when we evaluate on $Z(TQ)$ will be $\mathbb{R}$-linear combinations of vector fields from $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\}))$. Thus, to compute these vector fields, we need to figure out which vector fields need to be "removed" from $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\}))$. We present an algorithm which we shall prove determines exactly which $\mathbb{R}$-linear combinations from $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}\cup\{\mathrm{grad}\,V\}))$ we need to compute. We define *two* sequences of families of vector fields on Q which we shall denote by $\mathcal{C}_{ver}^{(k)}(\mathcal{Y},V)$ and $\mathcal{C}_{hor}^{(k)}(\mathcal{Y},V)$ where $k\in\mathbb{Z}^+$. In Figure 2 the algorithm is presented for computing these families. When we have computed these sequences we define

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

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

We may now state the form of the accessibility distribution $\overline{\mathrm{Lie}}(\mathcal{V}')$ for (5.4) when restricted to the zero section of TQ.

5.11 Proposition: *Let $q\in Q$. Then*

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

5.1 ALGORITHM:

```
For $i \in \mathbb{Z}^+$ do
    For $B \in \mathrm{Br}^{(i)}(\boldsymbol{X})$ primitive do
        If $\delta_{m+1}(B) = 0$ then
            If $B \in \mathrm{Br}_{-1}(\boldsymbol{X})$ then
                $U \in \mathcal{C}^{\frac{1}{2}(i+1)}_{ver}(\mathcal{Y}, V)$ where $\mathrm{Ev}(\phi)(B) = U^{lift}$
            else
                $U \in \mathcal{C}^{(i/2)}_{hor}(\mathcal{Y}, V)$ where $U(q) = \mathrm{Ev}_{0_q}(\phi)(B)$
            end
        else
            If $B$ has no components of the form $[X_0, X_{m+1}]$ then
                Compute $B' \in \mathrm{Br}(\boldsymbol{X})$ 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 $a = 1, \ldots, m$.
                Let $B'' = 0$.
                For $\tilde{B} \in \boldsymbol{S}(B') \cap (\mathrm{Br}_{-1}(\boldsymbol{X}) \cup \mathrm{Br}_0(\boldsymbol{X}))$ do
                    Write $\tilde{B}$ as a finite sum of primitive brackets in $\mathrm{Br}(\boldsymbol{X})$ by
                    Lemma 5.3.
                    $B'' = B'' + (-1)^{\delta_{m+1}(\tilde{B})}\tilde{B}$
                end
                If $B \in \mathrm{Br}_{-1}(\boldsymbol{X})$ then
                    $U \in \mathcal{C}^{\frac{1}{2}(i+1)}_{ver}(\mathcal{Y}, V)$ where $\mathrm{Ev}(\phi)(B'') = U^{lift}$
                else
                    $U \in \mathcal{C}^{(i/2)}_{hor}(\mathcal{Y}, V)$ where $U(q) = \mathrm{Ev}_{0_q}(\phi)(B'')$
                end
            end
        end
    end
end
END
```

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

and

$$D_{\overline{\mathrm{Lie}}(\mathcal{V}')}(0_q) \cap T_qQ = C_{hor}(\mathcal{Y}, V)(q).$$

Proof: Studying the algorithm that we have used to compute $\mathcal{C}_{ver}(\mathcal{Y}, V)$ and $\mathcal{C}_{hor}(\mathcal{Y}, V)$, the reader will notice that we have exactly taken each primitive bracket $B \in \mathrm{Br}(\boldsymbol{X})$ and computed which $\mathbb{R}$-linear combinations from $\mathrm{Br}(\boldsymbol{X})$ appear along with B in the decomposition of some $B' \in \mathrm{Br}(\boldsymbol{X}')$ given by Lemma 5.5. Since it is only these primitive brackets which appear in $\overline{\mathrm{Lie}}(\mathcal{V}') \mid Z(TQ)$, this will, by construction, generate $D_{\overline{\mathrm{Lie}}(\mathcal{V}')} \mid Z(TQ)$.

We need to prove that, as stated in the first step of the algorithm, if $\delta_{m+1}(B) = 0$, then $\mathrm{Ev}_{0_q}(\phi)(B) \in D_{\overline{\mathrm{Lie}}(\mathcal{V}')}(0_q)$. To show that this is in fact the case, let $B' \in \mathrm{Br}(\boldsymbol{X}')$ be the bracket obtained by replacing X_a with X'_a for $a = 0, \dots, m$. We claim that the only bracket in $\boldsymbol{S}(B')$ which contributes to $\mathrm{Ev}(\phi')(B')$ is B. This is true since any other brackets in $\boldsymbol{S}(B')$ 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 $\mathrm{Br}_{-l}(\boldsymbol{X})$ for $l \geq 2$. These brackets evaluate to zero by Lemma 5.8.

We also need to show that if B has components of the form $[X_0, X_{m+1}]$ then it will not contribute to $\overline{\mathrm{Lie}}(\mathcal{V}') \mid Z(TQ)$. 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. ■

It is perhaps useful to construct a few of the families $\mathcal{C}^{(k)}_{ver}(\mathcal{Y}, V)$ and $\mathcal{C}^{(k)}_{hor}(\mathcal{Y}, V)$ 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 $\mathrm{Br}^{(1)}(\boldsymbol{X})$ are $X_1, \dots, X_{m+1}$. For the brackets $B = X_a$, $a = 1, \dots, m$, $\delta_{m+1}(B) = 0$. Note that $\mathrm{Ev}(\phi)(B) = Y_a^{\mathit{lift}}$ so $Y_a \in \mathcal{C}^{(1)}_{ver}(\mathcal{Y}, V)$ 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 $\mathcal{C}^{(1)}_{ver}(\mathcal{Y}, V)$. If $B = X_{m+1}$ we compute $B' = X'_0$. Therefore, $\boldsymbol{S}(B') = \{X_0, X_{m+1}\}$. The only element in $\boldsymbol{S}(B')$ which is in $\mathrm{Br}_{-1}(\boldsymbol{X}) \cup \mathrm{Br}_0(\boldsymbol{X})$ is X_{m+1}. Therefore, $B'' = -X_{m+1}$. We then see that $\mathrm{Ev}(\phi)(B'') = -\operatorname{grad} V^{\mathit{lift}}$ from which we conclude that $\operatorname{grad} V \in \mathcal{C}^{(1)}_{ver}(\mathcal{Y}, V)$. In summary,

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

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

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

In a similar manner we may compute

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

and

$$\mathcal{C}^{(2)}_{hor}(\mathcal{Y}, V) = \mathcal{C}^{(2)}_{ver}(\mathcal{Y}, V) \cup \{[Y_a, Y_b] \mid \quad a, b = 1, \ldots, m\} \cup \\ \{2\,\langle Y_a : \operatorname{grad} V\rangle + [Y_a, \operatorname{grad} V] \mid \quad a = 1, \ldots, m\}\,.$$

To compute the terms $2\,\langle Y_a : \operatorname{grad} V\rangle + [Y_a, \operatorname{grad} V]$ in $\mathcal{C}^{(2)}_{hor}(\mathcal{Y}, V)$ we have used the computations of Example 5.4.

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

5.12 Remarks:

1. The generators we have written for $\mathcal{C}^{(k)}_{ver}(\mathcal{Y}, V)$ and $\mathcal{C}^{(k)}_{hor}(\mathcal{Y}, V)$ 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 $L(\boldsymbol{X}')$ and compute the image of these brackets under $\operatorname{Ev}(\phi')$. 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 $\operatorname{Ev}(\phi')$.

2. We claim that $C_{hor}(\mathcal{Y}, V)$ is involutive. Let $B'_1, B'_2 \in \operatorname{Br}(\boldsymbol{X}')$ be brackets which, when evaluated under $\operatorname{Ev}_{0_q}(\phi')$, give vector fields $U_1, U_2 \in \mathcal{C}_{hor}(\mathcal{Y}, V)$. Then the decomposition of B_i given by Lemma 5.5 has the form $B'_i = B_i + \tilde{B}_i$ where $B_i \in \operatorname{Br}_0(\boldsymbol{X})$ and $\tilde{B}_i$ is a sum of brackets in $\operatorname{Br}_j(\boldsymbol{X})$ for $j \geq 2$. Therefore, $[B'_1, B'_2] = [B_1, B_2] + B''$ where B'' is a sum of brackets in $\operatorname{Br}_j(\boldsymbol{X})$ 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 $\operatorname{Br}_{-j}(\boldsymbol{X})$ are taken to be zero for $j \geq 2$ (see Lemma 5.3).

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

$$\mathcal{C}_{ver}(\mathcal{Y}, V) = \overline{\operatorname{Sym}}(\mathcal{Y}), \qquad \mathcal{C}_{hor}(\mathcal{Y}, V) = \overline{\operatorname{Lie}}(\overline{\operatorname{Sym}}(\mathcal{Y})).$$

   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 5.3 remain valid if we replace $\operatorname{grad} V$ with an arbitrary vector field on Q. □

5.5. 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.

5.13 DEFINITION: A *solution* of (5.4) is a pair, (c, u), where $c\colon [0,T] \to Q$ is a piecewise smooth curve and $u \in \mathcal{U}$ such that

$$\nabla_{c'(t)} c'(t) = \operatorname{grad} V(c(t)) + u^a(t) Y_a(c(t)). \qquad \square$$

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

$$\mathcal{R}_Q^U(q_0, T) = \{q \in Q \mid \text{ there exists a solution } (c,u) \text{ of (5.4)} \\ \text{such that } c'(0) = 0_{q_0},\ c(t) \in U \text{ for } t \in [0,T], \text{ and } c'(T) \in T_q Q\}$$

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.

5.14 DEFINITION: We shall say that (5.4) 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 neighborhoods 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 (5.4) 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 neighborhood 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 (5.4) is *equilibrium controllable* if, for $q_1, q_2 \in \mathfrak{E}(L)$, there exists a solution (c, u) of (5.4) where $c\colon [0,T] \to Q$ 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.

5.6. Conditions for Controllability of Simple Mechanical Control Systems Lewis and Murray [1995] present sufficient conditions for local configuration accessibility. Here, since we have a complete description of $\overline{\mathrm{Lie}}(\mathcal{V}') \mid Z(TQ)$, we can give stronger results.

5.15 THEOREM: *The control system (5.4) is locally configuration accessible at q if $C_{hor}(\mathcal{Y}, V)(q) = T_q Q$.*

Proof: Let C denote the accessibility distribution. Since $C_{hor}(\mathcal{Y}, V)(q) \subset C(0_q)$ by Proposition 5.11, and $C_{hor}(\mathcal{Y}, V)(q) = T_q Q$ 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 (5.4) 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 neighborhood $\tilde{U} \subset \Lambda$ of 0_q and for every T sufficiently small. Now let U be a neighborhood of q and define a neighborhood 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 5.15 in the case when there is no potential energy.

5.16 THEOREM: *Suppose $V = 0$ and that (5.4) is locally configuration accessible. Then $C_{hor}(\mathcal{Y}, V)(q) = T_qQ$ for q in an open dense subset of Q.*

Proof: First note that if $C_{hor}(\mathcal{Y}, V)(q_0) = T_{q_0}Q$ then $C_{hor}(\mathcal{Y}, V)(q) = T_qQ$ in a neighborhood of q_0. This proves that the set of points q where $C_{hor}(\mathcal{Y}, V)(q) = T_qQ$ is open. Now suppose that $C_{hor}(\mathcal{Y}, V)(q) \subsetneq T_qQ$ in an open subset U of Q. Then there exists an open subset $\bar{U} \subset U$ so that $\operatorname{rank}(C_{hor}(\mathcal{Y}, V)(q)) = k < n$ for all $q \in \bar{U}$. However, this contradicts local configuration accessibility by Theorem 5.20. Therefore, there can be no open subset of Q on which $C_{hor}(\mathcal{Y}, V)(q) \subsetneq T_qQ$. Thus the set of points q where $C_{hor}(\mathcal{Y}, V)(q) = T_qQ$ 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 $\boldsymbol{Y} = \{X_1, \ldots, X_{m+1}\}$ and establish a bijection $\psi \colon \boldsymbol{Y} \to \mathcal{Y} \cup \{\operatorname{grad} V\}$ by asking that $\psi(X_a) = Y_a$ for $a = 1, \ldots, m$ and $\psi(X_{m+1}) = \operatorname{grad} V$. If $P \in \operatorname{Pr}(\boldsymbol{Y})$ we shall say that P is *bad* if $\gamma_a(P)$ is even for each $a = 1, \ldots, m$. 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 $P \in \operatorname{Pr}(\boldsymbol{Y})$ define $\bar{\pi}(P)$ to be the bracket obtained by fixing X_{m+1} and sending X_a to $X_{\pi(a)}$ for $a = 1, \ldots, m$. Now define

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

We may now state the sufficient conditions for STLCC.

5.17 THEOREM: *Suppose that $\mathcal{Y} \cup \{\operatorname{grad} V\}$ is such that every bad symmetric product in $\operatorname{Pr}(\boldsymbol{Y})$ has the property that*

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

where C_a are good symmetric products in $\operatorname{Pr}(\boldsymbol{Y})$ of lower degree than P and $\xi_a \in \mathbb{R}$ for $a = 1, \ldots, m$. Also, suppose that (5.4) is locally configuration accessible at q. Then (5.4) is STLCC at q.

Proof: First recall from the proof of Theorem 5.15 that if (5.4) 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, (5.4) is locally accessible. To show that (5.4) is STLCC at q, it clearly suffices to show that (5.4) is STLC at 0_q when restricted to Λ. We do this by showing that (5.4) satisfies the hypotheses of Theorem 4.2 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 $\mathrm{Br}(\boldsymbol{X}')$ and bad symmetric products in $\mathrm{Pr}(\boldsymbol{Y})$ and good brackets in $\mathrm{Br}(\boldsymbol{X}')$ and good symmetric products in $\mathrm{Pr}(\boldsymbol{Y})$.

Suppose that $B' \in \mathrm{Br}(\boldsymbol{X}')$ is bad. Thus $\delta_a(B')$ is even for $a = 1, \dots, m$ and $\delta_0(B')$ is odd. When we evaluate $\mathrm{Ev}_{0_q}(\phi')(B')$, the only terms that will remain in the decomposition of $\mathrm{Ev}(\phi')(B')$ given by Lemma 5.5 are the terms obtained from brackets in $\boldsymbol{S}(B')$ which are in $\mathrm{Br}_0(\boldsymbol{X}) \cup \mathrm{Br}_{-1}(\boldsymbol{X})$. Since B' is bad, we must have $\delta_a(B)$ even and $\delta_0(B) + \delta_{m+1}(B)$ odd for each $B \in \boldsymbol{S}(B')$. 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 $\boldsymbol{S}(B')$ which contribute to $\mathrm{Ev}(\phi')(B')$ must be in $\mathrm{Br}_{-1}(\boldsymbol{X})$. 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 $B \in \boldsymbol{S}(B')$. 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 $\boldsymbol{S}(B')$ which contribute to $\mathrm{Ev}(\phi')(B')$ must be in $\mathrm{Br}_{-1}(\boldsymbol{X})$. We then conclude that $\mathrm{Ev}_{0_q}(\phi')(B')$ must be of the form $(\mathrm{Ev}_q(\psi)(P))^{lift}$ where $P \in \mathrm{Pr}(\boldsymbol{Y})$ is bad.

Now suppose that $B' \in \mathrm{Br}(X')$ 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 $\boldsymbol{S}(B')$ which contribute to $\mathrm{Ev}(\phi')(B')$. 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 $\mathrm{Br}_0(\boldsymbol{X})$. 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 $B \in \mathrm{Br}_0(\boldsymbol{X})$. Therefore, good brackets from $\mathrm{Br}(\boldsymbol{X}')$ 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.

5.18 Corollary: *Suppose that the hypotheses of Theorem 5.17 hold. Then the system (5.4) is equilibrium controllable.*

5.19 Remarks:

1. Notice that Theorem 5.15 explains the example from Section 2. More precisely, 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 $C_{ver}(\mathcal{Y}, V)$ which need not span $V_{0_q}TQ$.

2. This result 5.18 may be made even stronger if we allow a point $q \in Q$ to be an equilibrium point if $\operatorname{grad} V(q)$ is in the span of the inputs at q. □

5.7. Decompositions for Simple Mechanical Control Systems Now we give decomposition results which mirror those for standard nonlinear control systems. Our first result gives a decomposition which is valid for systems with no potential energy.

5.20 THEOREM: *Suppose that $V = 0$ for the control system (5.4) and suppose that $C_{hor}(\mathcal{Y}, V)$ has constant rank k in a neighborhood of $q_0 \in Q$. Then there exists a coordinate chart, (U, ϕ), around q_0 such that the submanifold*

$$S_{q_0} = \{q \in U \mid \; q^i(q) = q^i(q_0), \;\; i = k+1, \dots, n\}$$

is an integral manifold of $C_{hor}(\mathcal{Y}, V)$. Then, for any neighborhood $W \subset U$ of q_0 and for all $T > 0$ sufficiently small, $\mathcal{R}_Q^W(q_0, T)$ is contained in S_{q_0}. Hence the system restricted to S_{q_0} is locally configuration accessible.

Proof: The coordinate decomposition exists since $C_{hor}(\mathcal{Y}, V)$ is integrable as pointed out in Remark 5.12(2). Since $V = 0$, we have $\mathcal{C}_{ver}(\mathcal{Y}, V) = \overline{\mathrm{Sym}}(\mathcal{Y})$ and $\mathcal{C}_{hor}(\mathcal{Y}, V) = \overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))$ as in Remark 5.12(3). This implies that $\mathcal{C}_{ver}(\mathcal{Y}, V) \subset \mathcal{C}_{hor}(\mathcal{Y}, V)$ and so all solutions of (5.4) which start on S_{q_0} with zero initial velocity will remain on S_{q_0}. Thus $\mathcal{R}_Q^W(q_0, T) \subset S_{q_0}$. It is also clear that the system is locally configuration accessible when restricted to initial conditions in S_{q_0} since $\dim(S_{q_0}) = \mathrm{rank}(C_{hor}(\mathcal{Y}, V) \mid S_{q_0})$. ■

Now we give a result which gives the form of the equations on the integral manifolds of $C_{hor}(\mathcal{Y}, V)$ when the potential energy is non-zero.

5.21 THEOREM: *Suppose that $C_{hor}(\mathcal{Y}, V)$ has constant rank k in a neighborhood of $q_0 \in Q$. Then there exists coordinates $(x^1, \dots, x^k, y^1, \dots, y^{n-k})$ so that the system (5.4) has the form*

$$\ddot{x}^i + \Gamma^i_{jk}(x,y)\dot{x}^j\dot{x}^k + \Gamma^i_{j\alpha}(x,y)\dot{x}^j\dot{y}^\alpha + \Gamma^i_{\alpha\beta}(x,y)\dot{y}^\alpha\dot{y}^\beta + g^{ij}\frac{\partial V}{\partial x^j} + g^{i\alpha}\frac{\partial V}{\partial y^\alpha} = u^a Y^i_a$$
$$\ddot{y}^\alpha + \Gamma^\alpha_{j\beta}(x,y)\dot{x}^j\dot{y}^\beta + \Gamma^\alpha_{\beta\gamma}(x,y)\dot{y}^\beta\dot{y}^\gamma + g^{\alpha j}\frac{\partial V}{\partial x^j} + g^{\alpha\beta}\frac{\partial V}{\partial y^\beta} = 0.$$

Furthermore, for each fixed value of y, the control system

$$\ddot{x}^i + \Gamma^i_{jk}(x,y)\dot{x}^j\dot{x}^k + g^{ij}\frac{\partial V}{\partial x^j} + g^{i\alpha}\frac{\partial V}{\partial y^\alpha} = u^a Y^i_a$$

is locally configuration accessible.

Proof: Since $C_{hor}(\mathcal{Y}, V)$ has constant rank in a neighborhood of q_0 and $C_{hor}(\mathcal{Y}, V)$ is integrable, by Frobenius' theorem we may find coordinates $(x^1, \dots, x^k, y^1, \dots, y^{n-k})$ for Q so that

$$C_{hor}(\mathcal{Y}, V)(q_0) = \left\langle \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k} \right\rangle_{\mathbb{R}}.$$

In general, the equations (5.4) in these coordinates will have the form

$$\ddot{x}^i + \Gamma^i_{jk}(x,y)\dot{x}^j\dot{x}^k + \Gamma^i_{j\alpha}(x,y)\dot{x}^j\dot{y}^\alpha + \Gamma^i_{\alpha\beta}(x,y)\dot{y}^\alpha\dot{y}^\beta + g^{ij}\frac{\partial V}{\partial x^j} + g^{i\alpha}\frac{\partial V}{\partial y^\alpha} = u^a Y^i_a \tag{5.10a}$$

$$\ddot{y}^\alpha + \Gamma^\alpha_{jk}(x,y)\dot{x}^j\dot{x}^k + \Gamma^\alpha_{j\beta}(x,y)\dot{x}^j\dot{y}^\beta + \Gamma^\alpha_{\beta\gamma}(x,y)\dot{y}^\beta\dot{y}^\gamma + g^{\alpha j}\frac{\partial V}{\partial x^j} + g^{\alpha\beta}\frac{\partial V}{\partial y^\beta} = 0. \tag{5.10b}$$

We claim that the term $\Gamma^{\alpha}_{jk}(x,y)\dot{x}^j\dot{x}^k$ in (5.10b) must be zero. This follows from Theorem 5.20 proving the given form of the decomposition. That the top system is locally configuration accessible follows from the fact that $\operatorname{rank}(C_{hor}(\mathcal{Y},V)) = k$. (It makes sense to speak of local configuration accessibility of this system by Remark 5.12(4) and the statement immediately following Definition 5.14.) ■

5.22 Remark: In Theorem 5.20 the act of restricting to S_{q_0} has specific meaning. We may pull-back the Riemannian metric to S_{q_0} since it is a submanifold of Q. Doing so defines a Riemannian metric on S_{q_0}. This defines a simple mechanical control system (with zero potential energy) on S_{q_0} and, as long as we begin with zero initial velocity, the trajectories of this control system will be the same as those of the larger system. □

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 5.15 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 In this section we return to the example discussed in Section 2. This example, although simple, exhibits much of the subtle behaviour that makes the study of mechanical systems interesting.

In the coordinates (θ,ψ,r) presented in Section 2, the Riemannian metric for the robotic leg is

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

the input one-forms are

$$F^1 = d\theta - d\psi, \qquad F^2 = dr,$$

and the potential energy function is zero. In Section 2 we computed the input vector fields to be

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

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

We will find the following computations to be sufficient:

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

The reader may wish to compare these calculations with the bracket calculations of Section 2.

We may now go ahead and determine the configuration controllability of the robotic leg for the following three combinations of inputs:

RL1. Inputs Y_1 and Y_2: In this case it is clear that the system is locally configuration accessible by Theorem 5.15 as the input vector fields and their Lie bracket generates the maximal distribution on Q. Also, the bad symmetric product $\langle Y_1 : Y_1 \rangle$ is a multiple of Y_2 so the system is STLCC by Theorem 5.17.

RL2. Input Y_1: In this case the system is again locally configuration accessible since the vector fields $\{Y_1, \langle Y_1 : Y_1 \rangle, [Y_1, \langle Y_1 : Y_1 \rangle]\}$ generate the maximal distribution on Q. Note that the bad symmetric product $\langle Y_1 : Y_1 \rangle$ does not lie in the span of the inputs. Therefore, with this input, the robotic leg violates the sufficient conditions of Theorem 5.17 for STLCC.

RL3. Input Y_2: In this case we only generate the direction Y_2 and so the system is not locally configuration accessible. Indeed, starting from rest and only applying force in the r-direction, the only behaviour that can be observed is motion back and forth of the mass on the end of the leg. The decomposition of Theorem 5.21 in this case is given by

$$\ddot{r} - r\dot{\psi}^2 = \frac{1}{m}u_1$$

$$\ddot{\theta} = 0$$
$$\ddot{\psi} + \frac{2}{r}\dot{r}\dot{\psi} = 0.$$

The top system is obviously locally configuration accessible and also STLCC.

RL4. 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.

6.1 Remarks:

1. The fact that the system is STLCC with both inputs (RL1) is not surprising given the discussion of Section 2. Here we have just verified the claim in that section using the formalism developed in Section 5.
2. Observe that the decomposition in RL3 is just as specified in Theorem 5.21. No inputs appear in the bottom two equations, and no terms which are quadratic in $\dot{r}$ appear in the bottom two equations.
3. Although the system only violates the *sufficient* conditions for STLCC in RL2, one may easily see by looking at the r-component of Lagrange's equations 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. □

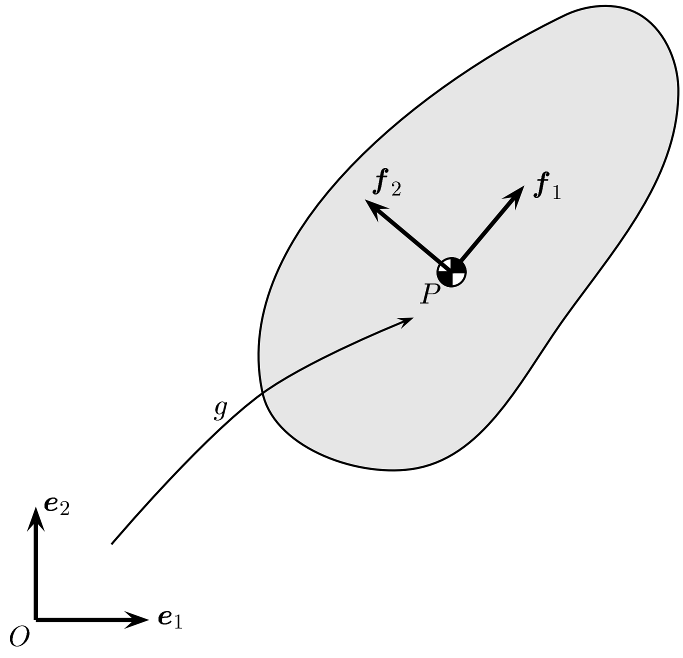

FIGURE 3. The configuration of a planar body as an element of $SE(2)$

6.2. The Forced Planar Rigid Body In this section we study the planar rigid body 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 $\{\boldsymbol{e}_1 = \frac{\partial}{\partial x}, \boldsymbol{e}_2 = \frac{\partial}{\partial y}\}$ be the standard orthonormal frame at that point. Let $\{\boldsymbol{f}_1, \boldsymbol{f}_2\}$ 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 $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ to the position, P, of the centre of mass of the body with its frame $\{\boldsymbol{f}_1, \boldsymbol{f}_2\}$. See Figure 3. 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 $\{\boldsymbol{f}_1, \boldsymbol{f}_2\}$) we may suppose that the point of application of the force is a distance h along the $\boldsymbol{f}_1$ body-axis from the centre of mass. The situation is illustrated in Figure 4.

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

$$g = mdx \otimes dx + mdy \otimes dy + Jd\theta \otimes d\theta.$$

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

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

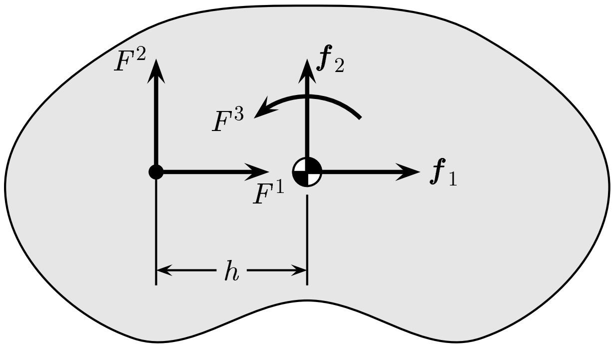

FIGURE 4. Positions for application of forces on a planar rigid body after simplifying assumptions

from which we compute the input vector fields as

$$Y_1 = \frac{\cos\theta}{m}\frac{\partial}{\partial x} + \frac{\sin\theta}{m}\frac{\partial}{\partial x},$$
$$Y_2 = -\frac{\sin\theta}{m}\frac{\partial}{\partial x} + \frac{\cos\theta}{m}\frac{\partial}{\partial y} - \frac{h}{J}\frac{\partial}{\partial\theta}, \qquad Y_3 = \frac{1}{J}\frac{\partial}{\partial\theta}.$$

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

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

$$\begin{aligned}
\langle Y_1 : Y_1\rangle &= 0,\\
\langle Y_1 : Y_2\rangle &= \frac{h\sin\theta}{mJ}\frac{\partial}{\partial x} - \frac{h\cos\theta}{mJ}\frac{\partial}{\partial y},\\
\langle Y_1 : Y_3\rangle &= -\frac{\sin\theta}{mJ}\frac{\partial}{\partial x} + \frac{\cos\theta}{mJ}\frac{\partial}{\partial y},\\
\langle Y_2 : Y_2\rangle &= \frac{2h\cos\theta}{mJ}\frac{\partial}{\partial x} + \frac{2h\sin\theta}{mJ}\frac{\partial}{\partial y},\\
\langle Y_2 : Y_3\rangle &= -\frac{\cos\theta}{mJ}\frac{\partial}{\partial x} - \frac{\sin\theta}{mJ}\frac{\partial}{\partial y},\\
\langle Y_3 : Y_3\rangle &= 0,\\
[Y_1, Y_2] &= -\frac{h\sin\theta}{mJ}\frac{\partial}{\partial x} + \frac{h\cos\theta}{mJ}\frac{\partial}{\partial y},\\
[Y_1, Y_3] &= \frac{\sin\theta}{mJ}\frac{\partial}{\partial x} - \frac{\cos\theta}{mJ}\frac{\partial}{\partial y},\\
[Y_2, Y_3] &= \frac{\cos\theta}{mJ}\frac{\partial}{\partial x} + \frac{\sin\theta}{mJ}\frac{\partial}{\partial y},\\
[Y_2, \langle Y_2 : Y_2\rangle] &= \frac{2h^2\sin\theta}{mJ^2}\frac{\partial}{\partial x} - \frac{2h^2\cos\theta}{mJ^2}\frac{\partial}{\partial y}.
\end{aligned}$$

With the computations done, we may proceed to determine configuration controllability for the planar rigid body with various combinations of inputs. Since the case where all inputs are present is trivial from the point of view of controllability, we do not present it.

PB1. Inputs Y_1 and Y_2: In this case the maximal distribution on Q is generated by the inputs and their Lie bracket. Therefore, the system in locally configuration accessible with these inputs by Theorem 5.15. Also, the bad symmetric product $\langle Y_2 : Y_2\rangle$, is a multiple of Y_1 so the system is STLCC by Theorem 5.17.

PB2. Inputs Y_1 and Y_3: It is easy to see that the vector fields $\{Y_1, Y_3, [Y_1, Y_3]\}$ generate the maximal distribution on Q and so the system is locally configuration accessible with these inputs. All bad symmetric products vanish so the system is also STLCC.

PB3. Input Y_1: The only direction generated by all symmetric products and Lie brackets is Y_1 itself. Thus the system is not locally configuration accessible. To use the decomposition of Theorem 5.21 we must make a change of coordinates. In the coordinates $(\xi, \eta, \psi) = (x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, \theta)$ the equations have the form

$$\begin{aligned}
&\ddot{\xi} + 2\left(\frac{m\eta^2}{J} - \frac{J+m\eta^2}{J}\right)\dot{\eta}\dot{\psi} + \left(\frac{m\xi\eta^2}{J} - \frac{\xi J + m\xi\eta^2}{J}\right)\dot{\psi}^2 = \\
&\qquad\qquad\left(\frac{J+m\eta^2}{J} - \frac{\eta^2}{J}\right)u_1 \\
&\ddot{\eta} + 2\left(\frac{J+m\xi^2}{J} - \frac{m\xi^2}{J}\right)\dot{\xi}\dot{\psi} + \left(\frac{m\eta\xi^2}{J} - \frac{\eta J + m\eta\xi^2}{J}\right)\dot{\psi}^2 = 0 \\
&\ddot{\psi} = 0.
\end{aligned}$$

The top system is locally configuration accessible and STLCC.

PB4. Inputs Y_2 and Y_3: With these inputs the maximal distribution on Q is generated by the input vector fields and their Lie bracket. Thus the system is locally configuration accessible. However, the bad symmetric product $\langle Y_2 : Y_2\rangle$ does not lie in the span of the inputs so the sufficient conditions of Theorem 5.17 are violated and the system may not be STLCC.

PB5. Input Y_2: With this input the maximal distribution on Q is generated by the vector fields $\{Y_2, \langle Y_2 : Y_2\rangle, [Y_2, \langle Y_2 : Y_2\rangle]\}$. Thus the system is locally configuration accessible by Theorem 5.15. The bad symmetric product $\langle Y_2 : Y_2\rangle$, is not a multiple of Y_2 so the system does not satisfy the sufficient conditions for STLCC.

PB6. Input Y_3: In this final case all symmetric products and Lie brackets are in the direction Y_3. Thus the system is not locally configuration accessible. We may use the coordinates (θ, x, y) to render the system in the form specified by Theorem 5.21. We obtain

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

The top system is clearly locally configuration accessible and STLCC.

6.2 REMARKS:

1. In this example, in the cases when the system fails to satisfy the sufficient conditions for STLCC of Theorem 5.17, we are not able to say whether the system is, in fact, not STLCC. In fact, in PB4, even though the system does not satisfy the sufficient conditions of Theorem 5.17, it is easy to see that it *is* STLCC.

2. On a related note, in the robotic leg we saw that it was "Coriolis forces" which caused the loss of STLCC in RL2. In this example the metric is flat so the same explanation does not work. It would be interesting to ascertain why STLCC may be lost in the cases where the metric is flat.

3. The reader should verify that the decompositions given in PB3 and PB6 are in fact of the form guaranteed by Theorem 5.21.

4. 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.

5. The planar rigid body we presented in this section is an example of a class of systems whose configuration manifold is a Lie group, and the Riemannian metric and the input one-forms are left-invariant. In this case the control vector fields will also be left-invariant. We may choose a basis, $\{\xi_1, \ldots, \xi_n\}$, for the Lie algebra of the group. Corresponding to this basis will be a basis of left-invariant vector fields, $\{X_1, \ldots, X_n\}$, obtained by left translating the Lie algebra basis to each point in the group. The covariant derivative $\nabla_{X_i} X_j$ will also be a left-invariant vector field and so we may write $\langle X_i : X_j \rangle = \gamma_{ij}^k X_k$ for some set of constants γ_{ij}^k. Similarly we may write $[X_i, X_j] = c_{ij}^k X_k$ where the constants c_{ij}^k are the *structure constants* for the Lie algebra relative to the given basis. The conditions for local configuration accessibility and STLCC may then be expressed in terms of the constants γ_{ij}^k and c_{ij}^k. □

6.3. The Pendulum on a Cart In this section we study 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 5. 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

$$V = m a_g l(1 - \cos\theta)$$

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

$$F^1 = dx.$$

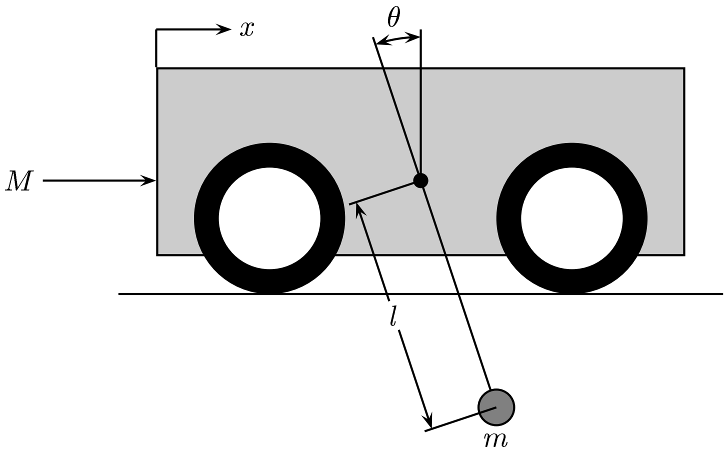

FIGURE 5. Pendulum suspended from a cart

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:

$$\langle Y_1 : Y_1\rangle = \frac{16m\cos^2\theta\sin\theta}{l(m+2M-m\cos 2\theta)^3}\frac{\partial}{\partial x} + \frac{8(M+m)\sin\theta}{l^2(m\cos 2\theta - m - 2M)^3}\frac{\partial}{\partial\theta},$$

$$\langle Y_1 : \operatorname{grad} V\rangle = \frac{4a_g m\cos\theta(m - m\cos 2\theta - 2M\cos 2\theta)}{l(m\cos 2\theta - m - 2M)^3}\frac{\partial}{\partial x} +$$
$$\frac{4a_g(2M^2\cos 2\theta + 3Mm\cos 2\theta + m^2\cos 2\theta - Mm - m^2)}{l^2(m\cos 2\theta - m - 2M)^3}\frac{\partial}{\partial\theta}.$$

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 $\{Y_1, \langle Y_1 : \operatorname{grad} V\rangle\}$ span T_qQ 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.

In determining these conditions, we came across a new geometric object: the symmetric product. Clearly a good understanding of the symmetric product will be an essential part of any further understanding of simple mechanical control systems. Nevertheless, from a computational point of view, the symmetric product is quite helpful.

In the examples in Section 6 some interesting circumstances may be observed. The most interesting of these is a comparison of the robotic leg in Case 2 and the planar rigid body in Case 4. In the former case the system does not satisfy the sufficient conditions for STLCC and is 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 5.5, of equilibrium controllability. If a system satisfies the hypotheses of Theorem 5.17 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. This work was partially funded by the Powell Foundation.

References

Bloch, Anthony M., and Crouch, Peter E. 1992. Kinematics and dynamics of nonholonomic control systems on Riemannian manifolds. *Pages 1–5 of: Proceedings of the 32nd IEEE Conference on Decision and Control.* Tucson, AZ: IEEE.

Bloch, Anthony M., Reyhanoglu, Mahmut, and McClamroch, N. Harris. 1992. Control and stabilization of nonholonomic dynamic systems. *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, **37**, 1746–1757.

Hermes, Henry. 1978. On local controllability. *SIAM Journal on Control and Optimization*, **20**, 211–220.

Lewis, Andrew D. 1995. *Aspects of Geometric Mechanics and Control of Mechanical Systems.* Ph.D. thesis, California Institute of Technology.

Lewis, Andrew D., and Murray, Richard M. 1995. *Equilibrium Controllability of a Class of Mechanical Systems.* Submitted to the 34th IEEE Conference on Decision and Control. Available electronically via `http://avalon.caltech.edu`.

Li, Zexiang, Montgomery, Richard, and Raibert, Mark. 1989. Dynamics and control of a legged robot in flight phase. *In: Proceedings of the 1989 IEEE International Conference on Robotics and Automation*. IEEE.

Murray, Richard M., and Sastry, Shankar S. 1993. Nonholonomic motion planning: Steering using sinusoids. *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, **38**(5), 700–716.

San Martin, L., and Crouch, Peter E. 1984. Controllability on principal fibre bundles with compact structure group. *Systems & Control Letters*, **5**, 35–40.

Serre, Jean-Pierre. 1992. *Lie Algebras and Lie Groups*. Lecture Notes in Mathematics, volume 1500. New York-Heidelberg-Berlin: Springer-Verlag.

Sussmann, Hector J. 1983. Lie brackets and local controllability: A sufficient condition. *SIAM Journal on Control and Optimization*, **21**, 686–713.

Sussmann, Hector J. 1987. A general theorem on local controllability. *SIAM Journal on Control and Optimization*, **25**(January), 158–194.