SIAM REVIEW
Vol. 41, No. 3, pp. 555–574

Configuration Controllability of Simple Mechanical Control Systems*

Andrew D. Lewis†
Richard M. Murray‡

Abstract. In this paper we present a definition of "configuration controllability" for mechanical systems whose Lagrangian is kinetic energy with respect to a Riemannian metric minus potential energy. A computable test for this new version of controllability is derived. This condition involves an object that 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.

Key words. mechanics, Riemannian geometry, controllability, symmetric product

AMS subject classifications. 53B20, 70H35, 70Q05, 93B03, 93B05, 93B29

PII. S0036144599351065

1. Introduction.

Motivated by applications in robotics, spacecraft dynamics, underwater vehicle dynamics, and other fields, there has been an recent upswell of interest in control theory for mechanical systems. Indeed, an upcoming special issue of *IEEE Transactions of Automatic Control* will be devoted to the subject. An early paper which suggested that such problems might be interesting is that of Brockett [5]. However, for the most part, Brockett's suggestions were not followed up aggressively by other researchers. When dealing with mechanical control systems, one wants to exploit the extra structure possessed by these systems. Just which structure one wishes to consider is, in a sense, a matter of taste. The Hamiltonian framework has received a great deal of attention and produces a "dual pair" interpretation of controllability decompositions. This theory is well-enough advanced to constitute a major part of Chapter 12 of [22]. With Hamiltonian control systems, one obviously wants to exploit the symplectic—or, more generally, Poisson—structure. In a Lagrangian framework, it is less clear what available structure ought best be utilized. A recent survey of Lagrangian control theory was provided by Murray [21]. A certain class of mechanical systems is invariant under the action of a Lie group, and this structure is employed by Bloch and Crouch [2] to obtain some controllability results. Here the

*Published electronically July 27, 1999. This paper originally appeared in *SIAM Journal on Control and Optimization*, Volume 35, Number 3, 1997, pages 766–790. This work was partially funded by the Powell Foundation, NSF grant CMS–9502224, and EPSRC grant GR/K 99893.
http://www.siam.org/journals/sirev/41-3/35106.html
†Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (andrew@mast.queensu.ca). While this paper was being prepared, this author enjoyed the support of a startup grant from Queen's University.
‡Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125 (murray@cds.caltech.edu).

authors rely on a result of San Martin and Crouch [26] concerning systems on principal fiber bundles. Systems with nonholonomic constraints are considered by Bloch, Reyhanoglu, and McClamroch [4]. Here the authors suppose that the inputs span a distribution complementary to the constraint distribution. With such an assumption one can essentially, by utilizing constraint and input forces, generate all motions compatible with the constraints. Systems with nonholonomic constraints *and* symmetry are considered by Ostrowski in joint work with Burdick [23, 24].

In this work we investigate, in the Lagrangian framework, "simple" mechanical systems that, by way of definition, are characterized by having "kinetic minus potential energy" Lagrangians. In the present communication of our results, we will simplify matters by supposing that the systems have no potential energy, a situation initially considered by Lewis and Murray [17]. Analogous results with the presence of potential are given by Lewis and Murray in [18], a paper that, further, and for the first time, thoroughly presents the methodology that we describe here. As we have suggested, the approach one takes to Lagrangian mechanical control systems reflects in large part the taste of the researcher. Our bias is toward a detailed consideration of the structure provided by the kinetic energy of a simple mechanical system. Let us be a bit more specific. One should think of kinetic energy as being provided by, and providing, a Riemannian metric on the system's configuration space. Associated with a Riemannian metric is a natural affine connection called the *Levi–Civita* connection. This affine connection may be used to succinctly write the equations of motion, as we shall see at the beginning of section 4. However, the value of the affine connection formalism goes far beyond this mundane and well-known virtue. Indeed, as Lewis and Murray [18] demonstrate, the Levi–Civita affine connection plays a fundamental role in the controllability analysis for simple mechanical control systems, even when potential energy is present. Interestingly, and motivated by work of Synge [30], Lewis [14] shows that the controllability analysis of [18] may be applied directly to simple mechanical systems with nonholonomic constraints linear in velocity.[1] We shall consider superficially an example of this type in section 5.3. In this case the Levi–Civita affine connection is replaced by a different affine connection that includes data from the nonholonomic constraints in its definition. We feel that all this, when combined with work of a somewhat different flavor, for example,[2] that of Rathinam and Murray [25], justifies the following statement:

> *Affine connections provide a valuable tool for studying simple mechanical control systems.*

It is to a justification of this statement that we devote this exposition.

2. Preliminary Statement of Results.

To gain a clear vision of where we are headed, it is perhaps useful to provide a preliminary statement of our results. We shall be somewhat more precise in sections 4.2 and 4.4. A truly precise formulation and proof of the results requires substantial development, and for this we refer to [18] and the dissertation of Lewis [12].

We begin with an example. Consider the planar rigid body system of Figure 2.1. On this body we consider two possible sets of forces. In one case we are able to apply a force in any direction to the body at a point away from the center of mass (case (a) in the figure). In the other case, we can only apply a force that is in a direction perpendicular to the line joining the point of application of the force with the center

[1] The first author wishes to acknowledge the work of [3] for motivating his interest in this approach.
[2] We refer to section 6 for a further discussion of related work.

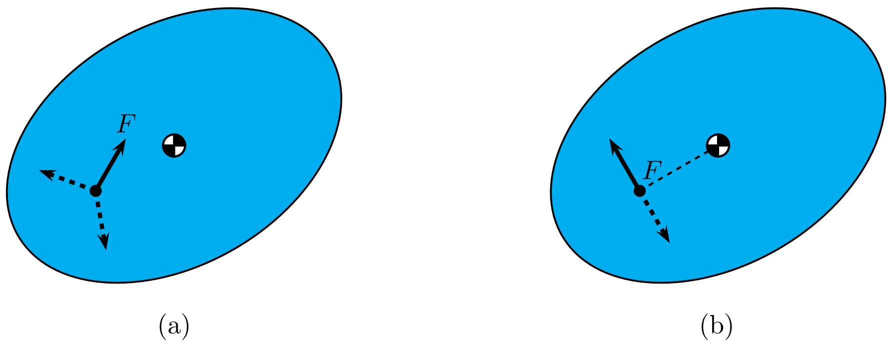

Fig. 2.1 *A planar rigid body with a variable direction thruster* (a) *and a fixed direction thruster* (b).

of mass (case (b) in the figure). The reader may wish to consider the former case as corresponding to having a thruster on the body whose direction may be varied, while in the second case the thruster can provide thrust only along a fixed line. In each of these cases one may ask certain questions about the controllability of this system. We list some of these questions below and in parentheses give the name of the general notion corresponding to each question.

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

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

Although we delay answering the above questions for the planar rigid body until section 5.2, we may state general results for a class of systems of which the planar rigid body is an example. Consistent with the outline of our approach in section 1, consider 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 $\mathcal{Y} = \{Y_1, \dots, Y_m\}$ on Q. We may define the *symmetric product* between two vector fields on Q by

$$\langle X : Y \rangle = \overset{g}{\nabla}_X Y + \overset{g}{\nabla}_Y X,$$

where $\overset{g}{\nabla}_X Y$ is the *covariant derivative* of Y with respect to X, taken with the Levi–Civita connection $\overset{g}{\nabla}$. If $\mathfrak{X}(Q)$ denotes the set of vector fields on Q, and if $\mathcal{V} \subset \mathfrak{X}(Q)$, we denote by $\overline{\mathrm{Sym}}(\mathcal{V})$ the distribution on Q obtained by taking iterated symmetric products of vector fields from $\mathcal{V}$. The usual involutive closure of $\mathcal{V}$ will be denoted $\overline{\mathrm{Lie}}(\mathcal{V})$. We shall say that a 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*. The *degree* of an iterated symmetric product of factors from $\mathcal{Y}$ will denote the total number of factors.

Notice that with the Lagrangian given by just kinetic energy, all states with zero velocity are equilibrium points for the unforced mechanical system. We shall say the system is *locally configuration accessible* at $q \in Q$ if the set of configurations 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 2.1. *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 $\mathcal{Y} = \{Y_1, \dots, Y_m\}$. Then*

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

To prove this result, one basically proceeds as follows. Compute the accessibility distribution on TQ for the mechanical control system and evaluate at zero velocity. This will describe the set of *states* accessible from points of zero velocity. However, since we are interested in controllability of the *configurations*, we can project the accessibility distribution to Q with $T\tau_Q$, the derivative of the tangent bundle projection. It turns out that this is exactly the distribution $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))$. In this way we see that the conditions in part (i) give local configuration accessibility. To prove part (ii), we apply the results of Sussmann [28] on local controllability to the systems we are considering.

The sections that follow formalize somewhat the above definitions and results. For a generalization to the case where the system has potential energy, see Lewis and Murray [18].

3. Machinery from Nonlinear Control Theory and Geometric Mechanics.

Our results bring together two fields: nonlinear control and geometric mechanics. Since the language of each field may be unfamiliar to researchers in the other, and since this paper is intended for a general audience, we present a brief review of applicable material from each subject. For a more thorough introduction to nonlinear control, we refer to [22], and for a thorough treatment of geometric mechanics, we refer to [1], especially section 3.7.

In this paper, "smooth" will mean analytic. Some of our results hold in the C^∞ category, but for all we say to be true, we need analyticity, so we make this a blanket assumption.

3.1. Nonlinear Control Theory.

In this section, let M be a finite-dimensional manifold and let $f_0, f_1, \dots, f_m$ be vector fields on M. We consider control systems of the form

$$\dot{x}(t) = f_0(x(t)) + u^a(t) f_a(x(t)). \tag{3.1}$$

We employ here the summation convention wherein summation over repeated raised and lowered indices is implied. The vector field f_0 is called the *drift* vector field and the vector fields $f_1, \dots, f_m$ are called the *control* or *input* vector fields. The m functions, $u^1, \dots, u^m$, are the *controls* or *inputs*. The idea is to design the inputs, as functions of x or t or both, to accomplish certain objectives. For example, one may wish to design the u^a's so as to make a point $x_0 \in M$ asymptotically stable. One typically specifies a class of allowable inputs when considering a control problem.

In this paper we shall denote by $\mathcal{U}$ the set of piecewise constant inputs and always suppose our inputs to be in this set. One may also consider inputs that are measurable and essentially bounded (for example).

As a first step in the analysis of a system of the form (3.1), one might wish to describe the set of reachable states. Let $x_0 \in M$, let V be a neighborhood of x_0, and let $T > 0$. We denote by $\mathcal{R}^V(x_0, T)$ the set of points that can be reached from x_0 in time T while remaining in V using inputs from $\mathcal{U}$. We also denote $\mathcal{R}^V(x_0, \leq T) = \cup_{t=0}^T \mathcal{R}^V(x_0, t)$. We say that the system (3.1) is *locally accessible* at x_0 if $\mathcal{R}^V(x_0, \leq T)$ contains a nonempty open subset of M for each V and for each T sufficiently small. Furthermore, we say that (3.1) is *small-time locally controllable* (STLC) if it is locally accessible and if x_0 is in the interior of $\mathcal{R}^V(x_0, \leq T)$ for each V and for each T sufficiently small.

Consulting Chapter 3 of [22], one sees that if the involutive closure of the vector fields $\{f_0, f_1, \dots, f_m\}$ has maximal rank at $x \in M$, then (3.1) is locally accessible at x. This condition is quite sharp. For analytic systems, it is necessary [29]. This condition is known as the local accessibility rank condition (LARC) at x.

Conditions for STLC of systems of the form (3.1) are difficult to obtain, and at the moment a useful statement of necessary and sufficient conditions is unavailable. However, a fairly strong sufficient condition is offered by Sussmann [28]. A precise statement of his results is beyond the scope of this paper. However, we can make use of a simpler result which we can state in a comprehensible, if not entirely precise, form.[3] A Lie bracket formed of combinations of vector fields from $\{f_0, f_1, \dots, f_m\}$ is *bad* if it contains an even number of each of the vector fields f_a, $a = 1, \dots, m$, and an odd number of f_0's. A like Lie bracket that is not bad is *good*. The *degree* of a bracket is the total number of vector fields of which it is comprised. This becomes clear with a few examples: the bracket $[[f_0, f_a], [f_0, f_b]]$ is good and of degree 4 for any $a, b \in \{1, \dots, m\}$, and the bracket $[f_a, [f_0, f_a]]$ is bad and of degree 3 for any $a \in \{1, \dots, m\}$. Let S_m denote the permutation group on m symbols. For $\pi \in S_m$ and B a Lie bracket of vector fields from $\{f_0, f_1, \dots, f_m\}$, define $\bar{\pi}(B)$ to be the bracket obtained by fixing f_0 and sending f_a to $f_{\pi(a)}$ for $a = 1, \dots, m$. Now define

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

We may state sufficient conditions for STLC.

THEOREM 3.1 (see Sussmann [28]). *Suppose that an analytic control system of the form* (3.1) *is such that every bad bracket B has the property that $\beta(B)(x)$ is an $\mathbb{R}$-linear combination of good brackets, evaluated at x, of lower degree than B. Also suppose that* (3.1) *satisfies the LARC at x. Then* (3.1) *is STLC at x.*

In practice, one comes up with a basis of vector fields made up of good brackets then checks to see that all bad brackets of degree not greater than the highest degree of a good bracket satisfy the hypothesis of the theorem.

3.2. Riemannian Geometry and Mechanics. A *Riemannian metric* on a manifold M is a smooth specification of an inner product on each tangent space of M. One may demonstrate that every manifold (with fairly weak topological hypotheses) possesses a Riemannian metric. More to the point, however, is the fact that Riemannian metrics are practically synonymous with simple mechanical systems. Indeed, if we let (x, v) denote natural coordinates for TM, then a kinetic energy function is nothing

[3] To make these statements precise, one needs the notion of a free Lie algebra (see [28] for details).

more than a function of (x, v), which is quadratic and positive definite in v. Since positive-definite quadratic forms are in one-to-one correspondence with inner products, this gives us the relationship between kinetic energy and a Riemannian metric. We shall denote a typical Riemannian metric by g.

Associated with a Riemannian metric is a natural affine connection. Let us first define what is meant by an affine connection in a general context. There are many excellent books to which one can refer for information on affine differential geometry. For example, the classic [11] presents an attractive approach. However, an excellent quick introduction may be found in section 2.7 of [1], and we shall distill this approach here. An *affine connection* assigns to each pair of vector fields X and Y on M a vector field $\nabla_X Y$, and this assignment satisfies the following properties.

AC1. The map $(X, Y) \mapsto \nabla_X Y$ is $\mathbb{R}$-bilinear.

AC2. $\nabla_{fX} Y = f\nabla_X Y$ for $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$.

AC3. $\nabla_X(fY) = f\nabla_X Y + (\mathcal{L}_X f)Y$ for $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$.

Here $\mathfrak{X}(M)$ denotes the set of vector fields on M, $C^\infty(M)$ denotes the set of smooth functions on M, and $\mathcal{L}_X f$ is the Lie derivative of f with respect to X. If we define $\nabla_X f = \mathcal{L}_X f$ for $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$, then we may extend ∇_X to a derivation on the entire tensor algebra on M. This means that we may define the covariant derivative $\nabla_X T$, where T is a tensor field of arbitrary type.

Locally an affine connection may be easily expressed. Let $(x^1, \ldots, x^n)$ be local coordinates for M, and for $i, j \in \{1, \ldots, n\}$ write

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij} \frac{\partial}{\partial x^k},$$

in this way defining n^3 local functions Γ^i_{jk}, $i, j, k = 1, \ldots, n$, called the *Christoffel symbols*. These functions are uniquely defined by the affine connection ∇ and the coordinates $(x^1, \ldots, x^n)$. The converse of this statement can be made true with the proviso that the functions should transform in a certain way when one changes from one chart to another. This transformation rule can be found in [1], but let us remark here that the Christoffel symbols are *not* the components of a $(1, 2)$ tensor field on M. An affine connection ∇ is *torsion free* if $\nabla_X Y - \nabla_Y X = [X, Y]$ for each $X, Y \in \mathfrak{X}(M)$.

If ∇ is an affine connection on M, a curve $c\colon [a, b] \to M$ is a *geodesic* for ∇ if $\nabla_{c'(t)} c'(t) = 0$. One must be careful how to interpret this equation since c' is not a vector field. However, when the appropriate care is taken, the condition for a curve $t \mapsto (x^1(t), \ldots, x^n(t))$ to be a geodesic takes the form

$$\ddot{x}^i + \Gamma^i_{jk} \dot{x}^j \dot{x}^k = 0, \quad i = 1, \ldots, n.$$

This is a second-order differential equation on M, and so it defines a first-order differential equation on TM. The vector field corresponding to this first-order differential equation is given in coordinates by

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

The vector field S is called the *geodesic spray* associated with the affine connection ∇.

Now we can assign to a Riemannian metric g an affine connection. We define $\overset{g}{\nabla}$ to be the unique torsion-free affine connection with the property that $\overset{g}{\nabla}_X g = 0$ for each vector field X. One may verify that this definition makes sense and implies that the Christoffel symbols are given in local coordinates by

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

This affine connection is known as the *Levi–Civita* connection, and [18] concerns itself solely with systems that utilize this affine connection. However, all the results stated hold for general affine connections. The relationship between the affine connection $\overset{g}{\nabla}$ and mechanics with the kinetic energy Lagrangian corresponding to g may be stated as follows:

> *The geodesics of the affine connection $\overset{g}{\nabla}$ are precisely the solutions of the Euler–Lagrange equations corresponding to the regular Lagrangian $v_x \mapsto \frac{1}{2} g(v_x, v_x)$.*

We shall use this correspondence to write the equations of motion for simple mechanical control systems in the next section.

The final object we need to discuss in Riemannian geometry seems innocuous enough but it plays a major role in the development of control theory for simple mechanical systems. Given two vector fields X and Y on M, and an affine connection ∇, we define their *symmetric product* to be the vector field $\langle X : Y \rangle = \nabla_X Y + \nabla_Y X$.

4. Controllability of Simple Mechanical Control Systems. As its title suggests, this section contains the important ideas in the paper. We begin by formulating the equations of motion for the systems we consider. We put the equations in the form of (3.1), so it becomes apparent how to treat the system as a nonlinear control system. However, we wish to ask questions that are germane to the special structure of mechanical control systems. In particular, we are interested only in initial states that have zero velocity and in the set of reachable configurations, rather than reachable states. This greatly simplifies the controllability analysis, as we shall see. We then turn to generating conditions for the special forms of controllability we consider. The approach we take in this paper is to make the results believable. We provide precise proofs in [18].

4.1. The Nonlinear Control Form of Equations of Motion for Simple Mechanical Control Systems. Let us first be precise about what systems we study. A *simple mechanical control system* is a quadruple $(Q, g, V, \mathcal{F})$, where (1) Q is a finite-dimensional (say, n-dimensional) manifold, (2) g is a Riemannian metric on Q, (3) V is a smooth function on Q, and (4) $\mathcal{F} = \{F^1, \dots, F^m\}$ is a collection of linearly independent one-forms on Q. The one-forms $\mathcal{F}$ form a basis for the available control forces. Consistent with our intentions expressed in the introduction, we shall take the potential function V to be zero unless otherwise stated. As we asserted in section 3.2, the equations of motion for the uncontrolled system are simply $\overset{g}{\nabla}_{c'(t)} c'(t) = 0$, the solutions of which are geodesics of the Levi–Civita connection. If one wishes to think in terms of Newtonian mechanics, where the governing equations are $ma = F$ (a is acceleration), then the term $\nabla_{c'(t)} c'(t)$ corresponds to a. Thus, for the forced equations, one should equate a with $\frac{1}{m} F$. This means that, rather than dealing directly with the forces $F^1, \dots, F^m$, we deal with the vector fields $Y_1, \dots, Y_m$, where, in coordinates, $Y_a^i = g^{ij} F_j^a$, with g^{ij} the components of the inverse of the matrix with components g_{ij}. We shall always deal directly with the vector fields $\mathcal{Y} = \{Y_1, \dots, Y_m\}$ rather than the one-forms $\mathcal{F}$. However, we wish to emphasize that forces are one-forms, not vector fields. In any event, the control equations may be written conveniently as

$$\nabla_{c'(t)}c'(t) = u^a(t)Y_a(c(t)). \tag{4.1}$$

We gain nothing by using the Levi–Civita connection, so we use a general affine connection ∇ in this equation. However, readers may wish always to think of ∇ as the Levi–Civita connection. In section 5 we shall consider one example where ∇ is *not* Levi–Civita.

Convenient though (4.1) may be, it is not in the form of (3.1). To convert it to this general control form, we need another bit of notation. Let X be a vector field on Q. The *vertical lift* of X is the vector field X^{lift} on TQ defined by

$$X^{\text{lift}}(v_q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (v_x + tX(x)).$$

In coordinates, if $X = X^i \frac{\partial}{\partial q^i}$, then $X^{\text{lift}} = X^i \frac{\partial}{\partial v^i}$. One may readily see, with a coordinate computation if necessary, that (4.1) is equivalent to the system

$$\dot{v}(t) = S(v(t)) + u^a(t)Y_a^{\text{lift}}(v(t)) \tag{4.2}$$

on TQ, where we recall that S is the geodesic spray associated with ∇. This equation is in the form of (3.1) with $f_0 = S$ and $f_a = Y_a^{\text{lift}}$, $a = 1, \ldots, m$. We are now in a position to perform a controllability analysis for the system (4.2), but first let us clearly state the notions of controllability we consider.

4.2. Controllability Definitions for Simple Mechanical Control Systems. It is possible simply to adopt the controllability definitions from nonlinear control theory, since our system may be written as a standard control system on TQ (this, after all, was the point of the previous section). However, since we are dealing with simple mechanical control 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 control theory for mechanical systems is, "What is the set of configurations that are reachable from a given configuration if we start at rest?" This is in fact exactly the question we pose.

Definition 4.1. A solution of (4.2) is a pair, (c, u), where $c\colon [0, T] \to Q$ is a piecewise smooth curve and $u \in \mathcal{U}$ such that (c', u) satisfies the first-order control system (4.2).

Note that since S is a second-order vector field on TQ, every solution of the control system (4.2) will be of the form (c', u) for some curve c on Q. We refer the reader to [1] for a discussion of second-order, and particularly Lagrangian, vector fields.

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

$$\begin{aligned}\mathcal{R}_Q^U(q_0, T) = \{q \in Q \mid\ & \text{there exists a solution } (c, u) \text{ of (4.2)} \\ & \text{such that } c'(0) = 0_{q_0},\ c(t) \in U \text{ for } t \in [0, T], \text{ and } c'(T) \in T_qQ\}\end{aligned}$$

and denote $\mathcal{R}_Q^U(q_0, \leq T) = \bigcup_{0 \leq t \leq T} \mathcal{R}_Q^U(q_0, t)$. Here 0_{q_0} is the zero tangent vector at q_0. 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 required, however, that the initial velocity be zero.

We now introduce our notions of controllability.

Definition 4.2. *We shall say that* (4.2) *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 nonempty 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 (4.2) *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 considered* STLCC.

We shall say that (4.2) *is* equilibrium controllable *if, for $q_1, q_2 \in Q$, there exists a solution (c, u) of* (4.2)*, 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.*

Remark 4.3.

1. Note that these definitions may be made to apply to any second-order control system that evolves on TQ.
2. Lewis and Murray [18], when considering systems with potential function V, define equilibrium controllability as the ability to steer between any two equilibrium points of the Lagrangian vector field corresponding to the Lagrangian $L(v_q) = \frac{1}{2} g(v_q, v_q) - V(q)$. Such equilibrium points occur exactly where $\boldsymbol{d}V = 0$. Thus, for systems without potential, all points in Q are equilibria, and so our notion here is consistent with that in [18].

4.3. The Structure of the Control Lie Algebra for Simple Mechanical Control Systems. Given our discussion in section 3.1, it seems reasonable that to derive conditions to test for the notions of controllability defined in the previous section, we would begin by looking at Lie brackets of vector fields from the set $\{S, Y_1^{\text{lift}}, \ldots, Y_m^{\text{lift}}\}$. This is indeed the correct thing to do because these calculations yield a great deal of structure. In this section we shall describe this structure, again making the assumption that the systems have no potential. The inclusion of potential makes the treatment [18] rather more involved than the one we give here.

Since we are interested only in initial states with zero velocity, we will be evaluating all brackets at such points. The $2n$-dimensional tangent space $T_{0_q}TQ$ admits a natural decomposition into the direct sum of two copies of T_qQ. This is accomplished as follows. The set $Z(TQ)$ of all zero vectors in TQ is an embedded submanifold of TQ that is naturally diffeomorphic to Q, with the diffeomorphism given by $0_q \mapsto q$. Thus the tangent space to $Z(TQ)$ at 0_q is a vector space that is naturally isomorphic to T_qQ. This gives us one part, which we call the *horizontal* part, in our proposed direct sum decomposition of $T_{0_q}TQ$. The other component in the direct sum decomposition comes from the fact that the tangent space to the fiber T_qQ, thought of as a submanifold of TQ, is naturally isomorphic to T_qQ by virtue of T_qQ being a vector space. Since the fiber T_qQ is transverse to $Z(TQ)$ at 0_q, this gives our natural decomposition $T_{0_q}TQ \simeq T_qQ \oplus T_qQ$ for each $q \in Q$. The first component we shall take to be the horizontal part, and we call the second component the *vertical* part. From now on, we may use this decomposition of $T_{0_q}TQ$ without warning.[4] Note that $Y_a^{\text{lift}}(0_q) = (0_q, Y_a(q))$ with respect to this decomposition.

Let us begin with a few example calculations that suggest how one might proceed. First, we immediately note that all brackets involving only the input vector fields $Y_1^{\text{lift}}, \ldots, Y_m^{\text{lift}}$ are identically zero. Also, $S(0_q)$ is zero (this, after all, is what it means for 0_q to be an equilibrium point of S). A few simple coordinate computations produce

[4]Given a second-order vector field X on TQ, it is possible to define, for each $v_q \in TQ$, a splitting $T_{v_q}TQ = T_qQ \oplus T_qQ$ which depends on X. If X is the geodesic spray associated with an affine connection, then this splitting agrees with the one we define when $v_q \in Z(TQ)$.

the following formulas:

$$[S, Y_a^{\text{lift}}](0_q) = (-Y_a(q), 0_q), \qquad [Y_a^{\text{lift}}, [S, Y_b^{\text{lift}}]](0_q) = (0_q, \langle Y_a : Y_b \rangle (q)),$$
$$[[S, Y_a^{\text{lift}}], [S, Y_a^{\text{lift}}]](0_q) = ([Y_a, Y_b](q), 0_q). \tag{4.3}$$

The second of these equalities, in fact, holds more generally; we have

$$[Y_a^{\text{lift}}, [S, Y_b^{\text{lift}}]] = (\langle Y_a : Y_b \rangle)^{\text{lift}}.$$

This suggests the importance of the symmetric product in our calculations. Indeed, the equalities (4.3) suggest that the accessibility algebra for (4.2), when evaluated at those states with zero velocity, maybe computable in terms of Lie brackets and symmetric products of vector fields from $\mathcal{Y} = \{Y_1, \ldots, Y_m\}$.

REMARK 4.4. A preliminary remark concerning generators for Lie algebras is helpful in simplifying the task of selecting which brackets to compute. If we have a set of vector fields $\{f_0, f_1, \ldots, f_m\}$, then any Lie bracket in these vector fields may be written as an $\mathbb{R}$-linear combination of brackets of the form

$$[X_1, [X_2, \ldots, [X_{k-1}, X_k]]], \tag{4.4}$$

where $X_\alpha \in \{f_0, f_1, \ldots, f_m\}$, $\alpha = 1, \ldots, k$. One proves this by induction and by using the Jacobi identity.

A few moments' consideration of (4.3) suggests how one might proceed to compute higher order brackets. To organize the calculations, it is convenient to introduce some notation. If B is a bracket formed from vector fields in $\mathcal{X} = \{S, Y_1^{\text{lift}}, \ldots, Y_m^{\text{lift}}\}$, then we denote by $\delta_0(B)$ the number of occurrences of S in B, and by $\delta_a(B)$ the number of occurrences of Y_a^{lift} in B for $a \in \{1, \ldots, m\}$. Let us denote by $\text{Br}_k(\mathcal{X})$ the set of brackets B in $\mathcal{X}$ for which

$$\delta_0(B) - \sum_{a=1}^{m} \delta_a(B) = k.$$

Thus $\text{Br}_k(\mathcal{X})$ is composed of brackets in which S appears k times more often than all the input vector fields combined.[5] Now we introduce the idea of the components of a bracket B formed from the vector fields $\mathcal{X}$. Any such bracket will itself be a bracket of two other brackets: $B = [B_1, B_2]$. One can then write $B_\alpha = [B_{\alpha 1}, B_{\alpha 2}]$ for $\alpha = 1, 2$, and may carry on this way until we end up with elements from $\mathcal{X}$. The collection of brackets $B_1, B_2, B_{11}, B_{12}, B_{21}, B_{22}, \ldots$ is called the *components* of B. A bracket B is called *primitive* if all of its components are brackets in $\text{Br}_{-1}(\mathcal{X}) \cup \text{Br}_0(\mathcal{X}) \cup \{S\}$.

It is perhaps illustrative to write a few primitive brackets so we know what they look like. Here is a list of the primitive brackets up to degree 4:

Degree 1: $\{Y_a^{\text{lift}} \mid a = 1, \ldots, m\}$,
Degree 2: $\{[S, Y_a^{\text{lift}}] \mid a = 1, \ldots, m\}$,
Degree 3: $\{[Y_a^{\text{lift}}, [S, Y_b^{\text{lift}}]] \mid a, b = 1, \ldots, m\}$,
Degree 4: $\{[S, [Y_a^{\text{lift}}, [S, Y_b^{\text{lift}}]]] \mid a, b = 1, \ldots, m\} \cup$
$\{[[S, Y_a^{\text{lift}}], [S, Y_b^{\text{lift}}]] \mid a, b = 1, \ldots, m\}$.

[5]The reader with even a mild tendency to pedantry is perhaps becoming uncomfortable with our unclear use of word "bracket" here. This is because, to make it clear, one needs to use free Lie algebras, as is done in [18].

It turns out that primitive brackets are the only brackets one needs to consider. The reasoning behind this goes as follows. One can show with an inductive calculation that all brackets B in $\mathrm{Br}_k(B)$, $k \leq 2$, are identically zero. Examples of such brackets include brackets that involve only the input vector fields. One may prove the following lemma by induction using the Jacobi identity.

LEMMA 4.5. *If $\mathcal{X}$ has the property that any bracket in $\mathrm{Br}_k(\mathcal{X})$, $k \leq 2$, is identically zero, then any bracket in $\mathrm{Br}_0(\mathcal{X}) \cup \mathrm{Br}_{-1}(\mathcal{X})$ is a finite sum of primitive brackets.*

As we have already asserted the hypotheses of the lemma, its conclusion must follow, and so all brackets in $\mathrm{Br}_0(\mathcal{X}) \cup \mathrm{Br}_{-1}(\mathcal{X})$, no matter where they are evaluated, are finite linear combinations of primitive brackets. For example, one may use the Jacobi identity to verify that

$$[Y_a^{\mathrm{lift}}, [S, [S, Y_b^{\mathrm{lift}}]]] = [[S, Y_b^{\mathrm{lift}}], [S, Y_a^{\mathrm{lift}}]] + [S, [Y_a^{\mathrm{lift}}, [S, Y_b^{\mathrm{lift}}]]].$$

The bracket on the left is not primitive, but it is the sum of two brackets that are.

This takes care of the brackets in $\mathrm{Br}_k(\mathcal{X})$, $k \leq 0$: they are either identically zero or a sum of primitive brackets. But what about the other brackets? They are *not*, it turns out, identically zero. However, they *are* zero when evaluated on $Z(TQ)$. This is because the local coordinate expressions for such vector fields produce components that are at least linear in the velocity variables. Bullo [7] explains this in terms of homogeneity.

So now we are at the point where the only brackets we need to consider for evaluation on $Z(TQ)$ are primitive brackets. By Remark 4.4, we need only consider those primitive brackets of the form (4.4). Given this, it becomes important to know just what such primitive brackets actually look like. We take our lead from the computations (4.3). Let us make a few preliminary observations based on these calculations. Primitive brackets in $\mathrm{Br}_{-1}(\mathcal{X})$ are vertical, and those in $\mathrm{Br}_0(\mathcal{X})$ are horizontal when evaluated on $Z(TQ)$. Note that primitive brackets in $\mathrm{Br}_{-1}(\mathcal{X})$ (and so all brackets in $\mathrm{Br}_{-1}(\mathcal{X})$, by Lemma 4.5) are vertical (in the sense that they vanish under the application of $T\tau_Q$) even at points away from $Z(TQ)$. In fact, primitive brackets in $\mathrm{Br}_{-1}(\mathcal{X})$ are exactly vertical lifts of symmetric products of vector fields in $\mathcal{Y}$. The precise meaning of this statement is made clear by a few examples to augment the second equality of (4.3):

$$(4.5)\qquad \begin{aligned} [Y_a^{\mathrm{lift}}, [[S, Y_b^{\mathrm{lift}}], [S, Y_c^{\mathrm{lift}}]]] &= (\langle Y_a : \langle Y_b : Y_c \rangle\rangle)^{\mathrm{lift}}, \\ [[Y_a^{\mathrm{lift}}, [S, Y_b^{\mathrm{lift}}]], [S, [Y_c^{\mathrm{lift}}, [S, Y_d^{\mathrm{lift}}]]]] &= (\langle\langle Y_a : Y_b \rangle : \langle Y_c : Y_d \rangle\rangle)^{\mathrm{lift}}. \end{aligned}$$

From a close examination of these examples, we hope it is clear how, at least in symbols, one may write the correspondence between primitive brackets in $\mathrm{Br}_{-1}(\mathcal{X})$ and symmetric products in $\mathcal{Y}$. We denote by $\overline{\mathrm{Sym}}(\mathcal{Y})$ the distribution obtained by closing the input distribution under symmetric product.

Let us follow a similar methodology to describe the appearance of primitive brackets in $\mathrm{Br}_0(\mathcal{X})$. That is, we shall provide a few examples and refer the reader to [18] and the dissertation [12] for details. One may verify the following equalities:

$$(4.6)\qquad \begin{aligned} [S, [Y_a^{\mathrm{lift}}, [[S, Y_b^{\mathrm{lift}}], [S, Y_c^{\mathrm{lift}}]]]](0_q) &= (-\langle Y_a : \langle Y_b : Y_c \rangle\rangle(q), 0_q), \\ [[S, [Y_a^{\mathrm{lift}}, [S, Y_b^{\mathrm{lift}}]]], [S, [Y_c^{\mathrm{lift}}, [S, Y_d^{\mathrm{lift}}]]]](0_q) &= ([\langle Y_a : Y_b \rangle, \langle Y_c : Y_d \rangle](q), 0_q). \end{aligned}$$

Thus one gleans that all primitive brackets in $\mathrm{Br}_0(\mathcal{X})$ give symmetric products in $\mathcal{Y}$, as do all Lie brackets between these symmetric products. We denote the distribution generated in this way by $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))$.

Interestingly, the drift vector field vanishes from the formulas (4.5) and (4.6), its role being taken up by the symmetric product.

To summarize the point of this section, we have the following result which is central to our methodology.

PROPOSITION 4.6. $\overline{\mathrm{Lie}}(\mathcal{X})_{0_q} = \overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))_q \oplus \overline{\mathrm{Sym}}(\mathcal{Y})_q$.

Roughly speaking, one can regard $\overline{\mathrm{Sym}}(\mathcal{Y})_q$ as the velocity directions that are accessible from 0_q, and $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))_q$ as the configuration directions accessible from 0_q.

4.4. Controllability Results for Simple Mechanical Control Systems. Since, by our discussion in section 3.1, local accessibility of (4.2) at 0_q is determined by computing the involutive closure of $\mathcal{X}$ at 0_q, from Proposition 4.6 we immediately ascertain that (4.2) is locally accessible at 0_q if $\overline{\mathrm{Sym}}(\mathcal{Y})_q$ has the dimension of Q. But this is *not* necessary for local *configuration* accessibility. Indeed, given that the horizontal component of $\overline{\mathrm{Lie}}(\mathcal{X})_{0_q}$ is $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))_q$, the following result is the obvious one to guess and is in fact correct.

THEOREM 4.1. *The control system* (4.2) *is locally configuration accessible at q if* $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))_q = T_qQ$.

The hypotheses of the theorem are necessary for analytic systems by virtue of the results of Sussmann and Jurdjevic [29]. For smooth systems, the conditions are necessary in that if (4.2) is locally configuration accessible at every $q \in Q$, then the hypotheses of Theorem 4.1 hold on an open, dense subset of Q.

REMARK 4.7. There are examples that are locally configuration accessible but not locally accessible (see section 5.1). Thus our controllability definitions are genuinely weaker than the standard ones.

It is a similarly simple matter to use our hard work of section 4.3 to adapt Theorem 3.1 to give a result for STLCC. If P is a symmetric product in the vector fields $\mathcal{Y}$,[6] we let $\gamma_a(P)$ denote the number of occurrences of Y_a in P, and we define the *degree* of P by $\gamma_1(P) + \cdots + \gamma_m(P)$. 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 a symmetric product in the vector fields $\mathcal{Y}$, define $\bar{\pi}(P)$ to be the bracket obtained by sending Y_a to $Y_{\pi(a)}$ 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.

THEOREM 4.2. *Suppose that $\mathcal{Y}$ is such that every bad symmetric product P in $\mathcal{Y}$ has the property that*

$$\rho(P)(q) = \sum_{a=1}^{m} \xi^a C_a(q),$$

where C_a are good symmetric products in $\mathcal{Y}$ of lower degree than P and $\xi^a \in \mathbb{R}$ for $a = 1, \ldots, m$. Also, suppose that $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathcal{Y}))_q$ has the dimension of Q. Then (4.2) *is STLCC at q.*

[6]To be precise when talking about "brackets" we really need to use free Lie algebras; to be precise about "symmetric products" we need to use free symmetric algebras, as we did in [18].

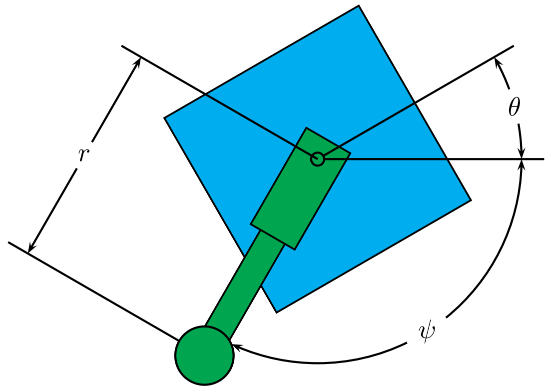

Fig. 5.1 *The robotic leg.*

REMARK 4.8.

1. The proof of this result follows from Theorem 3.1 and an examination of the bracket computations of section 4.3: one observes a one-to-one correspondence between bad brackets in $\mathcal{X}$ (when evaluated on $Z(TQ)$) and bad symmetric products in $\mathcal{Y}$.
2. A closer examination of the proof of Theorem 4.2 reveals the remarkable fact that if the hypotheses of the theorem hold at all points in Q, then (4.2) is in fact *equilibrium controllable*. This, it turns out, is a consequence of the system being STLC on the set of reachable *states* if the hypotheses are satisfied on all of Q.

5. 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 condition for local configuration accessibility given in Theorem 4.1 is that it lends itself to symbolic computation. Indeed, a Mathematica package was written to facilitate the computations in this section. All the examples we consider here are without potential. For a simple example with potential, see [18].

It is worth emphasizing that for each of these examples, and indeed for *all* examples of the form (4.2), the linearization at points of zero velocity is not controllable.

5.1. The Robotic Leg. This example, although simple, exhibits much of the behavior that makes the study of mechanical systems interesting. The example is a rigid body with inertia J that is pinned to the ground at its center of mass. The body has attached to it an extensible massless leg and the leg has a point mass with mass m at its tip. The coordinate θ describes the angle of the body and ψ describes the angle of the leg from an inertial reference frame. The coordinate r describes the extension of the leg. Thus the configuration space for this problem is $Q = \mathbb{T}^2 \times \mathbb{R}^+$ (see Figure 5.1). In the coordinates (θ, ψ, r) the Riemannian metric for the robotic leg is

$$g = J\mathrm{d}\theta \otimes \mathrm{d}\theta + mr^2\mathrm{d}\psi \otimes \mathrm{d}\psi + m\mathrm{d}r \otimes \mathrm{d}r,$$

Table 5.1 *Controllability results for the robotic leg. The first column displays which inputs are present, the second column indicates whether the system is locally configuration accessible with these inputs, the third column indicates whether the system with these inputs satisfies the sufficient conditions of Theorem 4.2 for STLCC, and the last column indicates whether the system with these inputs is actually STLCC.*

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

and the input one-forms are $F^1 = \mathrm{d}\theta - \mathrm{d}\psi$ and $F^2 = \mathrm{d}r$. We may compute the input vector fields to be

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

We find the following computations to be sufficient:

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

The controllability results for the robotic leg are displayed in Table 5.1.

Remark 5.1. Although the system only violates the *sufficient* conditions for STLCC with the input Y_1, one may easily determine that the system is, in fact, not STLCC. The reason for this is that, because of "centrifugal force," or whatever may be your favorite name for the related phenomenon, r will 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.

5.2. The Forced Planar Rigid Body. In this section we study the planar rigid body discussed in the introduction with various combinations of forces and torques. The configuration space for the system is the Lie group $SE(2)$. To establish the correspondence between the configuration of the body and $SE(2)$, fix a point $O \in \mathbb{R}^2$ and let $\{\boldsymbol{e}_1 = \frac{\partial}{\partial x}, \boldsymbol{e}_2 = \frac{\partial}{\partial y}\}$ 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 center 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 center of mass of the body with its frame $\{\boldsymbol{f}_1, \boldsymbol{f}_2\}$ (see Figure 5.2). The inputs for this problem consist of forces applied at an arbitrary point and a torque about the center 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 center of mass. The situation is illustrated in Figure 5.3.

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

$$g = m\mathrm{d}x \otimes \mathrm{d}x + m\mathrm{d}y \otimes \mathrm{d}y + J\mathrm{d}\theta \otimes \mathrm{d}\theta.$$

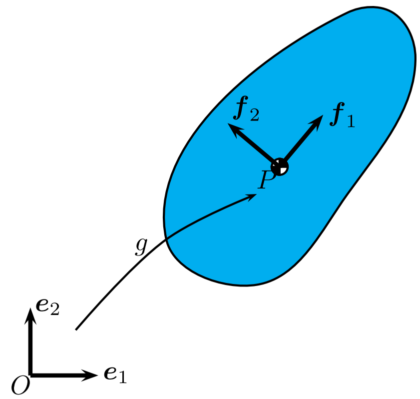

Fig. 5.2 *The configuration of a planar body as an element of SE(2).*

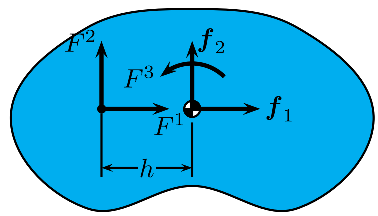

Fig. 5.3 *Positions for application of forces on a planar rigid body after simplifying assumptions.*

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

$$F^1 = \cos\theta \mathrm{d}x + \sin\theta \mathrm{d}y, \qquad F^2 = -\sin\theta \mathrm{d}x + \cos\theta \mathrm{d}y - h\mathrm{d}\theta, \qquad F^3 = \mathrm{d}\theta,$$

from which we compute the input vector fields as

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

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

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

With the computations done, we may proceed to determine configuration controllability for the planar rigid body with various combinations of inputs. The results are displayed in Table 5.2.

REMARK 5.2. For this example, in the cases when the system fails to satisfy the sufficient conditions for STLCC of Theorem 4.2, we are not able to say immediately whether the system is, in fact, not STLCC; further analysis is required.

1. When the inputs Y_2 and Y_3 are present, although the system does not satisfy the sufficient conditions of Theorem 4.2, one may readily show that it *is*

Table 5.2 *Controllability results for the planar rigid body. The first column displays which inputs are present, the second column indicates whether the system is locally configuration accessible with these inputs, the third column indicates whether the system with these inputs satisfies the sufficient conditions of Theorem 4.2 for STLCC, and the last column indicates whether the system with these inputs is actually STLCC.*

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

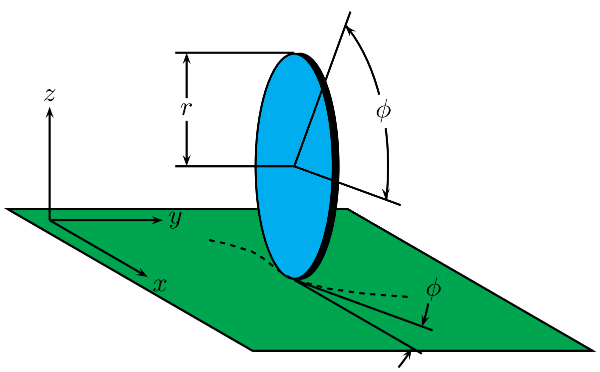

Fig. 5.4 *The rolling disk.*

STLCC. To do this, one makes a feedback transformation which changes the system into one that satisfies the hypotheses of Theorem 4.2.

2. When one has only the input Y_2 available, things are a bit less trivial. Nevertheless, the analysis of Lewis [13], following Sussmann [27], shows that the system is *not* STLCC.

5.3. The Upright Rolling Disk. Now we sketch an example for which we have not presented a means for writing the equations of motion in the form of (4.2). Nevertheless, the equations *are* of this form [14]. We shall simply write an affine connection whose geodesics, when restricted to the appropriate initial conditions, are the unforced solutions. We present this example to reinforce the utility of using a general geodesic spray and general vertically lifted vector fields in (4.2).

The example we consider is one with nonholonomic constraints. It is an upright rolling disk as depicted in Figure 5.4 and has $Q = SE(2) \times \mathbb{S}^1$ as its configuration manifold. The system has its natural kinetic energy defined by the Riemannian metric

$$g = m\mathrm{d}x \otimes \mathrm{d}x + m\mathrm{d}y \otimes \mathrm{d}y + J\mathrm{d}\theta \otimes \mathrm{d}\theta + I\mathrm{d}\phi \otimes \mathrm{d}\phi.$$

Here $m > 0$ is the mass of the disk, $I > 0$ is the moment of inertia of the disk about its center, and $J > 0$ is the moment of inertia of the disk about the "z-axis." However, the equations of motion are *not* the geodesics of the corresponding Levi–Civita connection. This is a consequence of the fact that the system is constrained.

Indeed, the condition that the disk roll without slipping is modeled by declaring that the velocities satisfy the relations

$$\dot{x} = r\cos\theta\dot{\phi}, \qquad \dot{y} = r\sin\theta\dot{\phi}.$$

It turns out that the constrained equations of motion, in accordance with the Lagrange–d'Alembert principle, are those geodesics, whose initial conditions satisfy the constraints, of a certain affine connection.[7] The affine connection has Christoffel symbols

$$\Gamma^x_{x\theta} = \frac{mr^2\sin 2\theta}{I+mr^2}, \qquad \Gamma^x_{y\theta} = -\frac{mr^2\cos 2\theta}{I+mr^2}, \qquad \Gamma^x_{\phi\theta} = \frac{Ir\sin\theta}{I+mr^2},$$
$$\Gamma^y_{x\theta} = -\frac{mr^2\cos 2\theta}{I+mr^2}, \qquad \Gamma^y_{y\theta} = -\frac{mr^2\sin 2\theta}{I+mr^2}, \qquad \Gamma^y_{\phi\theta} = -\frac{Ir\cos\theta}{I+mr^2},$$
$$\Gamma^\phi_{x\theta} = \frac{mr\sin\theta}{I+mr^2}, \qquad \Gamma^\phi_{y\theta} = -\frac{mr\cos\theta}{I+mr^2}.$$

This system has two natural inputs: a torque that makes the disk roll, and a torque that makes the disk spin. These inputs are modeled by the one-forms $F^1 = \mathrm{d}\phi$ and $F^2 = \mathrm{d}\theta$, and the input vector fields associated with these forces are

$$Y_1 = \frac{1}{I+mr^2}\left(r\cos\theta\frac{\partial}{\partial x} + r\sin\theta\frac{\partial}{\partial y} + \frac{\partial}{\partial\phi}\right), \qquad Y_2 = \frac{1}{J}\frac{\partial}{\partial\theta}.$$

Note that these vector fields are *not* obtained just by multiplying the force one-forms by the "inverse" of g. The theory outlined by Lewis [14] asks that we further g-orthogonally project these vector fields to the distribution D. The details are of no real consequence here; the point is that the upright rolling disk is a control system of the form (4.2).

We now perform the symmetric product and Lie bracket computations necessary to make conclusions about the controllability of the system. We compute

$$\langle Y_1 : Y_1\rangle = 0, \qquad \langle Y_1 : Y_2\rangle = 0, \qquad \langle Y_2 : Y_2\rangle = 0,$$
$$[Y_1, Y_2] = \frac{r}{J(I+mr^2)}\left(\sin\theta\frac{\partial}{\partial x} - \cos\theta\frac{\partial}{\partial y}\right),$$
$$[Y2, [Y1, Y2]] = \frac{r}{J^2(I+mr^2)}\left(\cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}\right).$$

We may now easily deduce some basic facts about the controllability of the upright rolling disk, and the results are displayed in Table 5.3.

6. Subsequent and Future Work. In this paper we were primarily concerned with presenting the essential features of the program initiated by the authors in [18]. In doing so, we have made passing reference to work that utilizes the results and methodology in that paper. Let us here summarize these contributions and mention some that we might have omitted.

The results of Lewis and Murray [18] provided a practical approach to controllability theory for simple mechanical control systems. However, they suggested a question whose answer was unknown at the time of that paper's publication: What is the "meaning" of the symmetric product? The answer is to be found in [15] and is

[7] Actually, there are many affine connections which will serve here.

Table 5.3 *Controllability results for the upright rolling disk. The first column displays which inputs are present, the second column indicates whether the system is locally configuration accessible with these inputs, the third column indicates whether the system with these inputs satisfies the sufficient conditions of Theorem* 4.2 *for STLCC, and the last column indicates whether the system with these inputs is actually STLCC.*

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

quite simple and revealing. Let D be a distribution on a manifold Q with an affine connection ∇. D is *geodesically invariant* under ∇ if for each geodesic $c\colon [a,b] \to Q$, $c'(a) \in D_{c(a)}$ implies that $c'(t) \in D_{c(t)}$ for $t \in (a,b]$. Lewis [15] shows that D is geodesically invariant if and only if $\langle X : Y \rangle$ is a section of D for all vector fields X and Y taking values in D. Thus the symmetric product performs for geodesically invariant distributions the same task the Lie bracket performs for integrable distributions. This interpretation is employed in [19] to describe a decomposition for the systems we consider in this paper.

As mentioned in the introduction, and assumed by the example of section 5.3, systems with nonholonomic constraints have equations of motion whose solutions are geodesics of a certain affine connection. This reinforces our view that the proper abstraction for the class of mechanical systems we consider is a system of the form (4.2), with S the geodesic spray of an arbitrary affine connection, and $Y_1, \dots, Y_m$ arbitrary vector fields on Q (i.e., not necessarily obtained from one-forms, as we describe in section 4.1). This is the approach taken by the authors [19] and by Lewis [14]. It is interesting to note that, at this point, there is actually nothing in the theory that distinguishes the results for Levi–Civita affine connections from those for general affine connections.

Our main controllability result, Theorem 4.2, is a sufficient condition. This suggests that further work might sharpen this condition. An example of when this may be done is in the single-input case [13]. In this case—and here it is essential that the systems are without potential—one may show that a single-input simple mechanical control system is STLCC if and only if $\dim(Q) = 1$, i.e., only in the trivial case when the system is fully actuated. This, for example, reveals that the planar body with the input perpendicular to the line joining the point of application of the force with the center of mass is not STLCC. This result allows Lynch and Mason [20] to prove the necessity of three unilateral forces to "dynamically grasp" a planar object. Lynch and Mason also use our multi-input sufficient condition, Theorem 4.2.

The single-input result referred to above, while seemingly innocuous, is perhaps suggestive of something nontrivial about simple mechanical control systems. The essential point of interest is that we have necessary and sufficient conditions for STLCC of simple mechanical control systems, in the absence of potential, with a single input. Results of this strength are not available for general single-input control systems (a fairly strong result is proved by Sussmann [27]), and this suggests that simple mechanical control systems have a very structured control Lie algebra—certainly the computations of section 4.3 bear this out. Perhaps it is possible to provide computable necessary and sufficient conditions for STLCC for multi-input simple mechanical control systems.

Our results provide a starting point for the analysis of a simple mechanical control system: if a system is not controllable, certain control tasks become impossible. However, our results go nowhere toward addressing the essential problems of controller design. Interestingly, Bullo, Leonard, and Lewis [8] provide a synthesis method that is applicable to invariant systems in Lie groups. (The planar rigid body of section 5.2 is a system of this type.) Here one uses averaging theory, along with the controllability conditions of Theorem 4.2, to design control laws to perform certain tasks. Systems without potential energy possess an interesting feature: while the lack of potential makes for easier statements of controllability results, it greatly increases the difficulty of control design. This is reflected, for example, by the fact that the absence of potential guarantees that linear control design methods are inapplicable. Another example of the difficulty that one encounters in control synthesis is the fact that asymptotic stabilization of an equilibrium point under continuous state feedback is impossible by a result of Brockett [6], and exponential stabilization is impossible with smooth, time-dependent feedback. Exponential stabilizers that are continuous and time dependent are provided in [8].

At this point we would like to emphasize that methods designed for trajectory generation for "nonholonomic" (i.e., driftless) control systems are *not* generally applicable to the systems we consider. That they are in some cases (for example, the leg of section 5.1 with both inputs) is a consequence of a special relationship between the inputs and the affine connection, as explained by Lewis [16].

Another approach to trajectory generation uses "differential flatness" as introduced by Fliess et al. [10]. The work of Rathinam and Murray [25] uses affine connections to describe conditions for "configuration flatness" for a class of simple mechanical control systems. Other work in mechanical control systems that utilises a Riemannian geometry framework includes that of Bullo and Murray [9].

Acknowledgments. We would like to thank Jerry Marsden and Jim Ostrowski for helpful conversations during the preparation of [18]. Francesco Bullo also deserves our appreciation for utilizing the methodology described here in synthesising controllers for simple mechanical systems. His work with Naomi Leonard has done much to vindicate our approach. We also express our sincere gratitude to the editors of *SIAM Review* for providing us with the opportunity to communicate our results to a general audience and possibly to obtain new subscribers to our view that mechanical control systems are interesting objects to study.

REFERENCES

[1] R. Abraham and J. E. Marsden, *Foundations of Mechanics*, 2nd ed., Addison-Wesley, Reading, MA, 1978.

[2] A. M. Bloch and P. E. Crouch, *Kinematics and dynamics of nonholonomic control systems on Riemannian manifolds*, in Proceedings of the 32nd IEEE Conference on Decision and Control, Tucson, AZ, 1992, pp. 1–5.

[3] A. M. Bloch and P. E. Crouch, *Another view of nonholonomic mechanical control systems*, in Proceedings of the 34th IEEE Conference on Decision and Control, New Orleans, LA, 1995, pp. 1066–1071.

[4] A. M. Bloch, M. Reyhanoglu, and N. H. McClamroch, *Control and stabilization of nonholonomic dynamic systems*, IEEE Trans. Automat. Control, 37 (1992), pp. 1746–1757.

[5] R. W. Brockett, *Control theory and analytical mechanics*, in Geometric Control Theory, C. Martin and R. Hermann, eds., Math Sci Press, Brookline, MA, 1977, pp. 1–48.

[6] R. W. Brockett, *Asymptotic stability and feedback stabilization*, in Differential Geometric Control Theory, R. W. Brockett, R. S. Millman, and H. J. Sussmann, eds., Birkhäuser, Boston, Basel, Stuttgart, 1983, pp. 181–191.

[7] F. Bullo, *A series describing the evolution of mechanical control systems*, Proceedings of IFAC '99, Beijing, China, International Federation of Automatic Control, submitted.
[8] F. Bullo, N. E. Leonard, and A. D. Lewis, *Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups*, IEEE Trans. Automat. Control, to appear.
[9] F. Bullo and R. M. Murray, *Tracking for fully actuated mechanical systems: A geometric framework*, Automatica, 1999 (35), pp. 17–34.
[10] M. Fliess, J. Lévine, P. Martin, and P. Rouchon, *Sur les systèmes non linéaires différentiellement plats*, C. R. Acad. Sci. Sér. I. Math., 315 (1992), pp. 619–624.
[11] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Interscience Tracts in Pure and Appl. Math. 15, Interscience, New York, 1963.
[12] A. D. Lewis, *Aspects of Geometric Mechanics and Control of Mechanical Systems*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1995.
[13] A. D. Lewis, *Local configuration controllability for a class of mechanical systems with a single input*, in Proceedings of the European Control Conference, Brussels, Belgium, European Control Systems Society, 1997.
[14] A. D. Lewis, *Simple mechanical control systems with constraints*, IEEE Trans. Automat. Control, to appear.
[15] A. D. Lewis, *Affine connections and distributions with applications to nonholonomic mechanics*, Rep. Math. Phys., 42 (1998), pp. 135–164.
[16] A. D. Lewis, *When Is a Mechanical Control System Kinematic?* preprint, 1998.
[17] A. D. Lewis and R. M. Murray, *Configuration controllability of a class of mechanical systems*, in Proceedings of the 34th IEEE Conference on Decision and Control, New Orleans, LA, 1995, pp. 1–5.
[18] A. D. Lewis and R. M. Murray, *Configuration controllability of simple mechanical control systems*, SIAM J. Control Optim., 35 (1997), pp. 766–790.
[19] A. D. Lewis and R. M. Murray, *Decompositions of control systems on manifolds with an affine connection*, Systems Control Lett., 31 (1997), pp. 199–205.
[20] K. M. Lynch and M. T. Mason, *Dynamic nonprehensile manipulation: Controllability, planning, and experiments*, Internat. J. of Robotics Research, 1999 (18), pp. 64–92.
[21] R. M. Murray, *Nonlinear control of mechanical systems: A Lagrangian perspective*, in IFAC Symposium on Nonlinear Control Systems (NOLCOS), Lake Tahoe, CA, International Federation of Automatic Control, 1995, pp. 378–389.
[22] H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, Heidelberg, Berlin, 1990.
[23] J. P. Ostrowski, *The Mechanics and Control of Undulatory Robotic Locomotion*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1995.
[24] J. P. Ostrowski and J. W. Burdick, *Controllability tests for mechanical systems with constraints and symmetries*, J. Appl. Math. Comput. Sci., 7 (1997), pp. 101–127.
[25] M. Rathinam and R. M. Murray, *Configuration flatness of Lagrangian systems underactuated by one control*, SIAM J. Control Optim., 36 (1998), pp. 164–179.
[26] L. San Martin and P. E. Crouch, *Controllability on principal fibre bundles with compact structure group*, Systems Control Lett., 5 (1984), pp. 35–40.
[27] H. J. Sussmann, *Lie brackets and local controllability: A sufficient condition for scalar-input systems*, SIAM J. Control Optim., 21 (1983), pp. 686–713.
[28] H. J. Sussmann, *A general theorem on local controllability*, SIAM J. Control. Optim., 25 (1987), pp. 158–194.
[29] H. J. Sussmann and V. Jurdjevic, *Controllability of nonlinear systems*, J. Differential Equations, 12 (1972), pp. 95–116.
[30] J. L. Synge, *Geodesics in nonholonomic geometry*, Math. Annal., 99 (1928), pp. 738–751.