Simple mechanical control systems with constraints

Andrew D. Lewis*

01/03/1998
Last updated: 02/03/1999

Abstract

We apply some recently developed control theoretic techniques to the analysis of a class of mechanical systems with constraints. Certain simple aspects of the theory of affine connections play an important part in our presentation. The necessary background is presented in order to illustrate how the methods may be applied. The bulk of the paper is devoted to a detailed analysis of some examples of nonholonomic mechanical control systems. We look at the Heisenberg system, the upright rolling penny, the roller racer, and the snakeboard.

Keywords. mechanics, nonholonomic constraints, controllability, affine connections.

AMS Subject Classifications. 53B05, 53B20, 53C22, 70D99, 70E15, 70H33, 70H03, 70Q05, 93B03, 93B05, 93B29.

1. Introduction

Lewis and Murray [1997a] provide a foundation for control theory for so-called simple mechanical control systems; mechanical control systems whose Lagrangian is kinetic energy minus potential energy, and whose inputs depend only upon configuration, not on velocity or time. In their work, Lewis and Murray introduced the notion of configuration controllability wherein one is interested in the control of only the configurations of the system, and not the velocities. Also, Lewis and Murray make the assumption that the initial condition is one with zero velocity. These restrictions of interest to configurations and initial states with zero velocity combine with the structured control Lie algebra for these systems to provide quite simple conditions for configuration accessibility of these systems. In the absence of potential energy, the geometry of these conditions was explained in [Lewis and Murray 1997b]. Configuration controllability (as opposed to configuration accessibility) is, as one expects, a more difficult problem. Lewis and Murray [1997a] provide a computable sufficient condition for configuration controllability for general simple mechanical control systems. These conditions are derived from results of Sussmann [1987] on controllability of general affine control systems. For multi-degree of freedom mechanical systems with no potential energy, Lewis [1997] demonstrated that *all* single-input simple mechanical

*Associate Professor, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
Email: andrew@mast.queensu.ca, URL: http://penelope.mast.queensu.ca/~andrew/

control systems are locally configuration *un*controllable. This result relies on another of Sussmann's results, this one a necessary condition which appeared in [Sussmann 1983]. Simple mechanical systems on Lie groups were studied by Bullo and Lewis [1996] where the controllability conditions were shown to take place entirely on the Lie algebra, and so are purely algebraic.

Mechanical systems with constraints have also received some attention in the control literature. We refer the reader to the work of Bloch and Crouch [1992], Bloch, Reyhanoglu, and McClamroch [1992], and Bloch and Crouch [1995b]. Constrained control systems in the presence of symmetry are the subject of the dissertation of Ostrowski [1995]. The paper of Ostrowski and Burdick [1997] provides an approach, utilising symmetry, complementary to that we take here.

One of the goals of the controllability analysis of Lewis and Murray [1997a] is to provide a framework for the development of motion control algorithms. In the kinematic context we refer to the work of Murray and Sastry [1993] and Leonard and Krishnaprasad [1995]. However, this analysis does not generally extend to systems where dynamical effects (mass and inertia) are taken into account. For dynamical analysis based on the controllability results of Lewis and Murray we refer to [Bullo, Leonard, and Lewis 2000].

Interestingly, the entire methodology of Lewis and Murray [1997a], and the work stemming from that paper, can be adapted to simple mechanical systems with **time-independent catastatic constraints**, simply meaning that the constraint is specified by a distribution on the configuration manifold (see [Pars 1965]). This results from an observation, first made by Synge [1928], that the equations for such a constrained system were the geodesics of an affine connection. Unconstrained simple mechanical systems are provided with a *natural* affine connection which arises from the kinetic energy (the **Levi-Civita** affine connection). This affine connection plays an essential rôle in the controllability conditions of Lewis and Murray [1997a] and in the geometric meaning of these conditions as provided by Lewis and Murray [1997b]. In fact, the context of the latter work was that of a system with a *general* affine connection, not just a Levi-Civita affine connection. This extends the applicability of the analysis of Lewis and Murray [1997a] to constrained systems *without modification*. In this paper we indicate in detail how this extension should be performed, and study some examples of simple mechanical control systems with constraints. The examples we study are the Heisenberg system (a constrained particle in $\mathbb{R}^3$), the upright rolling penny, the snakeboard, and the roller racer. For each example we provide a thorough controllability analysis in an attempt to illustrate the power of the basic ideas laid out in [Lewis and Murray 1997a] and subsequent papers.

The technology of affine and Riemannian differential geometry is not currently part of the lore of nonlinear control theory. It is our contention that for the control systems we are studying, simple mechanical control systems, affine connections play a vital rôle. This is borne out not only by the work mentioned above, but by the work of Rathinam and Murray [1998] where the affine connection is used to describe necessary and sufficient conditions for a simple mechanical control system to be "configuration flat."

The layout of the paper is as follows. In Section 2 we provide the necessary background for the reader unacquainted with affine connections. The presentation is intended to emphasise the fact that affine connections are simple to use in performing computations; as simple as, for example, the Lie bracket. The methodology for computing an affine connection for a constrained system is outlined in Section 3. With the constrained equations modelled

as a system whose drift vector field is associated with an affine connection, we apply the methods of Lewis and Murray [1997a] to provide conditions for configuration accessibility and configuration controllability. The rest of the paper is devoted to a detailed presentation of a few examples.

2. Background on affine connections

The theory of affine and Riemannian differential geometry is vast and well-developed. The classic reference is [Kobayashi and Nomizu 1963] which contains a fine presentation of the theory from a differential geometric perspective. A somewhat more "down to earth," but still agreeable, presentation of Riemannian geometry is given by Gallot, Hulin, and Lafontaine [1987]. However, we do not really need much equipment from the general theory — we are certainly able to present within the confines of the paper all the reader shall need to work out examples such as the ones we provide.

2.1. Definitions and basic properties. We shall follow the differential geometric conventions and notation of Abraham, Marsden, and Ratiu [1988].

Let Q be an n-dimensional manifold (the ***configuration manifold***) and denote by $C^\infty(Q)$ the set of C^∞ functions on Q, by $\mathscr{T}(Q)$ the set of C^∞ vector fields on Q, and by $\mathscr{T}^*(Q)$ the set of one-forms on Q. We denote the Lie derivative of $f \in C^\infty(Q)$ with respect to $X \in \mathscr{T}(Q)$ by $\mathscr{L}_X f$. An ***affine connection*** on Q assigns to a pair of vector fields $X, Y \in \mathscr{T}(Q)$ a vector field $\nabla_X Y \in \mathscr{T}(Q)$ (called the ***covariant derivative*** of Y with respect to X) and the assignment satisfies

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

AC2. $\nabla_{fX} Y = f(\nabla_X Y)$, and

AC3. $\nabla_X(fY) = f(\nabla_X Y) + (\mathscr{L}_X f)Y$

for $f \in C^\infty(Q)$ and $X, Y \in \mathscr{T}(Q)$. Notice that an affine connection differs from the Lie bracket of vector fields by, among other ways, being linear with respect to multiplication by functions in the first argument. It is important to note that, unlike the Lie bracket, an affine connection is *not* defined by the manifold structure; it is something extra which comes from problem data as we shall see below. For $f \in C^\infty(Q)$ *define* $\nabla_X f = \mathscr{L}_X f$. Then, just as is done for the Lie derivative $\mathscr{L}_X$, we may define $\nabla_X t$ where t is a tensor field of arbitrary type. As we shall have occasion to use this extension of ∇_X, we will recall (see [Abraham, Marsden, and Ratiu 1988, Theorem 5.3.2]) how it works. To obtain $\nabla_X \alpha$ for $\alpha \in \mathscr{T}^*(Q)$ we define

$$(\nabla_X \alpha)(Y) = \nabla_X(\alpha(Y)) - \alpha(\nabla_X Y) \tag{2.1}$$

for $Y \in \mathscr{T}(Q)$. Now, if t is a tensor field of type (r, s) on Q, then $\nabla_X t$ is defined by

$$(\nabla_X t)(\alpha^1, \ldots, \alpha^r, X_1, \ldots, X_s) = \mathscr{L}_X(t(\alpha^1, \ldots, \alpha^r, X_1, \ldots, X_s)) -$$
$$\sum_{i=1}^{r} t(\alpha^1, \ldots, \nabla_X \alpha^i, \ldots, \alpha^r, X_1, \ldots, X_s) - \sum_{j=1}^{s} t(\alpha^1, \ldots, \alpha^r, X_1, \ldots, \nabla_X X_j, \ldots, X_s). \tag{2.2}$$

Thus, to covariantly differentiate a one-form one uses (2.1) which involves only knowing how ∇_X acts on $\mathscr{T}(Q)$ and on $C^\infty(Q)$. Using this knowledge, we may then differentiate *arbitrary* type tensors. We shall only have occasion to use (2.2) when $(r, s) = (0, 2)$ and when $(r, s) = (1, 1)$. Recall that a $(1, 1)$ tensor field may simply be regarded as a base preserving, fibre-linear map from TQ to itself — that is, a linear map from each tangent space to itself. This identification is given explicitly by identifying a $(1, 1)$ tensor A with the fibre-linear map

$$T_q Q \ni v_q \mapsto \{\alpha_q \mapsto A_q(\alpha_q, X v_q)\} \in T_q Q.$$

On the right hand-side of the first "$\mapsto$" we are defining an element of $T_q Q$ by how it acts on $\alpha_q \in T_q Q$, this making sense since Q is finite-dimensional.

How does one compute in coordinates with an affine connection? Well, let (U, ϕ) be a chart for Q with coordinates $(q^1, \ldots, q^n)$. Let us write the covariant derivative of the basis vector field $\frac{\partial}{\partial q^j}$ with respect to the basis vector field $\frac{\partial}{\partial q^i}$ as

$$\nabla_{\frac{\partial}{\partial q^i}} \frac{\partial}{\partial q^j} = \Gamma_{ij}^k \frac{\partial}{\partial q^k}$$

which we may certainly do for some choice of coefficients Γ_{ij}^k, $i, j, k = 1, \ldots, n$, defined as functions on U. These coefficients are called the ***Christoffel symbols***.[1] One may verify, using the properties of affine connections, that if X and Y are vector fields whose components in coordinates we denote by $X^1, \ldots, X^n$ and $Y^1, \ldots, Y^n$, then the covariant derivative of Y with respect to X has the coordinate expression

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

Thus, to compute with the affine connection in coordinates, one simply needs its Christoffel symbols. We shall show in Section 2.3 how to compute these symbols for the systems we are considering. For now let us point out that if A is a $(1, 1)$ tensor field with components A_j^i, $i, j = 1, \ldots, n$, in a chart, then the formula

$$(\nabla_X A)_j^i = \frac{\partial A_j^i}{\partial q^k} X^k + \Gamma_{kl}^i A_j^l X^k - \Gamma_{kj}^l A_l^i X^k \tag{2.3}$$

follows directly from (2.1) and (2.2).

Besides the notion of covariant differentiation, affine connections are also responsible for the notion of geodesics. A curve $c \colon [a, b] \to Q$ is a ***geodesic*** for an affine connection ∇ if $\nabla_{\dot{c}(t)} \dot{c}(t) = 0$ for $t \in [a, b]$. One has to be somewhat careful that this definition makes sense, but that is taken care of by the general theory. For our purposes we shall simply need to know a coordinate expression for the equations defining a geodesic. If the curve c is given by $t \mapsto (q^1(t), \ldots, q^n(t))$ in coordinates, then it is a geodesic if and only if

$$\ddot{q}^i + \Gamma_{jk}^i \dot{q}^j \dot{q}^k = 0, \qquad i = 1, \ldots, n. \tag{2.4}$$

Thus geodesics are determined by a second-order differential equation on Q. Of course, corresponding to this is a first-order differential equation on TQ. If we denote natural

---

[1] The Christoffel symbols are *not* the components of a tensor field on Q.

coordinates on TQ by $(q^1, \ldots, q^n, v^1, \ldots, v^n)$ then the first-order differential equation is

$$\begin{aligned} \dot{q}^i &= v^i, \\ \dot{v}^i &= -\Gamma^i_{jk} v^j v^k, \end{aligned} \qquad i = 1, \ldots, n.$$

These first order equations define a vector field Z_g on TQ given in coordinates by

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

This vector field is known as the **_geodesic spray_**. Note that the integral curves of Z_g (they are curves on TQ) project to geodesics on Q.

Given an affine connection ∇ on Q we define the associated **_symmetric product_** by

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

Thus the symmetric product takes two vector fields and returns another. In Section 2.3 we shall use the symmetric product to say things about certain types of distributions. In Section 3.2 we shall simply use the symmetric product as a computational tool. One easily determines that in coordinates we have

$$\langle X : Y \rangle = \left(\frac{\partial X^i}{\partial q^j} Y^j + \frac{\partial Y^i}{\partial q^j} X^j + \Gamma^i_{jk} X^j Y^k + \Gamma^i_{jk} Y^j X^k \right) \frac{\partial}{\partial q^i}.$$

We shall have more to say about the symmetric product below.

The final general notion we have for affine connections is that of torsion. Given an affine connection ∇ we define its **_torsion tensor_** T by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

One may verify that this is indeed a $(1, 2)$ tensor field (meaning it is linear with respect to multiplication by functions in each argument). The components of the torsion tensor are readily seen to be

$$T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj}, \qquad i, j, k = 1, \ldots, n.$$

We shall not have _much_ to do with the torsion, but will use it in a couple of places. Here we make the observation that the geodesics of ∇, as may be seen from the governing equations (2.4), depend only upon the symmetric part of the lower indices of the Christoffel symbols. That is to say, since the torsion is the skew-symmetric part of these indices, the geodesics of an affine connection do not "depend" on its torsion. However, we will be discussing affine connections where their torsion _is_ important. This means that we must be interested in more than just the geodesics of these affine connections.

2.2. The Levi-Civita affine connection. In mechanics, a certain affine connection plays a distinguished rôle, and this affine connection will be basic to our presentation. A **_Riemannian metric_** on a manifold Q, typically denoted g, is a smooth specification of an inner product on each tangent space $T_q Q$. In mechanics, a Riemannian metric is specified by a system's kinetic energy — the inertia matrix precisely represents the components of a Riemannian metric in a set of coordinates. Note that a Riemannian metric is an example

of a $(0, 2)$ tensor field, and so we know how to covariantly differentiate it if we are given an affine connection (recall (2.2)).

For a given Riemannian metric g on Q there is a unique affine connection $\overset{g}{\nabla}$ on Q which satisfies the properties

LC1. $\overset{g}{\nabla}_X g = 0$ for every $X \in \mathscr{T}(Q)$ and

LC2. the torsion tensor for $\overset{g}{\nabla}$ vanishes.

The affine connection $\overset{g}{\nabla}$ is called the **_Levi-Civita_** affine connection associated with g. The two defining conditions of the Levi-Civita affine connection enable one to compute its Christoffel symbols explicitly in terms of the components of the Riemannian metric g (see [Gallot, Hulin, and Lafontaine 1987] for example). Doing so yields the formula

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

Here g^{ij}, $i, j = 1, \ldots, n$, are defined by $g_{ij} g^{jk} = \delta^k_i$. That is, the matrix with components g^{ij} is the inverse of the matrix with components g_{ij}.

Let us now see how the Levi-Civita affine connection is related to mechanics. Suppose that we have a Riemannian metric g on Q which models a system's kinetic energy. The Lagrangian for the system is then the function on TQ defined by $L(v_q) = \frac{1}{2} g(q)(v_q, v_q)$, i.e., the kinetic energy. In coordinates for TQ this reads $L(q, \dot{q}) = \frac{1}{2} g_{ij} \dot{q}^i \dot{q}^j$. One then readily computes the Euler-Lagrange equations to be

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} &= g_{ij} \left[\ddot{q}^j + g^{jk} \left(\frac{\partial g_{kl}}{\partial q^m} - \frac{1}{2} \frac{\partial g_{lm}}{\partial q^k} \right) \dot{q}^l \dot{q}^m \right] \\ &= g_{ij} \left[\ddot{q}^j + \overset{g}{\Gamma}^j_{lm} \dot{q}^l \dot{q}^m \right]. \end{aligned}$$

Here we have used the fact that the Christoffel symbols are the symmetric part of

$$\frac{\partial g_{kl}}{\partial q^m} - \frac{1}{2} \frac{\partial g_{lm}}{\partial q^k}$$

with respect to the indices l and m. This proves the following basic result.

2.1 Proposition: _The geodesics of the Levi-Civita affine connection_ $\overset{g}{\nabla}$ _are in one-to-one correspondence with the solutions of the Euler-Lagrange equations for the Lagrangian_ $L(v_q) = \frac{1}{2} g(q)(v_q, v_q)$.

2.3. Restricting the Levi-Civita affine connection to a distribution. If D is a distribution on Q, we denote by $\mathscr{D}$ the set of vector fields taking their values in D. If ∇ is an _arbitrary_ affine connection and D is a distribution on Q, we say ∇ **_restricts_** to D if $\nabla_X Y \in \mathscr{D}$ for $X \in \mathscr{T}(Q)$ and $Y \in \mathscr{D}$. Thus, if ∇ restricts to D, we have a well-defined notion of covariant differentiation in D itself and so we say ∇ defines a **_vector bundle connection_** in D. A distribution D on Q is **_geodesically invariant_** if for every geodesic $c : [a, b] \to Q$ of ∇, $\dot{c}(a) \in D_{c(a)}$ implies that $\dot{c}(t) \in D_{c(t)}$ for $t \in]a, b]$. That is, D is geodesically invariant if geodesics whose initial velocities lie in D evolve so that all subsequent velocities lie in D. It is also evident that D is geodesically invariant if and only if Z_g is tangent to $D \subset TQ$.

2.2 Theorem: ([Lewis 1998]) *Let ∇ be an affine connection on Q and let D be a distribution on Q. D is geodesically invariant if and only if $\langle X : Y \rangle \in \mathscr{D}$ for every two vector fields X and Y taking their values in D.*

Thus the symmetric product plays the same rôle for geodesically invariant distributions as the Lie bracket does for integrable distributions. Also note that if ∇ restricts to D then D is geodesically invariant.

Now we consider a given Levi-Civita affine connection $\overset{g}{\nabla}$ and we seek to restrict it to a distribution D. Thus we construct from $\overset{g}{\nabla}$ another affine connection $\widetilde{\nabla}$ which has the property that it restricts to a given distribution D. The motivation for doing so comes from mechanical systems with nonholonomic constraints. We consider here such systems whose Lagrangians are the kinetic energy associated with a Riemannian metric g on Q. The constraints are modelled by a distribution D on Q.

We formulate the equations of motion using the Lagrange-d'Alembert principle (see [Lanczos 1986]). There are other methods for deriving equations of motion which actually yield *different* equations. The Lagrange-d'Alembert principle has the property of agreeing with Newton's equations in instances where both methods apply. We think this seals the case that the Lagrange-d'Alembert principle is correct, and we refer the reader to [Lewis and Murray 1995] for further discussion of this. In any event, the Lagrange-d'Alembert principle for the systems we are dealing with states that the constrained solutions are those curves c which satisfy

$$\overset{g}{\nabla}_{\dot{c}(t)}\dot{c}(t) = \lambda(t) \tag{2.5a}$$
$$P'(\dot{c}(t)) = 0 \tag{2.5b}$$

where λ (the "Lagrange multiplier") is a section of $D^\perp$, the g-orthogonal complement to D, along c, and where $P' : TQ \to TQ$ is the orthogonal projection onto $D^\perp$. It is important that λ be a section of the *orthogonal* complement to D, and not an arbitrary complement. Not choosing the orthogonal complement misquotes the Lagrange-d'Alembert principle.

One may differentiate the constraint equation (2.5b) to obtain

$$\overset{g}{\nabla}_{\dot{c}(t)}(P'(\dot{c}(t))) = (\overset{g}{\nabla}_{\dot{c}(t)}P')(\dot{c}(t)) + P'(\overset{g}{\nabla}_{\dot{c}(t)}\dot{c}(t)) = 0.$$

In writing this equation we have used the differential operator property (2.2) for affine connections. Applying P' to the equations of motion (2.5a) we obtain

$$P'(\overset{g}{\nabla}_{\dot{c}(t)}\dot{c}(t)) = \lambda(t)$$

as $\lambda(t) \in D^\perp_{c(t)}$. Combining the two equations to eliminate λ one sees that c must be a geodesic of the affine connection $\widetilde{\nabla}$ which is defined by

$$\widetilde{\nabla}_X Y = \overset{g}{\nabla}_X Y + (\overset{g}{\nabla}_X P')(Y).$$

To verify that $\widetilde{\nabla}$ *is* an affine connection, one need only observe that the term $(\overset{g}{\nabla}_X P')(Y)$ is bilinear with respect to multiplication by functions in the arguments X and Y. We postpone a precise proof of the equivalence of geodesics of $\widetilde{\nabla}$ and solutions of (2.5a) until

Section 3.1 at which time we will also consider the addition of external forces. We mention that the construction of $\widetilde{\nabla}$ is related to a construction of Synge [1928]. Other authors have subsequently taken up various incarnations of this method. We mention in particular [Cattaneo-Gasparini 1963], [Cattaneo 1963], [Vershik 1984], and [Bloch and Crouch 1995a].

2.3 Remarks: The affine connection $\widetilde{\nabla}$ has some rather interesting properties which are explored in detail by Lewis [1998]. Let us make here a few relevant comments.

1. Note that it is important that $\widetilde{\nabla}$ be allowed to have torsion. If it is required to be torsion free, one readily ascertains that this implies D must be integrable — certainly not what one wants in general.

2. When we write the equations for a constrained system as the geodesic equations for the affine connection $\widetilde{\nabla}$, the condition that the solutions satisfy the constraints is reflected by the fact that the constraint distribution D is geodesically invariant with respect to $\widetilde{\nabla}$. This follows, as we shall see in Section 3.1, from the fact that $\widetilde{\nabla}$ restricts to D. ●

In practice computation of $\widetilde{\nabla}$ can be a bit troublesome. Often the computations are simplified by using instead the affine connection $\overset{A}{\nabla}$ defined by

$$\overset{A}{\nabla}_X Y = \overset{g}{\nabla}_X Y + A^{-1}((\overset{g}{\nabla}_X(AP'))(Y)) \tag{2.6}$$

where A is *any* invertible $(1,1)$ tensor field. Note that $\widetilde{\nabla}$ is one in the family of affine connections $\overset{A}{\nabla}$, namely that where $A = \mathrm{id}_{TQ}$, the identity map on TQ. We shall see in Section 3.1 that $\overset{A}{\nabla}$ serves as well as $\widetilde{\nabla}$ in determining the constrained equations of motion. The freedom to make this choice of A is a consequence of our only being interested in geodesics whose initial velocities are in D. Different choices of A will yield different geodesics if the initial velocity is *not* in D, but this is irrelevant as far as we are concerned. We will find it convenient to make certain choices of A in the examples later in the paper. Essentially one chooses A to cancel denominators in P' which will cause problems when it comes to differentiation.

Using (2.3) we may compute the Christoffel symbols of $\overset{A}{\nabla}$ to be

$$\overset{A}{\Gamma}{}^i_{jk} = \overset{g}{\Gamma}{}^i_{jk} + B^i_l \frac{\partial (AP')^l_j}{\partial q^k} + B^i_l \overset{g}{\Gamma}{}^l_{km}(AP')^m_j - B^i_l \overset{g}{\Gamma}{}^m_{kj}(AP')^l_m. \tag{2.7}$$

Here B^i_j are the components of A^{-1}. We keep A and P' together as that is often the whole point of introducing A — to make the product AP' simpler that P' by itself.

3. Constrained control systems

We now know enough about affine connections to allow us to formulate precisely the equations of motion for constrained systems in terms of an affine connection. Let us first formally define the mechanical systems we will be talking about. A ***simple mechanical system with constraints*** consists of the following data:

SMSC1. an n-dimensional manifold Q (the **configuration manifold**);

SMSC2. a Riemannian metric g on Q (the **kinetic energy Riemannian metric**);

SMSC3. a function V on Q (the **potential energy function**);

SMSC4. a distribution D on Q (the **constraint distribution**).

When we talk explicitly about control systems later in the section, we will say what we mean by a simple mechanical *control* system with constraints.

In this section we first demonstrate how one may write equations for a simple mechanical system with constraints with the inclusion of a single external force. Then we are able to easily write down the control problem and formulate controllability definitions and review existing results.

3.1. Adding a single force. Let us look at the case where $\overset{g}{\nabla}$ is a Levi-Civita affine connection on Q and D is a distribution on Q with orthogonal complement $D^\perp$. Recall that P' denotes the orthogonal projection onto $D^\perp$ and we shall denote by P the orthogonal projection onto D. We shall represent an external force by a one-form F on Q (for example, we may take $F = -\boldsymbol{d}V$ if we want to model potential forces). Associated with the one-form F is the vector field $Y = g^\sharp(F)$ where $g^\sharp \colon T^*Q \to TQ$ is the musical isomorphism associated with the Riemannian metric g. In coordinates $g^\sharp(F) = g^{ij}F_j \frac{\partial}{\partial q^i}$. It is the vector field Y which is the more convenient representation of external forces in the computations below. Following the presentation in Section 2.3, the equations of motion for a constrained system subject to the external force Y are

$$\overset{g}{\nabla}_{\dot{c}(t)}\dot{c}(t) = \lambda(t) + Y(c(t)) \tag{3.1a}$$

$$P'(\dot{c}(t)) = 0. \tag{3.1b}$$

We now verify that we may represent solutions of (3.1) by solutions of the forced geodesic equations with respect to the family of affine connections $\overset{A}{\nabla}$ introduced in Section 2.3. We state a result which follows easily from our construction, but is essential in understanding exactly how forces interact with the affine connection $\overset{A}{\nabla}$. In the proof of the result, we shall use the notion of the **vertical lift** of a vector field on Q. Let $X \in \mathscr{T}(Q)$. We define a vector field X^{lift} on TQ by

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

(see [Abraham and Marsden 1978, Definition 3.7.5]). In coordinates, if $X = X^i \frac{\partial}{\partial q^i}$, then we have

$$X^{\text{lift}} = X^i \frac{\partial}{\partial v^i}.$$

One may readily check that $X \in \mathscr{D}$ if and only if X^{lift} is tangent to $D \subset TQ$.

3.1 Proposition: *Let A be an invertible $(1,1)$ tensor field and let $\overset{A}{\nabla}$ be as defined by (2.6). A curve $c \colon [a,b] \to Q$ is a solution of (3.1) if and only if $\dot{c}(a) \in D_{c(a)}$ and c satisfies*

$$\overset{A}{\nabla}_{\dot{c}(t)}\dot{c}(t) = P(Y(c(t))) \tag{3.2}$$

thought of as a system on $D \subset TQ$.

Proof: We first need to show that the solutions of (3.2) are tangent to D. That is to say, thought of as a curve in TQ, we show that $\dot{c}(t)$ is tangent to $D \subset TQ$ if $\dot{c}(a) \in D_{c(a)}$. We first note that we can write the equation (3.2) in the form

$$\dot{v} = \overset{A}{Z}_g(v) - (P(Y))^{\text{lift}}$$

where $v \in TQ$ and $\overset{A}{Z}_g$ is the geodesic spray associated to the affine connection $\overset{A}{\nabla}$. This then defines a vector field on TQ given by $\overset{A}{Z}_g - (P(Y))^{\text{lift}}$. We claim that this vector field is tangent to D. To show that $\overset{A}{Z}_g$ is tangent to D it suffices to show that D is geodesically invariant with respect to $\overset{A}{\nabla}$. By Theorem 2.2 this will follow if we can show that $\overset{A}{\nabla}$ restricts to D. Let X_1 and X_2 be vector fields on Q. Then

$$P'(\overset{A}{\nabla}_{X_1}X_2) = P'(\overset{g}{\nabla}_{X_1}X_2) + P' \circ A^{-1}((\overset{g}{\nabla}_{X_1}(AP'))(X_2)). \tag{3.3}$$

If $X_2 \in \mathscr{D}$ then

$$P'(X_2) = 0$$
$$\implies AP'(X_2) = 0$$
$$\implies (\overset{g}{\nabla}_{X_1}(AP'))(X_2) + AP'(\overset{g}{\nabla}_{X_1}X_2) = 0$$
$$\implies A^{-1}((\overset{g}{\nabla}_{X_1}(AP'))(X_2)) + P'(\overset{g}{\nabla}_{X_1}X_2) = 0$$
$$\implies P' \circ A^{-1}((\overset{g}{\nabla}_{X_1}(AP'))(X_2)) + P'(\overset{g}{\nabla}_{X_1}X_2) = 0 \tag{3.4}$$

since $P' \circ P' = P'$. Substituting (3.4) into (3.3) we see that $P'(\overset{A}{\nabla}_{X_1}X_2) = 0$ for $X_1 \in \mathscr{T}(Q)$ and $X_2 \in \mathscr{D}$. Therefore, $\overset{A}{\nabla}_{X_1}X_2 \in \mathscr{D}$. This shows that $\overset{A}{Z}_g$ is tangent to D. Now, since $P(Y)$ takes its values in D, $(P(Y))^{\text{lift}}$ is tangent to D. Thus $\overset{A}{Z}_g - (P(Y))^{\text{lift}}$ is tangent to D and so solutions of (3.2) which start on D will remain on D.

Now suppose that c is a solution of (3.1). Since c satisfies the constraint equation (3.1b) we must have $\dot{c}(a) \in D_{c(a)}$. The constraint equation implies that $AP'(\dot{c}(t)) = 0$. Differentiating this relation and solving for λ gives

$$\lambda(t) = -A^{-1}((\overset{g}{\nabla}_{\dot{c}(t)}(AP'))(\dot{c}(t))) - P'(Y(c(t))). \tag{3.5}$$

Thus

$$\overset{g}{\nabla}_{\dot{c}(t)}\dot{c}(t) = -A^{-1}((\overset{g}{\nabla}_{\dot{c}(t)}(AP'))(\dot{c}(t))) - P'(Y(c(t))) + Y(c(t))$$

which is equivalent to (3.2) by the definition of $\overset{A}{\nabla}$ and since $P = \text{id}_{TQ} - P'$.

Now suppose that $\dot{c}(a) \in D_{c(a)}$ and that c satisfies (3.2). As we have shown above, if $\dot{c}(a) \in D_{c(a)}$ then $\dot{c}(t)$ is in D for all $t \in [a,b]$. Thus c satisfies (3.1b). Also, c satisfies (3.1a) with λ given by (3.5). Thus c is a solution of (3.1). $\blacksquare$

The essence of the proposition is that we may simply consider the system (3.2) restricted to $D \subset TQ$. The restriction make sense since D is geodesically invariant with respect to $\overset{A}{\nabla}$ and since $P(Y)$ takes its values in D. The proposition has the following important consequence. To treat the constrained system (3.1), we may eliminate the constraint equation by considering the system (3.2) on TQ and then restricting the solutions to those whose initial conditions are in D. Conceptually this is much easier than having to deal with the separate constraint equations. We also make the remark that it is only the component of the force Y in the D-direction which makes any contribution to the dynamics of the problem. The component of Y in the $D^{\perp}$-direction gets absorbed by the constraint force.

Because of the proposition, the following definition makes sense.

3.2 Definition: Let g be a Riemannian metric on Q with $\overset{g}{\nabla}$ the corresponding Levi-Civita affine connection, and let D be a distribution on Q. For an invertible $(1,1)$ tensor field A, the connection $\overset{A}{\nabla}$ restricted to D and thought of as a vector bundle connection in D is called the **constrained affine connection**.   •

Notice that this definition does not depend on A since we are only interested in the restriction of $\overset{A}{\nabla}$ to D. That is to say, two different choices for A will yield the same connection *in* D.

3.2. Controllability definitions and results. With the results of the previous section, we may easily extend our analysis to control systems. A **simple mechanical control system with constraints** is a simple mechanical system with constraints plus the additional data of m linearly independent one-forms $F^1, \ldots, F^m$ on Q which we shall consider as modelling our input forces. As in Section 3.1, we shall use the vector fields $Y_a = g^{\sharp}(F^a)$, $a = 1, \ldots, m$, in the computations. The control system we consider is

$$\nabla_{\dot{c}(t)} \dot{c}(t) = -\operatorname{grad} V(c(t)) + \lambda(t) + u^a(t) Y_a(c(t))$$
$$P'(c(t)) = 0$$

where we have included the effect of potential forces from the potential function V (here $\operatorname{grad} V = (\boldsymbol{d}V)^{\sharp}$). By virtue of our efforts of Section 3.1, this system is equivalent to the system

$$\overset{A}{\nabla}_{\dot{c}(t)} \dot{c}(t) = -P(\operatorname{grad} V(c(t))) + u^a(t) P(Y_a(c(t))) \tag{3.6}$$

restricted to D. Without loss of generality we can make the assumption that the external forces $Y_1, \ldots, Y_m$ lie in the distribution D. We shall do this in the remainder of this section to simplify the notation.

We shall assume that our control inputs come from the set

$$\mathscr{U} = \{u \colon [0, T] \to \mathbb{R}^m \mid T > 0 \text{ and } u \text{ is bounded and measurable}\}.$$

A **solution** of (3.6) is a pair (c, u) where $c \colon [0, T] \to Q$, a piecewise differentiable curve with $\dot{c}(0) \in D_{c(0)}$, and $u \in \mathscr{U}$ satisfy (3.6).

Now we formulate the controllability definitions for this system. Let $q_0 \in Q$ and let U be a neighbourhood of q_0 in Q. Define

$$\mathcal{R}_Q^U(q_0, T) = \{q \in Q \mid \text{ there exists a solution } (c, u) \text{ of } (3.6)$$
$$\text{such that } \dot{c}(0) = 0_{q_0}, c(t) \in U \text{ for } t \in [0, T], \text{ and } \dot{c}(T) \in T_q Q\}$$

and denote

$$\mathcal{R}_Q^U(q_0, \leq T) = \bigcup_{0 \leq t \leq T} \mathcal{R}_Q^U(q_0, t).$$

Here $0_q \in T_q Q$ denotes the zero vector in the tangent space. We may now state our versions of controllability.

3.3 Definition: We say that (3.6) 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 subset of Q for all neighbourhoods U of q_0 and all $0 < t \leq T$. If this holds for any $q_0 \in Q$ then the system is called **locally configuration accessible**.

We say that (3.6) is **locally configuration controllable** at q_0 if it is locally configuration accessible at q_0 and if there exists $T > 0$ such that q_0 is in the interior of $\mathcal{R}_Q^U(q_0, \leq t)$ for every neighbourhood U of q_0 and $0 < t \leq T$. If this holds for any $q_0 \in Q$ then the system is called **locally configuration controllable**.

We say that (3.6) is **equilibrium controllable** if, for $q_1, q_2 \in Q$, there exists a solution (c, u) of (3.6) where $c \colon [0, T] \to Q$ is such that $c(0) = q_1$, $c(T) = q_2$ and both $\dot{c}(0)$ and $\dot{c}(T)$ are zero.   •

For a discussion of these definitions we refer the reader to [Lewis and Murray 1997a]. We also refer the reader to that work for a treatment of systems of the form (3.6) where the affine connection $\overset{A}{\nabla}$ is replaced by the Levi-Civita affine connection on a Riemannian manifold. Note that there are no constraints present in the work of Lewis and Murray [1997a]. However, with constraints, the system (3.6) is much the same as the system studied in that paper with the exception that the affine connection is not the Levi-Civita affine connection. Moreover, we have the following statement:

The analysis of Lewis and Murray [1997a] may be directly applied to the system (3.6).

This may be easily seen by looking at the relevant formulas used by Lewis and Murray [1997a] to compute the accessibility distribution. This is explained in [Lewis and Murray 1997b].

Now we turn to stating the controllability results which are known for the system (3.6). In stating the results we shall make the assumption that $V = 0$. This is not essential — the results of [Lewis and Murray 1997a] are given in the presence of potential energy. However, its presence complicates the results tremendously, and the development to restate the results here would constitute a substantial diversion. Instead we direct the reader to [Lewis and Murray 1997a] and mention again that all conclusions of that paper hold for arbitrary affine connections.

Let $\mathscr{V}$ be a family of vector fields on Q. Denote by $\overline{\operatorname{Sym}}(\mathscr{V})$ the smallest subset of $\mathscr{T}(Q)$ containing $\mathscr{V}$ which is closed under the symmetric product associated with $\overset{A}{\nabla}$, and denote by $\overline{\operatorname{Lie}}(\mathscr{V})$ the smallest Lie subset of $\mathscr{T}(Q)$ containing $\mathscr{V}$ which is closed under Lie bracket. We call $\overline{\operatorname{Sym}}(\mathscr{V})$ the **symmetric closure** of $\mathscr{V}$ and $\overline{\operatorname{Lie}}(\mathscr{V})$ the **involutive closure** in $\mathscr{V}$. Let $\mathscr{Y}$ denote the family of vector fields $\{Y_1, \ldots, Y_m\}$. We shall suppose that the families

of vector fields $\overline{\mathrm{Sym}}(\mathscr{Y})$ and $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathscr{Y}))$ generate constant rank distributions which we denote by $C_{\mathrm{ver}}(\mathscr{Y})$ and $C_{\mathrm{hor}}(\mathscr{Y})$, respectively. As is shown in [Lewis and Murray 1997a], we think of $C_{\mathrm{ver}}(\mathscr{Y})(q)$ as representing the set of velocity directions accessible from 0_q and we think of $C_{\mathrm{hor}}(\mathscr{Y})(q)$ as the set of configuration directions accessible from 0_q. Of course, to compute $C_{\mathrm{ver}}(\mathscr{Y})$ in practice, one simply computes iterated symmetric products of vector fields from $\mathscr{Y}$ until new directions are no longer generated. Then, to compute $C_{\mathrm{hor}}(\mathscr{Y})$ one computes iterated Lie brackets of the vector fields which were used to generate $C_{\mathrm{ver}}(\mathscr{Y})$, again until new directions are no longer generated. This methodology is implicitly applied in the examples in Section 4.

We may now state the basic accessibility result, valid for $V = 0$.

3.4 THEOREM: ([Lewis and Murray 1997a]) *The system (3.6) is locally configuration accessible at $q \in Q$ if the rank of $C_{\mathrm{hor}}(\mathscr{Y})$ at q is equal to the dimension of Q. Conversely, if (3.6) is locally configuration accessible, then $C_{\mathrm{hor}}(\mathscr{Y})$ has maximal rank on an open dense subset of Q.*

3.5 REMARKS: We make some comments which relate to the plausibility of Theorem 3.4.

1. By Theorem 2.2, $C_{\mathrm{ver}}(\mathscr{Y})$ is geodesically invariant. Since we are assuming that $Y_a \in \mathscr{D}$, $a = 1, \ldots, m$, and since D is geodesically invariant by construction of $\overset{A}{\nabla}$, this implies that $C_{\mathrm{ver}}(\mathscr{Y}) \subset D$. Of course this makes sense given Proposition 3.1 which implies that solutions of (3.6) which start in D must remain in D.

2. Since $C_{\mathrm{ver}}(\mathscr{Y}) \subset D$, we must also have $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(\mathscr{Y})) \subset \overline{\mathrm{Lie}}(\mathscr{D})$. That is to say, the set of reachable configurations is contained in the smallest integrable distribution containing D. This agrees with purely kinematic considerations which indicate that the constraints restrict the system's movement to leaves of the smallest integrable distribution containing D. ●

Now we state a sufficient condition for local configuration controllability. For this result, we need the notion of good and bad symmetric products. An iterated symmetric product from $\mathscr{Y}$ is called **bad** if it contains an even number of each of the vector fields Y_a, $a = 1, \ldots, m$, and otherwise an iterated symmetric product is called **good**. For example, the symmetric product $\langle \langle Y_a : Y_b \rangle : \langle Y_a : Y_b \rangle \rangle$ is bad, but the symmetric product $\langle Y_a : \langle Y_a : Y_b \rangle \rangle$ is good. We call the **degree** of a symmetric product the total number of vector fields from $\mathscr{Y}$ which comprise it. Thus our bad example has degree four, and the good example has degree three.[2]

The following theorem happens to hold in the case when V is non-zero in (3.6).

3.6 THEOREM: ([Lewis and Murray 1997a]) *If $\dim(C_{\mathrm{hor}}(\mathscr{Y})_q) = n$ and if every bad symmetric product at $q \in Q$ is a $\mathbb{R}$-linear combination of good symmetric products at q of lower degree, then (3.6) is locally configuration controllable at q. If these conditions are satisfied at every point in Q, then (3.6) is equilibrium controllable.*

For single-input systems with no potential energy, we have the following sharp controllability result.

---
[2]We are being somewhat imprecise here for the sake of expediency. To make proper sense of bad, good, and degree, one must use the notion of a free symmetric algebra. See [Lewis and Murray 1997a] for details. It should, however, be clear what we mean here.

3.7 THEOREM: ([Lewis 1997]) *If $m = 1$, the system (3.6) is locally configuration controllable if and only if $\dim(Q) = 1$.*

This result is not generally true in the presence of potential energy.

Finally we turn to decompositions of control systems of the form (3.6) with $V = 0$.

3.8 THEOREM: ([Lewis and Murray 1997b]) *Let $(x^1, \ldots, x^k, y^1, \ldots, y^{n-k})$ be coordinates adapted to the integrable distribution $C_{\mathrm{hor}}(\mathscr{Y})$ (thus $C_{\mathrm{hor}}(\mathscr{Y}) = \mathrm{span}\left\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^k}\right\}$). The system (3.6) in these coordinates has the form*

$$\ddot{x}^\sigma + \Gamma^\sigma_{\rho\tau}\dot{x}^\rho\dot{x}^\tau + \Gamma^\sigma_{\rho\beta}\dot{x}^\rho\dot{y}^\beta + \Gamma^\sigma_{\beta\rho}\dot{y}^\beta\dot{x}^\rho + \Gamma^\sigma_{\beta\gamma}\dot{y}^\beta\dot{y}^\gamma = u^a Y^\sigma_a, \qquad \sigma = 1, \ldots, k$$

$$\ddot{y}^\alpha + \Gamma^\alpha_{\rho\tau}\dot{x}^\rho\dot{x}^\tau + \Gamma^\alpha_{\rho\beta}\dot{x}^\rho\dot{y}^\beta + \Gamma^\alpha_{\beta\rho}\dot{y}^\beta\dot{x}^\rho + \Gamma^\alpha_{\beta\gamma}\dot{y}^\beta\dot{y}^\gamma = 0, \qquad \alpha = 1, \ldots, n-k$$

where $\Gamma^\alpha_{\rho\tau}(x,y)\dot{x}^\rho\dot{x}^\tau = 0$ if $\dot{x} \in C_{\mathrm{ver}}(\mathscr{Y})_{(x,y)}$. In particular, if $\dot{x} = 0$ and $\dot{y} = 0$ then $\dot{y}(t) = 0$ for $t > 0$. Furthermore, the system restricted to the leaves of $C_{\mathrm{hor}}(\mathscr{Y})$ is locally configuration accessible.

The Christoffel symbols in this last theorem are defined by

$$\overset{A}{\nabla}_{\frac{\partial}{\partial x^\rho}}\frac{\partial}{\partial x^\tau} = \Gamma^\sigma_{\rho\tau}\frac{\partial}{\partial x^\sigma} + \Gamma^\alpha_{\rho\tau}\frac{\partial}{\partial y^\alpha}, \quad \overset{A}{\nabla}_{\frac{\partial}{\partial x^\rho}}\frac{\partial}{\partial y^\beta} = \Gamma^\sigma_{\rho\beta}\frac{\partial}{\partial x^\sigma} + \Gamma^\alpha_{\rho\beta}\frac{\partial}{\partial y^\alpha},$$

$$\overset{A}{\nabla}_{\frac{\partial}{\partial y^\beta}}\frac{\partial}{\partial x^\rho} = \Gamma^\sigma_{\beta\rho}\frac{\partial}{\partial x^\sigma} + \Gamma^\alpha_{\beta\rho}\frac{\partial}{\partial y^\alpha}, \quad \overset{A}{\nabla}_{\frac{\partial}{\partial y^\beta}}\frac{\partial}{\partial y^\gamma} = \Gamma^\sigma_{\beta\gamma}\frac{\partial}{\partial x^\sigma} + \Gamma^\alpha_{\beta\gamma}\frac{\partial}{\partial y^\alpha}.$$

4. Examples

The remainder of the paper we devote to the investigation of four examples: the Heisenberg system, the upright rolling penny, the snakeboard, and the roller racer. These systems have been presented previously in the literature (references are given with each example) with the emphasis on various dynamical or locomotive properties. For each example we explicitly give the data describing the problem (i.e., the kinetic energy metric, the constraints, and the input one-forms). Using this data, and (2.7), we compute the Christoffel symbols of the affine connection $\overset{A}{\nabla}$ (A is chosen to suit the particular problem). We then use Theorems 3.4 and Theorem 3.6 to determine the accessibility and controllability properties of the example for various combinations of the given inputs. Where applicable (and feasible), the decomposition of Theorem 3.8 is computed.

Some of the computations for the examples can be rather tedious, and so were performed using *Mathematica*. When doing distribution computations in the examples, we ignore points where distributions lose rank. To consider singularities of the distributions is important, but is not something we will take into account here.

4.1.   The Heisenberg system. This system is a "canonical" example and has been studied by many authors. For recent treatments we refer to [Bates and Śniatycki 1993] and [Bloch, Krishnaprasad, Marsden, and Murray 1996]. In both of these papers the emphasis is on dealing with the symmetries in the system which we here neglect.

We take a particle in $Q = \mathbb{R}^3$ and model the kinetic energy with the Riemannian metric

$$g = m\left(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z\right)$$

for $m > 0$ the mass of the particle. The inertia matrix associated with this Riemannian metric is

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

For the remaining examples we shall simply write the inertia matrix with the understanding that it represents the components of the kinetic energy Riemannian metric in the given coordinates. The system's velocities are required to satisfy the constraint

$$\dot{z} = y\dot{x}. \tag{4.1}$$

The constraint distribution D is then the span of the two vector fields

$$\frac{\partial}{\partial y}, \quad \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}$$

on $\mathbb{R}^3$. This system is sometimes called the "Heisenberg system" since the brackets of the vector fields satisfying the constraints obey commutator relations which are reminiscent of the Lie algebra of the Heisenberg group.

The constrained affine connection

The orthogonal complement of D is easily seen to be generated by the vector field

$$y\frac{\partial}{\partial x} - \frac{\partial}{\partial z}.$$

We may compute the projection P' to be

$$P'(x, y, z) \cdot (v_x, v_y, v_z) = \frac{1}{1 + y^2}\left(y^2 v_x - yv_z, 0, -yv_x + v_z\right).$$

It is convenient to choose $A = (1 + y^2)\,\mathrm{id}_{TQ}$. The reader will wish to recall that, as a consequence of Proposition 3.1, the choice of A does not affect the dynamics of the system restricted to D. We compute the non-zero Christoffel symbols of $\overset{A}{\nabla}$ to be

$$\Gamma^x_{xy} = \frac{2y}{1 + y^2}, \qquad \Gamma^x_{zy} = -\frac{1}{1 + y^2}, \qquad \Gamma^z_{xy} = -\frac{1}{1 + y^2}.$$

Remember that this affine connection is not torsion free and so does not have the indicial symmetries of a Levi-Civita affine connection.

Controllability analysis

Now we make the Heisenberg system a control system by consider the two input one-forms

$$F^1 = \mathrm{d}y, \qquad F^2 = \mathrm{d}x + y\mathrm{d}z.$$

The associated input vector fields are

$$Y_1 = \frac{1}{m}\frac{\partial}{\partial y}, \qquad Y_2 = \frac{1}{m}\left(\frac{\partial}{\partial x} + y\frac{\partial}{\partial z}\right).$$

Since the system is one of academic interest, its hard to motivate this choice of inputs, except to say that they span D. The equations for the control system are then

$$\ddot{x} + \frac{1}{1 + y^2}\left(2y\dot{x}\dot{y} - \dot{y}\dot{z}\right) = \frac{1}{m}u^2$$

$$\ddot{y} = \frac{1}{m}u^1$$

$$\ddot{z} - \frac{1}{1 + y^2}\dot{x}\dot{y} = \frac{y}{m}u^2.$$

Here $(u^1, u^2) \in \mathbb{R}^2$ are the inputs. These equations should be restricted to D since those are the only interesting velocities. Using (4.1) we compute the equations of motion restricted to D to be

$$\ddot{x} + \frac{y}{1 + y^2}\dot{x}\dot{y} = \frac{1}{m}u^2$$

$$\ddot{y} = \frac{1}{m}u^1$$

$$\ddot{z} - \frac{1}{1 + y^2}\dot{x}\dot{y} = \frac{y}{m}u^2.$$

To determine the controllability of the system with various combinations of inputs, we need to perform some symmetric product and Lie bracket computations. We compute

$$\langle Y_1 : Y_1 \rangle = 0, \qquad \langle Y_1 : Y_2 \rangle = \frac{1}{m^2}\left(\frac{y}{1 + y^2}\frac{\partial}{\partial x} + \frac{y^2}{1 + y^2}\frac{\partial}{\partial z}\right),$$

$$\langle Y_2 : Y_2 \rangle = 0, \qquad [Y_1, Y_2] = \frac{1}{m^2}\frac{\partial}{\partial y}.$$

With these computations we may easily deduce the configuration controllability of the Heisenberg system.

HS1. $\mathscr{Y} = \{Y_1\}$: In this case we see that $\overline{\mathrm{Sym}}(\mathscr{Y})$ is simply generated by the single vector field Y_1. Thus the distribution C_{hor} is also generated by Y_1. The system is, therefore, not locally configuration accessible. Note that $Y_1 \simeq \frac{\partial}{\partial y}$ and so the coordinates (x, y, z) are adapted to this integrable distribution. We may write the control system restricted to D as

$$\ddot{y} = \frac{1}{m}u_1$$

$$\ddot{x} + \frac{y}{1 + y^2}\dot{x}\dot{y} = 0$$

$$\ddot{z} - \frac{1}{1 + y^2}\dot{x}\dot{y} = 0.$$

With zero velocity initial conditions, the first equation decouples from the last two as stated by Theorem 3.8.

HS2. $\mathscr{Y} = \{Y_2\}$: In this case we again see that the system is not locally configuration accessible. The distribution C_{hor} is spanned by Y_2. To make the system decouple as

in Theorem 3.8, we must make a change of coordinates. The coordinates $(\xi, \eta, \zeta) = (x, y, z - xy)$ have the property that $C_{\text{hor}} = \text{span}\left\{\frac{\partial}{\partial \xi}\right\}$. One may verify that the non-zero Christoffel symbols of $\overset{A}{\nabla}$ in this set of coordinates are

$$\Gamma^\xi_{\eta\xi} = -\frac{\eta}{1+\eta^2}, \qquad \Gamma^\xi_{\eta\eta} = -\frac{\xi}{1+\eta^2}, \qquad \Gamma^\xi_{\zeta\eta} = -\frac{1}{1+\eta^2},$$

$$\Gamma^\zeta_{\xi\eta} = -\xi, \qquad \Gamma^\zeta_{\eta\xi} = 1 - \xi, \qquad \Gamma^\zeta_{\eta\eta} = \frac{\xi\eta}{1+\eta^2}, \qquad \Gamma^\zeta_{\zeta\eta} = \frac{\eta}{1+\eta^2}.$$

The constrained velocities in these coordinates must satisfy the equation

$$\dot\zeta = -\xi\dot\eta.$$

The equations restricted to D may then be computed as

$$\ddot\xi - \frac{\eta}{1+\eta^2}\dot\xi\dot\eta = \frac{1}{m}u_2$$

$$\ddot\eta = 0$$

$$\ddot\zeta + (1 - 2\xi)\dot\xi\dot\eta = 0.$$

As expected, if we have zero initial velocity, the first equation decouples from the last two as a consequence of Theorem 3.8.

HS3. $\mathscr{Y} = \{Y_1, Y_2\}$: Here we see that the vector fields $\{Y_1, Y_2, [Y_1, Y_2]\}$ generate TQ and so the system is locally configuration accessible by Theorem 3.4. In fact, by Theorem 3.6 it is also locally configuration controllable, in fact equilibrium controllable.

4.2. The upright rolling penny. Here we study the example of an upright rolling penny. This is another "canonical" example which has been studied in detail by many authors. The symmetries of the rolling penny are discussed, for example, by Bloch, Krishnaprasad, Marsden, and Murray [1996].

This object is as depicted in Figure 1 and has $Q = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ as its configuration manifold. The inertia matrix we consider is

$$\begin{bmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & J & 0 \\ 0 & 0 & 0 & I \end{bmatrix}.$$

Here $m > 0$ is the mass of the penny, $I > 0$ is the moment of inertia of the penny about its centre, and $J > 0$ is the moment of inertia of the penny about the "z-axis." The condition that the penny roll without slipping is modelled by declaring that the velocities satisfy the constraint equations

$$\dot x = r\cos\theta\dot\phi, \qquad \dot y = r\sin\theta\dot\phi. \tag{4.2}$$

The constraint distribution is then generated by the two vector fields

$$\frac{\partial}{\partial\theta}, \qquad \frac{\partial}{\partial\phi} + r\cos\theta\frac{\partial}{\partial x} + r\sin\theta\frac{\partial}{\partial y}.$$
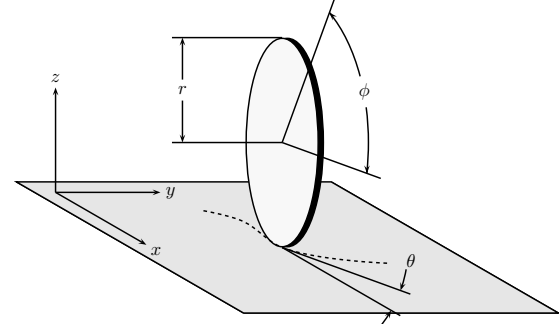

FIGURE 1. The rolling penny.

The constrained affine connection

The orthogonal complement to D is easily seen to be generated by the two vector fields

$$I\frac{\partial}{\partial x} - mr\cos\theta\frac{\partial}{\partial\phi}, \qquad I\frac{\partial}{\partial y} - mr\sin\theta\frac{\partial}{\partial\phi}.$$

Using this information we may compute the orthogonal projection onto $D^\perp$ to be

$$P'(x, y, \theta, \phi) \cdot (v_x, v_y, v_\theta, v_\phi) = \frac{1}{I + mr^2}\big((I + mr^2\sin^2\theta)v_x - mr^2\sin\theta\cos\theta v_y - Ir\cos\theta v_\phi,$$

$$-mr^2\sin\theta\cos\theta v_x + (I + mr^2\cos^2\theta)v_y - Ir\sin\theta v_\phi, -mr\cos\theta v_x - mr\sin\theta v_y + mr^2 v_\phi, 0\big).$$

We choose $A = (I + mr^2)\,\text{id}_{TQ}$ and compute the non-zero Christoffel symbols as

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

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

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

Controllability analysis

This system has two natural inputs: a torque which makes the penny roll, and a torque which makes the penny spin. These inputs are modelled by the one-forms

$$F^1 = \mathrm{d}\phi, \qquad F^2 = \mathrm{d}\theta$$

and the inputs vector fields associated with these forces are

$$Y'_1 = \frac{1}{I}\frac{\partial}{\partial\phi}, \qquad Y'_2 = \frac{1}{J}\frac{\partial}{\partial\theta}.$$

However, the theory discussed in Section 3 asks that these vector fields have their values in D, so we orthogonally project these to D to obtain the actual inputs vector fields we shall use in the computations as

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

Then we compute the forced equations of motion for the rolling penny to be

$$\ddot{x} + \frac{1}{I + mr^2}\left(mr^2\sin 2\theta\dot{x}\dot{\theta} - mr^2\cos 2\theta\dot{y}\dot{\theta} + Ir\sin\theta\dot{\theta}\dot{\phi}\right) = \frac{r\cos\theta}{I + mr^2}u^1$$

$$\ddot{y} - \frac{1}{I + mr^2}\left(mr^2\cos 2\theta\dot{x}\dot{\theta} + mr^2\sin 2\theta\dot{y}\dot{\theta} + Ir\cos\theta\dot{\theta}\dot{\phi}\right) = \frac{r\sin\theta}{I + mr^2}u^1$$

$$\ddot{\theta} = \frac{1}{J}u^2$$

$$\ddot{\phi} + \frac{1}{I + mr^2}\left(mr\sin\theta\dot{x}\dot{\theta} - mr\cos\theta\dot{y}\dot{\theta}\right) = \frac{1}{I + mr^2}u^1.$$

Using (4.2) the resulting equations of motion restricted to D are

$$\ddot{x} + r\sin\theta\dot{\phi}\dot{\theta} = \frac{r\cos\theta}{I + mr^2}u^1$$

$$\ddot{y} - r\cos\theta\dot{\phi}\dot{\theta} = \frac{r\sin\theta}{I + mr^2}u^1$$

$$\ddot{\theta} = \frac{1}{J}u^2$$

$$\ddot{\phi} = \frac{1}{I + mr^2}u^1.$$

One may readily ascertain that the unforced x and y equations are simply the time derivatives of the constraint equations.

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

$$\langle Y_1 : Y_1\rangle = 0, \qquad \langle Y_1 : Y_2\rangle = 0, \qquad \langle Y_2 : Y_2\rangle = 0,$$

$$[Y_1, Y_2] = \frac{r}{J(I + mr^2)}\left(\sin\theta\frac{\partial}{\partial x} - \cos\theta\frac{\partial}{\partial y}\right)$$

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

We may now easily deduce some basic facts about the controllability of the upright rolling penny.

RP1. $\mathscr{Y} = \{Y_1\}$: We see that C_{hor} is spanned by Y_1 and so the system is not locally configuration accessible. To get the equations to decouple as stated in Theorem 3.8, we must make a change of coordinates. To this end we introduce the coordinates

$$(\xi, \eta, \zeta, \psi) = (x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, x\cos\theta + y\sin\theta - r\phi, \theta).$$

We may verify that C_{hor} is generated by $\frac{\partial}{\partial\xi}$ in these coordinates. The non-zero Christoffel symbols for $\overset{A}{\nabla}$ in these coordinates are

$$\Gamma^\xi_{\eta\psi} = -\frac{I(1 + \eta) + 2mr^2}{I + mr^2}, \qquad \Gamma^\xi_{\psi\eta} = -\frac{I(1 + \eta) + mr^2}{I + mr^2},$$

$$\Gamma^\xi_{\psi\psi} = -\xi\frac{I(1 + \eta) + 2mr^2}{I + mr^2}, \qquad \Gamma^\eta_{\xi\psi} = -\xi, \qquad \Gamma^\eta_{\zeta\psi} = \frac{I}{I + mr^2},$$

$$\Gamma^\eta_{\psi\xi} = 1 - \xi, \qquad \Gamma^\eta_{\psi\psi} = \eta\frac{I(\xi - 1) + \xi mr^2}{I + mr^2},$$

$$\Gamma^\zeta_{\eta\psi} = -1 - \eta, \qquad \Gamma^\zeta_{\psi\eta} = -1 - \eta, \qquad \Gamma^\zeta_{\psi\psi} = -\xi(1 + \eta).$$

We may verify that the constraint equations (4.2) in these coordinates are given by

$$-\sin\psi\dot{\eta} + \cos\psi\dot{\zeta} - (\xi\sin\psi + \eta\cos\psi)\dot{\psi} = 0$$

$$\cos\psi\dot{\eta} + \sin\psi\dot{\zeta} + (\xi\cos\psi - \eta\sin\psi)\dot{\psi} = 0.$$

These may be simplified to give the equivalent relations

$$\dot{\eta} = -\xi\dot{\psi}, \qquad \dot{\zeta} = \eta\dot{\psi}.$$

Now we may derive the equations for the control system restricted to D in these coordinates to be

$$\ddot{\xi} + \xi\frac{I(1 + \eta) + mr^2}{I + mr^2}\dot{\psi}^2 = \frac{r}{I + mr^2}u_1$$

$$\ddot{\eta} + (1 - 2\xi)\dot{\xi}\dot{\psi} + \xi\eta\dot{\psi}^2 = 0$$

$$\ddot{\zeta} + \xi(1 + \eta)\dot{\psi}^2 = 0$$

$$\ddot{\psi} = 0.$$

Note that for zero initial velocities, the top equation decouples from the last three. Physically, this motion is simply a rolling of the penny in the direction ξ which is, as is illustrated in Figure 2, in the direction of the fixed angle θ.

RP2. $\mathscr{Y} = \{Y_2\}$: In this case the system is not locally configuration accessible as the distribution C_{hor} is spanned by the vector field Y_2. The coordinates (x, y, ϕ, θ) are adequate for decoupling the system. Indeed, the equations restricted to D in these coordinates with the input Y_2 have the form

$$\ddot{\theta} = \frac{1}{J}u_2$$

$$\ddot{x} + r\sin\theta\dot{\phi}\dot{\theta} = 0$$

$$\ddot{y} - r\cos\theta\dot{\phi}\dot{\theta} = 0$$

$$\ddot{\phi} = 0.$$

As in Theorem 3.8, we see that the top equation decouples from the last three equations when the initial velocities are zero. The physical motion in this case is simply a spinning of the penny about its vertical axis while the rest of the configuration variables remain unchanged.
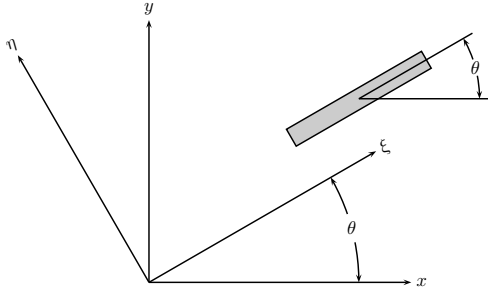
Figure 2. The motion of the rolling penny with the input Y_1

RP3. $\mathscr{Y} = \{Y_1, Y_2\}$: In this case the vector fields $\{Y_1, Y_2, [Y_1, Y_2], [Y_2, [Y_1, Y_2]]\}$ span TQ and so the system is locally configuration accessible. By Theorem 3.6 we also see that the system is locally configuration controllable and, in fact, equilibrium controllable.

4.3. The snakeboard. This example was first studied by Lewis, Ostrowski, Murray, and Burdick [1994]. The snakeboard was an instrumental example in the development of the theory of Bloch, Krishnaprasad, Marsden, and Murray [1996] and Ostrowski [1995]. The snakeboard as a control system is considered by Ostrowski and Burdick [1997].

The system is a derivative of the skateboard where the rider uses his own momentum, coupled with the constraints, to make the system "locomote." The mathematical model of the system is shown in Figure 3. The configuration manifold for the system is $Q = SE(2) \times \mathbb{T}^2$ and we use coordinates $(x, y, \theta, \psi, \phi)$ as shown in the figure. We constrain the front and rear wheel angles to be equal and opposite. This is both a natural thing to do physically, and simplifies the analysis considerably. Note that we are using slightly different coordinates than were used by Lewis, Ostrowski, Murray, and Burdick. Their choice of coordinates was motivated by ease of modelling the external forces, and our coordinates make the representation of the affine connection $\overset{A}{\nabla}$ easier. The inertia matrix representing the kinetic energy of the snakeboard is

$$\begin{bmatrix} m & 0 & 0 & 0 & 0 \\ 0 & m & 0 & 0 & 0 \\ 0 & 0 & J + 2J_1 & 0 & 0 \\ 0 & 0 & 0 & J_0 & 0 \\ 0 & 0 & 0 & 0 & 2J_1 \end{bmatrix}.$$

Here $m > 0$ is the total mass of the snakeboard, $J > 0$ is the moment of inertia of the body of the snakeboard about its centre of mass, $J_0 > 0$ is the moment of inertia of the rotor mounted on top of the body's centre of mass, and $J_1 > 0$ is the moment of inertia of each
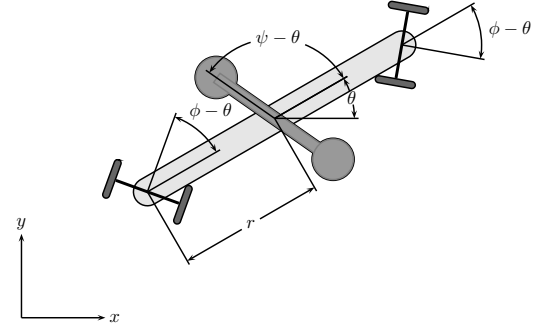
Figure 3. The mathematical model of the Snakeboard

of the wheel axles. The constraint equations for the snakeboard are

$$-\sin\phi\,\dot{x} + \cos\phi\,\dot{y} - r\cos(\theta - \phi)\dot{\theta} = 0$$
$$-\sin(2\theta - \phi)\dot{x} + \cos(2\theta - \phi)\dot{y} + r\cos(\theta - \phi)\dot{\theta} = 0 \tag{4.3}$$

from which we ascertain that the constraint distribution is generated by the vector fields

$$\frac{\partial}{\partial\psi}, \quad \frac{\partial}{\partial\phi}, \quad a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial\theta}$$

where

$$a = r\cos\theta\cos(\theta - \phi), \quad b = r\sin\theta\cos(\theta - \phi), \quad c = \sin(\theta - \phi).$$

The constrained affine connection

The orthogonal complement to D is generated by the vector fields

$$J'c\frac{\partial}{\partial x} - ma\frac{\partial}{\partial\theta}, \quad b\frac{\partial}{\partial x} - a\frac{\partial}{\partial y}.$$

Here $J' = J + 2J_1$. We then compute the orthogonal projection onto $D^\perp$ to be

$$P'(x, y, \theta, \psi, \phi) \cdot (v_x, v_y, v_\theta, v_\psi, v_\phi) = ((J'c^2 + mb^2)v_x - mabv_y - J'acv_\theta,$$
$$-mabv_x + (J'c^2 + ma^2)v_y - J'bcv_\theta, -macv_x - mbcv_y + m(a^2 + b^2)v_\theta)/(J'c^2 + m(a^2 + b^2)).$$

We choose $A = (J'c^2 + m(a^2 + b^2))\,\mathrm{id}_{TQ}$ and with this we compute the non-zero Christoffel symbols to be

$$\Gamma^x_{x\theta} = \frac{\cos(\phi - \theta)(-(mr^2\sin(\phi - 3\theta)) + (-2J' + mr^2)\sin(\phi - \theta))}{J'c^2 + m(a^2 + b^2)},$$

$$\Gamma^x_{x\phi} = \frac{(2J' - mr^2 + mr^2 \cos 2\theta)\sin(2(\phi - \theta))}{2(J'c^2 + m(a^2 + b^2))},$$

$$\Gamma^x_{y\theta} = -\frac{mr^2(\cos(2(\phi - 2\theta)) + \cos 2\theta)}{2(J'c^2 + m(a^2 + b^2))}, \qquad \Gamma^x_{y\phi} = \frac{mr^2 \sin(2(\phi - \theta))\sin(2\theta)}{2(J'c^2 + m(a^2 + b^2))},$$

$$\Gamma^x_{\theta\theta} = -\frac{J'r(3\cos(2\phi - 3\theta) + \cos(2\phi - \theta))}{4(J'c^2 + m(a^2 + b^2))}, \qquad \Gamma^x_{\theta\phi} = \frac{J'r\cos(2(\phi - \theta))\cos(\theta)}{J'c^2 + m(a^2 + b^2)},$$

$$\Gamma^y_{x\theta} = -\frac{mr^2(\cos(2(\phi - 2\theta)) + \cos 2\theta)}{2(J'c^2 + m(a^2 + b^2))}, \qquad \Gamma^y_{x\phi} = \frac{mr^2 \sin(2(\phi - \theta))\sin 2\theta}{2(J'c^2 + m(a^2 + b^2))},$$

$$\Gamma^y_{y\theta} = \frac{\cos(\phi - \theta)(mr^2 \sin(\phi - 3\theta) + (-2J' + mr^2)\sin(\phi - \theta))}{J'c^2 + m(a^2 + b^2)},$$

$$\Gamma^y_{y\phi} = -\frac{(-2J' + mr^2 + mr^2 \cos 2\theta)\sin(2(\phi - \theta))}{2(J'c^2 + m(a^2 + b^2))},$$

$$\Gamma^y_{\theta\theta} = -\frac{J'r(-3\sin(2\phi - 3\theta) + \sin(2\phi - \theta))}{4(J'c^2 + m(a^2 + b^2))}, \qquad \Gamma^y_{\theta\phi} = \frac{J'r\cos(2(\phi - \theta))\sin\theta}{J'c^2 + m(a^2 + b^2)},$$

$$\Gamma^\theta_{x\theta} = -\frac{mr(3\cos(2\phi - 3\theta) + \cos(2\phi - \theta))}{4(J'c^2 + m(a^2 + b^2))}, \qquad \Gamma^\theta_{x\phi} = \frac{mr\cos(2(\phi - \theta))\cos\theta}{J'c^2 + m(a^2 + b^2)},$$

$$\Gamma^\theta_{y\theta} = -\frac{mr(-3\sin(2\phi - 3\theta) + \sin(2\phi - \theta))}{4(J'c^2 + m(a^2 + b^2))}, \qquad \Gamma^\theta_{y\phi} = \frac{mr\cos(2(\phi - \theta))\sin\theta}{J'c^2 + m(a^2 + b^2)},$$

$$\Gamma^\theta_{\theta\theta} = \frac{mr^2 \sin(2\phi - 2\theta)}{J'c^2 + m(a^2 + b^2)}, \qquad \Gamma^\theta_{\theta\phi} = -\frac{mr^2 \sin(2\phi - 2\theta)}{J'c^2 + m(a^2 + b^2)}.$$

Controllability analysis

The control system we consider has the ability to apply torques to control the relative angle between the rotor and the body and the relative angle between the wheels and the body. The input one-forms are thus

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

The resulting input vector fields are

$$Y'_1 = \frac{1}{J'}\frac{\partial}{\partial\theta} - \frac{1}{J_0}\frac{\partial}{\partial\psi}, \qquad Y'_2 = \frac{1}{2J_1}\frac{\partial}{\partial\phi}.$$

Orthogonally projecting these vector fields to D gives the inputs vector fields

$$Y_1 = \left(-\frac{1}{2}r\cos\theta\sin(2(\phi - \theta))\frac{\partial}{\partial x} - \frac{1}{2}r\sin(2(\phi - \theta))\sin\theta\frac{\partial}{\partial y} + \right.$$
$$\left. \sin^2(\phi - \theta)\frac{\partial}{\partial\theta}\right)/(J'c^2 + m(a^2 + b^2)) - \frac{1}{J_0}\frac{\partial}{\partial\psi},$$

$$Y_2 = \frac{1}{2J_1}\frac{\partial}{\partial\phi}.$$

We may determine the equations of motion as follows.

$$\ddot{x} - \left(-\left(\cos(\phi - \theta)(-(mr^2 \sin(\phi - 3\theta)) + (-2J' + mr^2)\sin(\phi - \theta))\right)\dot{x}\dot{\theta} - \right.$$

$$\frac{1}{2}\left((2J' - mr^2 + mr^2 \cos 2\theta)\sin(2(\phi - \theta))\right)\dot{x}\dot{\phi} +$$

$$\frac{1}{2}\left(mr^2(\cos(2(\phi - 2\theta)) + \cos 2\theta)\right)\dot{y}\dot{\theta} - \frac{1}{2}\left(mr^2 \sin(2(\phi - \theta))\sin 2\theta\right)\dot{y}\dot{\phi} +$$

$$\frac{1}{4}\left(3J'r\cos(2\phi - 3\theta) + J'r\cos(2\phi - \theta)\right)\dot{\theta}^2 -$$

$$(J'r\cos(2(\phi - \theta))\cos\theta)\,\dot{\theta}\dot{\phi})/(J'c^2 + m(a^2 + b^2)) = -\frac{r\cos\theta\sin(2(\phi - \theta))}{2(J'c^2 + m(a^2 + b^2))}u^1$$

$$\ddot{y} - \left(\frac{1}{2}\left(mr^2(\cos(2(\phi - 2\theta)) + \cos 2\theta)\right)\dot{x}\dot{\theta} - \frac{1}{2}\left(mr^2 \sin(2(\phi - \theta))\sin 2\theta\right)\dot{x}\dot{\phi} - \right.$$

$$\left(\cos(\phi - \theta)(mr^2 \sin(\phi - 3\theta) + (-2J' + mr^2)\sin(\phi - \theta))\right)\dot{y}\dot{\theta} +$$

$$\frac{1}{2}\left((-2J' + mr^2 + mr^2 \cos 2\theta)\sin(2(\phi - \theta))\right)\dot{y}\dot{\phi} +$$

$$\frac{1}{4}\left(-3J'r\sin(2\phi - 3\theta) + J'r\sin(2\phi - \theta)\right)\dot{\theta}^2 -$$

$$(J'r\cos(2(\phi - \theta))\sin\theta)\,\dot{\theta}\dot{\phi})/(J'c^2 + m(a^2 + b^2)) = -\frac{r\sin(2(\phi - \theta))\sin\theta}{2(J'c^2 + m(a^2 + b^2))}u^1$$

$$\ddot{\theta} - \left(\frac{1}{4}\left(mr(3\cos(2\phi - 3\theta) + \cos(2\phi - \theta))\right)\dot{x}\dot{\theta} - (mr\cos(2(\phi - \theta))\cos\theta)\dot{x}\dot{\phi} + \right.$$

$$\frac{1}{4}\left(mr(-3\sin(2\phi - 3\theta) + \sin(2\phi - \theta))\right)\dot{y}\dot{\theta} - (mr\cos(2(\phi - \theta))\sin\theta)\,\dot{y}\dot{\phi} -$$

$$\left(mr^2 \sin(2(\phi - \theta))\right)\dot{\theta}^2 + \left(mr^2 \sin(2(\phi - \theta))\right)\dot{\theta}\dot{\phi})/(J'c^2 + m(a^2 + b^2)) =$$

$$\frac{\sin^2(\phi - \theta)}{J'c^2 + m(a^2 + b^2)}u^1$$

$$\ddot{\psi} = -\frac{1}{J_0}u^1$$

$$\ddot{\phi} = \frac{1}{2J_1}u^2.$$

These equations may be simplified by asking that the velocities satisfy the constraint equations (4.3). With some straightforward manipulation these equations may be reduced to

$$\sin(\phi - \theta)\dot{x} + r\cos\theta\cos(\theta - \phi)\dot{\theta} = 0, \qquad \sin(\phi - \theta)\dot{y} + r\sin\theta\cos(\theta - \phi)\dot{\theta} = 0.$$

We may then write the constrained equations of motion as

$$\ddot{x} - \left(\frac{1}{4}\left(3J'r\cos(2\phi - 3\theta) + J'r\cos(2\phi - \theta)\right)\dot{\theta}^2 + \right.$$

$$\frac{1}{4}\left(r(-2J'\cos(2\phi - 3\theta) - 2J'\cos(2\phi - \theta))\right)\dot{\theta}\dot{\phi})/(J'c^2 + m(a^2 + b^2)) =$$

$$-\frac{r\cos\theta\sin(2(\phi - \theta))}{2(J'c^2 + m(a^2 + b^2))}u^1$$

$$\ddot{y} - \left(\frac{1}{4}\left(-3J'r\sin(2\phi - 3\theta) + J'r\sin(2\phi - \theta)\right)\dot{\theta}^2 - \right.$$

$$(J'r\cos(2(\phi - \theta))\sin\theta)\,\dot{\theta}\dot{\phi})/(J'c^2 + m(a^2 + b^2)) =$$

$$-\frac{r\sin(2(\phi-\theta))\sin\theta}{2(J'c^2+m(a^2+b^2))}u^1$$

$$\ddot{\theta}-\left(\left(-(mr^2\cot(\phi-\theta))+mr^2\cos(2(\phi-\theta))\cot(\phi-\theta)\right)\dot{\theta}^2-\right.$$
$$\left.\left(mr^2(-1+\cos(2(\phi-\theta)))\cot(\phi-\theta)\right)\dot{\theta}\dot{\phi}\right)/(J'c^2+m(a^2+b^2))=$$
$$\frac{\sin^2(\phi-\theta)}{J'c^2+m(a^2+b^2)}u^1$$

$$\ddot{\psi}=-\frac{1}{J_0}u^1$$

$$\ddot{\phi}=\frac{1}{2J_1}u^2.$$

The results of the symmetric product and Lie bracket computations for the snakeboard are too lengthy to reproduce in their entirety here. Let us provide some of the expressions evaluated at the point $q_0=(0,0,0,0,0)$.

$$\langle Y_1:Y_1\rangle(q_0)=0,\quad\langle Y_2:Y_2\rangle(q_0)=0,\quad\langle Y_1:Y_2\rangle(q_0)=-\frac{1}{2mrJ_1}\frac{\partial}{\partial x}$$

$$\langle\langle Y_1:Y_2\rangle:\langle Y_1:Y_2\rangle\rangle(q_0)=0,\quad[Y_1,Y_2](q_0)=\frac{1}{2mrJ_1}\frac{\partial}{\partial x},$$

$$[Y_2,[Y_1,Y_2]](q_0)=-\frac{1}{2J_1^2mr^2}\frac{\partial}{\partial\theta},$$

$$[Y_2,[Y_1,[Y_2,[Y_1,Y_2]]]]](q_0)=-\frac{1}{4J_1^3m^2r^3}\frac{\partial}{\partial y}-\frac{1}{2J_1^3m^2r^4}\frac{\partial}{\partial\theta}.$$

With these computations we may determine the controllability of the snakeboard at the operating point $q_0=(0,0,0,0,0)$.

SB1. $\mathscr{Y}=\{Y_1\}$: In this case we note that C_{hor} is generated at q_0 by the vector field $\{Y_1\}$ and so the system is not locally configuration accessible with this input. The set of configurations reachable from q_0 will be a one-dimensional submanifold generated by the flow of Y_1 through q_0.

SB2. $\mathscr{Y}=\{Y_2\}$: The system is not locally configuration accessible and the distribution C_{hor} is generated by Y_2. In this case the equations decouple in the original coordinates. The physical motion of the control system is simply a rotation of the wheels while the rest of the body is at rest.

SB3. $\mathscr{Y}=\{Y_1,Y_2\}$: We see that the vector fields

$$\{Y_1,Y_2,[Y_1,Y_2],[Y_2,[Y_1,Y_2]],[Y_2,[Y_1,[Y_2,[Y_1,Y_2]]]]\}$$

span TQ at q_0 and so the system is locally configuration accessible with the inputs Y_1 and Y_2 by Theorem 3.4. Moreover, since the bad symmetric products $\langle Y_1:Y_1\rangle$ and $\langle Y_2:Y_2\rangle$ are zero at q_0, the system satisfies the sufficient condition test for local configuration controllability of Theorem 3.6.

4.1 REMARKS: 1. With the snakeboard computations, one must be aware of singularities in the rank of certain distributions, especially as the "natural" initial operating point $(x,y,\theta,\psi,\phi)=(0,0,0,0,0)$ can often be involved in the singularities. The vector fields we have chosen to span C_{hor} are selected so that they have maximal rank in a neighbourhood of $(0,0,0,0,0)$. Therefore, by Theorem 3.6 the snakeboard is equilibrium controllable in a neighbourhood of $(0,0,0,0,0)$.

2. There is an interesting, but still not perhaps completely understood correspondence between the brackets used to generate C_{hor} and the snakeboard "gaits" observed by Lewis, Ostrowski, Murray, and Burdick [1994]. Corresponding to the bracket $[Y_1,Y_2]$ there is the gait where the rotor and the wheels are forced periodically with the same frequency. This gives rise to a motion in the x-direction when starting from $(0,0,0,0,0)$. Corresponding to the bracket $[Y_2,[Y_1,Y_2]]$ there is the gait where the wheels are forced periodically at twice the frequency as the rotor. In this case a motion in the θ-direction is observed. And, corresponding to the bracket $[Y_2,[Y_1,[Y_2,[Y_1,Y_2]]]]]$, there is a gait where the wheels and the rotor are forced periodically at a $3:2$ ratio. This gives rise to a motion in the y-direction. In each case, note that the ratio of the forcing frequencies of the wheels and the rotor is matched by the relative number of terms of the forcing vector fields in the brackets. The same observations are made by Ostrowski and Burdick [1997] with regards to the controllability analysis they perform. This is doubtless related to sinusoidal motion planning methodology of Murray and Sastry [1993]. Further, it is also doubtless related to the averaging analysis of Leonard and Krishnaprasad [1995] in the kinematic case, and Bullo, Leonard, and Lewis [2000] in the dynamic case.

4.4. The roller racer. The roller racer is a toy in which the rider sits over an axle and propels himself using his feet to apply a cyclic motion to another pair of wheels which pivot in front of the vehicle. The mathematical model we study here was introduced by Krishnaprasad and Tsakiris [1995] (see Figure 4). The configuration manifold is $Q=SE(2)\times\mathbb{S}$ and for the moment we use coordinates (x,y,θ,ψ) as indicated in the figure. We make the assumption, not unreasonably, that the centre of mass of the body of the roller racer is located over the wheel axle. In this case the Riemannian metric describing the kinetic energy is

$$g=m(\mathrm{d}x\otimes\mathrm{d}x+\mathrm{d}y\otimes\mathrm{d}y)+I_1\mathrm{d}\theta\otimes\mathrm{d}\theta+I_2\mathrm{d}\psi\otimes\mathrm{d}\psi.$$

Here $m>0$ is the mass of the body of the roller racer, $I_1>0$ is the moment of inertia of the body about its centre of mass, and $I_2>0$ is the moment of inertia of the wheel assembly about the pivot point. The equations of constraint for the roller racer are

$$-\sin\theta\dot{x}+\cos\theta\dot{y}=0$$
$$-\sin\psi\dot{x}+\cos\psi\dot{y}+l_1\cos(\theta-\psi)\dot{\theta}+l_2\dot{\psi}=0.$$

It turns out to be convenient for the computations we do here to use in place of (x,y) the rotated coordinates

$$(\xi,\eta)=(x\cos\theta+y\sin\theta,-x\sin\theta+y\cos\theta),$$
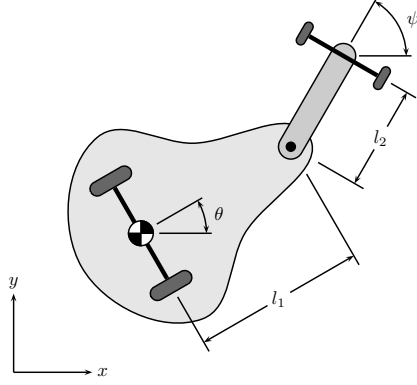
Figure 4. The mathematical model of the roller racer

and the relative angle $\phi = \psi - \theta$. In the coordinates $(\xi, \eta, \theta, \phi)$ the inertia matrix may be computed to be

$$\begin{bmatrix} m & 0 & -m\eta & 0 \\ 0 & m & m\xi & 0 \\ -m\eta & m\xi & I_1 + I_2 + m(\xi^2 + \eta^2) & I_2 \\ 0 & 0 & I_2 & I_2 \end{bmatrix}$$

and the constraint equations are

$$\dot{\eta} + \xi\dot{\theta} = 0$$
$$-\sin\phi\,\dot{\xi} + \cos\phi\,\dot{\eta} + (l_2 + (l_1 + \xi)\cos\phi + \eta\sin\phi)\dot{\theta} + l_2\dot{\phi} = 0. \tag{4.4}$$

The constraint distribution is then spanned by the two vector fields

$$l_2\frac{\partial}{\partial\xi} + \sin\phi\frac{\partial}{\partial\phi}, \qquad l_2\eta\frac{\partial}{\partial\xi} - l_2\xi\frac{\partial}{\partial\eta} + l_2\frac{\partial}{\partial\theta} - (l_2 + l_1\cos\phi)\frac{\partial}{\partial\phi}.$$

The constrained affine connection

The orthogonal complement to D is generated by the two vector fields

$$\frac{1}{m}\frac{\partial}{\partial\eta},$$
$$\left(\frac{l_1\eta\cos\phi}{I_1} - \frac{\sin\phi}{m}\right)\frac{\partial}{\partial\xi} + \frac{(I_1 - ml_1\xi)\cos\phi}{mI_1}\frac{\partial}{\partial\eta} + \frac{l_1\cos\phi}{I_1}\frac{\partial}{\partial\theta} + \left(\frac{l_2}{I_2} - \frac{l_1\cos\phi}{I_1}\right)\frac{\partial}{\partial\phi}.$$

The orthogonal projection is then computed to be

$$P'(\xi, \eta, \theta, \phi) \cdot (v_\xi, v_\eta, v_\theta, v_\phi) =$$

$$\Big((I_1 I_2 \sin^2\phi - ml_1\eta I_2\cos\phi\sin\phi)v_\xi + I_2(l_2 + l_1\cos\phi + \eta\sin\phi)(ml_1\eta\cos\phi - I_1\sin\phi)v_\theta +$$
$$(ml_1 l_2\eta I_2\cos\phi - l_2 I_1 I_2\sin\phi)v_\phi, ml_1\xi I_2\sin\phi\cos\phi v_\xi +$$
$$(ml_2^2 I_1 + ml_1^2 I_2\cos^2\phi + I_2 I_2\sin^2\phi)v_\eta +$$
$$(ml_2\xi(l_2 I_1 - l_1 I_2\cos\phi) + \xi I_1 I_2\sin^2\phi - ml_1\xi\eta I_2\sin\phi\cos\phi)v_\theta -$$
$$ml_1 l_2\xi I_2\cos\phi v_\phi, -ml_1 I_2\sin\phi\cos\phi v_x +$$
$$(ml_1 l_2 I_2\cos\phi + ml_1^2 I_2\cos^2\phi + ml_1\eta I_2\sin\phi\cos\phi)v_\theta + ml_1 l_2 I_2\cos\phi v_\phi,$$
$$(ml_1 I_2\sin\phi\cos\phi - ml_2 I_1\sin\phi)v_x + m(l_2 I_1 - l_1 I_2\cos\phi)(l_2 + l_1\cos\phi + \eta\sin\phi)v_\theta +$$
$$ml_2^2 I_1 - ml_1 l_2 I_2\cos\phi v_\phi\Big)/\Delta$$

where

$$\Delta = ml_2^2 I_2(1 + \cos^2\phi) + I_1 I_2\sin^2\phi.$$

When doing the controllability analysis for the roller racer, one may suppose, without loss of generality, that we have an initial configuration for which $(\xi, \eta) = (0, 0)$. Thus, to simplify our life, let us only write the Christoffel symbols at such a configuration point. We choose $A = \Delta\,\mathrm{id}_{TQ}$ and compute the non-zero Christoffel symbols of $\overset{A}{\nabla}$ (at $(\xi, \eta) = (0, 0)$) as

$$\Gamma^\xi_{\xi\phi} = \frac{I_1 I_2\sin 2\phi}{\Delta}, \qquad \Gamma^\xi_{\eta\theta} = -\frac{2m(l_2^2 I_1 + l_1^2 I_2) + I_1 I_2\sin\phi(2l_2 + l_1\cos\phi + \sin\phi)}{\Delta},$$
$$\Gamma^\xi_{\theta\eta} = -\frac{m(l_2^2 I_1 + l_1^2 I_2\cos^2\phi) + I_1 I_2\sin\phi(2l_2 + l_1\cos\phi + \sin\phi)}{\Delta},$$
$$\Gamma^\xi_{\theta\phi} = -\frac{I_1 I_2(l_1\cos 2\phi + l_2\cos\phi)}{\Delta}, \Gamma^\xi_{\phi\phi} = -\frac{l_1 I_1 I_2\cos\phi}{\Delta},$$
$$\Gamma^\eta_{\xi\theta} = \frac{I_1 I_2\sin^2\phi}{\Delta}, \qquad \Gamma^\eta_{\eta\phi} = \frac{I_2(I_1 + ml_1^2)\sin 2\phi}{\Delta},$$
$$\Gamma^\eta_{\theta\xi} = \frac{ml_2^2 I_1 + ml_1^2 I_1\cos^2\phi + I_1 I_2\sin^2\phi}{\Delta}, \qquad \Gamma^\eta_{\theta\theta} = -\frac{I_1 I_2\sin\phi(l_2 + l_1\cos\phi)}{\Delta},$$
$$\Gamma^\eta_{\phi\theta} = -\frac{l_2 I_1 I_2\sin\phi}{\Delta}, \qquad \Gamma^\theta_{\xi\phi} = -\frac{ml_1 I_2\cos 2\phi}{\Delta},$$
$$\Gamma^\theta_{\theta\phi} = \frac{ml_1 I_2\sin 2\phi}{2\Delta}, \qquad \Gamma^\theta_{\theta\phi} = -\frac{ml_1 I_2\sin\phi(l_2 + 2l_1\cos\phi)}{\Delta},$$
$$\Gamma^\theta_{\phi\phi} = -\frac{ml_2 l_2 I_2\sin\phi}{\Delta}, \qquad \Gamma^\phi_{\xi\phi} = \frac{m(l_1 I_2\cos 2\phi - l_2 I_1\cos\phi)}{\Delta},$$
$$\Gamma^\phi_{\theta\eta} = \frac{m\sin\phi(l_2 I_1 - l_1 I_2\cos\phi)}{\Delta}, \qquad \Gamma^\phi_{\theta\phi} = \frac{ml_1\sin\phi(l_2 I_2 - l_2 I_1 + 2l_1 I_2\cos\phi)}{\Delta},$$
$$\Gamma^\phi_{\phi\phi} = \frac{ml_1 l_2 I_2\sin\phi}{\Delta}.$$

The expressions for the Christoffel symbols at general values of ξ and η are significantly more imposing. Of course, when doing the actual analysis, one must make the evaluations $\xi = \eta = 0$ at the right time (i.e., *after* differentiation). We simply give these formulas in case the reader wishes to check their own computations.

Controllability analysis

The natural input for the roller racer is a torque which actuates the angle between the main body and the front wheels. The force is specified by the one-form

$$F^1 = \mathrm{d}\phi.$$

The control vector field corresponding to this input force is

$$Y_1' = -\frac{\eta}{I_1}\frac{\partial}{\partial\xi} + \frac{\xi}{I_1}\frac{\partial}{\partial\eta} - \frac{1}{I_1}\frac{\partial}{\partial\theta} + \frac{I_1+I_2}{I_1 I_2}\frac{\partial}{\partial\phi}.$$

As with the previous examples, we must project these vector fields onto D. Again, as we are only interested in controllability at a point where $(\xi, \eta) = (0, 0)$, let us write the vector field $Y_1 = P(Y_1')$ at such a configuration. We get

$$Y_1 = \Big((l_2 I_1 \sin\phi - \tfrac{1}{2}l_1 I_2 \sin 2\phi)\frac{\partial}{\partial\xi} - (\tfrac{1}{2}I_2 + ml_2^2 + ml_1 l_2 \cos\phi - \tfrac{1}{2}I_2 \cos 2\phi)\frac{\partial}{\partial\theta} +$$
$$\tfrac{1}{2}(I_1 + I_2 + ml_1^2 + 2ml_2^2 + 4ml_1 l_2 \cos\phi + (ml_1^2 - I_1 - I_2)\cos 2\phi)\frac{\partial}{\partial\phi} \Big)/\Delta$$

The symmetric product and Lie bracket formulas are too lengthy for us to write here so let us simply record the results of our investigations.

RR1. The local accessibility computations for the roller racer are rather complicated within our approach. Fortunately, Krishnaprasad and Tsakiris [1998] have performed a detailed analysis of the roller racer which includes local accessibility computations. A direct application of their results yields following results. Define a function f in Q in the coordinates (x, y, θ, ψ) by

$$f(x, y, \theta, \psi) = \frac{m(l_1 + l_2 \cos(\psi - \theta))(I_2 l_1 \cos(\psi - \theta) - I_1 l_2)}{(I_1 + I_2)\sin^2(\psi - \theta) + m(l_1 \cos(\psi - \theta) + l_2)^2}$$

Krishnaprasad and Tsakiris assert the following, although not in this exact language:

(a) $f(x, y, \theta, \psi) \neq 0$: In this case C_{hor} has full rank at (x, y, θ, ψ) and is spanned by the vector fields

$$\{Y_1, \langle Y_1 : Y_1\rangle, [Y_1, \langle Y_1 : Y_1\rangle], [\langle Y_1 : Y_1\rangle, [Y_1, \langle Y_1 : Y_1\rangle]]\}.$$

(b) $f(x, y, \theta, \psi) = 0$: In this case C_{hor} still has full rank at (x, y, θ, ψ) and is spanned by the vector fields

$$\{Y_1, \langle Y_1 : \langle Y_1 : Y_1\rangle\rangle, [Y_1, \langle Y_1 : \langle Y_1 : Y_1\rangle\rangle], [\langle Y_1 : \langle Y_1 : Y_1\rangle\rangle, [Y_1, \langle Y_1 : \langle Y_1 : Y_1\rangle\rangle]]\}.$$

Thus the roller racer is locally accessible from all initial configurations.

RR2. In any event, no matter what the accessibility properties of the roller racer, Theorem 3.7 asserts that it will not be locally configuration controllable. The issue of global controllability is an interesting one, but we do not address it here.

Note that even though the roller racer seems like a simple enough example, it is non-trivial to carry out the local configuration accessibility analysis of Section 3.2. One reason for this is that we have not taken into account the $SE(2)$ symmetry the problem possesses. The inclusion of symmetry will be a subject of future research, and will hopefully simplify the implementation of our results. Nonetheless, it is interesting to note that by casting the roller racer as a control system whose drift vector field is the geodesic spray of an affine connection, we are able to immediately employ the result of Lewis [1997] to state that the system is not locally configuration controllable. This immediately makes invalid any attempt to design controllers for the roller racer which are local in nature.

5. Discussion

The definitions and results of Lewis and Murray [1997a] are a valuable addition to the tools available for analysing simple mechanical control systems. In this paper we have extended these tools to systems with constraints. This extension greatly widens the applicability of the results of Lewis and Murray. Indeed, with these tools it is quite easy to provide a rather exhaustive analysis of some of the "standard" simple mechanical control systems with constraints. It is our hope that this analysis will stimulate interest in these methods as it is our opinion that they provide a solid basis from which to develop a fairly complete theory for simple mechanical systems. We refer to [Bullo, Leonard, and Lewis 2000] for an application of the ideas used in this paper to motion control. The methods of that paper should be readily applicable to the examples we study in this paper.

The representation of the equations of motion for constrained systems as geodesics of an affine connection may well offer insights beyond the control theoretic applications we have explored in this paper. In [Lewis 1998] some initial explorations are made into the subject of constrained systems with symmetries. The analysis in that paper provides, for example, some of the results obtained in [Bloch, Krishnaprasad, Marsden, and Murray 1996]. In particular, one can derive the "momentum equation" for constrained systems with symmetries using the methods of Lewis [1998]. It is likely there is much to explore in this area, as the geometry of affine connections brings to the table a new set of analysis tools. One interesting possibility is to make a connection with the work of Ostrowski [1995] and Bloch, Krishnaprasad, Marsden, and Murray [1996] on constrained systems with symmetry. In particular, it would be interesting to relate our controllability conditions to those of Ostrowski and Burdick [1997]. No doubt a proper account of symmetry (which all of our examples possess) would greatly simplify the computations in the examples we examined in this paper, especially those for the roller racer.

It is hoped that the paper provides some foundation for our belief that affine connections are important tools for analysing mechanical control systems, even in the presence of constraints. Although the application to a given example can be quite involved computationally, the geometric insight offered by the use of affine connections often provides a useful perspective. Further, as mentioned above, it is likely that a refinement of the methods we discuss here will make them easier to apply.

References

Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics*, second edition, Addison Wesley, Reading, MA, ISBN 0-8053-0102-X.

Abraham, R., Marsden, J. E., and Ratiu, T. S. [1988] *Manifolds, Tensor Analysis, and Applications*, second edition, number 75 in Applied Mathematical Sciences, Springer-Verlag, ISBN 0-387-96790-7.

Bates, L. M. and Śniatycki, J. Z. [1993] *Nonholonomic reduction*, Reports on Mathematical Physics, **32**(1), 444–452.

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

— [1995a] *Another view of nonholonomic mechanical control systems*, in *Proceedings of the 34th IEEE Conference on Decision and Control*, pages 1066–1071, Institute of Electrical and Electronics Engineers, New Orleans, LA.

— [1995b] *Nonholonomic and vakonomic control systems on Riemannian manifolds*, SIAM Journal on Control and Optimization, **33**(1), 126–148.

Bloch, A. M., Krishnaprasad, P. S., Marsden, J. E., and Murray, R. M. [1996] *Nonholonomic mechanical systems with symmetry*, Archive for Rational Mechanics and Analysis, **136**(1), 21–99.

Bloch, A. M., Reyhanoglu, M., and McClamroch, N. H. [1992] *Control and stabilization of nonholonomic dynamic systems*, Institute of Electrical and Electronics Engineers. Transactions on Automatic Control, **37**(11), 1746–1757.

Bullo, F., Leonard, N. E., and Lewis, A. D. [2000] *Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups*, Institute of Electrical and Electronics Engineers. Transactions on Automatic Control, **45**(8), 1437–1454.

Bullo, F. and Lewis, A. D. [1996] *Configuration controllability of mechanical systems on Lie groups*, Proceedings of MTNS '96.

Cattaneo, C. [1963] *Sulla struttura locale delle equazioni dinamiche di un sistema anholomo*, Atti della Accademia Nazionale dei Lincei. Rendiconti. Classe di Scienze Fisiche, Matematiche e Naturali. Serie VIII, **34**, 396–402.

Cattaneo-Gasparini, I. [1963] *Dérivée covariante "liée" dans une V_{n-1} riemannienne à structure presque produit*, Comptes Rendus de l'Académie des Sciences. Série I. Mathématique, **256**, 2089–2091.

Gallot, S., Hulin, D., and Lafontaine, J. [1987] *Riemannian Geometry*, Universitext, Springer-Verlag, New York–Heidelberg–Berlin, ISBN 0-387-17923-2.

Kobayashi, S. and Nomizu, K. [1963] *Foundations of Differential Geometry,* Volume I, number 15 in Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers, New York, ISBN 0-470-49647-9.

Krishnaprasad, P. S. and Tsakiris, D. P. [1995] *Oscillations, SE(2)-snakes and motion control*, in *Proceedings of the 34th IEEE Conference on Decision and Control*, New Orleans, LA.

— [1998] *Oscillations, SE(2)-snakes and motion control: a study of the roller racer*, Submitted to *Dynamics and Stability of Systems*.

Lanczos, C. [1986] *The Variational Principles of Mechanics*, Dover Publications, Inc., New York, ISBN 0-486-65067-7, reprint of fourth edition.

Leonard, N. E. and Krishnaprasad, P. S. [1995] *Motion control of drift-free, left-invariant systems on Lie groups*, Institute of Electrical and Electronics Engineers. Transactions on Automatic Control, **40**(9), 1539–1554.

Lewis, A. D. [1997] *Local configuration controllability for a class of mechanical systems with a single input*, Brussels, Belgium.

— [1998] *Affine connections and distributions with applications to nonholonomic mechanics*, Reports on Mathematical Physics, **42**(1/2), 135–164.

Lewis, A. D. and Murray, R. M. [1995] *Variational principles for constrained systems: Theory and experiment*, International Journal of Non-Linear Mechanics, **30**(6), 793–815.

— [1997a] *Controllability of simple mechanical control systems*, SIAM Journal on Control and Optimization, **35**(3), 766–790.

— [1997b] *Decompositions of control systems on manifolds with an affine connection*, Systems & Control Letters, **31**(4), 199–205.

Lewis, A. D., Ostrowski, J. P., Murray, R. M., and Burdick, J. W. [1994] *Nonholonomic mechanics and locomotion: the Snakeboard example*, pages 2391–2400, Institute of Electrical and Electronics Engineers, San Diego.

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

Ostrowski, J. P. [1995] *The Mechanics and Control of Undulatory Robotic Locomotion*, Ph.D. thesis, California Institute of Technology, Pasadena, California, USA. URL: http://www.cds.caltech.edu/

Ostrowski, J. P. and Burdick, J. W. [1997] *Controllability tests for mechanical systems with constraints and symmetries*, Journal of Applied Mathematics and Computer Science, **7**(2), 101–127.

Pars, L. A. [1965] *A Treatise on Analytical Dynamics*, John Wiley and Sons, New York, New York, ISBN 0-918024-07-2.

Rathinam, M. and Murray, R. M. [1998] *Configuration flatness of Lagrangian systems underactuated by one control*, SIAM Journal on Control and Optimization, **36**(1), 164–179.

Sussmann, H. J. [1983] *Lie brackets and local controllability: a sufficient condition for scalar-input systems*, SIAM Journal on Control and Optimization, **21**(5), 686–713.

— [1987] *A general theorem on local controllability*, SIAM Journal on Control and Optimization, **25**(1), 158–194.

Synge, J. L. [1928] *Geodesics in nonholonomic geometry*, Mathematische Annalen, **99**, 738–751.

Vershik, A. M. [1984] *Classical and Non-classical Dynamics with Constraints*, pages 278–301, number 1108 in Lecture Notes in Mathematics, Springer-Verlag, New York–Heidelberg–Berlin.