Poisson Reduction of Optimal Control Systems

Tomoki Ohsawa*

Abstract— **This paper explores the role of symmetries and reduction in nonlinear control and optimal control systems. We formulate symmetries in nonlinear control systems and then link it to symmetries in optimal control of such systems. We apply Poisson reduction to the Pontryagin maximum principle to reduce the optimal control system. Affine and kinematic optimal control systems are of particular interest: We show an example of kinematic optimal control to illustrate how the reduction simplifies optimal control systems.**

I. Introduction

A. Background

Many control systems, particularly those arising from mechanical systems, have symmetries. Such a symmetry is usually described as an invariance or equivariance under an action of a Lie group, and the system can be reduced to a lower-dimensional one or decoupled into subsystems by exploiting the symmetry. Nijmeijer and van der Schaft [1] and Grizzle and Marcus [2] formulated symmetries of nonlinear control systems, and also showed how one can reduce a control system with symmetry to a quotient space.

Likewise, optimal control systems also often have such symmetries as well, and various techniques are proposed to exploit such symmetries [3–8]. A certain class of optimal control problems has a rich geometric structure, and provides many interesting questions relating differential-geometric ideas with control-theoretic problems. Most notably, Montgomery [9, 10, 11, 12, 13], following the work of Shapere and Wilczek [14], explored optimal control of deformable bodies, such as the falling cat problem; in particular, principal bundles, along with connections on them defined by momentum maps, are identified as a natural geometric setting for such problems. See also [15–17] for applications of the same geometric idea to kinematic control of nonholonomic mechanical systems.

B. Main Results

We characterize symmetries in nonlinear control and optimal control systems, and apply the Poisson reduction of Cendra et al. [18] to the Hamiltonian system given as a necessary condition for optimality by the Pontryagin maximum principle. The result synthesizes some previous works, including optimal control of deformable bodies mentioned above and also the Lie–Poisson reduction of control systems on Lie groups in Krishnaprasad [19]

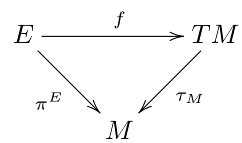
Many details, including proofs, are omitted for brevity; we would like to refer to [20] for them.

*Department of Mathematics, University of California, San Diego. e-mail: tohsawa@ucsd.edu

II. Symmetry and Reduction of Nonlinear Control Systems

A. Nonlinear Control Systems

Let M be a smooth manifold and $\tau_M : TM \to M$ be its tangent bundle; let $E := M \times \mathbb{R}^d$ and see $\pi^E : E \to M$ as a (trivial) vector bundle[1]; also let $f : E \to TM$ be a fiber-preserving smooth map, i.e., the diagram



commutes. Then, a *nonlinear control system* is defined by

$$\dot{x} = f(x, u). \tag{1}$$

B. Symmetry in Nonlinear Control Systems

Following [1, 2], we assume that the control system (1) has a symmetry in the following sense: Let G be a Lie group acting on M freely and properly; we have $\Phi : G \times M \to M$ or $\Phi_g : M \to M$ for any $g \in G$; as a result we have the principal bundle

$$\pi : M \to M/G.$$

The action Φ_g gives rise to the tangent lift $T\Phi_g : TM \to TM$. Let us also assume that we have a linear representation of G on the control space $\mathbb{R}^d$, i.e., we have a group homomorphism $\sigma_{(\cdot)} : G \to GL(d, \mathbb{R})$. Then, we define an action of G on $E = M \times \mathbb{R}^d$ as follows:

$$\Psi_g : E \to E; \quad (x, u) \mapsto (\Phi_g(x), \sigma_g(u)) = (gx, gu), \tag{2}$$

where we wrote $gx := \Phi_g(x)$ and $gu := \sigma_g(u)$.

We say that the nonlinear control system (1) has a G-symmetry if the map $f : E \to TM$ is equivariant under the G-actions on E and TM defined above, i.e., for any $g \in G$,

$$T\Phi_g \circ f = f \circ \Psi_g. \tag{3}$$

C. Reduced Control System

The equivariance of the map f shown above gives rise to the map $\bar{f} : E/G \to TM/G$, and hence defines the reduced control system. We can explicitly write down the reduced control system by introducing a principal bundle connection and also using the identification of the quotient TM/G with a bundle over M/G introduced in [21, Section 2.3] in the following way: Consider the Whitney sum $T(M/G) \oplus \tilde{\mathfrak{g}}$, where $\tilde{\mathfrak{g}}$ is the associated (vector) bundle defined as

$$\tilde{\mathfrak{g}} := M \times_G \mathfrak{g} = (M \times \mathfrak{g})/G$$

[1]More generally, we may take a fiber bundle for E (see, e.g., [1]).

with $\mathfrak{g}$ being the Lie algebra of the Lie group G; given a principal bundle connection

$$\mathcal{A} : TM \to \mathfrak{g}, \tag{4}$$

we introduce the identification (see [21, Section 2.4])

$$\begin{aligned} \alpha_{\mathcal{A}} : TM/G &\to T(M/G) \oplus \tilde{\mathfrak{g}} \\ [v_x]_G &\mapsto T_x\pi(v_x) \oplus [x, \mathcal{A}_x(v_x)]_G, \end{aligned} \tag{5}$$

where $[\,\cdot\,]_G$ stands for an equivalence class defined by the G-action. On the other hand, since $E = M \times \mathbb{R}^d$, the quotient E/G defines the associated bundle

$$E/G = (M \times \mathbb{R}^d)/G = M \times_G \mathbb{R}^d,$$

which is also a vector bundle over M/G [21, Section 2.3].

Therefore, we may introduce the maps $\bar{f}_{M/G} : E/G \to T(M/G)$ and $\bar{f}_{\tilde{\mathfrak{g}}} : E/G \to \tilde{\mathfrak{g}}$ defined by

$$\begin{aligned} \bar{f}_{M/G}([x,u]_G) &:= T_x\pi \circ f(x,u), \\ \bar{f}_{\tilde{\mathfrak{g}}}([x,u]_G) &:= [x, \mathcal{A}_x(f(x,u))]_G \end{aligned}$$

for any element $[x,u]_G \in E/G = M \times_G \mathbb{R}^d$. Then, we have $\alpha_{\mathcal{A}} \circ \bar{f} = \bar{f}_{M/G} \oplus \bar{f}_{\tilde{\mathfrak{g}}}$, and thus the reduced system is decoupled into two subsystems:

$$\dot{\bar{x}} = \bar{f}_{M/G}(\bar{u}_{\bar{x}}), \qquad \tilde{\xi}_{\bar{x}} = \bar{f}_{\tilde{\mathfrak{g}}}(\bar{u}_{\bar{x}}), \tag{6}$$

where $\bar{x} := \pi(x)$, $\bar{u}_{\bar{x}} := [x,u]_G$, and $\tilde{\xi}_{\bar{x}} := [x, \mathcal{A}_x(\dot{x})]_G$.

III. Symmetry and Reduction of Affine Control Systems

A. Symmetry in Affine Control Systems

Consider an *affine control system*, i.e., (1) with

$$f(x,u) = X_0(x) + \sum_{i=1}^{d} u^i X_i(x), \tag{7}$$

where $\{X_i\}_{i=0}^d$ are linearly independent vector fields on M. We assume that the vector field X_0 is G-invariant, i.e.,

$$T\Phi_g \circ X_0 = X_0 \circ \Phi_g \tag{8}$$

for any $g \in G$; also let $\mathcal{D} \subset TM$ be the distribution

$$\mathcal{D} = \operatorname{span}\{X_1, \dots, X_d\}, \tag{9}$$

and assume that it is invariant under the tangent lift of the G-action on Q, i.e.,

$$T\Phi_g(\mathcal{D}) = \mathcal{D} \tag{10}$$

for any $g \in G$. This implies that, for each vector field X_i for $i = 1, \dots, d$ and any $x \in M$ and $g \in G$, we have

$$T_x\Phi_g(X_i(x)) = \sum_{j=1}^{d} R_i^j(g)\, X_j(gx), \tag{11}$$

where $R(g)$ is an invertible $d \times d$ matrix, and is assumed to have no dependence on x; then the matrix $R^T(g)$ gives the representation $\sigma_{(\cdot)} : G \to GL(d, \mathbb{R})$, i.e., $\sigma_g = R^T(g)$.

This gives rise to an action of G on $E = M \times \mathbb{R}^d$, i.e., $\Psi_g : E \to E$ defined by $\Psi_g : (x,u) \mapsto \big(gx, R^T(g)u\big)$. Then, the symmetries of X_0 and $\mathcal{D}$, i.e., (8) and (10), imply that of f, i.e., Eq. (3). Therefore, this becomes a special case of the general setting considered in Section II.

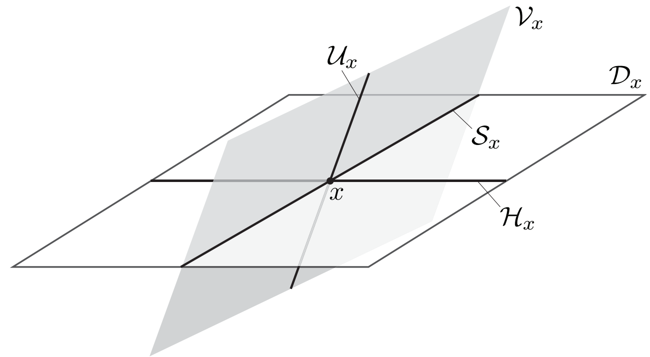


Fig. 1. Nonholonomic connection [22, 23]. $\mathcal{D}_x$ is spanned by the control vector fields $\{X_i\}_{i=1}^d$; $\mathcal{V}_x$ is the tangent space to the group orbit through $x \in M$; $\mathcal{H}_x$ defines a principal connection.

B. Nonholonomic Connection

Recall that we need to choose a connection $\mathcal{A}$ (see (4)). For affine control systems, the *nonholonomic connection* introduced in [22, Section 6.4] for reduction of nonholonomic mechanical systems turns out to be a natural choice.

Let $\mathcal{O}(x)$ be the orbit of the G-action on M through $x \in M$ and $\mathcal{V}_x$ be its tangent space at x, i.e.,

$$\mathcal{O}(x) := \{\Phi_g(x) \in M \mid g \in G\}, \qquad \mathcal{V}_x := T_x\mathcal{O}(x).$$

We make the "dimension assumption" [22], i.e., $T_xM = \mathcal{D}_x + \mathcal{V}_x$, and let

$$\mathcal{S}_x := \mathcal{D}_x \cap \mathcal{V}_x.$$

Then, we may write

$$\mathcal{D}_x = \mathcal{H}_x \oplus \mathcal{S}_x, \qquad \mathcal{V}_x = \mathcal{S}_x \oplus \mathcal{U}_x,$$

with complementary subspaces $\mathcal{H}_x$ and $\mathcal{U}_x$, and thus we have the following decomposition of the tangent space T_xM (see Fig. 1):

$$T_xM = \mathcal{H}_x \oplus \mathcal{V}_x = \mathcal{H}_x \oplus \mathcal{S}_x \oplus \mathcal{U}_x.$$

If, in addition, $\mathcal{H}$ is G-invariant, then it defines a G-invariant horizontal space in the principal bundle $\pi : M \to M/G$, and thus defines a connection $\mathcal{A} : TM \to \mathfrak{g}$ such that $\ker \mathcal{A}_x = \mathcal{H}_x$. This is called the *nonholonomic connection* [22, 23].

Using the nonholonomic connection, the reduced control system (6) can be written as

$$\begin{aligned} \dot{\bar{x}} &= \bar{X}_0(\bar{x}) + \sum_{i=1}^{d} u^i \bar{X}_i(\bar{x}), \\ \tilde{\xi}_{\bar{x}} &= [x, \mathcal{A}_x \cdot X_0(x)]_G + \sum_{i=1}^{d} u^i \,[x, \mathcal{A}_x \cdot X_i(x)]_G\,, \end{aligned} \tag{12}$$

where $\bar{X}_i := T\pi(X_i)$ for $i = 0, 1, \dots, d$.

C. Purely Kinematic Case

Consider the special case where the tangent space to the group orbit $\mathcal{V}_x = T_x\mathcal{O}(x)$ exactly complements the distribution $\mathcal{D}_x$, i.e., $\mathcal{S}_x = 0$ and thus $T_xM = \mathcal{D}_x \oplus \mathcal{V}_x$. This is the so-called "purely kinematic" case [22]. In this case, $\mathcal{D}_x$ itself gives the horizontal space and thus defines the connection $\mathcal{A}: TM \to \mathfrak{g}$ such that $\ker \mathcal{A}_x = \mathcal{D}_x$. This is the basic setting for control of deformable bodies (see, e.g., [11]) and also of robotic locomotion (see, e.g., [15–17]).

IV. Symmetry in Control Systems on Lie Groups

Consider, as a special case, the nonlinear control system (1) on a Lie group G, i.e., $M = G$, with symmetry under the action of G on itself by left translation $L_g : G \to G;\ h \mapsto gh$ for any $g \in G$. This case is particularly simple because we do not need a principal connection and the reduced system is defined on the Lie algebra $\mathfrak{g}$.

The associated bundle $E/G = M \times_G \mathbb{R}^d$ becomes a bundle over G/G, i.e., a point; hence $E/G \cong \mathbb{R}^d = \{u\}$. On the other hand, the quotient TM/G becomes $TG/G \cong \mathfrak{g}$. Therefore, we have $\bar{f}_{\mathfrak{g}} : \mathbb{R}^d \to \mathfrak{g}$ and the control system reduces to

$$\xi(t) = \bar{f}_{\mathfrak{g}}(\bar{u}(t)). \tag{13}$$

where $\xi := T_g L_{g^{-1}}(\dot{g})$.

In particular, for the affine control system (7), we obtain

$$\bar{f}_{\mathfrak{g}}(u) := \zeta_0 + \sum_{i=1}^{d} u^i \zeta_i, \tag{14}$$

where $\zeta_i \in \mathfrak{g}$ is defined such that $X_i(g) = T_e L_g(\zeta_i)$ for any $g \in G$ and $i = 0, 1, \dots, d$. This is the case considered by Krishnaprasad [19].

V. Symmetry in Optimal Control Systems

This section shows how the symmetry of a nonlinear control system implies that of the corresponding optimal control system under the assumption that the cost function is also G-invariant, as suggested by Grizzle and Marcus [3].

A. Pontryagin Maximum Principle and Symmetry in Optimal Control

Given a cost function $C : E \to \mathbb{R}$, fixed times t_0 and t_1 such that $t_0 < t_1$, and fixed points x_0 and x_1 in M, we formulate an *optimal control problem* as follows: Minimize the cost functional, i.e.,

$$\min_{u(\cdot)} \int_{t_0}^{t_1} C(x(t), u(t))\, dt,$$

subject to (1) and the endpoint constraints $x(t_0) = x_0$ and $x(t_1) = x_1$.

A Hamiltonian structure comes into play with the introduction of the augmented cost functional:

$$\begin{aligned}\hat{S} &:= \int_{t_0}^{t_1} \left[C(x(t), u(t)) + \langle \lambda(t), \dot{x}(t) - f(x(t), u(t))\rangle\right] dt \\ &= \int_{t_0}^{t_1} \Big[\langle \lambda(t), \dot{x}(t)\rangle - \hat{H}(x(t), \lambda(t), u(t))\Big]\, dt,\end{aligned}$$

where we introduced the costate $\lambda(t) \in T^*M$, and also defined the *control Hamiltonian* $\hat{H} : T^*M \oplus E \to \mathbb{R}$ by

$$\hat{H}(\lambda_x, u_x) = \hat{H}(x, \lambda, u) := \langle \lambda_x, f(u_x)\rangle - C(u_x), \tag{15}$$

where we wrote $\lambda_x := (x, \lambda) \in T^*_x M$ and $u_x := (x, u) \in E_x$ (recall that $E = M \times \mathbb{R}^d$ is a (trivial) vector bundle over M). If the cost function is invariant under the G-action Ψ defined in (2), i.e., for any $g \in G$,

$$C \circ \Psi_g = C, \tag{16}$$

then the control Hamiltonian $\hat{H}$ has a symmetry in the following sense: Define an action of G on the bundle $T^*M \oplus E$ as follows: For any $g \in G$,

$$\begin{gathered}\hat{\Psi}_g : T^*M \oplus E \to T^*M \oplus E; \\ (\lambda_x, u_x) \mapsto \left(T^*\Phi_{g^{-1}}(\lambda_x), \Psi_g(u_x)\right),\end{gathered}$$

where $T^*\Phi_{g^{-1}} : T^*M \to T^*M$ is the cotangent lift of Φ_g. Then, for any $g \in G$,

$$\hat{H} \circ \hat{\Psi}_g = \hat{H}. \tag{17}$$

Now, for an arbitrary fixed $\lambda_x \in T^*_x M$, define

$$\begin{gathered}\mathbb{F}_{\rm c}\hat{H}(\lambda_x, \,\cdot\,) : E_x \to E_x^*; \\ \left\langle \mathbb{F}_{\rm c}\hat{H}(\lambda_x, u_x), w_x \right\rangle = \left.\frac{d}{d\varepsilon}\hat{H}(\lambda_x, u_x + \varepsilon\, w_x)\right|_{\varepsilon=0}\end{gathered}$$

for an arbitrary $w_x \in E_x$. We assume that the optimal control $u_x^\star : T^*_x M \to E_x \cong \mathbb{R}^d$ is uniquely determined by

$$\mathbb{F}_{\rm c}\hat{H}(\lambda_x, u_x^\star(\lambda_x)) = 0$$

for any $\lambda_x \in T^*_x M$. This gives rise to the fiber-preserving bundle map

$$u^\star : T^*M \to E; \quad \lambda_x \mapsto u_x^\star(\lambda_x), \tag{18}$$

and so we may define the optimal Hamiltonian $H : T^*M \to \mathbb{R}$ by

$$\begin{aligned}H(\lambda_x) &:= \hat{H} \circ u^\star(\lambda_x) \\ &= \langle \lambda_x, f(u_x^\star(\lambda_x))\rangle - C(u_x^\star(\lambda_x)).\end{aligned} \tag{19}$$

It is easy to show that the optimal control $u^\star : T^*M \to E$ is equivariant under the G-actions, i.e., for any $g \in G$,

$$\Psi_g \circ u^\star = u^\star \circ T^*\Phi_{g^{-1}}. \tag{20}$$

Hence the optimal Hamiltonian H has a symmetry in the usual sense, i.e., for any $g \in G$,

$$H \circ T^*\Phi_{g^{-1}} = H. \tag{21}$$

VI. Reduction of Pontryagin Maximum Principle

A. Pontryagin Maximum Principle

The Pontryagin maximum principle says that the optimal flow on M of the control system (1) is necessarily the projection to M of the Hamiltonian flow on T^*M with the above Hamiltonian H. Specifically, let Ω be the standard symplectic form on T^*M, $\pi_M : T^*M \to M$ the cotangent bundle projection, and X_H the Hamiltonian vector field defined by

$$i_{X_H}\Omega = dH; \tag{22}$$

then there exists a solution $\lambda : [t_0, t_1] \to T^*M$ of the above Hamiltonian system with $\pi_M(\lambda(t_0)) = x_0$ and $\pi_M(\lambda(t_1)) = x_1$ such that its projection to M, $\pi_M \circ \lambda : [t_0, t_1] \to M$, is the optimal trajectory of the control system (see, e.g., Agrachev and Sachkov [24, Chapter 12] for more details).

B. Poisson Reduction and Hamilton–Poincaré Equations

We saw that the optimal Hamiltonian H has a symmetry under the G-action. It implies that we can apply the results of symmetry reduction of Hamiltonian systems to (22) to obtain a reduced Hamiltonian system for the optimal flow.

Reduction of Hamiltonian systems is a well-developed subject, whose roots go back to the symplectic reduction of Marsden and Weinstein [25]; there have been substantial subsequent developments (see [26] and references therein). In our case, the Poisson version of the cotangent bundle reduction (see [18] and [26, Section 2.3]; see also [27, 28]) turns out to be a natural choice for the following reason: Recall that we derived the reduced control system (6) using the identification $\alpha_{\mathcal{A}} : TM/G \to T(M/G) \oplus \tilde{\mathfrak{g}}$ defined in (5); hence it is natural to expect and also is desirable that the maximum principle, originally formulated on T^*M, reduces to the dual $T^*(M/G) \oplus \tilde{\mathfrak{g}}^*$. The Poisson version of the cotangent bundle reduction works precisely this way: The Poisson structure on T^*M reduces to that on $T^*(M/G) \oplus \tilde{\mathfrak{g}}^*$; accordingly, Hamilton's equations reduce to the Hamilton–Poincaré equations [18].

As shown in [26, Lemma 2.3.3], the identification of T^*M with $T^*(M/G) \oplus \tilde{\mathfrak{g}}^*$ is provided by the dual of the inverse of $\alpha_{\mathcal{A}}$:

$$
\begin{aligned}
(\alpha_{\mathcal{A}}^{-1})^* &: T^*M/G \to T^*(M/G) \oplus \tilde{\mathfrak{g}}^*; \\
&[\lambda_x]_G \mapsto \mathrm{hl}_x^*(\lambda_x) \oplus [x, \mathbf{J}(\lambda_x)]_G,
\end{aligned}
\tag{23}
$$

where $\mathrm{hl}_x^* : T_x^*M \to T_{\bar{x}}^*(M/G)$ is the adjoint of the horizontal lift $\mathrm{hl}_x : T_{\bar{x}}(M/G) \to T_xM$ associated with the connection $\mathcal{A} : TM \to \mathfrak{g}$, and $\mathbf{J} : T^*M \to \mathfrak{g}^*$ is the momentum map corresponding to the G-symmetry: Let $\xi \in \mathfrak{g}$ and $\xi_M \in \mathfrak{X}(M)$ its infinitesimal generator. Then,

$$
\langle \mathbf{J}(\lambda_x), \xi \rangle = \langle \lambda_x, \xi_M(x) \rangle .
$$

Noether's theorem (see, e.g., [29, Section 11.4]) says that the G-invariance of H implies that $\mathbf{J}$ is conserved along the flow of the Hamiltonian vector field X_H.

Cendra et al. [18] exploit this identification to reduce the Hamiltonian dynamics with a G-invariant Hamiltonian $H : T^*M \to \mathbb{R}$ as follows: The G-invariance implies that one can define the reduced Hamiltonian on T^*M/G, which is identified with $T^*(M/G) \oplus \tilde{\mathfrak{g}}^*$ by (23), i.e., one has $\bar{H} : T^*(M/G) \oplus \tilde{\mathfrak{g}}^* \to \mathbb{R}$. Then, through the reduction of Hamilton's phase space principle, i.e.,

$$
\delta \int_{t_0}^{t_1} [\langle p, \dot{q} \rangle - H(q,p)]\, dt = 0
$$

with $\delta q(t_0) = \delta q(t_1) = 0$, one obtains the Hamilton–Poincaré equations defined on $T^*(M/G) \oplus \tilde{\mathfrak{g}}^*$ (see [18] for details and an intrinsic expression):

$$
\begin{gathered}
\dot{\bar{x}}^\alpha = \frac{\partial \bar{H}}{\partial \bar{\lambda}^\alpha}, \qquad \tilde{\xi}^a = \frac{\partial \bar{H}}{\partial \tilde{\mu}_a}, \\
\dot{\bar{\lambda}}_\alpha = -\frac{\partial \bar{H}}{\partial \bar{x}^\alpha} - \tilde{\mu}_a \left(\mathcal{B}^a_{\beta\alpha} \dot{\bar{x}}^\beta + \mathcal{A}^b_\alpha C^a_{db} \frac{\partial \bar{H}}{\partial \tilde{\mu}_d} \right), \\
\dot{\tilde{\mu}}_a = \tilde{\mu}_b C^b_{da} \left(\frac{\partial \bar{H}}{\partial \tilde{\mu}_d} - \mathcal{A}^d_\alpha \dot{\bar{x}}^\alpha \right),
\end{gathered}
\tag{24}
$$

where $\bar{\lambda}_x := \mathrm{hl}_x^*(\lambda_x) \in T_x^*(M/G)$; $\tilde{\xi}_x \in \tilde{\mathfrak{g}}_x$ and $\tilde{\mu}_x \in \tilde{\mathfrak{g}}_x^*$ are the *locked body angular velocity* and its corresponding momentum (see [22, Section 5.3]) defined by

$$
\begin{aligned}
\tilde{\xi}^a &:= ([x, \mathcal{A}_x(\dot{x})]_G)^a = \xi^a + \mathcal{A}^a_\alpha \dot{\bar{x}}^\alpha, \\
\tilde{\mu}_a &:= ([x, \mathbf{J}(\lambda_x)]_G)_a = \left(\mathrm{Ad}_g^* \mathbf{J}(\lambda_x) \right)_a .
\end{aligned}
$$

with $\xi = T_g L_{g^{-1}}(\dot{g})$; the coefficients $\mathcal{A}^a_\alpha$ are defined in the coordinate expression for the connection $\mathcal{A}$ as follows:

$$
\mathcal{A}_{(\bar{x},g)}(\dot{\bar{x}}, \dot{g}) = \mathrm{Ad}_g(\xi^a + \mathcal{A}^a_\alpha \dot{\bar{x}}^\alpha)\, \mathbf{e}_a,
$$

where $\{\mathbf{e}_a\}_{a=1}^{\dim G}$ is a basis for the Lie algebra $\mathfrak{g}$. Also the coefficients $\mathcal{B}^a_{\beta\alpha}$ for the curvature are given by

$$
\mathcal{B}^a_{\beta\alpha} = \frac{\partial \mathcal{A}^a}{\partial \bar{x}^\alpha} - \frac{\partial \mathcal{A}^a}{\partial \bar{x}^\beta} - C^a_{bc} \mathcal{A}^b_\alpha \mathcal{A}^c_\beta .
$$

C. Poisson Reduction of Pontryagin Maximum Principle

Let us apply the above Poisson reduction to the Hamiltonian system (22) defined by the maximum principle. First calculate the reduced optimal Hamiltonian $\bar{H}$ corresponding to the optimal Hamiltonian (19). Using the identification in (23) and also the reduced optimal control $\bar{u}^\star : T^*M/G \to E/G$, which is well-defined by virtue of (20), we can rewrite the Hamiltonian H as follows:

$$
\begin{aligned}
H(\lambda_x) &= \left\langle (\alpha_{\mathcal{A}}^{-1})^*(\lambda_x), \alpha_{\mathcal{A}} \circ f(u_x^\star(\lambda_x)) \right\rangle - C(u_x^\star(\lambda_x)) \\
&= \left\langle \mathrm{hl}_x^*(\lambda_x), \bar{f}^\star_{M/G}([\lambda_x]_G) \right\rangle \\
&\quad + \left\langle [x, \mathbf{J}(\lambda_x)]_G, \bar{f}^\star_{\tilde{\mathfrak{g}}}([\lambda_x]_G) \right\rangle - \bar{C}^\star([\lambda_x]_G),
\end{aligned}
$$

where we defined the reduced cost function $\bar{C} : E/G \to \mathbb{R}$ by $\bar{C} \circ \pi_G^E = C$ and also

$$
\begin{aligned}
\bar{f}^\star_{M/G}([\lambda_x]_G) &:= \bar{f}_{M/G} \circ \bar{u}^\star_{\bar{x}}([\lambda_x]_G), \\
\bar{f}^\star_{\tilde{\mathfrak{g}}}([\lambda_x]_G) &:= \bar{f}_{\tilde{\mathfrak{g}}} \circ \bar{u}^\star_{\bar{x}}([\lambda_x]_G), \\
\bar{C}^\star([\lambda_x]_G) &:= \bar{C}(\bar{u}^\star_{\bar{x}}([\lambda_x]_G)).
\end{aligned}
$$

Define the reduced optimal Hamiltonian $\bar{H} : T^*(M/G) \oplus \tilde{\mathfrak{g}}^* \to \mathbb{R}$ by

$$
\begin{aligned}
\bar{H}(\bar{\lambda}_{\bar{x}} \oplus \tilde{\mu}_{\bar{x}}) &:= \left\langle \bar{\lambda}_{\bar{x}}, \bar{f}^\star_{M/G}(\bar{\lambda}_{\bar{x}} \oplus \tilde{\mu}_{\bar{x}}) \right\rangle \\
&\quad + \left\langle \tilde{\mu}_{\bar{x}}, \bar{f}^\star_{\tilde{\mathfrak{g}}}(\bar{\lambda}_{\bar{x}} \oplus \tilde{\mu}_{\bar{x}}) \right\rangle - \bar{C}^\star(\bar{\lambda}_{\bar{x}} \oplus \tilde{\mu}_{\bar{x}}) .
\end{aligned}
$$

Then, we have $H(\lambda_x) = \bar{H}(\bar{\lambda}_{\bar{x}} \oplus \tilde{\mu}_{\bar{x}})$ with $\bar{\lambda}_{\bar{x}} := \mathrm{hl}_x^*(\lambda_x)$ and $\tilde{\mu}_{\bar{x}} := [x, \mathbf{J}(\lambda_x)]_G$. In coordinates, we can write

$$
\begin{aligned}
\bar{H}(\bar{x}, \bar{\lambda}, \tilde{\mu}) &= \bar{\lambda}_\alpha \, \bar{f}^{\star,\alpha}_{M/G}(\bar{x}, \bar{\lambda}, \tilde{\mu}) \\
&\quad + \tilde{\mu}_a \, \bar{f}^{\star,a}_{\tilde{\mathfrak{g}}}(\bar{x}, \bar{\lambda}, \tilde{\mu}) - \bar{C}^\star(\bar{x}, \bar{\lambda}, \tilde{\mu}) .
\end{aligned}
\tag{25}
$$

Applying the Hamilton–Poincaré equations (24) to this particular choice of $\bar{H}$ gives the following:

Theorem 1: Suppose that the nonlinear control system (1) and the cost function have G-symmetries in the sense of Eqs. (3) and (16). Then, the necessary condition of the Pontryagin maximum principle reduces to the following set of equations (see [20] for an intrinsic expression):

$$
\begin{gathered}
\dot{\bar{x}}^\alpha = \bar{f}^{\star,\alpha}_{M/G}(\bar{x}, \bar{\lambda}, \tilde{\mu}), \qquad \tilde{\xi}^a = \bar{f}^{\star,a}_{\tilde{\mathfrak{g}}}(\bar{x}, \bar{\lambda}, \tilde{\mu}), \\
\dot{\bar{\lambda}}_\alpha = -\frac{\partial \bar{H}}{\partial \bar{x}^\alpha} - \tilde{\mu}_a \left(\mathcal{B}^a_{\beta\alpha} \dot{\bar{x}}^\beta + \mathcal{A}^b_\alpha C^a_{db} \bar{f}^{\star,d}_{\tilde{\mathfrak{g}}}(\bar{x}, \bar{\lambda}, \tilde{\mu}) \right), \\
\dot{\tilde{\mu}}_a = \tilde{\mu}_b C^b_{da} \left(\bar{f}^{\star,d}_{\tilde{\mathfrak{g}}}(\bar{x}, \bar{\lambda}, \tilde{\mu}) - \mathcal{A}^d_\alpha \dot{\bar{x}}^\alpha \right).
\end{gathered}
\tag{26}
$$

Remark 2: Notice that the equations for $(\bar{x}, \bar{\lambda}, \tilde{\mu})$ are decoupled from the second one. Thus one first solves this subsystem and then solve the second equation to "reconstruct" the dynamics in the group variables.

Remark 3: If the Lie group G is Abelian, then the structure constants C^a_{bc} vanish, and thus we have

$$\begin{aligned} \dot{\bar{x}}^\alpha &= \bar{f}^{\star,\alpha}_{M/G}(\bar{x}, \bar{\lambda}, \tilde{\mu}), \qquad \tilde{\xi}^a = \bar{f}^{\star,a}_{\tilde{\mathfrak{g}}}(\bar{x}, \bar{\lambda}, \tilde{\mu}), \\ \dot{\bar{\lambda}}_\alpha &= -\frac{\partial \bar{H}}{\partial \bar{x}^\alpha} - \tilde{\mu}_a \mathcal{B}^a_{\beta\alpha} \dot{\bar{x}}^\beta, \qquad \dot{\tilde{\mu}}_a = 0. \end{aligned} \tag{27}$$

In particular, the last equation directly gives a conservation of the momentum map $\mathbf{J}$, which simplifies the set of equations further.

The following kinematic optimal control problem illustrates the theory; note that this is non-purely kinematic case with an Abelian symmetry, and so Remark 3 applies here.

Example 4 (Snakeboard; see, e.g., [30], [22] and [31]): We consider a kinematic optimal control problem of the snakeboard shown in Fig. 2. The configuration space is

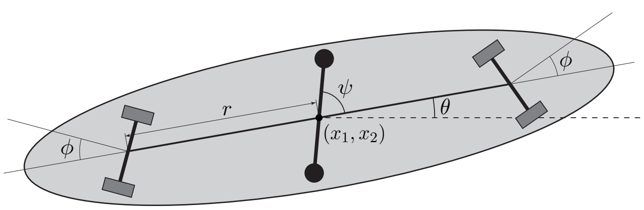


Fig. 2. The Snakeboard.

$M = SE(2) \times \mathbb{S}^1 \times \mathbb{S}^1 = \{(x_1, x_2, \theta, \psi, \phi)\}$. The velocity constraints are given by

$$\dot{x}_1 + (r\cos\theta\cot\phi)\,\dot{\theta} = 0, \quad \dot{x}_2 + (r\sin\theta\cot\phi)\,\dot{\theta} = 0,$$

and thus we have $\mathcal{D} = \operatorname{span}\{X_1, X_2, X_3\}$ with

$$\begin{aligned} X_1(x) &= \cos\theta\,\frac{\partial}{\partial x_1} + \sin\theta\,\frac{\partial}{\partial x_2} - \frac{\tan\phi}{r}\,\frac{\partial}{\partial\theta}, \\ X_2(x) &= \frac{\partial}{\partial\psi}, \qquad X_3(x) = \frac{\partial}{\partial\phi}, \end{aligned}$$

where $x = (x_1, x_2, \theta, \psi, \phi)$. Therefore, we may consider the following kinematic control system:

$$\dot{x} = f(x, u) := u_1 X_1(x) + u_2 X_2(x) + u_3 X_3(x),$$

or more explicitly,

$$\dot{x}_1 = u_1\cos\theta, \qquad \dot{x}_2 = u_1\sin\theta, \qquad \dot{\theta} = -u_1\frac{\tan\phi}{r}, \qquad \dot{\psi} = u_2, \qquad \dot{\phi} = u_3.$$

We define the cost function $C : SE(2)\times\mathbb{R}^3 \to \mathbb{R}$ as follows:

$$C(x, u) = \frac{1}{2}(u_1^2 + u_2^2 + u_3^2).$$

Then, the above control system has an $SE(2) \times SO(2)$-symmetry, where $SE(2)$ acting on the $SE(2)$ portion of M by left multiplication and $SO(2)$ acting on the first $\mathbb{S}^1$ in M, i.e., the variable ψ. We choose, however, the subgroup $G = \mathbb{R}^2 \times SO(2)$ of $SE(2)\times SO(2)$ since we are interested in an Abelian and non-purely kinematic case here.

Let $\Phi : G \times M \to M$ be the G-action on M, i.e.,

$$\Phi : ((a, b, \beta), (x_1, x_2, \theta, \psi, \phi)) \mapsto (x_1+a, x_2+b, \theta, \psi+\beta, \phi).$$

Also let $\sigma : G \times \mathbb{R}^3 \to \mathbb{R}^3$ be the trivial representation:

$$\sigma : ((a, b, \beta), (u_1, u_2, u_3)) \mapsto (u_1, u_2, u_3),$$

which induces the action $\Psi : G \times E \to E$ defined by

$$\begin{aligned} \Psi : ((a, b, \beta), &(x_1, x_2, \theta, \psi, \phi, u_1, u_2, u_3)) \\ &\mapsto (x_1 + a, x_2 + b, \theta, \psi + \beta, \phi, u_1, u_2, u_3). \end{aligned}$$

Then, it is straightforward to show that f and C satisfy the symmetry defined in Eqs. (3) and (16), respectively.

The optimal control $u^\star$ is given by

$$u_1^\star = \lambda_1\cos\theta + \lambda_2\sin\theta - \lambda_\theta\,\frac{\tan\phi}{r}, \quad u_2^\star = \lambda_\psi, \quad u_3^\star = \lambda_\phi,$$

and then the optimal Hamiltonian is

$$H(x, \lambda) = \frac{1}{2}\left(\lambda_1\cos\theta + \lambda_2\sin\theta - \lambda_\theta\frac{\tan\phi}{r}\right)^2 + \frac{\lambda_\psi^2}{2} + \frac{\lambda_\phi^2}{2},$$

which gives the optimal control system

$$\begin{aligned} \dot{x}_1 &= \frac{\cos\theta}{r}[r(\lambda_1\cos\theta + \lambda_2\sin\theta) - \lambda_\theta\tan\phi], \\ \dot{x}_2 &= \frac{\sin\theta}{r}[r(\lambda_1\cos\theta + \lambda_2\sin\theta) - \lambda_\theta\tan\phi], \\ \dot{\theta} &= -\frac{\tan\theta}{r^2}[r(\lambda_1\cos\theta + \lambda_2\sin\theta) - \lambda_\theta\tan\phi], \\ \dot{\psi} &= \lambda_\psi, \quad \dot{\phi} = \lambda_\phi, \quad \dot{\lambda}_1 = 0, \quad \dot{\lambda}_2 = 0, \quad \dot{\lambda}_\psi = 0, \\ \dot{\lambda}_\theta &= \frac{\lambda_1\sin\theta - \lambda_2\cos\theta}{r}[r(\lambda_1\cos\theta + \lambda_2\sin\theta) - \lambda_\theta\tan\phi], \\ \dot{\lambda}_\phi &= \frac{\lambda_\theta\sec^2\phi}{r^2}(r\lambda_1\cos\theta + r\lambda_2\sin\theta - \lambda_\theta\tan\phi). \end{aligned} \tag{28}$$

Now, let us perform the reduction. The connection $\mathcal{A} : TM \to \mathfrak{g}$ is given by (see [20, Appendix A] for details)

$$\begin{aligned} \mathcal{A}_{(\theta,\phi)} &= (dx_1 + r\cos\theta\cot\phi\,d\theta) \otimes \mathbf{e}_1 \\ &\quad + (dx_2 + r\sin\theta\cot\phi\,d\theta) \otimes \mathbf{e}_2 + d\psi \otimes \mathbf{e}_\psi, \end{aligned}$$

where $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_\psi\}$ is a basis for the Lie algebra $\mathfrak{g} = T_{(0,0)}\mathbb{R}^2 \times \mathfrak{so}(2) \cong \mathbb{R}^3$. We then identify the vertical space $\mathcal{U}$ as follows:

$$\mathcal{U} = \operatorname{span}\left\{\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial\phi}\right\}.$$

The reduced curvature form $\tilde{\mathcal{B}}$ at $\bar{x} = (\theta, \phi) \in M/G$ is then

$$\tilde{\mathcal{B}}_{\bar{x}} = r\cos\theta\csc^2\phi\,d\theta\wedge d\phi\otimes\mathbf{e}_1 + r\sin\theta\csc^2\phi\,d\theta\wedge d\phi\otimes\mathbf{e}_2.$$

Introducing $\bar{\lambda} \in T^*(M/G)$, $\tilde{\xi} \in \tilde{\mathfrak{g}}$, and $\tilde{\mu} \in \tilde{\mathfrak{g}}^*$ defined by

$$\begin{aligned} \bar{\lambda}_{\bar{x}} &= (\bar{\lambda}_\theta, \bar{\lambda}_\phi) := \mathrm{hl}_x^*(\lambda_x) \\ &= (\lambda_\theta - \lambda_1\,r\cot\phi\cos\theta - \lambda_2\,r\cot\phi\sin\theta,\ \lambda_\phi), \\ \tilde{\xi}_{\bar{x}} &= \left(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{\xi}_\psi\right) := [x, \mathcal{A}_x(\dot{x})]_G \\ &= \left(\dot{x}_1 - (r\cot\phi\,\cos\theta)\,\dot{\theta},\ \dot{x}_2 - (r\cot\phi\,\sin\theta)\,\dot{\theta},\ \dot{\psi}\right), \\ \tilde{\mu}_{\bar{x}} &= (\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\mu}_\psi) := [x, \mathbf{J}(\lambda_x)]_G = (\lambda_1, \lambda_2, \lambda_\psi), \end{aligned}$$

the reduced optimal Hamiltonian (25) is written as

$$\bar{H}(\bar{x}, \bar{\lambda}, \tilde{\mu}) = \frac{1}{2}\left(\frac{\bar{\lambda}_\theta^2 \tan^2\phi}{r^2} + \bar{\lambda}_\phi^2 + \tilde{\mu}_\psi^2\right).$$

As a result, the reduced optimal control system (27) gives

$$\dot{\theta} = \frac{\tan^2\phi}{r^2}\,\bar{\lambda}_\theta, \qquad \dot{\phi} = \bar{\lambda}_\phi,$$
$$\tilde{\xi}_1 = 0, \qquad \tilde{\xi}_2 = 0, \qquad \tilde{\xi}_\psi = \tilde{\mu}_\psi,$$
$$\dot{\bar{\lambda}}_\theta = \bar{\lambda}_\phi\, r \csc^2\phi\,(\tilde{\mu}_1\cos\theta + \tilde{\mu}_2\sin\theta),$$
$$\dot{\bar{\lambda}}_\phi = -\bar{\lambda}_\theta\,\sec^2\phi\left(\bar{\lambda}_\theta\tan\phi + \tilde{\mu}_1\, r\cos\theta + \tilde{\mu}_2\, r\sin\theta\right),$$
$$\dot{\tilde{\mu}}_1 = 0, \qquad \dot{\tilde{\mu}}_2 = 0, \qquad \dot{\tilde{\mu}}_\psi = 0.$$

This system is significantly simpler than the original optimal control system (28): Notice that we now have a decoupled subsystem for the variables $(\theta, \phi, \bar{\lambda}_\theta, \bar{\lambda}_\phi)$; so we first solve the subsystem and then can obtain the dynamics for (x, y, ψ) by quadrature.

D. Lie–Poisson Reduction of Pontryagin Maximum Principle

Consider the special case where $M = G$ and assume that the cost function $C : E \to \mathbb{R}$ is G-invariant, i.e., $C \circ \Psi_h = C$ for any $h \in G$; then, for any $g \in G$, we have $C(g, u) = C(e, u) = \bar{C}(u)$, where $\bar{C}$ is defined on $E/G \cong \mathbb{R}^d$.

In this case, the quotient M/G becomes a point and thus the bundle $T^*(M/G) \oplus \tilde{\mathfrak{g}}^*$ becomes just $\mathfrak{g}^*$; as a result, $\tilde{\xi}$ is equal to ξ. Notice also that, since the momentum map is given by $\mathbf{J}(\lambda_g) = T_e^* R_g(\lambda_g)$, we have $\tilde{\mu} \cong \mathrm{Ad}_g^* \mathbf{J}(\lambda_g) = T_e^* L_g(\lambda_g) \in \mathfrak{g}^*$. Therefore, the Hamilton–Poincaré equations (24) reduce to the Lie–Poisson equation [18], and so (26) becomes

$$\xi = \bar{f}_{\tilde{\mathfrak{g}}}^\star(\tilde{\mu}), \qquad \frac{d\tilde{\mu}}{dt} = \mathrm{ad}_\xi^*\,\tilde{\mu}.$$

This system with an affine control (14) and the cost function of the form

$$C(g, u) = \bar{C}(u) = \frac{1}{2}\sum_{i=1}^{d} I_i\,(u^i)^2$$

is the case considered by Krishnaprasad [19] (see also Koon and Marsden [32, Section 5.3]).

Acknowledgments

We would like to thank Anthony Bloch, Matthias Kawski, Melvin Leok, and Joris Vankerschaver for helpful comments and discussions. This work is partially supported by the National Science Foundation under the grant DMS-1010687.

References

[1] H. Nijmeijer and A. van der Schaft, "Controlled invariance for nonlinear systems," *Automatic Control, IEEE Transactions on*, vol. 27, no. 4, pp. 904–914, 1982.
[2] J. Grizzle and S. Marcus, "The structure of nonlinear control systems possessing symmetries," *Automatic Control, IEEE Transactions on*, vol. 30, no. 3, pp. 248–258, 1985.
[3] ——, "Optimal control of systems possessing symmetries," *Automatic Control, IEEE Transactions on*, vol. 29, no. 11, pp. 1037–1040, Nov 1984.
[4] A. J. van der Schaft, "Symmetries in optimal control," *SIAM Journal on Control and Optimization*, vol. 25, no. 2, pp. 245–259, 1987.
[5] M. de León, J. Cortés, D. Martín de Diego, and S. Martínez, "General symmetries in optimal control," *Reports on Mathematical Physics*, vol. 53, no. 1, pp. 55–78, 2004.
[6] A. Echeverrìa-Enrìquez, J. Marìn-Solano, M. C. Muñoz Lecanda, and N. Román-Roy, "Geometric reduction in optimal control theory with symmetries," *Reports on Mathematical Physics*, vol. 52, no. 1, pp. 89–113, 2003.
[7] G. Blankenstein and A. van der Schaft, "Optimal control and implicit Hamiltonian systems," in *Nonlinear control in the Year 2000*, A. Isidori, F. Lamnabhi-Lagarrigue, W. Respondek, G. Blankenstein, and A. van der Schaft, Eds. Springer, 2000, vol. 258, pp. 185–205.
[8] A. Ibort, T. R. de la Peña, and R. Salmoni, "Dirac structures and reduction of optiaml control problems with symmetries," *Preprint*, 2010.
[9] R. Montgomery, "Isoholonomic problems and some applications," *Communications in Mathematical Physics*, vol. 128, no. 3, pp. 565–592, 1990.
[10] ——, "Optimal control of deformable bodies and its relation to gauge theory," in *The Geometry of Hamiltonian Systems*, T. Ratiu, Ed. Springer, 1991, pp. 403–438.
[11] ——, "Nonholonomic control and gauge theory," in *Nonholonomic Motion Planning*, Z. Li and J. F. Canny, Eds. Kluwer, 1993.
[12] ——, "Gauge theory of the falling cat," *Fields Institute Communications*, vol. 1, pp. 193–218, 1993.
[13] ——, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. American Mathematical Society, 2002.
[14] A. Shapere and F. Wilczek, "Self-propulsion at low Reynolds number," *Physical Review Letters*, vol. 58, no. 20, 1987.
[15] S. D. Kelly and R. M. Murray, "Geometric phases and robotic locomotion," *Journal of Robotic Systems*, vol. 12, no. 6, pp. 417–431, 1995.
[16] R. M. Murray, Z. Li, and S. S. Sastry, *A mathematical introduction to robotic manipulation*. CRC Press, 1994.
[17] Z. Li and J. F. Canny, Eds., *Nonholonomic Motion Planning*. Kluwer, 1993.
[18] H. Cendra, J. E. Marsden, S. Pekarsky, and T. S. Ratiu, "Variational principles for Lie–Poisson and Hamilton–Poincaré equations," *Moscow Mathematical Journal*, vol. 3, no. 3, pp. 833–867, 2003.
[19] P. S. Krishnaprasad, "Optimal control and Poisson reduction," University of Maryland, Tech. Rep. T.R. 93-87, 1993.
[20] T. Ohsawa, "Symmetry reduction of optimal control systems," *Preprint (*`arXiv:1105.5174`*), Submitted.*
[21] H. Cendra, J. E. Marsden, and T. S. Ratiu, *Lagrangian Reduction by Stages*, ser. Memoirs of the American Mathematical Society. American Mathematical Society, 2001, vol. 152, no. 722.
[22] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, and R. M. Murray, "Nonholonomic mechanical systems with symmetry," *Archive for Rational Mechanics and Analysis*, vol. 136, pp. 21–99, 1996.
[23] H. Cendra, J. E. Marsden, and T. S. Ratiu, "Geometric mechanics, Lagrangian reduction, and nonholonomic systems," in *Mathematics Unlimited*, B. Engquist and W. Schmid, Eds. Springer, 2001.
[24] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*. Springer, 2004.
[25] J. E. Marsden and A. Weinstein, "Reduction of symplectic manifolds with symmetry," *Reports on Mathematical Physics*, vol. 5, no. 1, pp. 121–130, 1974.
[26] J. E. Marsden, G. Misiolek, J. P. Ortega, M. Perlmutter, and T. S. Ratiu, *Hamiltonian Reduction by Stages*. Springer, 2007.
[27] R. Montgomery, J. E. Marsden, and T. S. Ratiu, "Gauged Lie–Poisson structures," *Contemp. Math.*, vol. 28, pp. 101–114, 1984.
[28] R. Montgomery, "The bundle picture in mechanics," Ph.D. dissertation, University of California, Berkeley, 1986.
[29] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*. Springer, 1999.
[30] J. Ostrowski, A. Lewis, R. Murray, and J. Burdick, "Nonholonomic mechanics and locomotion: The snakeboard example," *Robotics and Automation, Proceedings., 1994 IEEE International Conference on*, pp. 2391–2397 vol.3, 1994.
[31] W. S. Koon and J. E. Marsden, "The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems," *Reports on Mathematical Physics*, vol. 40, no. 1, pp. 21–62, 1997.
[32] ——, "Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction," *SIAM Journal on Control and Optimization*, vol. 35, no. 3, pp. 901–929, 1997.