Optimal Control for Holonomic and Nonholonomic Mechanical Systems with Symmetry and Lagrangian Reduction

Wang-Sang Koon*
Department of Mathematics
University of California
Berkeley, CA 94720-3840
koon@math.berkeley.edu

Jerrold E. Marsden†
Control and Dynamical Systems
California Institute of Technology 104-44
Pasadena, CA 91125
marsden@cds.caltech.edu

March, 1995; this version: March 8, 1996
CDS Technical Report 95-022, California Institute of Technology
URL http://avalon.caltech.edu/cds/reports/cds95-022.ps
To appear in *SIAM J. of Control and Optimization*

Abstract

In this paper we establish necessary conditions for optimal control using the ideas of Lagrangian reduction in the sense of reduction under a symmetry group. The techniques developed here are designed for Lagrangian mechanical control systems with symmetry. The benefit of such an approach is that it makes use of the special structure of the system, especially its symmetry structure and thus it leads rather directly to the desired conclusions for such systems.

Lagrangian reduction can do in one step what one can alternatively do by applying the Pontryagin Maximum Principle followed by an application of Poisson reduction. The idea of using Lagrangian reduction in the sense of symmetry reduction was also obtained by Bloch and Crouch [1995a,b] in a somewhat different context and the general idea is closely related to those in Montgomery [1990] and Vershik and Gershkovich [1994]. Here we develop this idea further and apply it to some known examples, such as optimal control on Lie groups and principal bundles (such as the ball and plate problem) and reorientation examples with zero angular momentum (such as the satellite with moveable masses). However, one of our main goals is to extend the method to the case of nonholonomic systems with a nontrivial momentum equation in the context of the work of Bloch, Krishnaprasad, Marsden and Murray [1995]. The snakeboard is used to illustrate the method.

*Currently visiting Control and Dynamical Systems, California Institute of Technology 104-44, Pasadena, CA 91125
†Research partially supported by NSF and DOE.

Contents

1 Introduction

Recently several papers have appeared exploring the symmetry reduction of optimal control problems on configuration spaces such as Lie groups and principal bundles. The mechanical systems which they have modeled vary widely: ranging from the falling cat, the rigid body with two oscillators, to the plate-ball system as well as the (airport) landing tower problem. Since the Pontryagin Maximum Principle is such an important and powerful tool in optimal control theory, it is frequently employed as a first step in finding necessary conditions for the optimal controls. Finally, different variants of Poisson reduction on the cotangent bundle T^*Q of the configuration space Q are used to obtain the reduced equations of motion for the optimal trajectories.

This paper develops a Lagrangian alternative to the method of Pontryagin Maximum Principle and Poisson reduction used in many of the above studies. More importantly, our method can handle the optimal control of nonholonomic mechanical system such as the snakeboard which has a nontrivial evolution equation for its nonholonomic momentum. Our key idea is to link the method of Lagrange multipliers with Lagrangian reduction. This procedure which will be referred to as "reduced Lagrangian optimization", is able to handle all the above cases including the snakeboard. We hope that it will complement other existing methods and may also have the advantage that it is easier to use in many situations and can solve many new problems. In the optimal control problems we deal with in this paper, one encounters degenerate Lagrangians; fortunately this does not cause problems with the technique of Lagrangian reduction. For more information on these degeneracies, see Bloch and Crouch [1995a,b].

Our objectives in this paper are limited to presenting reduced Lagrangian optimization in the context of both holonomic and nonholonomic systems that may have conservation laws or nontrival momentum equations. We use this approach as an alternative to the Pontryagin Maximum Principle and Poisson reduction. Although an assumption of controllability underlies most optimal control problems, we are concerned here with finding necessary conditions for optimality and so do not discuss controllability explicitly. We do not extensively develop the geometry of the situation in much detail and we restrict our attention to regular extremals throughout the paper without explicit mention. Of course all of these points are of interest in themselves.

In the course of working on this paper, we have found some related ideas in Montgomery [1990], Vershik and Gershkovich [1994] and Bloch and Crouch [1994, 1995a,b]. The paper Bloch, Krishnaprasad, Marsden and Murray [1995] provides a useful framework for the present work.

Outline of the Paper

In §2, we recall some basic facts about both holonomic and nonholonomic mechanical systems with symmetry. We set up a class of optimal control problems for holonomic mechanical systems on a (trivial) principal bundle as was done in Montgomery [1990] and Krishnaprasad, Yang and Dayawansa [1991]. We also set up the corresponding problems for nonholonomic systems. We will call these "Lagrangian optimal control problems".

In §3 we review some aspects of the theory of Lagrangian reduction and use it to solve the Lagrangian optimal control problem in the holonomic case, showing that an optimal trajectory is a solution of Wong's equations (at least for regular extremals). This provides an alternative derivation to the approach (based on methods of subriemannian geometry) in Montgomery [1990] and the approach (based on the Pontryagin maximum principle and Poisson reduction) in Krishnaprasad, Yang and Dayawansa [1991].

In §4 we generalize these results to the case of nonholonomic systems. Notice in particular that our techniques allow for nonzero values of the momentum map, which is interesting even for the holonomic case. In §5 we consider a number of examples, such as the ball on a plate (as in Bloch, Krishnaprasad, Marsden and Murray [1995]), and the snakeboard. We also consider optimal control problems for systems on Lie groups such as the landing tower problem (see Krishnaprasad [1993] and Walsh, Montgomery and Sastry [1994]) and the plate ball problem considered in Jurdjevic [1993].

In the conclusions, we give a few remarks on future research directions.

2 Lagrangian Mechanical Systems with Symmetry

In this section we shall review, for the convenience of the reader, some notation and results for mechanical systems with symmetry. We will begin with the case of holonomic systems and then study the nonholonomic case.

2.1 Holonomic systems with symmetry

Notation

A simple Lagrangian system with symmetry consists of a configuration manifold Q, a metric tensor (the mass matrix) $\langle\!\langle\, , \rangle\!\rangle$, a symmetry group G (a Lie group) and a Lagrangian L. Assume that G acts on Q by isometries and that the Lagrangian L is of the form kinetic minus potential energy, *i.e.*,

$$L(q, v) = \frac{1}{2}\|v\|_q^2 - V(q)$$

where $\|\cdot\|_q$ denotes the norm on T_qQ and V is a G-invariant potential. For more information, see for example, Marsden [1992] and Marsden and Ratiu [1994]. Examples of such systems are the falling cat (Montgomery [1990, 1991]) and the rigid body with 2 oscillators (Krishnaprasad, Yang and Dayawansa [1991]).

The associated equivariant momentum map $J : TQ \to \mathfrak{g}^*$ for a simple Lagrangian system with symmetry is given by

$$\langle J(q, v), \xi \rangle = \langle\!\langle v, \xi_Q(q) \rangle\!\rangle = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q)^i, \tag{2.1.1}$$

where $\mathfrak{g}^*$ is the dual of the Lie algebra $\mathfrak{g}$ of G, ξ_Q is the infinitesimal generator of $\xi \in \mathfrak{g}$ on Q, and $\langle \, , \rangle$ is the pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$ (other natural pairings between spaces and their duals are also denoted $\langle \, , \rangle$ in this paper).

Assume that G acts freely and properly on Q, so we can regard $Q \to Q/G$ as a principal G-bundle (Q, B, π, G) where $B = Q/G$ is called the base (or shape) space and $\pi : Q \to B$ is the bundle projection. On this bundle, we construct the mechanical connection $\mathcal{A}$ as follows: for each $q \in Q$, let the locked inertia tensor be the map $\mathbb{I}(q) : \mathfrak{g} \to \mathfrak{g}^*$ defined by

$$\langle \mathbb{I}(q)\eta, \xi \rangle = \langle\!\langle \eta_Q(q), \xi_Q(q) \rangle\!\rangle.$$

The terminology comes from the fact that for a coupled rigid body, particle, or elastic system, $\mathbb{I}(q)$ is the classical moment of inertia tensor of the instantaneous rigid system. The mechanical connection is the map $\mathcal{A} : TQ \to \mathfrak{g}$ that assigns to each (q, v) the "angular velocity of the locked system"

$$\mathcal{A}(q, v) = \mathbb{I}(q)^{-1} J(q, v). \tag{2.1.2}$$

When there is danger of confusion, we will write the mechanical connection as $\mathcal{A}^{\mathrm{mec}}$ (additional connections will be introduced later in the paper). The map $\mathcal{A}$ is a connection on the principal G-bundle $Q \to Q/G$; that is, $\mathcal{A}$ is G-equivariant and satisfies $\mathcal{A}(\xi_Q(q)) = \xi$, both of which are readily verified. The horizontal space of the connection $\mathcal{A}$ is given by

$$\mathrm{hor}_q = \{(q, v) \mid J(q, v) = 0\},$$

i.e., the space orthogonal to the G-orbits. The vertical space consists of vectors that are tangent to the group orbits, *i.e.,*

$$\mathrm{ver}_q = T_q(\mathrm{Orb}(q)) = \{\xi_Q(q) \mid \xi \in \mathfrak{g}\}.$$

For later use, we would like to say a few words about a general principal connection and its expression in a local trivialization. As stated above, a principal connection is a $\mathfrak{g}$-valued 1-form $\mathcal{A} : TQ \to \mathfrak{g}$ such that $\mathcal{A}(g \cdot v) = \mathrm{Ad}_g \mathcal{A}(v)$, and $\mathcal{A}(\xi_Q(q)) = \xi$ for each $\xi \in \mathfrak{g}$. For example, if $Q = G$, there is a cannonical connection given by the right invariant 1-form which equals the identity at $g = e$. That is, for $v \in T_g G$, we let $\mathcal{A}_G : TG \to \mathfrak{g}$, $\mathcal{A}_G(v) = TR_{g^{-1}} \cdot v$. In a local trivialization where we can locally write $Q = B \times G$ and the action of G is given by left translation on the second factor, a connection $\mathcal{A}$ as a 1-form has the form

$$\mathcal{A}(r, g) = \mathcal{A}_{\mathrm{loc}}(r, g) dr + \mathcal{A}_G$$

and

$$\mathcal{A}(r, g)(\dot{r}, \dot{g}) = \mathcal{A}_{\mathrm{loc}}(r, g)\dot{r} + \dot{g}g^{-1} = \mathrm{Ad}_g(\mathcal{A}_{\mathrm{loc}}(r, e)\dot{r} + g^{-1}\dot{g}),$$

where $(\dot{r}, \dot{g})$ is the tangent vector at each point $q = (r, g)$. With abuse of notation, we denote $\mathcal{A}_{\mathrm{loc}}(r, e) = \mathcal{A}_{\mathrm{loc}}(r)$. Hence, for a principal connection, we can write

$$\mathcal{A}(r, g)(\dot{r}, \dot{g}) = \mathrm{Ad}_g(g^{-1}\dot{g} + \mathcal{A}_{\mathrm{loc}}(r)\dot{r}). \tag{2.1.3}$$

Holonomic Optimal Control Problems

Now we are ready to formulate an optimal control problem for a holonomic system on a trivial bundle $(B \times G, B, \pi, G)$. As in Montgomery [1990, 1991] and Krishnaprasad, Yang and Dayawansa [1991], let us assume that the control is internal to the system, which leaves invariant the conserved momentum map J, and that there is no drift, *i.e.,* $\mu = J(q, v) = 0$. Assume further that the velocity $\dot{r}$ of the path in the base space B can be directly controlled; then an associated control problem can be set up as

$$\left.\begin{aligned} \dot{r} &= u \\ g^{-1}\dot{g} &= -\mathcal{A}_{\mathrm{loc}}(r)u, \end{aligned}\right\} \tag{2.1.4}$$

because, from the results above, the constraint that $\mu = 0$ is nothing but $(\dot{r}, \dot{g}) \in \mathrm{hor}_{(r,g)}$ which is equivalent to $g^{-1}\dot{g} + \mathcal{A}_{\mathrm{loc}}(r)\dot{r} = 0$. Here $u(\cdot)$ is a vector-valued function.

Let C be a cost function which usually is a positive definite quadratic function in u and hence C can be written as the square of a metric on B. Then we can formulate an optimal control problem on $Q = B \times G$ as follows:

> **Optimal Control Problem for Holonomic Systems** Given two points q_0, q_1 in Q, find the optimal controls $u(\cdot)$ which steer from q_0 to q_1 and minimize $\int_0^1 C(u)dt$ subject to the constraints $\dot{r} = u$, $g^{-1}\dot{g} = -\mathcal{A}_{\mathrm{loc}}(r)u$.

Clearly the above optimal control problem is equivalent to the following constrained variational problem:

> **Constrained Variational Problem for Holonomic Systems** Among all curves $q(t)$ such that $\dot{q}(t) \in \mathrm{hor}_{q(t)}, q(0) = q_0, q(1) = q_1$, find the optimal curves $q(t)$ such that $\int_0^1 C(\dot{r})dt$ is minimized, where $r = \pi(q)$.

For example in Krishnaprasad, Yang and Dayawansa [1991], they considered a rigid body with 2 (driven) oscillators, which was used to model the drift observed in the Hubble Space Telescope due to thermo-elasticially driven shape changes of the solar panels arising from the day-night thermal cycling during orbit. The bundle used was $(\mathbb{R}^2 \times SO(3), \mathbb{R}^2, \pi, SO(3))$ and the corresponding optimal control problem was

> **Optimal Control for a Rigid Body with Two Oscillators** Find the control $u(\cdot) = (u^1(\cdot), u^2(\cdot))$ that minimizes $\int_0^1 ((u^1)^2 + (u^2)^2)dt$, subject to $\dot{r} = u$, $\dot{g} = -g\mathcal{A}_{\mathrm{loc}}(r)u$, for $r^1(0) = r^1(1) = r^2(0) = r^2(1) = 0, g(0) = g_0$, and $g(1) = g_1 \in SO(3)$.

For more details on the derivation of this model, see Krishnaprasad, Yang and Dayawansa [1991]. Below we will take this optimal control problem as given and focus on finding the necessary conditions for its optimal trajectories. See Montgomery [1990, 1991] for additional examples.

2.2 Simple Nonholonomic Mechanical Systems with Symmetry

Next, we recall some basic ideas and results from Bloch, Krishnaprasad, Marsden and Murray [1995] which will help to set the overall context for the optimal control of a simple nonholonomic system. Assume that we have data as before, namely a configuration manifold Q, a Lagrangian of the form kinetic minus potential, and a symmetry group G that leaves the Lagrangian invariant. However, now we also assume we have a distribution $\mathcal{D}$ that describes the kinematic nonholonomic constraints. Thus, $\mathcal{D}$ is a collection of linear subspaces denoted $\mathcal{D}_q \subset T_qQ$, one for each $q \in Q$. We assume that G acts on Q by isometries and leaves the distribution invariant, *i.e.,* the tangent of the group action maps $\mathcal{D}_q$ to $\mathcal{D}_{gq}$. Moreover, we assume that we are in the principal case where the constraints and the orbit directions span the entire tangent space to the configuration space: $\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_qQ$ for each $q \in Q$.

As discussed in Bloch, Krishnaprasad, Marsden and Murray [1995], the dynamics of a nonholonomically constrained mechanical system is governed by the Lagrange d'Alembert principle. This principle states that (at least in the case of homogeneous linear constraints) the equations of motion of a curve $q(t)$ in configuration space are obtained by setting to zero the variations in the integral of the Lagrangian subject to variations lying in the constraint distribution vanish and that the velocity of the curve $q(t)$ itself satisfies the constraints.

The Momentum Equation

In the case of a simple holonomic mechanical system, setting up an optimal control problem uses the momentum map J, the mechanical connection $\mathcal{A}$ as well as the reconstruction of path on Q given a path in Q/G. For the case of a simple nonholonomic mechanical system, we shall need similar notions and they are recalled in the following discussion.

Let the intersection of the tangent to the group orbit and the distribution at a point $q \in Q$ be denoted

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)).$$

Define, for each $q \in Q$, the vector subspace $\mathfrak{g}^q$ to be the set of Lie algebra elements in $\mathfrak{g}$ whose infinitesimal generators evaluated at q lie in $\mathcal{S}_q$:

$$\mathfrak{g}^q = \{\xi \in \mathfrak{g} : \xi_Q(q) \in S_q\}.$$

Then $\mathfrak{g}^{\mathcal{D}}$ is the corresponding bundle over Q whose fiber at the point q is given by $\mathfrak{g}^q$. The nonholonomic momentum map J^{nh} is the bundle map taking TQ to the bundle $(\mathfrak{g}^{\mathcal{D}})^*$ (whose fiber over the point q is the dual of the vector space $\mathfrak{g}^q$) that is defined by

$$\langle J^{\mathrm{nh}}(v_q), \xi\rangle = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q)^i, \tag{2.2.1}$$

where $\xi \in \mathfrak{g}^q$.

As the examples like the snakeboard show, in general the tangent space to the group orbit through q intersects the constraint distribution at q nontrivially.

Notice that the nonholonomic momentum map may be viewed as giving just some of the components of the ordinary momentum map, namely along those symmetry directions that are consistent with the constraints.

It is proven in Bloch, Krishnaprasad, Marsden and Murray [1995] that if the Lagrangian L is invariant under the group action and that if ξ^q is a section of the bundle $\mathfrak{g}^{\mathcal{D}}$, then any solution $q(t)$ of the Lagrange d'Alembert equations for a nonholonomic system must satisfy, in addition to the given kinematic constraints, the momentum equation:

$$\frac{d}{dt}\left(J^{\mathrm{nh}}(\xi^{q(t)})\right) = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}(\xi^{q(t)})\right]_Q^i. \tag{2.2.2}$$

When the momentum map is paired with a section in this way, we will just refer to it as the momentum. Examples show that the nonholonomic momentum map may or may not be conserved.

The Momentum Equation in an Orthogonal Body Frame

Let a local trivialization (r, g) be chosen on the principal bundle $\pi : Q \to Q/G$. Let $\eta \in \mathfrak{g}^q$ and $\xi = g^{-1}\dot{g}$. Since L is G-invariant, we can define a new function l by writing $L(r, g, \dot{r}, \dot{g}) = l(r, \dot{r}, \xi)$. Define $J^{\mathrm{nh}}_{\mathrm{loc}} : TQ/G \to (\mathfrak{g}^{\mathcal{D}})^*$ by

$$\left\langle J^{\mathrm{nh}}_{\mathrm{loc}}(r, \dot{r}, \xi), \eta\right\rangle = \left\langle \frac{\partial l}{\partial \xi}, \eta\right\rangle.$$

As with connections, J^{nh} and its version in a local trivialization are related by the Ad map; *i.e.,* $J^{\mathrm{nh}}(r, g, \dot{r}, \dot{g}) = \mathrm{Ad}^*_{g^{-1}} J^{\mathrm{nh}}_{\mathrm{loc}}(r, \dot{r}, \xi)$.

Choose a q-dependent basis $e_a(q)$ for the Lie algebra such that the first m elements span the subspace $\mathfrak{g}^q$. In a local trivialization, one chooses, for each r, such a basis at the identity element, say

$$e_1(r), e_2(r), \ldots, e_m(r), e_{m+1}(r), \ldots, e_k(r).$$

We may require the basis to be such that the infinitesimal generators of the first m basis elements are orthogonal in the kinetic energy metric and whose last $k-m$ generators are also orthogonal. Note, however, that elements from the first batch and elements from the second batch might not be orthogonal; indeed, the subspaces $T_q\mathrm{Orb}$ and $\mathcal{D}_q$ need not be orthogonal subspaces, so this is not possible in general. Define the orthogonal body frame by

$$e_a(r,g) = \mathrm{Ad}_g \cdot e_a(r);$$

thus, by G invariance, the first m elements span the subspace $\mathfrak{g}^q$. In this basis, we have

$$\left\langle J^{\mathrm{nh}}(r,g,\dot r,\dot g), e_b(r,g)\right\rangle = \left\langle \frac{\partial l}{\partial \xi}, e_b(r)\right\rangle := p_b, \tag{2.2.3}$$

which defines p_b, a function of r, $\dot r$ and ξ. It is proven in Bloch, Krishnaprasad, Marsden and Murray [1995] that in such an orthogonal body frame, the momentum equation can be written in the following form:

$$\dot p = \dot r^T H(r)\dot r + \dot r^T K(r)p + p^T D(r)p. \tag{2.2.4}$$

Note that in this body representation, the functions p_b are *invariant* rather than equivariant, as is usually the case with the momentum map, and the momentum equation is independent of, that is, decouples from, the group variables g.

The Nonholonomic Connection

Recall that in the case of holonomic mechanical systems, the mechanical connection $\mathcal{A}$ is defined by $\mathcal{A}(v_q) = \mathbb{I}(q)^{-1}J(v_q)$ or equivalently by the fact that its horizontal space at q is orthogonal to the group orbit at q. For the case of a simple nonholonomic mechanical system where the Lagrangian is of the form kinetic minus potential energy and G acts on Q by isometries and leaves $\mathcal{D}$ invariant, the result turns out to be quite similar.

As Bloch, Krishnaprasad, Marsden and Murray [1995] points out, in the principal case where the constraints and the orbit directions span the entire tangent space to the configuration space (that is, $\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_qQ$), the nonholonomic connection $\mathcal{A}^{\mathrm{nh}}$ is a principal connection on the bundle $Q \to Q/G$ whose horizontal space at the point $q \in Q$ is given by the orthogonal complement to the space $\mathcal{S}_q$ within the space $\mathcal{D}_q$. Moreover, Bloch, Krishnaprasad, Marsden and Murray [1995] develop formulas for $\mathcal{A}^{\mathrm{nh}}$ similar to those for the mechanical connection, namely

$$\mathcal{A}^{\mathrm{nh}}(v_q) = \mathbb{I}^{\mathrm{nh}}(q)^{-1}J^{\mathrm{nh}}(v_q) \tag{2.2.5}$$

where $\mathbb{I}^{\mathrm{nh}} : \mathfrak{g}^{\mathcal{D}} \to (\mathfrak{g}^{\mathcal{D}})^*$ is the locked inertia tensor defined in a way similar to that given above for holonomic systems. In an orthogonal body frame, (2.2.5) can be written as

$$\mathrm{Ad}_g(g^{-1}\dot g + \mathcal{A}^{\mathrm{nh}}_{\mathrm{loc}}(r)\dot r) = \mathrm{Ad}_g(\mathbb{I}^{\mathrm{nh}}_{\mathrm{loc}}(r)^{-1}p), \tag{2.2.6}$$

where $\mathcal{A}^{\mathrm{nh}}_{\mathrm{loc}}$ and $\mathbb{I}^{\mathrm{nh}}_{\mathrm{loc}}$ are the representations of $\mathcal{A}^{\mathrm{nh}}$ and $\mathbb{I}^{\mathrm{nh}}$ in a local trivialization. For simplicity in what follows, we shall omit the subscript "loc".

Control Systems in Momentum Equation Form

With the help of the momentum equations and the nonholonomic mechanical connection, Bloch, Krishnaprasad, Marsden and Murray [1995] provides a framework for studying the general form of nonholonomic mechanical control systems with symmetry that may have a nontrivial evolution of their nonholonomic momentum. The dynamics of such a system can be described by a system of equations of the form of a reconstruction equation for a group element g, an equation for the nonholonomic momentum p (no longer conserved in the general case), and the equations of motion for the reduced variables r which describe the "shape" of the system. In terms of these variables, the equations of motion have the functional form

$$\left.\begin{aligned} g^{-1}\dot{g} &= -\mathcal{A}^{\mathrm{nh}}(r)\dot{r} + \Gamma(r)p \\ \dot{p} &= \dot{r}^T H(r)\dot{r} + \dot{r}^T K(r)p + p^T D(r)p \\ M(r)\ddot{r} &= \delta(r, \dot{r}, p) + \tau, \end{aligned}\right\} \tag{2.2.7}$$

where (where $\Gamma(r) = \mathbb{I}^{\mathrm{nh}}(r)$).

The first equation describes the motion in the group variables as the flow of a left invariant vector field determined by the internal shape r, its velocity $\dot{r}$, as well as the generalized momentum p. The term $g^{-1}\dot{g} + \mathcal{A}^{\mathrm{nh}}(r)\dot{r} = \Gamma(r)^{-1}p$ is interpreted as the local representation of the body angular velocity. This is nothing more than the vertical part of the bundle velocity. The momentum equation describes the evolution of p and as was mentioned earlier, is bilinear in $(\dot{r}, p)$. Finally, the bottom (second-order) equation for $\ddot{r}$ describes the motion of the variables which describe the configuration up to a symmetry (*i.e.,* the shape). The variable τ represents the external forces applied to the system, which we assume here only affect the shape variables, *i.e.,* the external forces are G-invariant. Note that the evolution of the momentum p and the shape r decouple from the group variables.

The Optimal Control Problem for Nonholonomic Systems on a Trivial Bundle

Assume that we have a simple nonholonomic mechanical system with symmetry; thus, assume we have data $(Q, \mathcal{D}, \langle\!\langle\, , \rangle\!\rangle, G, L)$ where the Lagrangian L is G-invariant and of the form kinetic minus potential energy, the distribution $\mathcal{D}$ is G-invariant, and we are in the principal case where the constraints and the orbit directions span the tangent space to the configuration space. Let us also assume in this section that the principal bundle $\pi : Q \to Q/G$ is trivial; all the examples we consider (including the snakeboard) have a trivial principal bundle structure. We consider this simplification as a first step to the general case because in a local trivialization any principal bundle is a trivial bundle $(B \times G, B, \pi, G)$. Furthermore, we will assume that

1. Any control forces applied to the system affect only the shape variables which leaves the generalized momenta and the momentum equation unchanged. Indeed, such forces would be invariant under the action of the Lie group G and so would be annihilated by the variations taken to derive the momentum equation.

2. We have full control of the shape variables; that is, the curve $r(t)$ in the shape space B can be specified arbitrarily using a suitable control force τ.

Given a cost function C which is a positive definite quadratic function of $\dot{r}(t)$ (so can be written as the square of a metric on the shape space B), we can formulate an optimal control problem on $Q = B \times G$ as follows:

> **Optimal Control Problem for Nonholonomic Systems** Given two points $q_0, q_1 \in Q$, find the curves $r(t) \in B$ which steer the system from q_0 to q_1, and which minimize the total cost $\int_0^1 C(\dot{r})dt$, where $r = \pi(q)$, subject to the constraints $g^{-1}\dot{g} = -\mathcal{A}^{\mathrm{nh}}(r)\dot{r} + \Gamma(r)p$, and to the momentum equation $\dot{p} = \dot{r}^T H(r)\dot{r} + \dot{r}^T K(r)p + p^T D(r)p$.

This optimal control problem is clearly equivalent to the following constrained variational problem:

> **Constrained Variational Problem for Nonholonomic Systems** Among all curves $q(t)$ with $q(0) = q_0, q(1) = q_1$ and satisfying $g^{-1}\dot{g} = -\mathcal{A}^{\mathrm{nh}}(r)\dot{r} + \Gamma(r)p$, where $\dot{p} = \dot{r}^T H(r)\dot{r} + \dot{r}^T K(r)p + p^T D(r)p$, find the curves $q(t)$ such that $\int_0^1 C(\dot{r})dt$ is minimized, where $r = \pi(q)$.

Now we are ready to use the method of Lagrange multipliers and Lagrangian reduction to find necessary conditions for optimal trajectories.

3 Optimal Control and Lagrangian Reduction for Holonomic Systems

In this section we consider reduced Lagrangian optimization in the context of holonomic systems.

3.1 A Review of Lagrangian Reduction

We first recall some facts about Lagrangian reduction theory for systems with holonomic constraints (see Marsden and Scheurle [1993a,b].)

Rigid Body Reduction

Let $R \in SO(3)$ denote the time dependent rotation that gives the current configuration of a rigid body. The body angular velocity Ω is defined in terms of R by

$$R^{-1}\dot{R} = \hat{\Omega},$$

where $\hat{\Omega}$ is the three by three skew matrix defined by $\hat{\Omega}v := \Omega \times v$. Denoting by I the (time independent) moment of inertia tensor, the Lagrangian thought of as a function of R and $\dot{R}$ is given by $L(R, \dot{R}) = \frac{1}{2}\langle I\Omega, \Omega\rangle$ and when we think of it as a function of Ω alone, we write $l(\Omega) = \frac{1}{2}\langle I\Omega, \Omega\rangle$.

The following statements are equivalent:

1. $(R, \dot{R})$ satisfies the Euler-Lagrange equations on $SO(3)$ for L,
2. Hamilton's principle on $SO(3)$ holds:
$$\delta \int L\, dt = 0,$$
3. Ω satisfies the Euler equations
$$I\dot{\Omega} = I\Omega \times \Omega,$$
4. the reduced variational principle holds on $\mathbb{R}^3$:
$$\delta \int l\, dt = 0,$$
where variations in Ω are restricted to be of the form $\delta\Omega = \dot{\eta} + \eta \times \Omega$, with η an arbitrary curve in $\mathbb{R}^3$ satisfying $\eta = 0$ at the temporal endpoints.

An important point is that when one reduces the standard variational principle from $SO(3)$ to its Lie algebra $\mathfrak{so}(3)$, one ends up with a variational principle in which the *variations are constrained*; that is, one has a principle of Lagrange d'Alembert type. In this case, the term η represents the infinitesimal displacement of particles in the rigid body. Note that the same phenomenon of constrained variations occurs in the case of nonholonomic systems.

The Euler-Poincaré Equations

Let $\mathfrak{g}$ be a Lie algebra and let $l : \mathfrak{g} \to \mathbb{R}$ be a given Lagrangian. Then the Euler-Poincaré equations are:

$$\frac{d}{dt}\frac{\partial l}{\partial \xi} = \mathrm{ad}^*_\xi \frac{\partial l}{\partial \xi}$$

or, in coordinates,

$$\frac{d}{dt}\frac{\partial l}{\partial \xi^a} = C^b_{da}\xi^d \frac{\partial l}{\partial \xi^b},$$

where the structure constants are defined by $[\xi, \eta]^a = C^a_{de}\xi^d\eta^e$. If G is a Lie group with Lie algebra $\mathfrak{g}$, we let $L : TG \to \mathbb{R}$ be the left invariant extension of l and let $\xi = g^{-1}\dot{g}$. In the case of the rigid body, ξ is $\hat{\Omega}$, where Ω is the body angular velocity.

The basic fact regarding the Lagrangian reduction leading to these equations is:

Theorem 3.1 Euler-Poincaré reduction. *A curve $(g(t), \dot{g}(t)) \in TG$ satisfies the Euler-Lagrange equations for L if and only if ξ satisfies the Euler-Poincaré equations for l.*

In this situation, the reduction is implemented by the map $(g, \dot{g}) \in TG \mapsto g^{-1}\dot{g} =: \xi \in \mathfrak{g}$.

One proof of this theorem is of special interest, as it shows how to drop variational principles to the quotient (see Marsden and Scheurle [1993b] and Bloch, Krishnaprasad, Marsden and Ratiu [1994] for more details). Namely, we transform

$$\delta \int L \, dt = 0$$

under the map $(g, \dot{g}) \mapsto g^{-1}\dot{g}$ to give the reduced variational principle for the Euler-Poincaré equations: ξ satisfies the Euler-Poincaré equations if and only if

$$\delta \int l \, dt = 0,$$

where the variations are all those of the form

$$\delta\xi = \dot{\eta} + [\xi, \eta]$$

and where η is an arbitrary curve in the Lie algebra satisfying $\eta = 0$ at the endpoints. Variations of this form are obtained by calculating what variations are induced by variations on the Lie group itself.

One obtains the Lie-Poisson equations on $\mathfrak{g}^*$ by the Legendre transformation:

$$\mu = \frac{\partial l}{\partial \xi}, \quad h(\mu) = \mu \cdot \xi - l(\xi).$$

Dropping the variational principle this way is the analogue of Lie-Poisson reduction in which one drops the Poisson bracket from T^*G to the Lie-Poisson bracket on $\mathfrak{g}^*$.

The Reduced Euler-Lagrange Equations

The Euler-Poincaré equations can be generalized to the situation in which G acts freely on a configuration space Q to obtain the *reduced Euler-Lagrange equations*. This process starts with a G-invariant Lagrangian $L : TQ \to \mathbb{R}$, which induces a reduced Lagrangian $l : TQ/G \to \mathbb{R}$. The Euler-Lagrange equations for L induce the reduced Euler-Lagrange equations on TQ/G. To compute them in coordinates, it is useful to introduce a principal connection on the bundle $Q \to Q/G$. Although any can be picked, a common choice is the mechanical connection.

Thus, assume that the bundle $Q \to Q/G$ has a given (principal) connection $\mathcal{A}$. Divide variations into horizontal and vertical parts — this breaks up the Euler-Lagrange equations on Q into 2 sets of equations that we now describe. Let r^α be coordinates on shape space Q/G and Ω^a be coordinates for vertical vectors in a local bundle chart. Drop L to TQ/G to obtain a reduced Lagrangian $l : TQ/G \to \mathbb{R}$ in which the group coordinates are eliminated. We can represent this reduced Lagrangian in a couple of ways. First, if we choose a local trivialization as we have described earlier, we obtain l as a function of the variables $(r^\alpha, \dot{r}^\alpha, \xi^a)$. However, it will also be convenient to change variables from ξ^a to the local version of the locked angular velocity, *i.e.,* the body angular velocity, namely $\Omega = \xi + \mathcal{A}_{\mathrm{loc}}\dot{r}$, or in coordinates,

$$\Omega^a = \xi^a + \mathcal{A}^a_\alpha(r)\dot{r}^\alpha.$$

We will write $l(r^\alpha, \dot{r}^\alpha, \Omega^a)$ for the local representation of l in these variables.

Theorem 3.2 Lagrangian Reduction Theorem. *A curve $(q^i, \dot{q}^i) \in TQ$, satisfies the Euler-Lagrange equations if and only if the induced curve in TQ/G with coordinates given in a local trivialization by $(r^\alpha, \dot{r}^\alpha, \Omega^a)$ satisfies the* ***reduced Euler-Lagrange equations****:*

$$\begin{aligned} \frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} &= \frac{\partial l}{\partial \Omega^a}\left(-\mathcal{B}^a_{\alpha\beta}\dot{r}^\beta + \mathcal{E}^a_{\alpha d}\Omega^d\right) && (3.1.1)\\ \frac{d}{dt}\frac{\partial l}{\partial \Omega^b} &= \frac{\partial l}{\partial \Omega^a}(-\mathcal{E}^a_{\alpha b}\dot{r}^\alpha + C^a_{db}\Omega^d) && (3.1.2) \end{aligned}$$

where

$$\mathcal{B}^b_{\alpha\beta} = \frac{\partial \mathcal{A}^b_\alpha}{\partial r^\beta} - \frac{\partial \mathcal{A}^b_\beta}{\partial r^\alpha} - C^b_{ac}\mathcal{A}^a_\beta \mathcal{A}^c_\alpha,$$

are the coordinates of the curvature $\mathcal{B}$ of $\mathcal{A}$, and $\mathcal{E}^a_{\alpha d} = C^a_{bd}\mathcal{A}^b_\alpha$.

The first of these equations is similar to the Lagrange d'Alembert equations for a nonholonomic system written in terms of the constrained Lagrangian and the second is similar to the momentum equation. It is useful to note that the first set of equations results from Hamilton's principle by restricting the variations to be horizontal relative to the given connection.

If one uses the variables, $(r^\alpha, \dot{r}^\alpha, p_a)$, where p is the body angular momentum, so that $p = \mathbb{I}_{\mathrm{loc}}(r)\Omega = \partial l/\partial \Omega$, then the equations become (using the same letter l for the reduced lagrangian, an admitted abuse of notation):

$$\begin{aligned} \frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} &= p_a\left(-\mathcal{B}^a_{\alpha\beta}\dot{r}^\beta + \mathcal{E}^a_{\alpha d}I^{de}p_e\right) - p_d\frac{\partial I^{de}}{\partial r^\alpha}p_e && (3.1.3)\\ \frac{d}{dt}p_b &= p_a(-\mathcal{E}^a_{\alpha b}\dot{r}^\alpha + C^a_{db}I^{de}p_e), && (3.1.4) \end{aligned}$$

where I^{de} denotes the inverse of the matrix I_{ab}.

Connections are also useful in control problems with feedback. For example, Bloch, Krishnaprasad, Marsden and Sánchez de Alvarez [1992] found a feedback control that stabilizes rigid body dynamics about its middle axis using an internal rotor. This feedback controlled system can be described in terms of connections (Marsden and Sánchez de Alvarez [1995]): a shift in velocity (change of connection) turns the free Euler-Poincaré equations into the feedback controlled Euler-Poincaré equations.

3.2 Reduced Lagrangian Optimization for Holonomic Systems

Let us assume for the moment that we are dealing with a holonomic system on a trivial bundle and that the momentum map vanishes. Since we would like to use the method of Lagrange multipliers to relax the constraints, we define a new Lagrangian by $\mathcal{L}$

$$\mathcal{L} = C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}_{\mathrm{loc}}(r)\dot{r}\rangle \tag{3.2.1}$$

for some $\lambda(t) \in \mathfrak{g}^*$, where $\xi = g^{-1}\dot{g} \in \mathfrak{g}$. Clearly $\mathcal{L}$ is G-invariant and induces a function l on $(TQ/G) \times \mathfrak{g}^*$ where

$$l = C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}_{\text{loc}}(r)\dot{r} \rangle \tag{3.2.2}$$

Theorem 3.3 Reduced Lagrangian Optimization for Holonomic Systems. *Assume that $q(t) = (r(t), g(t))$ is a (regular) optimal trajectory for the above optimal control problem, then there exists a $\lambda(t) \in \mathfrak{g}^*$ such that the reduced curve $(r(t), \dot{r}(t), \xi(t)) \in TQ/G$ with coordinates given by $(r^\alpha, \dot{r}^\alpha, \xi^a)$ satisfies the constraints $\xi = -\mathcal{A}_{\text{loc}}(r)\dot{r}$, as well as the reduced Euler Lagrange equations*

$$\frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} = 0 \tag{3.2.3}$$

$$\frac{d}{dt}\frac{\partial l}{\partial \xi^b} = \frac{\partial l}{\partial \xi^a} C^a_{db} \xi^d \tag{3.2.4}$$

where $l = C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}_{\text{loc}}(r)\dot{r} \rangle$.

Proof If $(r(t), g(t))$ is a (regular) optimal trajectory, then by the method of Lagrange multipliers, it solves the following variational problem

$$\delta \int_0^1 \mathcal{L} dt = \delta \int_0^1 (C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}_{\text{loc}}(r)\dot{r} \rangle) dt = 0$$

for some $\lambda(t) \in \mathfrak{g}^*$.

Since $B \times G \to B$ is trivial, we can put a trivial connection on this bundle and use it to split the variations into the horizontal and vertical parts. Then by the Lagrangian reduction method recalled above, the reduced curve $(r(t), \dot{r}(t), \xi(t)) \in TQ/G$ with coordinates given by $(r^\alpha, \dot{r}^\alpha, \xi^a)$ satisfies the reduced Euler Lagrange equations stated above. (When using a trivial connection, the coefficients of $\mathcal{A}$ and $\mathcal{B}$ vanish and the reduced Euler-Lagrange equations are called Hamel's equations). ■

Now we are ready to generalize one of the results in Krishnaprasad, Yang and Dayawansa [1991]. Define the components $\mathcal{A}^a_\alpha$ of the mechanical connection by $\mathcal{A}_{\text{loc}}(r)\dot{r} = \mathcal{A}^a_\alpha \dot{r}^\alpha e_a$, where $\{e_a\}$ is the basis of $\mathfrak{g}$ and $\{e^a\}$ is its dual basis. Here α runs from 1 to $n-k$ and a runs from 1 to k where $n-k$ is the dimension of the base space B and k is the dimension of the Lie algebra $\mathfrak{g}$. The result deals with the following problem.

Isoholonomic Problem for Trivial Bundles Minimize $\int_0^1 C(\dot{r})\, dt$, subject to $\dot{r} = u, \dot{g} = -g\mathcal{A}_{\text{loc}} u = -g\mathcal{A}^a_\alpha(r) u^\alpha e_a$, for given boundary conditions

$$(r(0), g(0)) = (\mathbf{0}, g_0), (r(1), g(1)) = (\mathbf{0}, g_1).$$

Corollary 3.4 *Let the cost function C be quadratic in u, say $C = \sum_1^{n-k} c_\alpha (u^\alpha)^2$. If $(r(t), g(t))$ is a (regular) optimal trajectory with the control $\bar{u}(t)$ for the isoholonomic (falling cat) problem, then there exist $\rho(t) \in T^*B$, and $\lambda(t) \in \mathfrak{g}^*$ satisfying $\dot{r}^\alpha = \bar{u}^\alpha$, $\xi^a = -\mathcal{A}^a_\alpha(x)\bar{u}^\alpha$ and the following ordinary differential equations*

$$\begin{aligned} \dot{\rho}_\beta &= \lambda_a \frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta} \bar{u}^\alpha \\ \dot{\lambda}_b &= -C^a_{db} \lambda_a \mathcal{A}^d_\alpha \bar{u}^\alpha \end{aligned}$$

where

$$\bar{u}_\beta = \frac{1}{2c_\beta}(\rho_\beta - \lambda_a \mathcal{A}^a_\beta)$$

with boundary conditions $r(0) = \mathbf{0}, g(0) = g_0, r(1) = \mathbf{0}, g(1) = g_1$.

Proof According to Theorem 3.3, there exists some $\lambda(t) \in \mathfrak{g}^*$ such that the reduced curve $(r(t), \dot{r}(t), \xi(t))$ satisfies the reduced Euler-Lagrange equations for

$$l = c_\alpha(\dot{r}^\alpha)^2 + \langle \lambda_a e^a, (\xi^a + \mathcal{A}^a_\alpha \dot{r}^\alpha) e_a \rangle = c_\alpha(\dot{r}^\alpha)^2 + \lambda_a(\xi^a + \mathcal{A}^a_\alpha \dot{r}^\alpha).$$

After some computations, we find

$$\begin{aligned} \frac{\partial l}{\partial \dot{r}^\beta} &= 2c_\beta \dot{r}^\beta + \lambda_a \mathcal{A}^a_\beta \\ \frac{\partial l}{\partial r^\beta} &= \lambda_a \frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta} \dot{r}^\alpha \\ \frac{\partial l}{\partial \xi^b} &= \lambda_b. \end{aligned}$$

Now let

$$\rho_\beta = \frac{\partial l}{\partial \dot{r}^\beta} = 2c_\beta \dot{r}^\beta + \lambda_a \mathcal{A}^a_\beta$$

and solve for $\dot{r}$, to give

$$\dot{r}_\beta = \frac{1}{2c_\beta}(\rho_\beta - \lambda_a \mathcal{A}^a_\beta).$$

Moreover, the reduced Euler-Lagrange equations (3.2.3) and (3.2.4) give

$$\begin{aligned} \dot{\rho}_\beta &= \frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\beta} = \frac{\partial l}{\partial r^\beta} = \lambda_a \frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta} \dot{r}^\alpha \\ \dot{\lambda}_b &= \frac{d}{dt}\frac{\partial l}{\partial \xi^b} = \frac{\partial l}{\partial \xi^a} C^a_{db} \xi^d = C^a_{db} \lambda_a \xi^d. \end{aligned}$$

After substituting

$$\dot{r}^\alpha = \bar{u}^\alpha \qquad \xi^d = -\mathcal{A}^d_\alpha \bar{u}^\alpha,$$

we get the desired equations. ■

Remarks

1. This Corollary generalizes the result of Krishnaprasad, Yang and Dayawansa [1991] for the trivial principal bundle $(\mathbb{R} \times \mathbb{R} \times SO(3), \mathbb{R} \times \mathbb{R}, \pi, SO(3))$ (see Theorem **3.3** and Remark 3.2 in Krishnaprasad, Yang and Dayawansa [1991]).

2. The reduced equations of motion for ρ_β and λ_b can be written in intrinsic form as a special case of Wong's equations in r_β and λ_b (see the following section).

3.3 Optimal Control of a Holonomic System on a Principal Bundle

While the above method seems to work only for the case where the principle bundle is trivial, it can be easily generalized to an arbitrary principle bundle. In fact, the proof of the Lagrangian reduction theorem stated above provides all the necessary techniques. Recall that Marsden and Scheurle [1993b] arrived at the general reduced Euler-Lagrange equations in two steps:

1. one first gets the Hamel equations in a local bundle trivialization:

$$\frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} = 0$$

$$\frac{d}{dt}\frac{\partial l}{\partial \xi^b} = \frac{\partial l}{\partial \xi^a} C^a_{db} \xi^d,$$

2. one introduces an arbitrary principal connection $\mathcal{A}$ (which is not necessarily the mechanical connection) to split the original variational principle intrinsically and globally relative to horizontal and vertical parts of the variation δq, and derived the general form from the above form by means of a velocity shift replacing ξ by the vertical part relative to this connection:

$$\Omega^a = \mathcal{A}^a_\alpha \dot{r}^\alpha + \xi^a$$

Here, $\mathcal{A}^a_\alpha$ are the local coordinates of the connection $\mathcal{A}$. The resulting reduced Euler- Lagrange equations are then as given earlier.

Now we are ready to state a general theorem for the constrained variational problem on a principal bundle. This problem is as follows:

> **Isoholonomic Problem for General Bundles (The Falling Cat Problem)** Among all curves $q(t)$ such that $q(0) = q_0, q(1) = q_1$ and $\dot{q}(t) \in \mathrm{hor}_{q(t)}$ (horizontal with respect to the mechanical connection $\mathcal{A}^{\mathrm{mec}}$), find the optimal curves $q(t)$ such that $\int_0^1 C(\dot{r})dt$ is minimized, where $r = \pi(q)$.

Observe that while this problem is set up using the mechanical connection $\mathcal{A}^{\mathrm{mec}}$, when applying the Lagrangian reduction theorem, one may use an arbitrary connection $\mathcal{A}$ to split the variational principle. This observation is used in the proof of the following result.

Theorem 3.5 *If $q(t)$ is a (regular) optimal trajectory for the isoholonomic problem for general bundles, then there exists a $\lambda(t) \in \mathfrak{g}^*$ such that the reduced curve in TQ/G with coordinates given in a local trivializtion by $(r^\alpha, \dot{r}^\alpha, \Omega^\alpha)$ satisfies the constraints $\xi^a = -(\mathcal{A}^{\mathrm{mec}})^a_\alpha \dot{r}^\alpha$ as well as the reduced Euler-Lagrange equations (3.1.1) and (3.1.2), where*

$$l = C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}_{\mathrm{loc}}(r)\dot{r} \rangle$$

and

$$\Omega^a = \mathcal{A}^a_\alpha \dot{r}^\alpha + \xi^a$$

Proof The proof proceeds as in the proof in Marsden and Scheurle [1993b] in the present context. The needed modifications of what we have done before are minor, so are omitted. ■

Corollary 3.6 *In the preceding Theorem, if we use the mechanical connection $\mathcal{A}^{\mathrm{mec}}$ to split the variational principle, then the reduced Euler-Lagrange equations coincide with Wong's equations (see Montgomery [1984] and references therein):*

$$\begin{aligned} \dot{p}_\alpha &= -\lambda_a \mathcal{B}^a_{\alpha\beta} \dot{r}^\beta - \frac{1}{2}\frac{\partial g^{\beta\gamma}}{\partial r^\alpha} p_\beta p_\gamma \\ \dot{\lambda}_b &= -\lambda_a C^a_{db} \mathcal{A}^d_\alpha \dot{r}^\alpha \end{aligned}$$

where $g_{\alpha\beta}$ is the local representation of the metric on the base space B, that is

$$C(\dot{r}) = \frac{1}{2} g_{\alpha\beta} \dot{r}^\alpha \dot{r}^\beta,$$

$g^{\beta,\gamma}$ is the inverse of the matrix $g_{\alpha,\beta}$, p_α is defined by

$$p_\alpha = \frac{\partial C}{\partial \dot{r}^\alpha} = g_{\alpha\beta} \dot{r}^\beta$$

and where we write the components of $\mathcal{A}^{\mathrm{mec}}$ simply as $\mathcal{A}^b_\alpha$ and similarly for its curvature.

Proof Applying Theorem 3.5 to the function l where

$$\begin{aligned} l &= C(\dot r) + \langle \lambda(t), \xi + A_{\mathrm{loc}}(r)\dot r\rangle \\ &= C(\dot r) + \langle \lambda(t), \Omega\rangle \\ &= C(\dot r^\alpha) + \lambda_a \Omega^a. \end{aligned}$$

Clearly,

$$\begin{aligned} \frac{\partial l}{\partial \dot r^\alpha} &= \frac{\partial C}{\partial \dot r^\alpha} = g_{\alpha\beta}\dot r^\beta \\ \frac{\partial l}{\partial r^\alpha} &= \frac{\partial C}{\partial r^\alpha} = \frac{1}{2}\frac{\partial g^{\beta\gamma}}{\partial r^\alpha}\dot r^\beta \dot r^\gamma \\ \frac{\partial l}{\partial \Omega^a} &= \lambda_a. \end{aligned}$$

Since $\xi^a = -\mathcal{A}^a_\alpha \dot r^\alpha$ (the constraints) and $\Omega^a = \mathcal{A}^a_\alpha \dot r^\alpha + \xi^a$, we have $\Omega^a = 0$ and the reduced Euler-Lagrange equations become

$$\begin{aligned} \frac{d}{dt}\frac{\partial C}{\partial \dot r^\alpha} - \frac{\partial C}{\partial r^\alpha} &= -\lambda_a(\mathcal{B}^a_{\alpha\beta}\dot r^\beta) \\ \frac{d}{dt}\lambda_b &= -\lambda_a(\mathcal{E}^a_{\alpha b}\dot r^\alpha) = -\lambda_a C^a_{db}\mathcal{A}^d_\alpha \dot r^\alpha. \end{aligned}$$

But

$$\begin{aligned} \frac{d}{dt}\frac{\partial C}{\partial \dot r^\alpha} - \frac{\partial C}{\partial r^\alpha} &= \dot p_\alpha - \frac{1}{2}\frac{\partial g_{\beta\gamma}}{\partial r^\alpha}\dot r^\beta \dot r^\gamma \\ &= \dot p_\alpha + \frac{1}{2}\frac{\partial g^{\kappa\sigma}}{\partial r^\alpha} g_{\kappa\beta} g_{\sigma\gamma}\dot r^\beta \dot r^\gamma \\ &= \dot p_\alpha + \frac{1}{2}\frac{\partial g^{\kappa\sigma}}{\partial r^\alpha} p_\kappa p_\sigma \\ &= \dot p_\alpha + \frac{1}{2}\frac{\partial g^{\beta\gamma}}{\partial r^\alpha} p_\beta p_\gamma, \end{aligned}$$

and so we have the desired equations. ■

Remark

Recall that in Corollary 3.4, we have the reduced equations:

$$\begin{aligned} \dot\rho_\beta &= \lambda_a \frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta}\dot r^\alpha \\ \dot\lambda_b &= -C^a_{db}\lambda_a \mathcal{A}^d_\alpha \dot r^\alpha. \end{aligned}$$

But $\rho_\beta = \lambda_a \mathcal{A}^a_\alpha + 2c_\beta \dot r^\beta$ and hence

$$\dot\rho_\beta = 2c_\beta \ddot r^\beta + \dot\lambda_a \mathcal{A}^a_\beta + \lambda_a \frac{\partial \mathcal{A}^a_\beta}{\partial r^\alpha}\dot r^\alpha = \lambda_a \frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta}\dot r^\alpha.$$

Therefore,

$$\begin{aligned} 2c_\beta \ddot r^\beta &= \lambda_a \frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta}\dot r^\alpha - \lambda_a \frac{\partial \mathcal{A}^a_\beta}{\partial r^\alpha}\dot r^\alpha - (-C^a_{db}\lambda_a \mathcal{A}^d_\alpha \dot r^\alpha)\mathcal{A}^b_\beta \\ &= \lambda_a\left(\frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta} - \frac{\partial \mathcal{A}^a_\beta}{\partial r^\alpha} - C^a_{bd}\mathcal{A}^d_\alpha \mathcal{A}^b_\beta\right)\dot r^\alpha \\ &= -\lambda_a \mathcal{B}^a_{\beta\alpha}\dot r^\alpha. \end{aligned}$$

That is, the reduced equations in Corollary 3.4 (and those in Krishnaprasad, Yang and Dayawansa [1991]) can be written intrinsically as Wong's equations after a change of variables. This should not surprise us because Marsden and Scheurle derived the general reduced Euler-Lagrange equations from the Hamel equations using a suitable change of variables from local trivialization variables to those in which the Lie algebra variable is replaced by the vertical part of the bundle velocity.

4 Optimal Control and Lagrangian Reduction for Nonholonomic Systems

Now we are ready to use the method of Lagrange multipliers and Lagrangian reduction to find the necessary conditions for optimal trajectories of nonholonomic systems in the case of a trivial bundle.

4.1 The General Theorem for Optimization

In Bloch, Krishnaprasad, Marsden and Murray [1995], the reconstruction process may be seen in a two step fashion: given an initial condition and a path $r(t)$ in the base space, we first integrate the momentum equation to determine $p(t)$ for all time and then use $r(t)$ and $p(t)$ jointly to determine the motion $g(t)$ in the fiber. But in studying the optimal control problem, it is better to treat p as a set of independent variables and the momentum equation as an additional set of constraints. With this viewpoint, it is possible to write down the reduced equations of motion for the optimal trajectories.

Since we would like to use the method of Lagrange multipliers to relax the constraints, we define a new Lagrangian $\mathcal{L}$:

$$\mathcal{L} = C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}(r)\dot{r} - \Gamma(r)p \rangle + \left\langle \kappa(t), \dot{p} - \dot{r}^T H(r)\dot{r} - \dot{r}^T K(r)p - p^T D(r)p \right\rangle \tag{4.1.1}$$

for some $\lambda(t) \in \mathfrak{g}^*$ and for some $\kappa(t) \in \mathbb{R}^m$, where m is the number of momentum functions p_b. For simplicity of notation we have written $\mathcal{A}$ for $\mathcal{A}^{\mathrm{nh}}$. Clearly $\mathcal{L}$ is G-invariant and induces a function on $(T(Q \times \mathbb{R}^m)/G) \times \mathfrak{g}^* \times \mathbb{R}^m$ which is also denoted $\mathcal{L}$.

We formulate the main problem to be studied as follows.

> **Isoholonomic Problem for Nonholonomic Systems** Among all curves $q(t)$ such that $q(0) = q_0, q(1) = q_1$, $\dot{q}(t) \in \mathcal{D}_{q(t)}$ and that satisfy $g^{-1}\dot{g} + \mathcal{A}(r)\dot{r} = \Gamma(r)p$ and the momentum equation, find the optimal curves $q(t)$ such that $\int_0^1 C(\dot{r})dt$ is minimized, where $r = \pi(q)$.

Before we state the theorem and do some computations, we want to make sure that the readers understand the index convention used in this section:

1. The first batch of indices is denoted $a, b, c, ...$ and range from 1 to k corresponding to the symmetry direction ($k = \dim \mathfrak{g}$).
2. The second batch of indices will be denoted $i, j, k, ...$ and range from 1 to m corresponding to the symmetry direction along constraint space (m is the number of momentum functions).
3. The indices $\alpha, \beta, ...$ on the shape variables r range from 1 to $n-k$ ($n-k = \dim (Q/G)$, i.e., the dimension of the shape space).

Theorem 4.1 Reduced Lagrangian Optimization for Nonholonomic Systems *If $q(t) = (r(t), g(t))$ is a (regular) optimal trajectory for the above optimal control problem, then there exist* *a* $\lambda(t) \in \mathfrak{g}^*$ *and a* $\kappa(t) \in \mathbb{R}^m$ *such that the reduced curve* $(r(t), \dot r(t), \xi(t)) \in TQ/G$ *with coordinates* $(r^\alpha, \dot r^\alpha, \xi^\alpha)$ *satisfies the reduced Euler Lagrange equations*

$$
\begin{aligned}
\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot r^\beta} &- \frac{\partial \mathcal{L}}{\partial r^\beta} = 0 \\
\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \xi^b} &= \frac{\partial \mathcal{L}}{\partial \xi^a} C^a_{db}\xi^d \\
\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot p^j} &- \frac{\partial \mathcal{L}}{\partial p^j} = 0,
\end{aligned}
$$

as well as

$$
\begin{aligned}
\xi &= -\mathcal{A}(r)\dot r + \Gamma(r)p \\
\dot p &= \dot r^T H(r)\dot r + \dot r^T K(r)p + p^T D(r)p.
\end{aligned}
$$

Here C^a_{db} *are the structure coefficients of the Lie algebra* $\mathfrak{g}$ *and*

$$
\mathcal{L} = C(\dot r) + \langle \lambda(t), \xi + \mathcal{A}(r)\dot r - \Gamma(r)p\rangle + \left\langle \kappa(t), \dot p - \dot r^T H(r)\dot r - \dot r^T K(r)p - p^T D(r)p\right\rangle \tag{4.1.2}
$$

Proof If $(r(t), g(t))$ is a (regular) optimal trajectory, then by the method of Lagrange multipliers, it solves the following variational problem

$$
\delta \int_0^1 \mathcal{L}dt = 0
$$

for some $\lambda(t) \in \mathfrak{g}^*$ and some $\kappa(t) \in \mathbb{R}^m$. Since the bundle is trivial, we can put a flat connection on this bundle and use it to split the variations into horizontal and vertical parts. Then by the Lagrange reduction theorem, the reduced curve $(r(t), \dot r(t), \xi(t)) \in TQ/G$ satisfies the reduced Euler Lagrange equations stated above. ■

4.2 The Optimality Conditions in Coordinates

Now let us work out everything in detail in bundle coordinates. Since

$$
\mathcal{L} = \frac{1}{2}C_\alpha(\dot r^\alpha)^2 + \lambda_a(\xi^a + \mathcal{A}^a_\alpha \dot r^\alpha - \Gamma^{ai}p_i) + \kappa^i(\dot p_i - H_{\alpha\gamma i}\dot r^\alpha \dot r^\gamma - K^l_{i\alpha}\dot r^\alpha p_l - D^{lk}_i p_l p_k), \tag{4.2.1}
$$

we find after some computations that

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \dot r^\beta} &= C_\beta \dot r^\beta + \lambda_a \mathcal{A}^a_\beta - \kappa^i(2H_{\alpha\beta i}\dot r^\alpha + K^l_{i\beta}p_l) \\
\frac{\partial \mathcal{L}}{\partial r^\beta} &= \lambda_a\left(\frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta}\dot r^\alpha - \frac{\partial \Gamma^{ai}}{\partial r^\beta}p_i\right) - \kappa^i\left(\frac{\partial H_{\alpha\gamma i}}{\partial r^\beta}\dot r^\alpha \dot r^\gamma + \frac{\partial K^l_{i\alpha}}{\partial r^\beta}\dot r^\alpha p_l + \frac{\partial D^{lk}_i}{\partial r^\beta}p_l p_k\right).
\end{aligned}
$$

Also we have

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \xi^b} &= \lambda_b \\
\frac{\partial \mathcal{L}}{\partial \dot p^j} &= \kappa^j \\
\frac{\partial \mathcal{L}}{\partial p_j} &= -\lambda_a \Gamma^{aj} - \kappa^i(K^j_{i\alpha}\dot r^\alpha + 2D^{lj}_i p_l).
\end{aligned}
$$

By Theorem 4.1, we know that the reduced curve $(r(t), \dot r(t), \xi(t))$ must satisfy the following system of differential equations for the given boundary conditions $q(0) = (r_0, g_0), q(1) = (r_1, g_1)$:

$$\frac{d}{dt}[C_\beta \dot r^\beta + \lambda_a \mathcal{A}^a_\beta - \kappa^i (2H_{\alpha\beta i}\dot r^\alpha + K^l_{i\beta} p_l)]$$
$$= \lambda_a \left(\frac{\partial \mathcal{A}^a_\alpha}{\partial r^\beta}\dot r^\alpha - \frac{\partial \Gamma^{ai}}{\partial r^\beta} p_i\right) - \kappa^i\left(\frac{\partial H_{\alpha\gamma i}}{\partial r^\beta}\dot r^\alpha \dot r^\gamma + \frac{\partial K^l_{i\alpha}}{\partial r^\beta}\dot r^\alpha p_l + \frac{\partial D^{lk}_i}{\partial r^\beta} p_l p_k\right)$$

and

$$\begin{aligned}
\dot\kappa^j &= -\lambda_a \Gamma^{aj} - \kappa^i (K^j_{i\alpha}\dot r^\alpha + 2D^{lj}_i p_l) \\
\dot\lambda_b &= C^a_{db}\lambda_a \xi^d = C^a_{db}\lambda_a(-\mathcal{A}^d_\alpha \dot r^\alpha + \Gamma^{di} p_i) \\
\dot p_i &= H_{\alpha\gamma i}\dot r^\alpha \dot r^\gamma + K^l_{i\alpha}\dot r^\alpha p_l + D^{lk}_i p_l p_k.
\end{aligned}$$

Remarks

1. The first set of equations can be simplified somewhat as follows:

$$\frac{d}{dt}\left[C_\beta \dot r^\beta - \kappa^i(2H_{\alpha\beta i}\dot r^\alpha + K^l_{i\beta}p_l)\right]$$
$$= \lambda_a \mathcal{B}^a_{\beta\alpha}\dot r^\alpha - \lambda_a\left(\frac{\partial \Gamma^{ai}}{\partial r^\beta} + C^a_{db}\mathcal{A}^b_\beta \Gamma^{di}\right)p_i - \kappa^i\left(\frac{\partial H_{\alpha\gamma i}}{\partial r^\beta}\dot r^\alpha \dot r^\gamma + \frac{\partial K^l_{i\alpha}}{\partial r^\beta}\dot r^\alpha p_l + \frac{\partial D^{lk}_i}{\partial r^\beta}p_l p_k\right).$$

   where $\mathcal{B}^a_{\beta\alpha}$ are the coordinates of the curvature $\mathcal{B}$ of the nonholonomic connection $\mathcal{A}$, which is used to set up the constrained variational problem. Clearly more work is needed to establish a better form of the first set of equations as well as the geometry behind them. However, for the snakeboard, the reduced equations of motion for the optimal trajectories turn out to be rather simple.

2. In proving the above theorem, while variations with fixed endpoints for $r(t)$ can be used, we generally can only hold the initial endpoint fixed for the variations of $p(t)$ and leave their final endpoints free (which is called "free endpoint problem" in the language of calculus of variations). However, we will obtain the same system of differential equations (namely the reduced Euler- Lagrange equations) except the need to impose some kind of transversality condition at $t = 1$, e.g., in this case we need to have $\kappa(1) = 0$.

In the following section, we will apply the method of reduced Lagrangian optimization developed in this section to some examples, especially the snakeboard.

5 Examples

5.1 Optimal Control of a Homogeneous Ball on a Rotating Plate

Bloch, Krishnaprasad, Marsden and Murray [1995] also studies a well-known example, namely the model of a homogeneous ball on a rotating plate (for more informations, also see Neimark and Fufaev [1972] and Yang [1992] for the affine case and Bloch and Crouch [1992], Brockett and Dai [1992] and Jurdjevic [1993] for the linear case) and writes down its equations of motion in a form that is suitable for the application of control theory.

Fix coordinates in inertial space and let the plane rotate with constant angular velocity Ω about the z-axis. The configuration space of the sphere is $Q = \mathbb{R}^2 \times SO(3)$, parameterized by $(x, y, g), g \in SO(3)$, all measured with respect to the inertial frame. Let $\omega = (\omega_x, \omega_y, \omega_z)$ be the angular velocity vector of the sphere measured also with respect to the inertial frame, let m be the mass of the sphere, mk^2 its inertia about any axis, and let a be its radius.

The Lagrangian of the system is

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mk^2({\omega_x}^2 + {\omega_y}^2 + {\omega_z}^2)$$

with the affine nonholonomic constraints

$$\begin{aligned} \dot{x} - a\omega_y &= -\Omega y \\ \dot{y} + a\omega_x &= \Omega x. \end{aligned}$$

Note that the Lagrangian here is a metric on Q which is bi-invariant on $SO(3)$ as the ball is homogeneous. Note also that $\mathbb{R}^2 \times SO(3)$ is a principal bundle over $\mathbb{R}^2$ with respect to the right $SO(3)$ action on Q given by

$$(x, y, g) \mapsto (x, y, gh)$$

for $h \in SO(3)$. The action is on the *right* since the symmetry is a material symmetry.

After some computations, it can be shown that (for details, see Bloch, Krishnaprasad, Marsden and Murray [1995]) the equations of motion are:

$$\begin{aligned} \omega_x &+ \frac{1}{a}\dot{y} = \frac{\Omega x}{a} \\ \omega_y &- \frac{1}{a}\dot{x} = \frac{\Omega y}{a} \\ \omega_z &= c, \end{aligned}$$

(where c is a constant), together with

$$\begin{aligned} \ddot{x} + \frac{k^2\Omega}{a^2 + k^2}\dot{y} &= 0 \\ \ddot{y} - \frac{k^2\Omega}{a^2 + k^2}\dot{x} &= 0. \end{aligned}$$

Notice that the first set of three equations has the form

$$\dot{g}g^{-1} = -\mathcal{A}_{\mathrm{loc}}(r)\dot{r} + \Gamma_{\mathrm{loc}}(r),$$

where

$$\mathcal{A}_{\mathrm{loc}} = \frac{1}{a}e_1 dy - \frac{1}{a}e_2 dx$$

and

$$\Gamma_{\mathrm{loc}} = \frac{\Omega}{a}xe_1 + \frac{\Omega}{a}ye_2 + ce_3.$$

Here, $r^1 = x, r^2 = y$ and e_1, e_2, e_3 is the standard basis of $so(3)_-$. Also, $\mathcal{A}_{\mathrm{loc}}$ is the expression of nonholonomic connection relative to the (global) trivialization and Γ_{loc} is the expression of the affine piece of the constraints with respect to the same trivialization (see Bloch, Krishnaprasad, Marsden and Murray [1995]).

Now we are ready to apply reduced Lagrangian optimization to find the optimal trajectories for a homogeneous ball. Clearly the homogeneous ball on a rotating plate is a simple nonholonomic mechanical system with symmetry as defined earlier, which also has a trivial principal bundle structure (except that the constraint is affine which can be dealt with in the same way). Also we can assume that we have full control over the motion of the center of the ball, i.e., over the shape variables. Now let the cost function be $C(\dot{r}) = \frac{1}{2}[(\dot{x})^2 + (\dot{y}^2)]$ and set $a = 1$ for simplicity, then we can use the method of Lagrange multipliers and Lagrangian reduction to find the necessary conditions for the optimal trajectories of the following optimal control problem:

Plate Ball Problem Given two points $q_0, q_1 \in \mathbb{R}^2 \times SO(3)$, find the optimal control curves $(x(t), y(t)) \in \mathbb{R}^2$ that steer the system from q_0 to q_1 and minimizes $\int_0^1 \frac{1}{2}[(\dot{x})^2 + (\dot{y})^2]dt$, subject to the constraints

$$\dot{g}g^{-1} = -\dot{y}e_1 + \dot{x}e_2 + ce_3 + \Omega x e_1 + \Omega y e_2,$$

where, again, e_a is the standard basis of $so(3)_-$.

Following the reduced Lagrangian optimization method developed in the preceding section, we define a new Lagrangian $\mathcal{L}$ by

$$\mathcal{L} = \frac{1}{2}[(\dot{x})^2 + (\dot{y})^2] + \lambda_a \xi^a + \lambda_1 \dot{y} - \lambda_2 \dot{x} - \lambda_3 c - \Omega\lambda_1 x - \Omega\lambda_2 y,$$

where $\lambda(t) \in so(3)^*_-$ (note that we use the negagive Lie-Poisson structure because the right action is used).

By the preceding Theorem, we know that any the reduced optimal curve $(x(t), y(t), \dot{x}(t), \dot{y}(t), \xi^a(t))$ must satisfy the reduced Euler Lagrangian equations. Simple computations show that

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \dot{x}} &= \dot{x} - \lambda_2 = \rho_1 \\
\frac{\partial \mathcal{L}}{\partial x} &= -\Omega\lambda_1 \\
\frac{\partial \mathcal{L}}{\partial \dot{y}} &= \dot{y} + \lambda_1 = \rho_2 \\
\frac{\partial \mathcal{L}}{\partial y} &= -\Omega\lambda_2 \\
\frac{\partial \mathcal{L}}{\partial \xi^b} &= \lambda_b.
\end{aligned}$$

Therefore

$$\begin{aligned}
\dot{\rho}_1 &= -\Omega\lambda_1 \\
\dot{\rho}_2 &= -\Omega\lambda_2,
\end{aligned}$$

and

$$\dot{\lambda}_b = C^a_{db}\lambda_a\xi^d,$$

that is:

$$\begin{aligned}
\dot{\lambda}_1 &= \lambda_3\xi^2 - \lambda_2\xi^3 = \lambda_3(\rho_1 + \lambda_2 + \Omega y) - c\lambda_2 \\
\dot{\lambda}_2 &= -\lambda_3\xi^1 + \lambda_1\xi^3 = \lambda_3(\rho_2 - \lambda_1 - \Omega x) + c\lambda_1 \\
\dot{\lambda}_3 &= \lambda_2\xi^1 - \lambda_1\xi^2 = -(\lambda_1\rho_1 + \lambda_2\rho_2) + \Omega(\lambda_2 x - \lambda_1 y).
\end{aligned}$$

In the special case where $c = 0$ (no drift) and $\Omega = 0$ (no rotation) studied in Jurdjevic [1993], we have

$$\begin{aligned}
\dot{\rho}_1 &= 0 \\
\dot{\rho}_2 &= 0 \\
\dot{\lambda}_1 &= \lambda_3(\rho_1 + \lambda_2) \\
\dot{\lambda}_2 &= \lambda_3(\rho_2 - \lambda_1) \\
\dot{\lambda}_3 &= -(\lambda_1\rho_1 + \lambda_2\rho_2).
\end{aligned}$$

which gives the same result as in Jurdjevic [1993] obtained through the application of the Pontryagin Maximum Principle.

5.2 Optimal Control of the Snakeboard

The snakeboard is a modified version of a skateboard in which the front and back pairs of wheels are independently actuated. The extra degree of freedom enables the rider to generate forward motion by twisting their body back and forth, while simultaneously moving the wheels with the proper phase relationship. For details, see Bloch, Krishnaprasad, Marsden and Murray [1995] and the references listed there. Here we will include the computations shown in that paper both for completeness as well as to make concrete the nonholonomic theory.

The snakeboard is modeled as a rigid body (the board) with two sets of independently actuated wheels, one on each end of the board. The human rider is modeled as a momentum wheel which sits in the middle of the board and is allowed to spin about the vertical axis. Spinning the momentum wheel causes a counter-torque to be exerted on the board. The configuration of the board is given by the position and orientation of the board in the plane, the angle of the momentum wheel, and the angles of the back and front wheels. Thus the configuration space is $Q = SE(2) \times S^1 \times S^1 \times S^1$. Let (x, y, θ) represent the position and orientation of the center of the board, ψ the angle of the momentum wheel relative to the board, and ϕ_1 and ϕ_2 the angles of the back and front wheels, also relative to the board. Take the distance between the center of the board and the wheels to be r.

The Lagrangian for the snakeboard consists only of kinetic energy terms and can be written as

$$L(q, \dot{q}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}J\dot{\theta}^2 + \frac{1}{2}J_0(\dot{\theta} + \dot{\psi})^2 + \frac{1}{2}J_1(\dot{\theta} + \dot{\phi}_1)^2 + \frac{1}{2}J_2(\dot{\theta} + \dot{\phi}_2)^2,$$

where m is the total mass of the board, J is the inertia of the board, J_0 is the inertia of the rotor and J_i, $i = 1, 2$, is the inertia corresponding to ϕ_i. The Lagrangian is independent of the configuration of the board and hence it is invariant to all possible group actions.

The rolling of the front and rear wheels of the snakeboard is modeled using nonholonomic constraints which allow the wheels to spin about the vertical axis and roll in the direction that they are pointing. The wheels are not allowed to slide in the sideways direction. This gives constraint one forms

$$\begin{aligned}
\omega_1(q) &= -\sin(\theta + \phi_1)dx + \cos(\theta + \phi_1)dy - r\cos\phi_1 d\theta \\
\omega_2(q) &= -\sin(\theta + \phi_2)dx + \cos(\theta + \phi_2)dy + r\cos\phi_2 d\theta.
\end{aligned}$$

These constraints are invariant under the $SE(2)$ action given by

$$(x, y, \theta, \psi, \phi_1, \phi_2) \mapsto (x\cos\alpha - y\sin\alpha + a, x\sin\alpha + y\cos\alpha + b, \theta + \alpha, \psi, \phi_1, \phi_2),$$

where $(a, b, \alpha) \in SE(2)$. The constraints determine the kinematic distribution $\mathcal{D}_q$:

$$\mathcal{D}_q = \operatorname{span}\left\{\frac{\partial}{\partial\psi}, \frac{\partial}{\partial\phi_1}, \frac{\partial}{\partial\phi_2}, a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial\theta}\right\},$$

where a, b, and c, are given by

$$\begin{aligned}
a &= -r(\cos\phi_1\cos(\theta + \phi_2) + \cos\phi_2\cos(\theta + \phi_1)) \\
b &= -r(\cos\phi_1\sin(\theta + \phi_2) + \cos\phi_2\sin(\theta + \phi_1)) \\
c &= \sin(\phi_1 - \phi_2).
\end{aligned}$$

The tangent space to the orbits of the $SE(2)$ action is given by

$$T_q(\mathrm{Orb}(q)) = \operatorname{span}\left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial\theta}\right\}$$

The intersection between the tangent space to the group orbits and the constraint distribution is thus given by

$$\mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial \theta}.$$

The momentum can be constructed by choosing a section of $\mathcal{D} \cap T\mathrm{Orb}$ regarded as a bundle over Q. Since $\mathcal{D}_q \cap T_q\mathrm{Orb}(q)$ is one-dimensional, the section can be chosen to be

$$\xi_Q^q = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial \theta},$$

which is invariant under the action of $SE(2)$ on Q. The corresponding Lie algebra element in $se(2)$, ξ^q, is

$$\xi^q = (a + yc)e_x + (b - xc)e_y + ce_\theta$$

where e_x is the basis element of the Lie algebra corresponding to translations in the x direction (and whose corresponding infinitesimal generator is $\partial/\partial x$), etc. The nonholonomic momentum map is thus given by

$$\begin{aligned} p &= J^{\mathrm{nh}}(\xi^q) = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i \\ &= ma\dot{x} + mb\dot{y} + Jc\dot{\theta} + J_0 c(\dot{\theta} + \dot{\psi}) + J_1 c(\dot{\theta} + \dot{\phi}_1) + J_2 c(\dot{\theta} + \dot{\phi}_2). \end{aligned}$$

In Bloch, Krishnaprasad, Marsden and Murray [1995] a simplification is made which we shall also assume in this paper, namely $\phi_1 = -\phi_2$, $J_1 = J_2$. The parameters are also chosen such that $J + J_0 + J_1 + J_2 = mr^2$ (which eliminates some terms in the derivation but does not affect the essential geometry of the problem). Setting $\phi = \phi_1 = -\phi_2$, the constraints plus the momentum are given by

$$\begin{aligned} 0 &= -\sin(\theta + \phi)\dot{x} + \cos(\theta + \phi)\dot{y} - r\cos\phi\dot{\theta} \\ 0 &= -\sin(\theta - \phi)\dot{x} + \cos(\theta - \phi)\dot{y} + r\cos\phi\dot{\theta} \\ p &= -2mr\cos^2(\phi)\cos(\theta)\dot{x} - 2mr\cos^2(\phi)\sin(\theta)\dot{y} \\ &\quad + mr^2\sin(2\phi)\dot{\theta} + J_0\sin(2\phi)\dot{\psi}. \end{aligned}$$

Adding, subtracting, and scaling these equations, we can write (away from $\phi = \pi/2$),

$$\begin{bmatrix} \cos(\theta)\dot{x} + \sin(\theta)\dot{y} \\ -\sin(\theta)\dot{x} + \cos(\theta)\dot{y} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} -\dfrac{J_0}{2mr}\sin(2\phi)\dot{\psi} \\ 0 \\ \dfrac{J_0}{mr^2}\sin^2(\phi)\dot{\psi} \end{bmatrix} = \begin{bmatrix} \dfrac{-1}{2mr}p \\ 0 \\ \dfrac{\tan\phi}{2mr^2}p \end{bmatrix}. \tag{5.2.1}$$

These equations have the form

$$g^{-1}\dot{g} + \mathcal{A}_{\mathrm{loc}}(r)\dot{r} = \Gamma(r)p$$

where

$$\begin{aligned} \mathcal{A}_{\mathrm{loc}} &= -\frac{J_0}{2mr}\sin(2\phi)e_x\,d\psi + \frac{J_0}{mr^2}\sin^2(\phi)e_\theta\,d\psi \\ \Gamma(r) &= \frac{-1}{2mr}e_x + \frac{1}{2mr^2}\tan(\phi)\,e_\theta. \end{aligned}$$

These are precisely the terms which appear in the nonholonomic connection relative to the (global) trivialization (r, g). The momentum equation, which governs the evolution of p, is given by

$$\dot{p} = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}\xi^q\right]_Q^i$$

$$
\begin{aligned}
= \quad & 4mr\cos(\theta)\cos(\phi)\sin(\phi)\dot{x}\dot{\phi} + 4mr\sin(\theta)\cos(\phi)\sin(\phi)\dot{y}\dot{\phi} \\
& +2J_0\cos(2\phi)\dot{\phi}\dot{\psi} + 2mr^2\cos(2\phi)\dot{\theta}\dot{\phi} \\
& -2mr\cos(\theta)\cos^2(\phi)\dot{y}\dot{\theta} + 2mr\sin(\theta)\cos^2(\phi)\dot{x}\dot{\theta}
\end{aligned}
$$

Solving for the group velocities $\dot{x}, \dot{y}, \dot{\theta}$ from the equations which define the nonholonomic connection, the momentum equation can be rewritten as

$$\dot{p} = 2J_0\cos^2(\phi)\ \dot{\phi}\dot{\psi} - \tan(\phi)\ p\dot{\phi}$$

This version of the momentum equation corresponds to the coordinate form in body representation but it contains no terms which are quadratic in p, due to the fact that $\mathfrak{g}^q$ is one dimensional.

These equations describe how paths in the base space, parameterized by $r \in S^1 \times S^1 \times S^1$ (in fact, the base space is $S^1 \times S^1$ if we assume $\phi_1 = -\phi_2$), are lifted to the fiber $SE(2)$. The utility of these equations is that they greatly simplify the process of solving for the motion of the system given the base space trajectory.

Now we are ready to apply the method of reduced Lagrangian optimization to find the optimal trajectories for the snakeboard. Clearly the snakeboard is a simple nonholonomic mechanical system with symmetry as defined earlier and which also has a trivial principal bundle structure. Moreover, the control forces are only applied to the shape variables which we have full control of. Let the cost function be $C(\dot{r}) = \frac{1}{2}[(\dot{\psi})^2 + (\dot{\phi})^2]$ for simplicity. We can use the method of Lagrange multipliers and Lagrangian reduction to find the necessary conditions for the optimal trajectories of the following optimal control problem:

> **Optimal Control Problem for the Snakeboard** Given two points $q_0, q_1 \in SE(2) \times S^1 \times S^1$, find the optimal control curves $(\psi(t), \phi(t)) \in S^1 \times S^1$ that steer from q_0 to q_1 and minimize $\int_0^1 \frac{1}{2}((\dot{\psi})^2 + (\dot{\phi})^2)\mathrm{dt}$, subject to the constraints
>
> $$
\begin{aligned}
g^{-1}\dot{g} + \mathcal{A}_{\mathrm{loc}}(r)\dot{r} &= \Gamma(r)p \\
\dot{p} = 2J_0\cos^2(\phi)\dot{\phi}\dot{\psi} &- \tan(\phi)p\dot{\phi}
\end{aligned}
$$
>
> where
>
> $$
\begin{aligned}
\mathcal{A}_{\mathrm{loc}} &= -\frac{J_0}{2mr}\sin(2\phi)e_x\, d\psi + \frac{J_0}{mr^2}\sin^2(\phi)e_\theta\, d\psi \\
\Gamma(r) &= \frac{-1}{2mr}e_x + \frac{1}{2mr^2}\tan(\phi)\, e_\theta .
\end{aligned}
$$

Following the general procedures in the previous section, we define a new $\mathcal{L}$ by

$$
\begin{aligned}
\mathcal{L} \quad = \quad & \frac{1}{2}((\dot{\psi})^2 + (\dot{\phi})^2) + \lambda_a\xi^a - \frac{J_0}{2mr}\lambda_1\sin(2\phi)\dot{\psi} + \frac{J_0}{mr^2}\lambda_3\sin^2(\phi)\dot{\psi} \\
+ \quad & \frac{1}{2mr}\lambda_1 p - \frac{1}{2mr^2}\lambda_3\tan(\phi)p + \kappa\dot{p} - 2J_0\kappa\cos^2(\phi)\dot{\phi}\dot{\psi} + \kappa\tan(\phi)p\dot{\phi}
\end{aligned}
$$

where $\xi = g^{-1}\dot{g} \in \mathfrak{g}$, $\lambda(t) \in \mathfrak{g}^*$ and $\kappa(t) \in \mathbb{R}^1$ are Lagrange multipliers. Here ξ^a and λ_a are the components of ξ and λ in the standard basis of $se(2)$ and $se(2)^*$ respectively.

By Theorem 4.1, we know that the reduced optimal curves $(\psi(t), \phi(t), \dot{\psi}(t), \dot{\phi}(t), \xi^a(t))$ must satisfy the reduced Euler Lagrangian equations for $\mathcal{L}$. After some computations, we find

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \dot{\psi}} &= \dot{\psi} - \frac{J_0}{2mr}\lambda_1\sin(2\phi) + \frac{J_0}{mr^2}\lambda_3\sin^2(\phi) - 2J_0\kappa\cos^2(\phi)\dot{\phi} \\
\frac{\partial \mathcal{L}}{\partial \psi} &= 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \dot\phi} &= \dot\phi - 2J_0\kappa\cos^2(\phi)\dot\psi + \kappa\tan(\phi)p \\
\frac{\partial \mathcal{L}}{\partial \phi} &= -\frac{J_0}{mr}\lambda_1\cos(2\phi)\dot\psi + \frac{J_0}{mr^2}\lambda_3\sin(2\phi)\dot\psi - \frac{1}{2mr^2}\lambda_3\sec^2(\phi)p \\
&\quad +2J_0\kappa\sin(2\phi)\dot\phi\dot\psi + \kappa\sec^2(\phi)p\dot\phi \\
\frac{\partial \mathcal{L}}{\partial \dot p} &= \kappa \\
\frac{\partial \mathcal{L}}{\partial p} &= \frac{1}{2mr}\lambda_1 - \frac{1}{2mr^2}\lambda_3\tan(\phi) + \kappa\tan(\phi)\dot\phi \\
\frac{\partial \mathcal{L}}{\partial \xi^b} &= \lambda_b.
\end{aligned}
$$

Substitute the above calculations into the reduced Euler Lagrangian equations and simplify, giving

$$
\begin{aligned}
\ddot\psi &- \frac{J_0}{2mr}\dot\lambda_1\sin(2\phi) - \frac{J_0}{mr}\lambda_1\cos(2\phi)\dot\phi + \frac{J_0}{mr^2}\lambda_3\sin(2\phi)\dot\phi \\
&+ \frac{J_0}{mr^2}\dot\lambda_3\sin^2(\phi) - 2J_0\dot\kappa\cos^2\phi\,\dot\phi + 2J_0\kappa\sin(2\phi)(\dot\phi)^2 - 2J_0\kappa\cos^2(\phi)\ddot\phi = 0 \\
\ddot\phi &- 2J_0\dot\kappa\cos^2(\phi)\dot\psi - 2J_0\kappa\cos^2(\phi)\ddot\psi + \dot\kappa\tan(\phi)p + \kappa\tan(\phi)\dot p \\
&= -\frac{J_0}{mr}\lambda_1\cos(2\phi)\dot\psi + \frac{J_0}{mr^2}\lambda_3\sin(2\phi)\dot\psi - \frac{1}{2mr^2}\lambda_3\sec^2(\phi)p.
\end{aligned}
$$

Also, we have

$$
\begin{aligned}
\dot\kappa &= \frac{1}{2mr}\lambda_1 - \frac{1}{2mr^2}\lambda_3\tan(\phi) + \kappa\tan(\phi)\dot\phi \\
\dot\lambda_1 &= \lambda_2\xi^3 = \lambda_2\left(-\frac{J_0}{mr^2}\sin^2(\phi)\dot\psi + \frac{1}{2mr^2}\tan(\phi)p\right) \\
\dot\lambda_2 &= -\lambda_1\xi^3 = -\lambda_1\left(-\frac{J_0}{mr^2}\sin^2(\phi)\dot\psi + \frac{1}{2mr^2}\tan(\phi)p\right) \\
\dot\lambda_3 &= -\lambda_2\xi^1 = -\lambda_2\left(\frac{J_0}{2mr}\sin(2\phi)\dot\psi - \frac{1}{2mr}p\right) \\
\dot p &= 2J_0\cos^2(\phi)\,\dot\phi\dot\psi - \tan(\phi)\,p\dot\phi.
\end{aligned}
$$

After eliminating $\dot\lambda_1, \dot\lambda_3, \dot\kappa$ and $\dot p$ from the first set of two equations, we finally obtain

$$
\begin{aligned}
\ddot\psi &- \frac{J_0}{2mr}\lambda_1(1+3\cos(2\phi))\dot\phi + \frac{3J_0}{2mr^2}\lambda_3\sin(2\phi)\dot\phi + J_0\kappa\sin(2\phi)(\dot\phi)^2 - 2J_0\kappa\cos^2(\phi)\ddot\phi = 0 \\
\ddot\phi &- \frac{J_0}{mr}\lambda_1\sin^2\phi\,\dot\psi + \frac{1}{2mr}\lambda_1\tan(\phi)p + \frac{1}{2mr^2}\lambda_3 p - \frac{J_0}{2mr^2}\lambda_3\sin(2\phi)\dot\psi - 2J_0\kappa\cos^2(\phi)\ddot\psi = 0.
\end{aligned}
$$

5.3 Optimal Control on a Lie Group

Krishnaprasad [1994] considered the following optimal control problem on a finite dimensional Lie group G which has been used to model various problems in several other papers (e.g. the plate-ball problem in Jurdjevic [1993], and the landing tower problem in Walsh, Montgomery and Sastry [1994]). While it is possible to model this class of problems as a special case of the optimal control of nonholonomic system on a trivial principal bundle and apply reduced Lagrangian optimization, it may be useful to provide in this section a more direct proof that uses simpler machinery.

Optimal Control Problem for a Lie Group Given a left invariant control system on G, $\dot g = g\cdot\xi_u$, where $\xi_u = e_0 + \sum_{i=1}^m u^i(t)e_i$, find the optimal controls $u(\cdot)$ that steer from g_0 to g_1 and minimize $\int_0^1 L(u)dt$.

Here $\{e_0, e_1, \ldots, e_m\}$ spans an $(m+1)$-dimensional subspace of the whole Lie algebra $\mathfrak{g}$ of G, $m+1 \leq n = \dim(\mathfrak{g})$, $u(\cdot)$ is a vector valued control function with $u^i(t) \in \mathbb{R}$, L is a cost function on $\mathbb{R}^m$ which is the space of values of controls, and $L(u) = \frac{1}{2}\sum_{i=1}^m I_i(u^i)^2$ with $I_i > 0$.

To apply the method of Lagrangian reduction, we recast the above optimal control problem as a constrained variational problem. For simplicity of exposition, we will deal with the vector space case first where there is no e_0 term and will take up the affine case later.

Let $\mathcal{C}$ be the m-dimensional subspace of $\mathfrak{g}$ spanned by $\{e_1, \ldots, e_m\}$. We make the following points

(i) $\xi_u = \sum_{i=1}^m u^i(t) e_i$ lies in $\mathcal{C}$;

(ii) if we define $L_1 = L \circ \phi$ where $L = \frac{1}{2}\sum_{i=1}^m I_i(u^i)^2$ with $I_i > 0$ and $\phi = (e^1, \ldots, e^m)$ with $\{e^1, \ldots, e^m\}$ as the dual basis of $\{e_1, \ldots, e_m\}$, then $L_1 : \mathcal{C} \to \mathbb{R}$ is nothing but $\frac{1}{2}$ of the square of a metric on $\mathcal{C}$ which is intrinsically defined and does not depend on the basis chosen;

(iii) we can extend L_1 to be half of the square of a metric $\bar{L}$ on $\mathfrak{g}$ such that $\bar{L} = L_1$ on $\mathcal{C}$. As we will see, the necessary conditions for an optimal control do not depend on how the extension is done.

(iv) For the affine case, we will simply set $\xi_u - e_0 = \sum_{i=1}^m u^i(t) e_i$.

Now it should be clear that the original problem is equivalent to the following constrained variational problem:

> **Constrained Variational Problem for Optimal Control on Lie Groups** Given an m-dimensional subspace $\mathcal{C}$ of $\mathfrak{g}$, find the optimal control curves $\xi - e_0 \in \mathcal{C}$ such that $g(0) = g_0$, $g(1) = g_1$ and minimize $\int_0^1 \bar{L}(\xi - e_0)dt$.

Since we want to use the method of Langrange multipliers to relax the constraint on the variations, we define a new Langrangian

$$\mathcal{L} = \bar{L}(\xi - e_0) + \lambda(t)(\xi - e_0) = \tilde{L}(\xi) + \tilde{\lambda}(t)(\xi) \tag{5.3.1}$$

where $\lambda(t)$ lies in the annihilator $\mathcal{C}^0$ of $\mathcal{C}$; furthermore $\tau(\xi) = \xi - e_0$, $\tilde{L} = \bar{L} \circ \tau$ and $\tilde{\lambda} = \lambda \circ \tau$.

Theorem 5.1 Optimization Theorem for Nonholonomic Systems on Lie Groups. *If ξ is a (regular) optimal control curve in $\mathcal{C} + e_0 = \{\xi \in \mathfrak{g} : \xi = \xi_c + e_0, \xi_c \in \mathcal{C}\}$, then there exists a $\lambda(t) \in \mathfrak{g}^*$ such that ξ satisfies the Euler-Poincare equation:*

$$\frac{d}{dt}\left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right) = \mathrm{ad}_\xi^*\left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right) \tag{5.3.2}$$

Proof If $\xi(t)$ is an optimal control curve in $\mathcal{C} + e_0$, then by the Lagrangian reduction method, $\xi(t)$ is a solution of the following variational problem

$$\delta \int_0^1 \mathcal{L}(\xi)dt = \delta \int_0^1 (\tilde{L}(\xi) + \tilde{\lambda}(\xi))dt = 0$$

for some $\lambda \in \mathfrak{g}^*$, where the variations take the form $\delta\xi = \dot{\Omega} + [\xi, \Omega]$ with $\Omega = g^{-1} \cdot \delta g$ arbitrary except vanishing at the endpoints. Since

$$\begin{aligned} 0 &= \delta \int_0^1 (\tilde{L}(\xi) + \tilde{\lambda}(\xi))dt \\ &= \int_0^1 \left(\frac{\delta \tilde{L}}{\delta \xi}\delta\xi + \lambda(\delta\xi)\right) dt \end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 \left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right) \left(\dot{\Omega} + [\xi, \Omega]\right) dt \\
&= \int_0^1 \left(-\frac{d}{dt}\left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right) + \mathrm{ad}_\xi^* \left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right)\right) \Omega dt,
\end{aligned}
$$

we conclude that $\bar{\xi}(t)$ satisfies

$$
\frac{d}{dt}\left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right) = \mathrm{ad}_\xi^* \left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right). \quad \blacksquare
$$

Corollary 5.2 *Given a left invariant control system on G, $\dot{g} = g \cdot \xi_u$ where*

$$
\xi_u = e_0 + \sum_{i=1}^m u^i(t) e_i .
$$

If $\bar{u}(\cdot)$ is an optimal control, then

$$
\bar{u}^i(t) = \frac{\mu_i(t)}{I_i}
$$

where $i = 1, \ldots, m$, and μ_i, $i = 1, \ldots, m$ is the solution of the following system of differential equations

$$
\dot{\mu}_i = C_{ji}^k \mu_k \xi_u^j
$$

where $i, j, k = 0, \ldots, n-1$, and where C_{ij}^k are the structure constants of $\mathfrak{g}$.

Proof Extend $\{e_0, e_1, \ldots, e_m\}$ to a basis $\{e_0, \ldots, e_{n-1}\}$ and let $\{e^0, \ldots, e^{n-1}\}$ be its dual basis.

(i) For $i = 1, \ldots, m$, and $\xi_u \in e_0 + \mathcal{C}$, we have

$$
\frac{\delta \tilde{L}}{\delta \xi_u^i} = \frac{\partial L}{\partial u^i} = I_i u^i
$$

because $\tilde{L}(\xi_u) = L \circ \phi \circ \tau(\xi_u) = L(u)$ and $\xi_u^i = u^i$; furthermore,

$$
\lambda_i = 0, \qquad i = 1, \ldots, m
$$

because λ lies in the annihilator $\mathcal{C}^0$.

(ii) If we set

$$
\mu_i = \frac{\delta \tilde{L}}{\delta \xi_u^i}, \; i = 1, \ldots, m,
$$

and

$$
\mu_i = \frac{\delta \tilde{L}}{\delta \xi_u^i} + \lambda_i, \; i = m+1, \ldots, n-1, 0,
$$

and write out the Euler-Poincare equation using the above coordinates, we will get the desired system of differential equations. $\blacksquare$

Remarks

1. From the above computations, we can see that the necessary conditions for an optimal control $\bar{u}(\cdot)$ depend only on L and have nothing to do with how the extension is done, because not only $u^i(t) = \mu_i(t)/I_i$, but also $\dot{\mu}_i = C^k_{ji}\mu_k\xi^j_u$ do not depend on $\bar{L}$.

2. The necessary conditions given in the above Corollary are the same as those in Krishnaprasad [1994]:

$$\begin{aligned} u^i &= \frac{\mu_i}{I_i} & i = 1, \ldots, m, \\ \dot{\mu}_i &= -\mu_k C^k_{ij}\frac{\delta h}{\delta \mu_j} & i, j, k = 0, \ldots, n-1, \end{aligned}$$

   where

$$h = \mu_0 + \frac{1}{2}\sum_{i=1}^{m}\frac{\mu_i^2}{I_i}.$$

   This is because $C^k_{ji} = -C^k_{ij}$ and

$$\frac{\delta h}{\delta \mu_j} = \left\{ \begin{array}{cl} 1 & j = 0 \\ \dfrac{\mu_j}{I_j} = u^j & j = 1, \ldots, m \\ 0 & j = m+1, \ldots, n-1 \end{array} \right\} = \xi^j_u$$

Conclusions

We have found a procedure based on reduced Lagrangian optimization that can be used to directly establish results on

1. optimal control for left invariant system on Lie group with velocity constraint,

2. optimal control for holonomic system on principal bundle with the constraint of the vanishing of the momentum map, and

3. optimal control for nonholonomic system on (trivial) principal bundles that may have a non-trivial evolution of its nonholonomic momentum.

In fact, the first two results can be seen as special cases of the last result even though we have derived each of them in a parallel way. Recall that in the nonholonomic case, we have

$$\mathcal{L} = C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}(r)\dot{r} - \Gamma(r)p \rangle + \left\langle \kappa(t), \dot{p} - \dot{r}^T H(r)\dot{r} - \dot{r}^T K(r)p - p^T D(r)p \right\rangle. \qquad (5.3.3)$$

In the driftless holonomic case, $\mathcal{D}_q = T_qQ$ for each $q \in Q$, the momentum is conserved and assumed to be zero, so the above Lagrangian $\mathcal{L}$ will be reduced to

$$\mathcal{L} = C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}(r)\dot{r} \rangle,$$

which is exactly the same Lagrangian used in the second case. As for system on Lie group G with velocity constraint (say, $g^{-1}\dot{g} = \sum_{i=1}^m u^i e_i$ for simplicity), it can be seen as system on (trivial) principal bundle $G \times \mathbb{R}^m$ whose (nonholonomic) connection is independent of the shape variable r, i.e.,

$$\xi^a = \mathcal{A}^a_\alpha \dot{r}^\alpha$$

where $\mathcal{A}^a_\alpha = 1$ and $\dot{r}^\alpha = u^\alpha$.

Topics for Future Work

1. In the nonholonomic case, we have only stated the result for the case of a trivial principal bundle. While it is true that all examples known to us have only trivial bundle structure, it is of interest to generalize the reduced Lagrangian optimization theorem to the case of an arbitrary principal bundle. Also, we need to understand better the geometry underlying this procedure. We hope to address all of these issues in a follow-up paper.

2. We need to construct algorithms that can effectively find approximate solutions to the system of differential equations that are obtained through reduced Lagrangian optimization. For example, finite element techniques appear to be appropriate and will be explored.

Acknowledgments

We thank Tony Bloch, P.S. Krishnaprasad, Naomi Leonard, Jim Ostrowski and Richard Montgomery for helpful comments on this paper.

References

Bloch, A.M. and P. Crouch [1992] On the dynamics and control of nonholonomic systems on Riemannian Manifolds. *Proceedings of NOLCOS '92*, Bordeaux, 368–372.

Bloch, A.M. and P. Crouch [1994] Nonholonomic control systems on Riemannian manifolds. *SIAM J. on Control* (to appear).

Bloch, A.M. and P. Crouch [1995a] Reduction of Euler Lagrange problems and relation with optimal control problems. *Proc 33rd CDC* (to appear).

Bloch, A.M. and P. Crouch [1995b] Reduction of Euler Lagrange problems for constrained variational problems and relation with optimal control problems. *33rd CDC, to appear.*

Bloch, A.M., P. Crouch and T.S. Ratiu [1994] Subriemannian optimal control problems. *Fields Inst. Comm.* **3**, 35–48.

Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and R. Murray [1995] Nonholonomic mechanical systems with symmetry. *Arch. Rat. Mech. An.* (to appear). CDS Technical Report 94-013, California Institute of Technology, URL http://avalon.caltech.edu/cds/reports/cds94-013.ps.

Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and T.S. Ratiu [1995] The Euler-Poincaré equations and double bracket dissipation. *Comm. Math. Phys.* (to appear).

Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and G. Sánchez de Alvarez [1992] Stabilization of rigid body dynamics by internal and external torques, *Automatica* **28**, 745–756.

Brockett, R.W. and L. Dai [1992] Nonholonomic kinematics and the role of elliptic functions in constructive controllability, in *Nonholonomic Motion Planning*, eds. Z. Li and J. F. Canny, Kluwer, 1–22, 1993.

Jurdjevic, V. [1993] The geometry of the plate-ball problem. *Arch. Rat. Mech. An.* **124**, 305–328.

Jurdjevic, V. [1994] Optimal control problems on Lie groups: crossroads between geometry and mechanics. *in Geometry and Feedback Control*, Marcel Dekker, (to appear).

Kelly, S.D. and R.M. Murray [1995] Geometric phases and robotic locomotion. *Journal of Robotic Systems* (to appear). Also available as Caltech technical report CIT/CDS 94-014.

Krishnaprasad, P.S. [1994] Optimal control and Poisson reduction, *Institute for System Research Technical Report*, **93-87**, 16 pages.

Krishnaprasad, P.S., R. Yang and W. Dayawansa [1991] Control problems on principal bundles and nonholonomic mechanics, *Proc. 30th CDC*, 1133–1138.

Marsden, J.E. [1992], *Lectures on Mechanics* London Mathematical Society Lecture Note Series. **174**, Cambridge University Press.

Marsden, J.E. and T.S. Ratiu [1992] *An Introduction to Mechanics and Symmetry.* Texts in Appl. Math. **17**, Springer-Verlag.

Marsden, J.E. and G. Sánchez de Alvarez [1993] Connections and feedback control. *In preparation*.

Marsden, J.E. and J. Scheurle [1993a] Lagrangian reduction and the double spherical pendulum, *ZAMP* **44**, 17–43.

Marsden, J.E. and J. Scheurle [1993b] The reduced Euler-Lagrange equations, *Fields Institute Comm.* **1**, 139–164.

Montgomery, R. [1984] Canonical formulations of a particle in a Yang-Mills field, *Lett. Math. Phys.* **8**, 59–67.

Montgomery, R. [1990] Isoholonomic problems and some applications, *Comm. Math Phys.* **128**, 565–592.

Montgomery, R. [1991] *Optimal Control of Deformable Bodies and Its Relation to Gauge Theory* in The Geometry of Hamiltonian Systems, T. Ratiu ed., Springer-Verlag.

Montgomery, R. [1993] Gauge theory of the falling cat, *Fields Inst. Comm.* **1**, 193–218.

Vershik, A.M. and V.Ya. Gershkovich [1994] Nonholonomic dynamical systems, geometry of distributions and variational problems, *Dynamical Systems VII, V. Arnold and S.P. Novikov, eds.*, 1–81. Springer-Verlag, New York.

Walsh, G., R. Montgomery and S. Sastry [1994] Optimal path planning on matrix Lie groups, *preprint*.

Yang, R. [1992] *Nonholonomic Geometry, Mechanics and Control.* Institute for Systmes Research Technical Report Ph.D. 92-14, Univ. of Maryland.