Reduction, Reconstruction and Optimal Control for Nonholonomic Mechanical Systems with Symmetry

by

Wang Sang Koon

B.A. (University of California, Berkeley) 1968
M.A. (University of California, Berkeley) 1972

A dissertation submitted in partial satisfaction of the requirements for the degree of
Doctor of Philosophy

in

Mathematics

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Jerrold E. Marsden, Chair
Professor Alan D. Weinstein
Professor S. Shankar Sastry

1997

The dissertation of Wang Sang Koon is approved:

Chair Date

Date

Date

University of California at Berkeley

1997

Reduction, Reconstruction and Optimal Control for Nonholonomic Mechanical Systems with Symmetry

Abstract

Reduction, Reconstruction and Optimal Control
for Nonholonomic Mechanical Systems with Symmetry

by

Wang Sang Koon
Doctor of Philosophy in Mathematics

University of California at Berkeley

Professor Jerrold E. Marsden, Chair

Many problems in robotics, dynamics of wheeled vehicles and motion generation, involve nonholonomic mechanics. Despite considerable advances in both Hamiltonian and Lagrangian sides of the theory, there remains much to be done, and this thesis makes contributions in three important areas.

First, we establish necessary conditions for optimal control using the ideas of Lagrangian reduction. The techniques developed here are designed for Lagrangian mechanical control systems with symmetry. Lagrangian reduction can do in one step what one can alternatively do by applying Pontryagin Maximum Principle followed by Poisson reduction. We apply the techniques to some known examples of optimal control on Lie groups and principal bundles. More importantly, we extend the method to the case of nonholonomic systems with a nontrivial momentum equation, such as the snakeboard.

Second, we compare the Hamiltonian (symplectic) approach to nonholonomic systems with Lagrangian approach. There are many differences between these approaches, and it was not obvious how they were equivalent. For example, Bloch, Krishnaprasad, Marsden and Murray [1996] developed the momentum equation, the reconstruction equation and the reduced Lagrange-d'Alembert equations, which are important for control applications, and it is not obvious how these correspond to the developments in Bates and Sniatycki [1993]. Our second result establishes specific links between these two sides and uses the ideas and results of each to shed light on the other, deepening our understanding of both approaches. We treat a simplified model of the bicycle and obtain new and interesting results.

We also develop the Poisson point of view for nonholonomic systems. Some of this theory has been started in van der Schaft and Maschke [1994]. In our third result, we develop the Poisson reduction for nonholonomic systems with symmetry, which enables us to obtain specific formulas for the Hamiltonian dynamics. Moreover, we show that the equations given by the Poisson reduction are equivalent to those given by the Lagrangian reduction.

We hope that these results will help lay a firm foundation for further developments of control, stability and bifurcation theories for such systems.

Professor Jerrold E. Marsden
Dissertation Committee Chair

To my parents

Soo Wu and Ziang Yuen Kuen

Contents

List of Figures

Acknowledgements

First, I would like to thank my parents for their sacrifices under extreme economic hardship to provide me an excellent education in Hong Kong when others in similar situation probably would have opted for trade school for their son.

I would also like to thank my wife, Marilyn, for her unwavering faith in me as I pursue my dreams.

To my two wonderful daughters, Jidan and Danfeng, for their constant reminder that a better world is in the making.

To Professor Jack Wagoner, for his crucial support in my readmission to UC Berkeley after an absence of some years.

To my advisor, Professor Jerrold Marsden, for his pivotal guidance and encouragement. He broadens my horizon and deepens my appreciation for an interdisciplinary approach to applied mathematics.

To Professors Alan Weinstein and Shankar Sastry who serve on my thesis committee, for their interest and timely discussions of my work.

To Professors Anthony Bloch, P.S. Krishnaprasad, Naomi Leonard, Richard Montgomery, Richard Murray, and Jim Ostrowski, for their valuable critique and suggestion during various stages of my research.

To the staff and researchers in the Control and Dynamical Systems at Caltech, for the wonderful opportunity to interact with a dynamic and interdisciplinary team where applied mathematics comes alive.

Also my special thanks to friends and colleagues, Hendra Adiwidjaja, Anthony Blaom, Francesco Bullo, Hung Wen Cheung, Sameer Jalnakupar, Scott Kelly, Andrew Lewis, Martin Lo, Matt Perlmutter, Muruhan Rathinam, Sebastian Shriver, Willem Sluis, Yong Wang, Jeffrey Wendlandt, Dong Yan, and Simon Yeung, for their help and friendship.

Chapter 1

Introduction

Many important problems in robotics, the dynamics of wheeled vehicles and motion generation, involve nonholonomic mechanics, which typically means mechanical systems with rolling constraints. Some of the important issues are trajectory tracking, dynamic stability and feedback stabilization (including nonminimum phase systems), bifurcation and control. Many of these systems have symmetry, such as the group of Euclidean motions in the plane or in space and this symmetry plays an important role in the theory.

Recently there have been considerable advances in the study of such systems from both the Hamiltonian and the Lagrangian sides of the theory. We refer to papers such as Weber [1986], Koiller [1992], Bloch and Crouch [1992], Krishnaprasad, Dayawansa and Yang [1992, 1993], Bates and Sniatycki [1993], van der Schaft and Maschke [1994], Herman [1995], Marle [1995], Ostrowski [1996] and Bloch, Krishnaprasad, Marsden and Murray [1996] and references therein.

While these and other references have made considerable progress in recent years, there remains much to be done. This thesis makes contributions in three important areas. First, we have developed a "reduced Lagrangian optimization" procedure to find the optimal controls for such systems. Second, our work has clarified the relation between the Hamiltonian reduction (from the symplectic viewpoint) and Lagrangian reduction of nonholonomic systems, which enables us to write the Hamiltonian dynamics of such systems as a reconstruction equation, a momentum equation and a set of reduced Hamilton equations. Third, we have developed the Poisson reduction for nonholonomic systems with symmetry, which enables us to obtain specific formulas for the Hamiltonian dynamics. We have also shown that the equations given by the Poisson reduction are equivalent to those given by the

Lagrangian reduction. All these help to lay a firm foundation for the further development of control, stability and bifurcation theories for such systems.

The basic setting is a configuration space Q with a distribution (usually nonintegrable) $\mathcal{D} \subset TQ$ describing the constraints of interest. For simplicity, we consider only homogeneous velocity constraints. We are given a Lagrangian L on TQ and a Lie group G acting on the configuration space and leaving the constraints and the Lagrangian invariant. In many example, the group encodes the position and orientation information. For example, for the snakeboard, the group is $SE(2)$ of rotations and translations in the plane. The quotient space Q/G is called the shape space.

1.1 Motivation and Results

1.1.1 Optimal Control and Lagrangian Reduction

Bloch, Krishnaprasad, Marsden and Murray [1996], hereafter denoted [BKMM], applied the method of geometric mechanics to the Lagrange-d'Alembert formulation, generalizing the use of connections and momentum maps associated with a given symmetry group to this case. With the help of the generalized momentum and the nonholonomic mechanical connection, [BKMM] provided a framework for studying the general form of nonholonomic mechanical control systems with symmetry that might have a nontrivial evolution of their nonholonomic momentum. The dynamics of such a system was 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

$$\left.\begin{aligned} g^{-1}\dot{g} &= -\mathcal{A}^{\mathrm{nh}}(r)\dot{r} + \Gamma(r)p \\ \dot{p} &= \dot{r}^T H(r)\dot{r} + \dot{r}^T K(r)p + p^T D(r)p \\ M(r)\ddot{r} &= \delta(r, \dot{r}, p) + \tau. \end{aligned}\right\} \tag{1.1}$$

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, and can be considered as the reconstruction equation for the group variables. The momentum equation describes the evolution of p and 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, and is assumed to affect only 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.

It is natural to ask how can we develop the optimal control of nonholonomic systems, such as the snakeboard, using this Lagrangian framework.

In this thesis, we have established necessary conditions for optimal control of both holonomic and nonholonomic systems, using the ideas of Lagrangian reduction. The techniques developed here are designed for Lagrangian mechanical control systems with symmetry and will be referred to as "reduced Lagrangian optimization" procedures. The benefit of such an approach is that it makes use of special structure of the system, especially its symmetry structure and thus it leads rather directly to the desired conclusions for such systems.

Reduced Lagrangian optimization can do in one step what one can alternatively do by applying the Pontryagin Maximum Principle followed by an application of Poisson reduction. We have applied it to some known examples, such as optimal control on Lie groups and principal bundles (such as the ball and plate problem) and reorientation example with zero angular momentum (such as the satellite with movable masses). More important, we have extended the method to the case of nonholonomic systems with a nontrivial momentum equation in the context of the work of [BKMM]. The snakeboard is used to illustrate the method.

1.1.2 Symplectic and Poisson Geometry of Nonholonomic Systems

Bates and Sniatycki [1993], hereafter denoted [BS], developed the Hamiltonian (symplectic geometry) side of nonholonomic systems, while [BKMM] has explored the Lagrangian side. It was not obvious how these two approaches were equivalent. For example, [BKMM] developed the momentum equation, the reduced Lagrange-d'Alembert equations and the reconstruction equation on the Lagrangian side, which are important for the control theory of these systems. It is not obvious how these correspond to the developments in [BS] on the Hamiltonian side. Moreover, the two approaches are conceptually very different, especially as the mechanics of the computations are so dissimilar, and the resulting differential equations appear very different.

Our second main result in this thesis establishes the specific links between these two sides and uses the ideas and results of each to shed light on the other, deepening our understanding of both points of view. For example, in proving the equivalence of the Lagrangian reduction and the symplectic reduction, we have shown where the momentum equation is lurking on the Hamiltonian side and how this is related to the breaking up the dynamics of nonholonomic systems with symmetry into three parts: a reconstruction equation for the group element g, an equation for the nonholonomic momentum p and the reduced Hamilton equations in the shape variables r, p_r (and p). This way of breaking up the dynamics may have the same significance for the control theory as what has already been noted in [BKMM]. Moreover, we have also explored the reduced Lagrange-d'Alembert equations in greater detail than was done previously.

We have illustrated the basic theory with the snakeboard, a well known example treated in [BKMM]. We have also treated a simplified model of the bicycle (introduced in Getz [1994] and Getz and Marsden [1995]) and have obtained some new and interesting results. This is an important prototype control system because it is an underactuated balance system.

On the Hamiltonian side, besides the symplectic point of view, one can also develop the Poisson point of view. Some of this theory has been started in van der Schaft and Maschke [1994], hereafter denoted [VM]. In our third main result in this thesis, we have built on the work of [VM] and have developed the Poisson reduction for nonholonomic systems with symmetry. We have used this Poisson reduction procedure to obtain specific formulas for the nonholonomic Hamiltonian dynamics and have shown that the equations given by Poisson reduction are equivalent to those given by Lagrangian reduction via a reduced constrained Legendre transform.

This Poisson reduction is important for the future development of the stability theory for nonholonomic mechanical systems with symmetry. In particular, it will be required for the development of the powerful block diagonalization properties of the energy-momentum method developed by Simo, Lewis and Marsden [1991]. This technique is very important for the development of systematic methods for stability analysis.

1.2 Outline of the Thesis

Chapter2 Optimal Control and Lagrangian Reduction

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 [1996]), 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 additional remarks.

Chapter 3 Symplectic Geometry of Nonholonomic Systems

General Nonholonomic Mechanical Systems. We first consider mechanics in the presence of homogeneous linear nonholonomic velocity constraints. No symmetry assumptions are made. In this section,

1. we recall the basic ideas and results of [BKMM] on general nonholonomic systems: in particular, how to describe constraints using Ehresmann connection and how to write the Lagrange d'Alembert equations of motion using the curvature of this connection.

2. We review the geometric structure of Hamiltonian systems with nonholonomic constraints in [BS], including a general procedure for finding the equations of motion for

nonholonomic systems from the symplectic point of view.

3. We construct the geometric objects on the Lagrangian side corresponding to those on the Hamiltonian side using the Legendre transformation in the context of nonholonomic constraints.

4. We prove that these dual procedures give us the same Lagrange d'Alembert equations as in [BKMM]. Since this proof is done in coordinates, it also provides a concrete coordinate based procedure for finding the equations of motion on the Hamiltonian side.

5. We use the symplectic procedure to work out the example of snakeboard taken from [BKMM].

Nonholonomic Mechanical Systems with Symmetry. Now we add the hypothesis of symmetry to the preceding development. In this section,

1. we recall the basic ideas and results of [BKMM] on simple nonholonomic mechanical systems, especially on how it extend the Lagrangian reduction theory of Marsden and Scheurle [1993a,b] to the context of nonholonomic systems. We shall describe briefly how [BKMM] modifies the Ehresmann connection associated with the constraints to a new connection, called the *nonholonomic connection*, that also takes into account the symmetries, and how the reduced equations, relative to this new connection, break up into *two* sets: a set of reduced Lagrange-d'Alembert equations, and a momentum equation. When the reconstruction equation is added, one recovers the full set of equations of motion for the system.

2. We summarize the symplectic reduction formulation of [BS] on finding the reduced equations of motion for nonholonomic systems with symmetry.

3. We restate the reduction procedure on the Lagrangian side corresponding to those on the Hamiltonian side using the Legendre transformation.

4. We prove that these dual procedures give us the same reduced Lagrange-d'Alembert equations as in [BKMM]. Since this proof is done in coordinates, it does provide a systematic way to carry out the computations on the Hamiltonian side. Also, the proof clarifies which construction in [BS] corresponds to the momentum equation of [BKMM]

and how this is related to breaking up the dynamics of the nonholonomic system into a reconstruction equation for a group element g, an equation for the nonholonomic momentum p and the reduced Hamilton equations in the shape variables r, p_r (and p).

5. We apply the symplectic reduction procedure to the examples of the snakeboard, the bicycle and a nonholonomically constrained particle.

Chapter 4 Poisson Geometry of Nonholonomic Systems

General Nonholonomic Mechanical Systems. As in Chapter 3, we first consider general nonholonomic systems without symmetry assumptions. In this section,

1. we review the Poisson formulation of nonholonomic systems in [VM], including a procedure for finding the equations of motion for nonholonomic systems from the Poisson point of view.

2. With the help of the Ehresmann connection, we use the Poisson procedure to write a compact formula for the equations of motion of the nonholonomic Hamiltonian dynamics.

3. We prove the equivalence of the Poisson and Lagrange-d'Alembert formulations for nonholonomic mechanics.

4. We apply the Poisson procedure to the example of the snakeboard.

Nonholonomic Mechanical Systems with Symmetry. Now we add the hypothesis of symmetry to the preceding development. In this section,

1. we build on the work of [VM] and develop the Poisson reduction, using the tools like the nonholonomic connection and nonholonomic momentum. We write the equations of motion for the reduced constrained Hamiltonian dynamics using a reduced Poisson bracket. This Poisson reduction procedure breaks the Hamiltonian nonholonomic dynamics into a reconstruction equation, a momentum equation and a set of reduced Hamilton equations.

2. We prove that the set of equations given by Poisson reduction is equivalent to those given by Lagrangian reduction via a reduced Legendre transform.

3. We apply the Poisson reduction procedure to the example of the snakeboard.

Chapter 5 Conclusions

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

Chapter 2

Optimal Control and Lagrangian Reduction

2.1 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.

In this chapter, we develop 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 chapter, 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 chapter are limited to presenting reduced Lagrangian optimization in the context of both holonomic and nonholonomic systems that may have conservation laws or nontrivial 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 chapter without explicit mention. Of course all of these points are of interest in themselves.

In the course of working on this chapter, 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 [1996] provides a useful framework for the present work.

The materials in this chapter has appeared in *SIAM Journal of Control and Optimization*, Volume 35, Number 3, May 1997, and as the article "Optimal Control for Holonomic and Nonholonomic Mechanical Systems with Symmetry and Lagrangian Reduction" by W.S. Koon and J.E. Marsden.

Outline of the Chapter

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 [1996]), 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 additional remarks.

2.2 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.

2.2.1 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

$$\langle J(q, v), \xi\rangle = \langle\langle v, \xi_Q(q)\rangle\rangle = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q)^i, \tag{2.1}$$

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 chapter).

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"

$$\mathcal{A}(q, v) = \mathbb{I}(q)^{-1} J(q, v). \tag{2.2}$$

When there is danger of confusion, we will write the mechanical connection as $\mathcal{A}^{\text{mec}}$ (additional connections will be introduced later in the chapter). 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

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

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

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

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 $\mathcal{A}(g \cdot v) = \text{Ad}_g \mathcal{A}(v)$, and $\mathcal{A}(\xi_Q(q)) = \xi$ for each $\xi \in \mathfrak{g}$. For example, if $Q = G$, there is a canonical 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

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

and

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

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

$$\mathcal{A}(r,g)(\dot{r},\dot{g}) = \text{Ad}_g(g^{-1}\dot{g} + \mathcal{A}_{\text{loc}}(r)\dot{r}). \tag{2.3}$$

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

$$\left.\begin{aligned} \dot{r} &= u \\ g^{-1}\dot{g} &= -\mathcal{A}_{\text{loc}}(r)u, \end{aligned}\right\} \tag{2.4}$$

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

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 $\dot{q}(t) \in \text{hor}_{q(t)}, q(0) = q_0, q(1) = q_1$, 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-elastically 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.

2.2.2 Simple Nonholonomic Mechanical Systems with Symmetry

Next, we recall some basic ideas and results from Bloch, Krishnaprasad, Marsden and Murray [1996] 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 $q \in Q$. We also assume throughtout the thesis that all the relevant spaces have constant dimensions.

As discussed in Bloch, Krishnaprasad, Marsden and Murray [1996], 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

$$\langle J^{\mathrm{nh}}(v_q), \xi \rangle = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q)^i, \tag{2.5}$$

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 [1996] 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:

$$\frac{d}{dt}\left(J^{\mathrm{nh}}(\xi^{q(t)})\right) = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}(\xi^{q(t)})\right]_Q^i. \tag{2.6}$$

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 $J^{\rm nh}_{\rm loc} : TQ/G \to (\mathfrak{g}^{\mathcal{D}})^*$ by

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

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

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 orthogonal, that is, their corresponding infinitesimal generators are orthogonal in the given kinetic energy metric. Keep in mind that the subspaces $\mathcal{D}_q$ and $T_q\mathrm{Orb}$ need not be orthogonal but here we are choosing a basis corresponding only to the subspace $T_q\mathrm{Orb}$. (Actually, all we need is that the vectors in the set of basis vectors corresponding to the subspace $\mathcal{S}_q$ be orthogonal to the remaining basis vectors.) 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

$$\left\langle J^{\rm nh}(r, g, \dot{r}, \dot{g}), e_b(r, g) \right\rangle = \left\langle \frac{\partial l}{\partial \xi}, e_b(r) \right\rangle := p_b, \tag{2.7}$$

which defines p_b, a function of r, $\dot{r}$ and ξ. It is proven in Bloch, Krishnaprasad, Marsden and Murray [1996] that in such an orthogonal body frame, the momentum equation can be written in the following form:

$$\dot{p} = \dot{r}^T H(r)\dot{r} + \dot{r}^T K(r)p + p^T D(r)p. \tag{2.8}$$

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 [1996] 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 [1996] develop formulas for $\mathcal{A}^{\mathrm{nh}}$ similar to those for the mechanical connection, namely

$$\mathcal{A}^{\mathrm{nh}}(v_q) = \mathbb{I}^{\mathrm{nh}}(q)^{-1}J^{\mathrm{nh}}(v_q) \tag{2.9}$$

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, (2.9) can be written as

$$\mathrm{Ad}_g(g^{-1}\dot{g} + \mathcal{A}^{\mathrm{nh}}_{\mathrm{loc}}(r)\dot{r}) = \mathrm{Ad}_g(\mathbb{I}^{\mathrm{nh}}_{\mathrm{loc}}(r)^{-1}p), \tag{2.10}$$

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 [1996] 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

$$\left.\begin{aligned} g^{-1}\dot{g} &= -\mathcal{A}^{\mathrm{nh}}(r)\dot{r} + \Gamma(r)p \\ \dot{p} &= \dot{r}^T H(r)\dot{r} + \dot{r}^T K(r)p + p^T D(r)p \\ M(r)\ddot{r} &= \delta(r, \dot{r}, p) + \tau, \end{aligned}\right\} \tag{2.11}$$

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.

2.3 Optimal Control and Lagrangian Reduction for Holonomic Systems

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

2.3.1 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 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 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:*

$$\frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} = \frac{\partial l}{\partial \Omega^a}\left(-\mathcal{B}^a_{\alpha\beta}\dot{r}^\beta + \mathcal{E}^a_{\alpha d}\Omega^d\right) \tag{2.12}$$

$$\frac{d}{dt}\frac{\partial l}{\partial \Omega^b} = \frac{\partial l}{\partial \Omega^a}(-\mathcal{E}^a_{\alpha b}\dot{r}^\alpha + C^a_{db}\Omega^d) \tag{2.13}$$

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):

$$\frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} = p_a\left(-\mathcal{B}^a_{\alpha\beta}\dot{r}^\beta + \mathcal{E}^a_{\alpha d}I^{de}p_e\right) - p_d\frac{\partial I^{de}}{\partial r^\alpha}p_e \tag{2.14}$$

$$\frac{d}{dt}p_b = p_a(-\mathcal{E}^a_{\alpha b}\dot{r}^\alpha + C^a_{db}I^{de}p_e), \tag{2.15}$$

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.

2.3.2 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 $\mathcal{L}$

$$\mathcal{L} = C(\dot{r}) + \langle\lambda(t), \xi + \mathcal{A}_{\mathrm{loc}}(r)\dot{r}\rangle \tag{2.16}$$

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

$$l = C(\dot{r}) + \langle\lambda(t), \xi + \mathcal{A}_{\mathrm{loc}}(r)\dot{r}\rangle \tag{2.17}$$

Theorem 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* $\xi = -\mathcal{A}_{\rm loc}(r)\dot{r}$, *as well as the reduced Euler Lagrange equations*

$$\frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} = 0 \tag{2.18}$$

$$\frac{d}{dt}\frac{\partial l}{\partial \xi^b} = \frac{\partial l}{\partial \xi^a} C^a_{db}\xi^d \tag{2.19}$$

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

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

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

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 $\mathcal{A}_{\rm loc}(r)\dot{r} = \mathcal{A}^a_\alpha \dot{r}^\alpha e_a$, 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.

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

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

Corollary 1 *Let the cost function* $C = \sum_1^{n-k} c_\alpha (u^\alpha)^2$. *be quadratic in* u. *If* $(r(t), g(t))$ *is a (regular) optimal trajectory with the control* $\bar{u}(t)$ *for the iso-holonomic (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, 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 (2.18) and (2.19) 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** 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).

2.3.3 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:

Iso-holonomic 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 $\mathcal{A}^{\rm mec}$, 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 4 *If $q(t)$ is a (regular) optimal trajectory for the iso-holonomic 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 trivialization by $(r^\alpha, \dot{r}^\alpha, \Omega^\alpha)$ satisfies the constraints $\xi^a = -(\mathcal{A}^{\rm mec})^a_\alpha \dot{r}^\alpha$ as well as the reduced Euler-Lagrange equations (2.12) and (2.13), where*

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

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 2 *In the preceding Theorem, if we use the mechanical connection $\mathcal{A}^{\rm mec}$ 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 $\mathcal{A}^{\rm mec}$ simply as $\mathcal{A}^b_\alpha$ and similarly for its curvature.

Proof Applying Theorem 4 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.

2.4 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.

2.4.1 The General Theorem for Optimization

In Bloch, Krishnaprasad, Marsden and Murray [1996], 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 $\mathcal{L}$:

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

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 $\mathcal{A}^{\text{nh}}$. 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.

Iso-holonomic 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 5 Reduced Lagrangian Optimization for the 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*

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

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. ■

2.4.2 The Optimality Conditions in Coordinates

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

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

we find after some computations that

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

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 5, 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.

2.5 Examples

2.5.1 Optimal Control of a Homogeneous Ball on a Rotating Plate

Bloch, Krishnaprasad, Marsden and Murray [1996] 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 [1996]) 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

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

where

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

and

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

Here, $r^1 = x, r^2 = y$ and e_1, e_2, e_3 is the standard basis of $so(3)_-$. Also, $\mathcal{A}_{\text{loc}}$ 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 [1996]).

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

$$\dot{g}g^{-1} = -\dot{y}e_1 + \dot{x}e_2 + ce_3 + \Omega xe_1 + \Omega ye_2,$$

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 negative Lie-Poisson structure because the right action is used).

By the preceding Theorem in the last section, we know that any of 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.

2.5.2 Optimal Control of the Snakeboard

The snakeboard is a modified version of a skate-board 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 [1996] 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 $\mathcal{D}_q$:

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

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

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

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

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

The momentum can be constructed by choosing a section of $\mathcal{D}\cap T\text{Orb}$ regarded as a bundle over Q. Since $\mathcal{D}_q \cap T_q\text{Orb}(q)$ 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

$$\begin{aligned} p &= J^{\mathrm{nh}}(\xi^q) = \frac{\partial L}{\partial \dot{q}^i}(\xi^q_Q)^i \\ &= ma\dot{x} + mb\dot{y} + Jc\dot{\theta} + J_0c(\dot{\theta} + \dot{\psi}) + J_1c(\dot{\theta} + \dot{\phi}_1) + J_2c(\dot{\theta} + \dot{\phi}_2). \end{aligned}$$

In Bloch, Krishnaprasad, Marsden and Murray [1996] a simplification is made which we shall also assume in this chapter, 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 $\phi = \pi/2$),

$$\begin{bmatrix} \cos(\theta)\dot{x} + \sin(\theta)\dot{y} \\ -\sin(\theta)\dot{x} + \cos(\theta)\dot{y} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} -\dfrac{J_0}{2mr}\sin(2\phi)\dot{\psi} \\ 0 \\ \dfrac{J_0}{mr^2}\sin^2(\phi)\dot{\psi} \end{bmatrix} = \begin{bmatrix} \dfrac{-1}{2mr}p \\ 0 \\ \dfrac{\tan\phi}{2mr^2}p \end{bmatrix}. \tag{2.20}$$

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

$$\begin{aligned}
\dot{p} &= \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}\xi^q\right]^i_Q \\
&= 4mr\cos(\theta)\cos(\phi)\sin(\phi)\dot{x}\dot{\phi} + 4mr\sin(\theta)\cos(\phi)\sin(\phi)\dot{y}\dot{\phi} \\
&\quad +2J_0\cos(2\phi)\dot{\phi}\dot{\psi} + 2mr^2\cos(2\phi)\dot{\theta}\dot{\phi} \\
&\quad -2mr\cos(\theta)\cos^2(\phi)\dot{y}\dot{\theta} + 2mr\sin(\theta)\cos^2(\phi)\dot{x}\dot{\theta}
\end{aligned}$$

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

$$\begin{aligned}
g^{-1}\dot{g} + \mathcal{A}_{\text{loc}}(r)\dot{r} &= \Gamma(r)p \\
\dot{p} = 2J_0\cos^2(\phi)\dot{\phi}\dot{\psi} &- \tan(\phi)p\dot{\phi}
\end{aligned}$$

where

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

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

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

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 5, we know that 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

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\dot\psi} &= \dot\psi - \frac{J_0}{2mr}\lambda_1\sin(2\phi) + \frac{J_0}{mr^2}\lambda_3\sin^2(\phi) - 2J_0\kappa\cos^2(\phi)\dot\phi \\ \frac{\partial\mathcal{L}}{\partial\psi} &= 0 \\ \frac{\partial\mathcal{L}}{\partial\dot\phi} &= \dot\phi - 2J_0\kappa\cos^2(\phi)\dot\psi + \kappa\tan(\phi)p \\ \frac{\partial\mathcal{L}}{\partial\phi} &= -\frac{J_0}{mr}\lambda_1\cos(2\phi)\dot\psi + \frac{J_0}{mr^2}\lambda_3\sin(2\phi)\dot\psi - \frac{1}{2mr^2}\lambda_3\sec^2(\phi)p \\ &\quad +2J_0\kappa\sin(2\phi)\dot\phi\dot\psi + \kappa\sec^2(\phi)p\dot\phi \\ \frac{\partial\mathcal{L}}{\partial\dot p} &= \kappa \\ \frac{\partial\mathcal{L}}{\partial p} &= \frac{1}{2mr}\lambda_1 - \frac{1}{2mr^2}\lambda_3\tan(\phi) + \kappa\tan(\phi)\dot\phi \\ \frac{\partial\mathcal{L}}{\partial\xi^b} &= \lambda_b.\end{aligned}$$

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

$$\ddot{\psi} - \frac{J_0}{2mr}\lambda_1(1 + 3\cos(2\phi))\dot{\phi} + \frac{3J_0}{2mr^2}\lambda_3 \sin(2\phi)\dot{\phi} + J_0\kappa \sin(2\phi)(\dot{\phi})^2 - 2J_0\kappa \cos^2(\phi)\ddot{\phi} = 0$$

and

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

2.5.3 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 $\{e_0, e_1, \dots, e_m\}$ 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 Lagrange multipliers to relax the constraint on the variations, we define a new Lagrangian

$$\mathcal{L} = \bar{L}(\xi - e_0) + \lambda(t)(\xi - e_0) = \tilde{L}(\xi) + \tilde{\lambda}(t)(\xi) \tag{2.21}$$

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 6 Optimization Theorem for Nonholonomic Systems on Lie Groups. *If $\bar{\xi}$ 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 $\bar{\xi}$ satisfies the Euler-Poincare equation:*

$$\frac{d}{dt}\left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right) = \mathrm{ad}^*_{\xi}\left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda\right) \tag{2.22}$$

Proof If $\bar{\xi}(t)$ is an optimal control curve in $\mathcal{C} + e_0$, then by the Lagrangian reduction method, $\bar{\xi}(t)$ 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

$$\begin{aligned} 0 &= \delta \int_0^1 (\tilde{L}(\xi) + \tilde{\lambda}(\xi))dt \\ &= \int_0^1 \left(\frac{\delta \tilde{L}}{\delta \xi} \delta\xi + \lambda\left(\delta\xi\right) \right) dt \\ &= \int_0^1 \left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda \right) \left(\dot{\Omega} + [\xi, \Omega] \right) dt \\ &= \int_0^1 \left(-\frac{d}{dt} \left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda \right) + \mathrm{ad}_\xi^* \left(\frac{\delta \tilde{L}}{\delta \xi} + \lambda \right) \right) \Omega dt, \end{aligned}$$

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 3 *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 $i = 1, \dots, m$, and μ_i, $i = 1, \dots, m$ is the solution of the following system of differential equations

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

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

Proof Extend $\{e_0, e_1, \dots, e_m\}$ to a basis $\{e_0, \dots, e_{n-1}\}$ and let $\{e^0, \dots, e^{n-1}\}$ 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. ■

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 $\dot{\mu}_i = C_{ji}^k \mu_k \xi_u^j$ do not depend on $\bar{L}$.

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

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

where

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

This is because $C_{ji}^k = -C_{ij}^k$ and

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

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 nontrivial 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

$$\mathcal{L} = C(\dot{r}) + \langle \lambda(t), \xi + \mathcal{A}(r)\dot{r} - \Gamma(r)p \rangle + \left\langle \kappa(t), \dot{p} - \dot{r}^T H(r)\dot{r} - \dot{r}^T K(r)p - p^T D(r)p \right\rangle . \tag{2.23}$$

In the drift-less 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 $\dot{r}^\alpha = u^\alpha$.

Chapter 3

Symplectic Geometry of Nonholonomic Systems

3.1 Introduction

The General Setting. Many important problems in robotics, the dynamics of wheeled vehicles and motion generation, involve nonholonomic mechanics, which typically means mechanical systems with rolling constraints. Some of the important issues are trajectory tracking, dynamic stability and feedback stabilization (including non-minimum phase systems), bifurcation and control. Many of these systems have symmetry, such as the group of Euclidean motions in the plane or in space and this symmetry plays an important role in the theory.

In the last several years, several basic works have been done on both the Hamiltonian and the Lagrangian sides of the theory. Papers like Weber [1986], Koiller [1992], Bloch and Crouch [1992], Krishnaprasad, Dayawansa and Yang [1992, 1993], Bates and Sniatycki [1993], van der Schaft and Maschke [1994], Hermans [1995], Marle [1995],Ostrowski [1996] and Bloch, Krishnaprasad, Marsden and Murray [1996] have laid a firm foundation for understanding nonholonomic mechanical systems with symmetry.

Bates and Sniatycki [1993], hereafter denoted [BS], developed the Hamiltonian side, while Bloch, Krishnaprasad, Marsden and Murray [1996], hereafter denoted [BKMM], has explored the Lagrangian side. It was not obvious how these two approaches were equivalent because, for example, [BKMM] developed the momentum equation and the reduced

Lagrange-d'Alembert equations and it is not obvious how these correspond to the developments in [BS]. Our aim is to establish links between these two sides and use the ideas and results of each to shed light on the other, with the goal of deepening our understanding of both points of view. We hope that it will aid related efforts such as extending the results of energy-momentum method for stability of relative equilibria and the theory of Hamiltonian bifurcations to nonholonomic mechanics. In the spirit of [BKMM], we do many of the calculations in coordinates to help in the study of examples.

We illustrate the basic theory with the snakeboard, the well known example treated in [BKMM]. We also treat a simplified model of the bicycle (introduced in Getz [1994] and Getz and Marsden [1995]) and obtain results that were not known previously. This is an important prototype control system because it is an under-actuated balance system.

Outline of the Chapter. We begin in §3.2 by recalling some of the main results of [BKMM] and of [BS] in the general context of nonholonomically constrained systems. In that section, we establish the precise link between them. The snakeboard example is begun in this section.

In §3.3, we treat systems with symmetry and study the momentum equation, the reconstruction equation and the reduced Lagrange-d'Alembert equations from both the Hamiltonian and the Lagrangian points of view. This clarifies which construction in [BS] corresponds to the momentum equation of [BKMM]. This section also continues the snakeboard example and treats the bicycle.

Summary of the Main Results. The main results of the present work are as follows:

- The precise relation between the constructions in the papers [BS] and [BKMM] are given.
- The reduced Lagrange-d'Alembert equations established in [BKMM] are shown to be equivalent to the reduced nonholonomic Hamilton equations implicitly given in [BS].
- The relation between the constructions is illustrated with two specific examples, the snakeboard and the nonholonomically constrained particle.
- A simplified model of the bicycle is treated.

3.2 General Nonholonomic Mechanical Systems

Following the approaches of both [BS] and [BKMM], we first consider mechanics in the presence of homogeneous linear nonholonomic velocity constraints. For now, no symmetry assumptions are made; we add such assumptions in the following sections.

In this section,

1. we recall the basic ideas and results from [BKMM] on general nonholonomic systems: in particular, how to describe constraints using Ehresmann connections and how to write the Lagrange-d'Alembert equations of motion using the curvature of this connection.

2. We review the geometric structure of Hamiltonian systems with nonholonomic constraints from [BS], including a general procedure for finding the equations of motion for nonholonomic systems from the Hamiltonian point of view.

3. We construct the geometric objects on the Lagrangian side corresponding to those on the Hamiltonian side using the Legendre transformation in the context of nonholonomic constraints.

4. We prove that these dual procedures gives us the same Lagrange d'Alembert equations as in [BKMM]. Since this proof is done in coordinates, it also provides a concrete coordinate based procedure for finding the equations of motion on the Hamiltonian side.

5. We will use the Hamiltonian procedure to work out the example of snakeboard taken from [BKMM].

3.2.1 Review of the Lagrangian Approach

We start with a configuration space Q with local coordinates denoted $q^i, i = 1, \ldots, n$ and a distribution $\mathcal{D}$ on Q that describes the kinematic nonholonomic constraints. The distribution is given by the specification of a linear subspace $\mathcal{D}_q \subset T_qQ$ of the tangent space to Q at each point $q \in Q$.

In this chapter we consider only homogeneous velocity constraints. The extension to affine constraints is straightforward, as in [BKMM].

The dynamics of a nonholonomically constrained mechanical system is governed by the Lagrange-d'Alembert principle. The principle states that 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 and that the velocity of the curve $q(t)$ itself satisfies the constraints; that is, $\dot{q}(t) \in \mathcal{D}_{q(t)}$. Standard arguments in the calculus of variations show that this "constrained variational principle" is equivalent to the equations

$$-\delta L := \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right) \delta q^i = 0, \tag{3.1}$$

for all variations δq such that $\delta q \in \mathcal{D}_q$ at each point of the underlying curve $q(t)$. These equations are often equivalently written as

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \lambda_i,$$

where λ_i is a set of Lagrange multipliers ($i = 1, \ldots, n$), representing the force of constraint. Intrinsically, this multiplier λ is a section of the cotangent bundle over $q(t)$ that annihilates the constraint distribution. The Lagrange multipliers are often determined by using the condition that $\dot{q}(t)$ lies in the distribution.

In Bloch and Crouch [1992] and Lewis [1996], the Lagrange-d'Alembert equations are shown to have the form of a generalized acceleration condition

$$\nabla_{\dot{q}} \dot{q} = 0$$

for a suitable affine connection on Q and the force of constraint λ is interpreted as a generalized second fundamental form (as is well known for systems with holonomic constraints; see Abraham and Marsden [1978], for example). In this form of the equations, one can add external forces directly to the right hand sides so that the equations now become in the form of a generalized Newton law. This form is convenient for control purposes.

To explore the structure of the Lagrange-d'Alembert equations in more detail, let $\{\omega^a\}$, $a = 1, \ldots, k$ be a set of k independent one forms whose vanishing describes the constraints; i.e., the distribution $\mathcal{D}$. One can introduce local coordinates $q^i = (r^\alpha, s^a)$ where $\alpha = 1, \ldots n - k$, in which ω^a has the form

$$\omega^a(q) = ds^a + A^a_\alpha(r, s) dr^\alpha$$

where the summation convention is in force. In other words, we are locally writing the distribution as

$$\mathcal{D} = \{(r, s, \dot{r}, \dot{s}) \in TQ \mid \dot{s} + A^a_\alpha \dot{r}^\alpha = 0\}.$$

The equations of motion, (3.1) may be rewritten by noting that the allowed variations $\delta q^i = (\delta r^\alpha, \delta s^a)$ satisfy $\delta s^a + A^a_\alpha \delta r^\alpha = 0$. Substitution into (3.1) gives

$$\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{r}^\alpha} - \frac{\partial L}{\partial r^\alpha}\right) = A^a_\alpha \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a}\right). \tag{3.2}$$

Equation (3.2) combined with the constraint equations

$$\dot{s}^a = -A^a_\alpha \dot{r}^\alpha \tag{3.3}$$

gives a complete description of the equations of motion of the system; this procedure may be viewed as one way of eliminating the Lagrange multipliers. Using this notation, one finds that $\lambda = \lambda_a \omega^a$, where $\lambda_a = \frac{d}{dt}\frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a}$.

Equations (3.2) can be written in the following way:

$$\frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot{s}^b} B^b_{\alpha\beta} \dot{r}^\beta, \tag{3.4}$$

where

$$L_c(r^\alpha, s^a, \dot{r}^\alpha) = L(r^\alpha, s^a, \dot{r}^\alpha, -A^a_\alpha(r, s)\dot{r}^\alpha).$$

is the coordinate expression of the constrained Lagrangian defined by $L_c = L|\mathcal{D}$ and where

$$B^b_{\alpha\beta} = \left(\frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial A^b_\beta}{\partial s^a} - A^a_\beta \frac{\partial A^b_\alpha}{\partial s^a}\right). \tag{3.5}$$

Letting $d\omega^b$ be the exterior derivative of ω^b, a computation shows that

$$d\omega^b(\dot{q}, \cdot) = B^b_{\alpha\beta} \dot{r}^\alpha dr^\beta$$

and hence the equations of motion have the form

$$-\delta L_c = \left(\frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a}\right)\delta r^\alpha = -\frac{\partial L}{\partial \dot{s}^b} d\omega^b(\dot{q}, \delta r).$$

This form of the equations isolates the effects of the constraints, and shows, in particular, that in the case where the constraints are integrable (*i.e.,* $\mathbf{d}\omega = 0$), the equations of motion are obtained by substituting the constraints into the Lagrangian and then setting the variation of L_c to zero. However in the non-integrable case the constraints generate extra (curvature) terms, which must be taken into account.

The above coordinate results can be put into an interesting and useful intrinsic geometric framework. The intrinsically given information is the distribution and the Lagrangian. *Assume* that there is a bundle structure $\pi_{Q,R} : Q \to R$ for our space Q, where R is the *base* manifold and $\pi_{Q,R}$ is a submersion and the kernel of $T_q\pi_{Q,R}$ at any point $q \in Q$ is called the *vertical space* V_q. One can always do this locally. An *Ehresmann connection* A is a vertical valued one form on Q such that

1. $A_q : T_qQ \to V_q$ is a linear map and
2. A is a projection: $A(v_q) = v_q$ for all $v_q \in V_q$.

Hence, $T_qQ = V_q \oplus H_q$ where $H_q = \ker A_q$ is the *horizontal space at* q, sometimes denoted hor_q. Thus, an Ehresmann connection gives us a way to split the tangent space to Q at each point into a horizontal and vertical part.

If the Ehresmann connection is chosen in such a way that the given constraint distribution $\mathcal{D}$ is the horizontal space of the connection; that is, $H_q = \mathcal{D}_q$, then in the bundle coordinates $q^i = (r^\alpha, s^a)$, the map $\pi_{Q,R}$ is just projection onto the factor r and the connection A can be represented locally by a vector valued differential form ω^a:

$$A = \omega^a \frac{\partial}{\partial s^a}, \qquad \omega^a(q) = ds^a + A^a_\alpha(r,s)dr^\alpha,$$

and the horizontal projection is the map

$$(\dot{r}^\alpha, \dot{s}^a) \mapsto (\dot{r}^\alpha, -A^a_\alpha(r,s)\dot{r}^\alpha).$$

The curvature of an Ehresmann connection A is the vertical valued two form defined by its action on two vector fields X and Y on Q as

$$B(X,Y) = -A([\mathrm{hor}\, X, \mathrm{hor}\, Y])$$

where the bracket on the right hand side is the Jacobi-Lie bracket of vector fields obtained by extending the stated vectors to vector fields. This definition shows the sense in which the curvature measures the failure of the constraint distribution to be integrable.

In coordinates, one can evaluate the curvature B of the connection A by the following formula:

$$B(X,Y) = \mathbf{d}\omega^a(\mathrm{hor}\, X, \mathrm{hor}\, Y)\frac{\partial}{\partial s^a},$$

so that the local expression for curvature is given by

$$B(X,Y)^a = B^a_{\alpha\beta}X^\alpha Y^\beta$$

where the coefficients $B^a_{\alpha\beta}$ are given by (3.5).

The Lagrange d'Alembert equations may be written intrinsically as

$$\delta L_c = \langle \mathbb{F}L, B(\dot{q}, \delta q)\rangle,$$

in which δq is a horizontal variation (*i.e.,* it takes values in the horizontal space) and B is the curvature regarded as a vertical valued two form, in addition to the constraint equations

$$A(q) \cdot \dot{q} = 0.$$

Here $\langle \, , \rangle$ denotes the pairing between a vector and a dual vector and

$$\delta L_c = \left\langle \delta r^\alpha, \frac{\partial L_c}{\partial r^\alpha} - \frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - A^a_\alpha \frac{\partial L_c}{\partial s^a} \right\rangle.$$

As shown in [BKMM], when there is a symmetry group G present, there is a natural bundle one can work with and put a connection on, namely the bundle $Q \to Q/G$. In the generality of the preceding discussion, one can get away with just the distribution itself and can introduce the corresponding Ehresmann connection locally. In fact, the bundle structure $Q \to R$ is really a "red herring". The notion of curvature as a $T_qQ/\mathcal{D}_q$ valued form makes good sense and is given locally by the same expressions as above. However, keeping in mind that we eventually want to deal with symmetries and in that case there is a natural bundle, the Ehresmann assumption is nevertheless a reasonable bridge to the more interesting case with symmetries.

3.2.2 Review of the Hamiltonian Formulation

The approach of [BS] starts on the Lagrangian side with a configuration space Q and a Lagrangian L of the form kinetic energy minus potential energy, *i.e.,*

$$L(q, \dot{q}) = \frac{1}{2}\langle\langle \dot{q}, \dot{q}\rangle\rangle - V(q),$$

where $\langle\langle\ ,\ \rangle\rangle$ is a metric on Q defining the kinetic energy and V is a potential energy function. We do not restrict ourselves to Lagrangians of this form.

As above, our nonholonomic constraints are given by a distribution $\mathcal{D} \subset TQ$. We also let $\mathcal{D}^{\mathrm{o}} \subset T^*Q$ denote the annihilator of this distribution.

As above, the basic equations are given by the Lagrange-d'Alembert principle.

The Legendre transformation $\mathbb{F}L : TQ \to T^*Q$, assuming that it is a diffeomorphism, is used to define the Hamiltonian $H : T^*Q \to \mathbb{R}$ in the standard fashion (ignoring the constraints for the moment):

$$H = \langle p, \dot{q} \rangle - L = p_i \dot{q}^i - L.$$

Here, the momentum is $p = \mathbb{F}L(v_q) = \partial L / \partial \dot{q}$. Under this change of variables, the equations of motion are written in the Hamiltonian form as

$$\begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i}. \\ \dot{p}_i &= -\frac{\partial H}{\partial q^i} + \lambda_a \omega^a_i, \end{aligned}$$

where $i = 1, \dots, n$, together with the constraint equations.

The preceding Hamiltonian equations can be rewritten as

$$X \mathbin{\lrcorner} \Omega = dH + \lambda_a \pi_Q^* \omega^a,$$

where X is the vector field on T^*Q governing the dynamics, Ω is the canonical symplectic form on T^*Q, and $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection. We may write X in coordinates as $X = \dot{q}^i \partial_{q^i} + \dot{p}_i \partial_{p_i}$.

On Lagrangian side, we saw that one can get rid of the Lagrange multipliers. On the Hamiltonian side, it is also desirable to model the Hamiltonian equations without the Lagrange multipliers by a vector field on a submanifold of T^*Q. We do this in what follows.

First of all, we define the set $\mathcal{M} = \mathbb{F}L(\mathcal{D}) \subset T^*Q$, so that the constraints on the Hamiltonian side are given by

$$p \in \mathcal{M}.$$

Besides $\mathcal{M}$, another basic object we deal with is defined as

$$\mathcal{F} = (T\pi_Q)^{-1}(\mathcal{D}) \subset TT^*Q.$$

Using a basis ω^a of the annihilator $\mathcal{D}^{\mathrm{o}}$, we can write these spaces as

$$\mathcal{M} = \{p \in T^*Q \mid \omega^a((\mathbb{F}L)^{-1}(p)) = 0\}, \tag{3.6}$$

and

$$\mathcal{F} = \{u \in TT^*Q \mid \langle \pi_Q^* \omega^a, u \rangle = 0\}. \tag{3.7}$$

Finally, we define

$$\mathcal{H} = \mathcal{F} \cap T\mathcal{M}.$$

Using natural coordinates $(q^i, p_i, \dot{q}^i, \dot{p}_i)$ on TT^*Q, we see that the distribution $\mathcal{F}$ naturally lifts the constraint on $\dot{q}$ from TQ to TT^*Q. On the other hand, the space $\mathcal{M}$ puts the associated constraints on the variable p and therefore the intersection $\mathcal{H}$ puts the constraints on both variables.

To eliminate the Lagrange multipliers, we regard the Hamiltonian equations as a vector field on the constraint submanifold $\mathcal{M} \subset T^*Q$ which takes values in the constraint distribution $\mathcal{H}$. Next we recall from [BS] how to construct these equations intrinsically using the ideas of symplectic geometry.

A result of [BS] is that $\Omega_{\mathcal{H}}$, the restriction of the canonical two-form Ω of T^*Q fiberwise to the distribution $\mathcal{H}$ of the constraint submanifold $\mathcal{M}$, is nondegenerate. Note that $\Omega_{\mathcal{H}}$ is not a true two form on a manifold, so it does not make sense to speak about it being closed. We speak of it as a fiber-restricted two form to avoid any confusion. Of course it still makes sense to talk about it being nondegenerate; it just means nondegenerate as a bilinear form on each fiber of $\mathcal{H}$. The dynamics is then given by the vector field $X_{\mathcal{H}}$ on $\mathcal{M}$ which takes values in the constraint distribution $\mathcal{H}$ and is determined by the condition

$$X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{H}} = dH_{\mathcal{H}} \tag{3.8}$$

where $dH_{\mathcal{H}}$ is the restriction of $dH_{\mathcal{M}}$ to $\mathcal{H}$. We will be exploring the coordinate meaning of this condition and its comparison with the Lagrangian formulation in the subsequent sections.

3.2.3 Lagrangian Side

We now construct the geometric structures on the tangent bundle TQ corresponding to those on the Hamiltonian side from the preceding subsection and formulate a similar procedure for obtaining the equations of motion. By doing this, it will be easier to made comparison with the geometric constructions and analytic formulations in [BKMM].

First of all, we can define the energy function E simply as $E = H \circ \mathbb{F}L$ and pull back to TQ the canonical two-form on T^*Q and denote it by Ω_L.

We define the distribution

$$\mathcal{C} = (T\tau_Q)^{-1}(\mathcal{D}) \subset TTQ,$$

where $\tau_Q : TQ \to Q$. In coordinates, the distribution $\mathcal{C}$ consists of vectors annihilated by the form $\tau_Q^*\omega^a$:

$$\mathcal{C} = \{u \in TTQ \mid \langle \tau_Q^*\omega^a, u \rangle = 0\}. \tag{3.9}$$

When $\mathcal{C}$ is restricted to the constraint submanifold $\mathcal{D} \subset TQ$, we obtain the constraint distribution $\mathcal{K}$:

$$\mathcal{K} = \mathcal{C} \cap T\mathcal{D}. \tag{3.10}$$

Clearly $\mathcal{M} = \mathbb{F}L(\mathcal{D})$ and $\mathcal{H} = T\mathbb{F}L(\mathcal{K})$.

The dynamics is given by a vector field $X_\mathcal{K}$ on the manifold $\mathcal{D}$ which takes values in $\mathcal{K}$ and satisfies the equation

$$X_\mathcal{K} \,\lrcorner\, \Omega_\mathcal{K} = dE_\mathcal{K}, \tag{3.11}$$

where $dE_\mathcal{K}$ and $\Omega_\mathcal{K}$ are the restrictions of $dE_\mathcal{D}$ and $\Omega_\mathcal{D}$ respectively to the distribution $\mathcal{K}$ and where $E_\mathcal{D}$ and $\Omega_\mathcal{D}$ are the restrictions of E and Ω_L to $\mathcal{D}$.

3.2.4 The equivalence of the Hamiltonian and the Lagrange-d'Alembert formulations

The Lagrangian procedure on TQ formulated in the preceding subsection acts as a bridge between [BS] and [BKMM]. We can show the correctness of the Lagrangian procedure given above by (carefully) invoking the results of [BS] (generalized to arbitrary Lagrangians and with some gaps filled in), or by checking the methods against the results of [BKMM]. We choose the latter method.

Theorem 7 *Consider a configuration space Q, a hyper-regular Lagrangian L and a distribution $\mathcal{D}$ that describes the kinematic nonholonomic constraints. The $\mathcal{K}$-valued vector field*

$X_{\mathcal{K}}$ on $\mathcal{D}$ given by the equation

$$X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{K}} = dE_{\mathcal{K}}$$

defines dynamics that are equivalent to the Lagrange-d'Alembert equations together with the constraints.

Proof To keep things concrete and to provide additional insight, we shall give a coordinate based proof. Introduce local coordinates $(r^\alpha, s^a, \dot{r}^\alpha, \dot{s}^a)$ for TQ as described earlier. Local coordinates for the manifold $\mathcal{D}$ are given by $(r^\alpha, s^a, \dot{r}^\alpha)$.

Let us first compute $dE_{\mathcal{D}}$ and $X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{D}}$. We claim that

$$E_{\mathcal{D}} = \frac{\partial L_c}{\partial \dot{r}^\beta}\dot{r}^\beta - L_c,$$

where $L_c = L|\mathcal{D}$ is the constrained Lagrangian. This is because $E = \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L$ and so restricting it to $\mathcal{D}$ we get

$$\begin{aligned} E_{\mathcal{D}} &= \left(\frac{\partial L}{\partial \dot{r}^\alpha}\dot{r}^\alpha + \frac{\partial L}{\partial \dot{s}^a}\dot{s}^a\right)\Bigg|_{\mathcal{D}} - L_c \\ &= \frac{\partial L_c}{\partial \dot{r}^\alpha}\dot{r}^\alpha + A^a_\alpha \frac{\partial L}{\partial \dot{s}^a}\dot{r}^\alpha - A^a_\alpha \frac{\partial L}{\partial \dot{s}^a}\dot{r}^\alpha - L_c \\ &= \frac{\partial L_c}{\partial \dot{r}^\alpha}\dot{r}^\alpha - L_c. \end{aligned}$$

The differential of $E_{\mathcal{D}}$ is then computed to be

$$\begin{aligned} dE_{\mathcal{D}} &= \dot{r}^\beta \frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} dr^\alpha + \dot{r}^\beta \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha + \dot{r}^\beta \frac{\partial^2 L_c}{\partial s^b \partial \dot{r}^\beta} ds^b - \frac{\partial L_c}{\partial r^\alpha} dr^\alpha - \frac{\partial L_c}{\partial s^b} ds^b \\ &= \dot{r}^\beta \frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} dr^\alpha + \dot{r}^\beta \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha - \dot{r}^\beta A^b_\alpha \frac{\partial^2 L_c}{\partial s^b \partial \dot{r}^\beta} dr^\alpha + \dot{r}^\beta \frac{\partial^2 L_c}{\partial s^b \partial \dot{r}^\beta}(ds^b + A^b_\alpha dr^\alpha) \\ &\quad - \frac{\partial L_c}{\partial r^\alpha} dr^\alpha + A^b_\alpha \frac{\partial L_c}{\partial s^b} dr^\alpha - \frac{\partial L_c}{\partial s^b}(ds^b + A^b_\alpha dr^\alpha) \end{aligned}$$

As for $\Omega_{\mathcal{D}}$, we have

$$\begin{aligned}
\Omega_{\mathcal{D}} &= -d\left(\left.\frac{\partial L}{\partial \dot{q}^i}\right|_{\mathcal{D}} dq^i\right) \\
&= -d\left\{\left(\frac{\partial L_c}{\partial \dot{r}^\beta} + A^b_\beta \frac{\partial L}{\partial \dot{s}^b}\right) dr^\beta + \frac{\partial L}{\partial \dot{s}^b} ds^b\right\} \\
&= -d\left(\frac{\partial L_c}{\partial \dot{r}^\beta}\right) \wedge dr^\beta - d\left(A^b_\beta \frac{\partial L}{\partial \dot{s}^b}\right) \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge ds^b \\
&= -d\left(\frac{\partial L_c}{\partial \dot{r}^\beta}\right) \wedge dr^\beta - d\left(A^b_\beta \frac{\partial L}{\partial \dot{s}^b}\right) \wedge dr^\beta \\
&\quad + A^b_\beta d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge (ds^b + A^b_\beta dr^\beta) \\
&= -d\left(\frac{\partial L_c}{\partial \dot{r}^\beta}\right) \wedge dr^\beta - \frac{\partial L}{\partial \dot{s}^b} d(A^b_\beta) \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge (ds^b + A^b_\beta dr^\beta) \\
&= -\frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} dr^\alpha \wedge dr^\beta - \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha \wedge dr^\beta - \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^\beta} ds^a \wedge dr^\beta \\
&\quad - \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_\beta}{\partial r^\alpha} dr^\alpha \wedge dr^\beta - \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_\beta}{\partial s^a} ds^a \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge (ds^b + A^b_\beta dr^\beta) \\
&= \left(-\frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} + A^a_\alpha \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^\beta} - \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_\beta}{\partial s^a}\right) dr^\alpha \wedge dr^\beta \\
&\quad - \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha \wedge dr^\beta - \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^\beta} (ds^a + A^a_\alpha dr^\alpha) \wedge dr^\beta \\
&\quad - \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_\beta}{\partial s^a} (ds^a + A^a_\alpha dr^\alpha) \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge (ds^b + A^b_\beta dr^\beta),
\end{aligned}$$

where there is a sum on all α, β.

Now we are ready to find the equations of motion. Any vector field $X_{\mathcal{D}}$ on $\mathcal{D}$ has the following coordinate form:

$$X_{\mathcal{D}} = \dot{r}^\alpha \partial_{r^\alpha} + \dot{s}^a \partial_{s^a} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha}.$$

Since $X_{\mathcal{K}}$ lies in the distribution $\mathcal{K}$, it is annihilated by the one-forms $ds^a + A^a_\alpha dr^\alpha$ and hence must be of the form

$$X_{\mathcal{K}} = \dot{r}^\alpha \partial_{r^\alpha} - A^a_\alpha \dot{r}^\alpha \partial_{s^a} + \ddot{r}^\alpha \partial \dot{r}^\alpha,$$

i.e., for the vector field $X_{\mathcal{K}}$,

$$\dot{s}^a = -A^a_\alpha \dot{r}^\alpha.$$

As for $X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{K}}$, we have

$$\begin{aligned} X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{K}} &= \dot{r}^{\alpha} \left(-\frac{\partial^2 L_c}{\partial r^{\alpha} \partial \dot{r}^{\beta}} + A^a_{\alpha} \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^{\beta}} - \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_{\beta}}{\partial r^{\alpha}} + A^a_{\alpha} \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_{\beta}}{\partial s^a} \right) dr^{\beta} \\ &\quad - \dot{r}^{\beta} \left(-\frac{\partial^2 L_c}{\partial r^{\alpha} \partial \dot{r}^{\beta}} + A^a_{\alpha} \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^{\beta}} - \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_{\beta}}{\partial r^{\alpha}} + A^a_{\alpha} \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_{\beta}}{\partial s^a} \right) dr^{\alpha} \\ &\quad + \dot{r}^{\beta} \frac{\partial^2 L_c}{\partial \dot{r}^{\alpha} \partial \dot{r}^{\beta}} d\dot{r}^{\alpha} - \ddot{r}^{\alpha} \frac{\partial^2 L_c}{\partial \dot{r}^{\alpha} \partial \dot{r}^{\beta}} dr^{\beta} \end{aligned}$$

This is because $X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{K}}$ is the restriction of $X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{D}}$ to the distribution of $\mathcal{K}$ and hence all the terms in $\Omega_{\mathcal{D}}$ which involve $(ds^b + A^b_{\beta} dr^{\beta})$ vanish. The same is true for $dE_{\mathcal{D}}$.

Equating $X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{K}}$ with $dE_{\mathcal{K}}$ and recalling that we have already obtained $\dot{s}^a = -A^a_{\alpha} \dot{r}^{\alpha}$, we get the following set of equations

$$\begin{aligned} &\dot{r}^{\beta} \frac{\partial^2 L_c}{\partial r^{\alpha} \partial \dot{r}^{\beta}} - \dot{r}^{\beta} A^b_{\alpha} \frac{\partial^2 L_c}{\partial s^b \partial \dot{r}^{\beta}} - \frac{\partial L_c}{\partial r^{\alpha}} + A^b_{\alpha} \frac{\partial L_c}{\partial s^b} \\ &= \dot{r}^{\beta} \left(-\frac{\partial^2 L_c}{\partial r^{\beta} \partial \dot{r}^{\alpha}} + A^a_{\beta} \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^{\alpha}} - \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_{\alpha}}{\partial r^{\beta}} + A^a_{\beta} \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_{\alpha}}{\partial s^a} \right) \\ &\quad - \dot{r}^{\beta} \left(-\frac{\partial^2 L_c}{\partial r^{\alpha} \partial \dot{r}^{\beta}} + A^a_{\alpha} \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^{\beta}} - \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_{\beta}}{\partial r^{\alpha}} + A^a_{\alpha} \frac{\partial L}{\partial \dot{s}^b} \frac{\partial A^b_{\beta}}{\partial s^a} \right) - \ddot{r}^{\beta} \frac{\partial^2 L_c}{\partial \dot{r}^{\beta} \partial \dot{r}^{\alpha}} \end{aligned}$$

After simplification, we have

$$\begin{aligned} &\ddot{r}^{\beta} \frac{\partial^2 L_c}{\partial \dot{r}^{\beta} \partial \dot{r}^{\alpha}} + \dot{r}^{\beta} \frac{\partial^2 L_c}{\partial r^{\beta} \partial \dot{r}^{\alpha}} - \dot{r}^{\beta} A^a_{\beta} \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^{\alpha}} - \frac{\partial L_c}{\partial r^{\alpha}} + A^b_{\alpha} \frac{\partial L_c}{\partial s^b} \\ &= \dot{r}^{\beta} \frac{\partial L}{\partial \dot{s}^b} \left(-\frac{\partial A^b_{\alpha}}{\partial r^{\beta}} + A^a_{\beta} \frac{\partial A^b_{\alpha}}{\partial s^a} + \frac{\partial A^b_{\beta}}{\partial r^{\alpha}} - A^a_{\alpha} \frac{\partial A^b_{\beta}}{\partial s^a} \right), \end{aligned}$$

which indeed gives the Lagrange-d'Alembert equations in [BKMM]:

$$\frac{d}{dt} \frac{\partial L_c}{\partial \dot{r}^{\alpha}} - \frac{\partial L_c}{\partial r^{\alpha}} + A^b_{\alpha} \frac{\partial L_c}{\partial s^b} = -\frac{\partial L}{\partial \dot{s}^b} B^b_{\alpha\beta} \dot{r}^{\beta}. \quad \blacksquare$$

Remarks.

Here is another way of viewing the preceding theorem. Consider the following form of the equations:

$$X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} = dH_{\mathcal{M}} \qquad \text{on } \mathcal{H};$$

that is,

$$\langle X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}, u \rangle = \langle dH_{\mathcal{M}}, u \rangle ,$$

for all $u \in \mathcal{H}$. If we rewrite this in the form

$$\langle dH_{\mathcal{M}} - X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}}, u\rangle = 0,$$

then on the Lagrangian side, this is nothing but

$$\langle dE_{\mathcal{D}} - X_{\mathcal{K}} \lrcorner\, (\Omega_L)_{\mathcal{D}}, v\rangle = 0,$$

where $v \in \mathcal{K}$. With appropriate interpretations, this is equivalent to Lagrange-d'Alembert principle:

$$\begin{aligned} \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i}\right)(\delta q^i) &= 0 \\ \omega^a(\dot{q}) &= 0 \end{aligned}$$

where $\omega(\delta q) = 0$.

3.2.5 Example: 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 the body back and forth, while simultaneously moving the wheels with the proper phase relationship. For details, see [BKMM] and the references listed there. Here we will include some of 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. 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. See figure 3.1.

In [BKMM], a simplification is made which we shall also assume in this chapter, namely $\phi_1 = -\phi_2$, $J_1 = J_2$. The parameters are also chosen such that $J + J_0 + J_1 + J_2 = mr^2$,

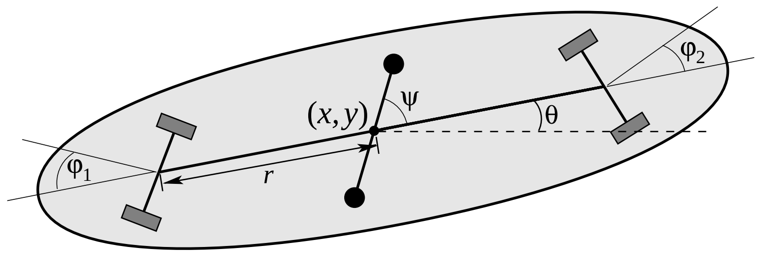

Figure 3.1: The geometry of the snakeboard.

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_1, J_2 are the inertia of the wheels. This simplification eliminates some terms in the derivation but does not affect the essential geometry of the problem. Setting $\phi = \phi_1 = -\phi_2$, then the configuration space becomes $Q = SE(2) \times S^1 \times S^1$ and the Lagrangian $L : TQ \to \mathbb{R}$ is the total kinetic energy of the system and is given by

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}J_0\dot{\psi}^2 + J_0\dot{\psi}\dot{\theta} + J_1\dot{\phi}^2.$$

The Constraints. 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. The constraints are defined by

$$-\sin(\theta + \phi)\dot{x} + \cos(\theta + \phi)\dot{y} - r\cos\phi\dot{\theta} = 0 \tag{3.12}$$

$$-\sin(\theta - \phi)\dot{x} + \cos(\theta - \phi)\dot{y} + r\cos\phi\dot{\theta} = 0 \tag{3.13}$$

and can be simplified as

$$\begin{aligned} \dot{x} &= -r\cot\phi\cos\theta\dot{\theta} \\ \dot{y} &= -r\cot\phi\sin\theta\dot{\theta}. \end{aligned}$$

Since the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ on the constraint submanifold $\mathcal{D}$

and its inverse are given by

$$\begin{aligned}
p_x &= -mr\cot\phi\cos\theta\dot{\theta} \\
p_y &= -mr\cot\phi\sin\theta\dot{\theta} \\
p_\theta &= mr^2\dot{\theta} + J_0\dot{\psi} \\
p_\psi &= J_0\dot{\psi} + J_0\dot{\theta} \\
p_\phi &= 2J_1\dot{\phi} \\
\dot{x} &= -\frac{r}{mr^2 - J_0}\cot\phi\cos\theta(p_\theta - p_\psi) \\
\dot{y} &= -\frac{r}{mr^2 - J_0}\cot\phi\sin\theta(p_\theta - p_\psi) \\
\dot{\theta} &= \frac{p_\theta - p_\psi}{mr^2 - J_0} \\
\dot{\psi} &= \frac{mr^2 p_\psi - J_0 p_\theta}{J_0(mr^2 - J_0)} \\
\dot{\phi} &= \frac{p_\phi}{2J_1},
\end{aligned}$$

the constraint submanifold $\mathcal{M}$ is defined by

$$\begin{aligned}
\mathcal{M} = & \{(x, y, \theta, \psi, \phi, p_x, p_y, p_\theta, p_\psi, p_\phi) \mid \\
& p_x = -\frac{mr}{mr^2 - J_0}\cot\phi\cos\theta(p_\theta - p_\psi), p_y = -\frac{mr}{mr^2 - J_0}\cot\phi\sin\theta(p_\theta - p_\psi).\}
\end{aligned}$$

Notice that $\mathcal{M}$ may be thought of as a graph in T^*Q and we can use the induced coordinates $(x, y, \theta, \psi, \phi, p_\theta, p_\psi, p_\phi)$ as its local coordinates. Hence the distribution $\mathcal{H}$ of $\mathcal{M}$ is

$$\begin{aligned}
\mathcal{H} &= \ker\{dx + r\cot\phi\cos\theta d\theta, dy + r\cot\phi\sin\theta d\theta\} \\
&= \operatorname{span}\{-r\cot\phi\cos\theta\partial_x - r\cot\phi\sin\theta\partial_y + \partial_\theta, \partial_\psi, \partial_\phi, \partial_{p_\theta}, \partial_{p_\psi}, \partial_{p_\phi}\}.
\end{aligned}$$

The Hamiltonian. The corresponding Hamiltonian is given via the Legendre transform by

$$H = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{1}{2J_0}p_\psi^2 + \frac{1}{2(mr^2 - J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1}p_\phi^2.$$

Now if we restrict the Hamiltonian H to the submanifold $\mathcal{M}$, we get

$$H_{\mathcal{M}} = \frac{mr^2}{2(mr^2 - J_0)^2}\cot^2\phi(p_\theta - p_\psi)^2 + \frac{1}{2J_0}p_\psi^2 + \frac{1}{2(mr^2 - J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1}p_\phi^2.$$

After computing its differential $dH_{\mathcal{M}}$ and restricting it to $\mathcal{H}$, we have

$$\begin{aligned} dH_{\mathcal{H}} &= -\frac{mr^2}{(mr^2-J_0)^2}\cot\phi\csc^2\phi(p_\theta-p_\psi)^2 d\phi \\ &\quad +\frac{mr^2}{(mr^2-J_0)^2}\cot^2\phi(p_\theta-p_\psi)(dp_\theta-dp_\psi) \\ &\quad +\frac{1}{J_0}p_\psi dp_\psi+\frac{1}{(mr^2-J_0)}(p_\theta-p_\psi)(dp_\theta-dp_\psi)+\frac{1}{2J_1}p_\phi dp_\phi. \end{aligned}$$

The Two Form. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned} \Omega_{\mathcal{M}} &= dx\wedge dp_x+dy\wedge dp_y+d\theta\wedge dp_\theta+d\psi\wedge dp_\psi+d\phi\wedge dp_\phi \\ &= kdx\wedge[\csc^2\phi\cos\theta(p_\theta-p_\psi)d\phi+\cot\phi\sin\theta(p_\theta-p_\psi)d\theta-\cot\phi\cos\theta(dp_\theta-dp_\psi)] \\ &\quad +kdy\wedge[\csc^2\phi\sin\theta(p_\theta-p_\psi)d\phi-\cot\phi\cos\theta(p_\theta-p_\psi)d\theta-\cot\phi\sin\theta(dp_\theta-dp_\psi)] \\ &\quad +d\theta\wedge dp_\theta+d\psi\wedge dp_\psi+d\phi\wedge dp_\phi, \end{aligned}$$

where $k=mr/(mr^2-J_0)$. If we restrict $\Omega_{\mathcal{M}}$ to the distribution $\mathcal{H}$, we get

$$\begin{aligned} \Omega_{\mathcal{H}} &= -kr\cot\phi\cos\theta d\theta\wedge \\ &\qquad [\csc^2\phi\cos\theta(p_\theta-p_\psi)d\phi+\cot\phi\sin\theta(p_\theta-p_\psi)d\theta-\cot\phi\cos\theta(dp_\theta-dp_\psi)] \\ &\quad -kr\cot\phi\sin\theta d\theta\wedge \\ &\qquad [\csc^2\phi\sin\theta(p_\theta-p_\psi)d\phi-\cot\phi\cos\theta(p_\theta-p_\psi)d\theta-\cot\phi\sin\theta(dp_\theta-dp_\psi)] \\ &\quad +d\theta\wedge dp_\theta+d\psi\wedge dp_\psi+d\phi\wedge dp_\phi \\ &= d\theta\wedge[-kr\cot\phi\csc^2\phi(p_\theta-p_\psi)d\phi+kr\cot^2\phi(dp_\theta-dp_\psi)+dp_\theta] \\ &\quad +d\psi\wedge dp_\psi+d\phi\wedge dp_\phi. \end{aligned}$$

The Equations of Motion. Notice that any vector field $X_{\mathcal{M}}$ is of the form

$$X_{\mathcal{M}}=\dot{x}\partial_x+\dot{y}\partial_y+\dot{\theta}\partial_\theta+\dot{\psi}\partial_\psi+\dot{\phi}\partial_\phi+\dot{p}_\theta\partial_{p_\theta}+\dot{p}_\psi\partial_{p_\psi}+\dot{p}_\phi\partial_{p_\phi}.$$

But $X_{\mathcal{H}}$ also lies in $\mathcal{H}=\ker\{dx+r\cot\phi\cos\theta d\theta, dy+r\cot\phi\sin\theta d\theta\}$ and hence must be of the form

$$X_{\mathcal{H}}=\dot{\theta}(-r\cot\phi\cos\theta\partial_x-r\cot\phi\sin\theta\partial_y+\partial_\theta)+\dot{\psi}\partial_\psi+\dot{\phi}\partial_\phi+\dot{p}_\theta\partial_{p_\theta}+\dot{p}_\psi\partial_{p_\psi}+\dot{p}_\phi\partial_{p_\phi},$$

which gives us the first set of relationships

$$\begin{aligned} \dot{x} &= -r\cot\phi\cos\theta\,\dot{\theta} \\ \dot{y} &= -r\cot\phi\sin\theta\,\dot{\theta}. \end{aligned}$$

Moreover,

$$\begin{aligned} X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{H}} &= - kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\dot\theta\, d\phi + kr\cot^2\phi\dot\theta(dp_\theta - dp_\psi) + \dot\theta\, dp_\theta \\ &\quad + \dot\psi\, dp_\psi + kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\dot\phi\, d\theta + \dot\phi\, dp_\phi - kr\cot^2\phi\dot p_\theta\, d\theta - \dot p_\theta\, d\theta \\ &\quad + kr\cot^2\phi\dot p_\psi\, d\theta - \dot p_\psi\, d\psi - \dot p_\phi\, d\phi, \end{aligned}$$

and if equated with $dH_{\mathcal{H}}$, we will get the following set of equations:

$$kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\,\dot\phi - kr\cot^2\phi\,\dot p_\theta - \dot p_\theta + kr\cot^2\phi\,\dot p_\psi = 0$$

$$\begin{aligned} -\dot p_\psi &= 0 \\ -\dot p_\phi - kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\dot\theta &= -\frac{mr^2}{(mr^2 - J_0)^2}\cot\phi\csc^2\phi(p_\theta - p_\psi)^2 \\ kr\cot^2\phi\dot\theta + \dot\theta &= \frac{mr^2}{(mr^2 - J_0)^2}\cot^2\phi(p_\theta - p_\psi) + \frac{1}{(mr^2 - J_0)}(p_\theta - p_\psi) \\ -kr\cot^2\phi\,\dot\theta + \dot\psi &= -\frac{mr^2}{(mr^2 - J_0)^2}\cot^2\phi(p_\theta - p_\psi) + \frac{1}{J_0}p_\psi \\ &\quad -\frac{1}{(mr^2 - J_0)}(p_\theta - p_\psi) \\ \dot\phi &= \frac{1}{2J_1}p_\phi. \end{aligned}$$

After simplification, we have

$$\dot p_\theta = \frac{\cot\phi}{2J_1(1 - \frac{J_0}{mr^2}\sin^2\phi)}p_\phi(p_\theta - p_\psi) \tag{3.14}$$

$$\dot p_\psi = 0 \tag{3.15}$$

$$\dot p_\phi = 0 \tag{3.16}$$

$$\dot\theta = \frac{p_\theta - p_\psi}{mr^2 - J_0} \tag{3.17}$$

$$\dot\psi = \frac{mr^2 p_\psi - J_0 p_\theta}{J_0(mr^2 - J_0)} \tag{3.18}$$

$$\dot\phi = \frac{p_\phi}{2J_1}. \tag{3.19}$$

Notice that the last 3 equations are nothing but the inverse of the constrained Legendre transformation $\mathbb{F}L|\mathcal{D}$ written in local coordinates. The first equation is equivalent to the momentum equation (discussed below and in [BKMM]) written in Hamiltonian form and the 2nd and 3rd equations are the reduced equations on the shape space, again in their Hamiltonian forms.

Moreover, the corresponding Lagrangian procedure gives the equations of the motion on the Lagrangian side as

$$\begin{aligned} \ddot{\theta} - \cot\phi\dot{\phi}\dot{\theta} + \frac{J_0}{mr^2}\sin^2\phi\ddot{\psi} &= 0 \\ J_0\ddot{\psi} + J_0\ddot{\theta} &= 0 \\ 2J_1\ddot{\phi} &= 0 \end{aligned}$$

and it can be shown that both systems of equations are equivalent via the Legendre transform $\mathbb{F}L|\mathcal{D}$.

3.3 Nonholonomic Mechanical Systems with Symmetry

Now we add the hypothesis of symmetry to the preceding development. Assume that we have a configuration manifold Q, a Lagrangian of the form kinetic minus potential, and a distribution $\mathcal{D}$ that describes the kinematic nonholonomic constraints. We also assume there is a symmetry group G (a Lie group) that leaves the Lagrangian invariant, and that acts on Q (by isometries) and also leaves the distribution invariant, *i.e.,* the tangent of the group action maps $\mathcal{D}_q$ to $\mathcal{D}_{gq}$ (for more details, see [BKMM].) Later, we shall refer this as a **simple nonholonomic mechanical system**.

In this section,

1. we recall the basic ideas and results from [BKMM] on simple nonholonomic mechanical systems, especially on how it extend the Lagrangian reduction theory of Marsden and Scheurle [1993a,b] to the context of nonholonomic systems. We shall describe briefly how [BKMM] modifies the Ehresmann connection associated with the constraints to a new connection, called the *nonholonomic connection*, that also takes into account the symmetries, and how the reduced equations, relative to this new connection, break up into *two* sets: a set of reduced Lagrange-d'Alembert equations, and a momentum equation. When the reconstruction equations are added, one recovers the full set of equations of motion for the system.

2. We summarize the Hamiltonian reduction formulation of [BS] on finding the reduced equations of motion for nonholonomic systems with symmetry.

3. We restate the reduction procedure on the Lagrangian side corresponding to those on the Hamiltonian side using the Legendre transformation.

4. We prove that these dual procedures give us the same reduced Lagrange-d'Alembert equations in [BKMM]. Since this proof is done in coordinates, it does provide a systematic way to carry out the computations on the Hamiltonian side. Also, the proof shows where the momentum equation is lurking on the Hamiltonian side and how this is related to breaking up the dynamics of the nonholonomic system into 3 parts: a reconstruction equation for a group element g, an equation for the nonholonomic momentum p and the reduced Hamilton equations in the shape variables r, p_r (and p). This way of breaking up the dynamics may have the same significance for the control theory as what has already been noted in [BKMM].

5. We apply the Hamiltonian reduction procedure to the examples of the snakeboard, the bicycle and a nonholonomically constrained particle.

3.3.1 Review of Lagrangian Reduction

We first recall how [BKMM] explains in general terms how one constructs reduced systems by eliminating the group variables.

Proposition 1 *Under the assumptions that both the Lagrangian L and the distribution $\mathcal{D}$ are G-invariant, we can form the reduced velocity phase space TQ/G and the constrained reduced velocity phase space $\mathcal{D}/G$. The Lagrangian L induces well defined functions, the reduced Lagrangian*

$$l : TQ/G \to \mathbb{R}$$

satisfying $L = l \circ \pi_{TQ}$ where $\pi_{TQ} : TQ \to TQ/G$ is the projection, and the constrained reduced Lagrangian

$$l_c : \mathcal{D}/G \to \mathbb{R},$$

which satisfies $L|\mathcal{D} = l_c \circ \pi_{\mathcal{D}}$ where $\pi_{\mathcal{D}} : \mathcal{D} \to \mathcal{D}/G$ is the projection. Also, the Lagrange-d'Alembert equations induce well defined reduced Lagrange-d'Alembert equations on $\mathcal{D}/G$. That is, the vector field on the manifold $\mathcal{D}$ determined by the Lagrange-d'Alembert equations (including the constraints) is G-invariant, and so defines a reduced vector field on the quotient manifold $\mathcal{D}/G$.

This proposition follows from general symmetry considerations, but to compute the associated reduced equations explicitly and to reconstruct the group variables, one defines the nonholonomic momentum map J^{nh}, and extends the *Noether Theorem* to nonholonomic system and synthesizes, out of the mechanical connection and the Ehresmann connection, a nonholonomic connection $\mathcal{A}^{\text{nh}}$ which is a connection on the principal bundle $Q \to Q/G$.

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

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

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 S_q:

$$\mathfrak{g}^q = \{\xi \in \mathfrak{g} \mid \xi_Q(q) \in \mathcal{S}_q\}.$$

We let $\mathfrak{g}^{\mathcal{D}}$ denote 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

$$\langle J^{\text{nh}}(v_q), \xi \rangle = \frac{\partial L}{\partial \dot{q}^i} (\xi_Q)^i, \tag{3.20}$$

where $\xi \in \mathfrak{g}^q$. Notice that the nonholonomic momentum map may be viewed as encoding *some* of the components of the ordinary momentum map, namely the projection along those symmetry directions that are consistent with the constraints.

[BKMM] extends the Noether Theorem to nonholonomic systems by deriving the equation for the momentum map that replace the usual conservation law. It is proven that if the Lagrangian L is invariant under the group action and that ξ^q is a section of the bundle $\mathfrak{g}^{\mathcal{D}}$, then any solution $q(t)$ of the Lagrange d'Alembert equations must satisfy, in addition to the given kinematic constraints, the momentum equation:

$$\frac{d}{dt}\left(J^{\text{nh}}(\xi^{q(t)})\right) = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}(\xi^{q(t)})\right]_Q^i. \tag{3.21}$$

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 Body Representation 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

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

Define the **body fixed basis** 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

$$\left\langle J^{\mathrm{nh}}(r, g, \dot{r}, \dot{g}), e_b(r, g) \right\rangle = \left\langle \frac{\partial l}{\partial \xi}, e_b(r) \right\rangle := p_b, \tag{3.22}$$

which defines p_b, a function of r, $\dot{r}$ and ξ. Note that in this body representation, the functions p_b are *invariant* rather than equivariant, as is usually the case with the momentum map. It is shown in [BKMM] that in this body representation, the momentum equation is given by

$$\frac{d}{dt} p_i = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_i] + \frac{\partial e_i}{\partial r^\alpha} \dot{r}^\alpha \right\rangle , \tag{3.23}$$

where the range of i is 1 to m. Moreover, the momentum equation in this representation is independent of, that is, decouples from, the group variables g.

The Nonholonomic Connection Recall that in the case of simple holonomic mechanical system, the mechanical connection $\mathcal{A}$ is defined by $\mathcal{A}(v_q) = \mathbb{I}(q)^{-1} J(v_q)$ where J is the associated momentum map and $\mathbb{I}(q)$ is the locked inertia tensor of the system. Equivalently the mechanical connection can also be defined by the fact that its horizontal space at q

is orthogonal to the group orbit at q with respect to the kinetic energy metric. For more information, see for example, Marsden [1992] and Marsden and Ratiu [1994].

As [BKMM] points out, in the *principal* case where the constraints and the orbit directions span the entire tangent space to the configuration space, that is,

$$\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_qQ, \tag{3.24}$$

the definition of the momentum map can be used to augment the constraints and provide a connection on $Q \to Q/G$. Let J^{nh} be the nonholonomic momentum map and define similarly as above a map $A_q^{\mathrm{sym}} : T_qQ \to \mathcal{S}_q$ given by

$$A^{\mathrm{sym}}(v_q) = (\mathbb{I}^{\mathrm{nh}}(q)^{-1} J^{\mathrm{nh}}(v_q))_Q$$

(this defines the momentum "constraints") where $\mathbb{I}^{\mathrm{nh}} : \mathfrak{g}^{\mathcal{D}} \to (\mathfrak{g}^{\mathcal{D}})^*$ is the locked inertia tensor defined in a similar way as in holonomic systems.

Choose a complementary space to $\mathcal{S}_q$ by writing $T_q(\mathrm{Orb}(q)) = \mathcal{S}_q \oplus \mathcal{U}_q$. Let $A_q^{\mathrm{kin}} : T_qQ \to \mathcal{U}_q$ be a $\mathcal{U}_q$ valued form that projects $\mathcal{U}_q$ onto itself and maps $\mathcal{D}_q$ to zero. Then the kinematic constraints are defined by the equation

$$A^{\mathrm{kin}}(q)\dot{q} = 0.$$

This kinematic constraints equation plus the momentum "constraints" equation can be used to synthesis a nonholonomic connection $\mathcal{A}^{\mathrm{nh}}$ which is a principal connection on the bundle $Q \to Q/G$ and 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,

$$\mathcal{A}^{\mathrm{nh}}(v_q) = \mathbb{I}^{\mathrm{nh}}(q)^{-1} J^{\mathrm{nh}}(v_q). \tag{3.25}$$

In a body fixed basis, (3.25) can be written as

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

Hence, the constraints can be represented in a nice way by

$$g^{-1}\dot{g} = \xi = -\mathcal{A}(r)\dot{r} + \Gamma(r)p, \tag{3.26}$$

where $\mathcal{A}(r)$ is the abbreviation for $\mathcal{A}^{\mathrm{nh}}_{\mathrm{loc}}(r)$ and $\Gamma(r) = \mathbb{I}^{\mathrm{nh}}_{\mathrm{loc}}(r)^{-1}$.

Moreover, with the help of the nonholonomic mechanical connection, the Lagrange-d'Alembert principle may be broken up into two principles by breaking the variations δq into two parts, namely parts that are horizontal with respect to the nonholonomic connection and parts that are vertical (but still in $\mathcal{D}$), and the reduced equations break up into *two* sets: a set of reduced Lagrange-d'Alembert equations (which have curvature terms appearing as 'forcing'), and a momentum equation, which have a form generalizing the components of the Euler-Poincaré equations along the symmetry directions consistent with the constraints. When one supplements these equations with the reconstruction equations, one recovers the full set of equations of motion for the system.

3.3.2 Hamiltonian Reduction

In working out the nonholonomic Hamiltonian reduction, [BS] also starts out with a *simple nonholonomic mechanical system*. Recall from Section 2 that the Legendre transformation $\mathbb{F}L : TQ \to T^*Q$ is used to define the constraint submanifold $\mathcal{M} \subset T^*Q$ where

$$\mathcal{M} = \mathbb{F}L(\mathcal{D}).$$

On this manifold, there is a distribution $\mathcal{H}$

$$\mathcal{H} = \mathcal{F} \cap T\mathcal{M},$$

where

$$\mathcal{F} = (T\pi)^{-1}(\mathcal{D}),$$

and $\pi : T^*Q \to Q$. Also recall that $\Omega_{\mathcal{H}}$, the restriction of the canonical two-form Ω of T^*Q to the distribution $\mathcal{H}$ of the constraint submanifold $\mathcal{M}$, is nondegenerate and that the dynamics is given by a vector field $X_{\mathcal{H}}$ on $\mathcal{M}$ taking values in $\mathcal{H}$ and satisfies the equation

$$X_{\mathcal{H}} \lrcorner \, \Omega_{\mathcal{H}} = dH_{\mathcal{H}} \tag{3.27}$$

where $dH_{\mathcal{H}}$ is the (fiberwise) restriction of $dH_{\mathcal{M}}$ to $\mathcal{H}$.

Now let G be the symmetry group of this system and assume that the quotient space $\overline{\mathcal{M}} = \mathcal{M}/G$ of the G-orbit in $\mathcal{M}$ is a quotient manifold with projection map $\rho : \mathcal{M} \to \overline{\mathcal{M}}$. Since G is a symmetric group, all intrinsically defined vector fields and distributions push down to $\overline{\mathcal{M}}$. In particular, the vector field $X_{\mathcal{M}}$ on $\mathcal{M}$ pushes down to a vector field $\overline{X}_{\overline{\mathcal{M}}} = \rho_* X_{\mathcal{M}}$, and the distribution $\mathcal{H}$ pushes down to a distribution $\rho_*\mathcal{H}$ on $\overline{\mathcal{M}}$.

However, $\Omega_{\mathcal{H}}$ need not push down to a two-form defined on $\rho_*\mathcal{H}$, despite the fact that $\Omega_{\mathcal{H}}$ is G-invariant. This is because there may be infinitesimal symmetry $\xi_{\mathcal{M}}$ that lies in $\mathcal{H}$ such that

$$\xi_{\mathcal{M}} \lrcorner\, \Omega_{\mathcal{H}} \neq 0,$$

To eliminate this difficulty, [BS] restricts $\Omega_{\mathcal{H}}$ to a subdistribution $\mathcal{U}$ of $\mathcal{H}$ defined by

$$\mathcal{U} = \{u \in \mathcal{H} \mid \Omega_{\mathcal{H}}(u,v) = 0 \quad \text{for all} \quad v \in \mathcal{V} \cap \mathcal{H}\} = \mathcal{H} \cap (\mathcal{V} \cap \mathcal{H})^{\perp}, \tag{3.28}$$

where $\mathcal{V}$ is the distribution on $\mathcal{M}$ tangent to the orbits of G in $\mathcal{M}$ and is spanned by the infinitesimal symmetries and $(\mathcal{V}\cap\mathcal{H})^{\perp}$ is the $\Omega_{\mathcal{H}}$-orthogonal complement of $(\mathcal{V}\cap\mathcal{H})$. Clearly, $\mathcal{U}$ and $\mathcal{V}$ are both G-invariant, project down to $\overline{\mathcal{M}}$ and $\rho_* \mathcal{V} = 0$. Define $\overline{\mathcal{H}}$ by

$$\overline{\mathcal{H}} = \rho_* \mathcal{U}.$$

It is proven in [BS] that

1. The vector field $X_{\mathcal{H}}$ which satisfies the above Hamiltonian equation of motion (3.27) lies in the distribution $\mathcal{U}$.

2. The restriction $\Omega_{\mathcal{U}}$ of Ω to the distribution $\mathcal{U}$ pushes down to a nondegenerate 2-form $\Omega_{\overline{\mathcal{H}}} = \rho_* \Omega_{\mathcal{U}}$ on $\overline{\mathcal{H}}$, which is modeled by the symplectic space $(\mathcal{V}\cap\mathcal{H})^{\perp}/(\mathcal{V}\cap\mathcal{H})\cap(\mathcal{V}\cap \mathcal{H})^{\perp}$.

3. Furthermore,

$$\overline{X}_{\overline{\mathcal{H}}} \lrcorner\, \Omega_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{H}}}, \tag{3.29}$$

   where $h_{\overline{\mathcal{M}}} = \rho_* H_{\mathcal{M}}$ is the pushdown of the restriction to $\mathcal{M}$ of the Hamiltonian H and $dh_{\overline{\mathcal{H}}}$ is the restriction of $dh_{\overline{\mathcal{M}}}$ to $\overline{\mathcal{H}}$. This is because the equation $X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{H}} = dH_{\mathcal{H}}$, restricted to $\mathcal{U} \subset \mathcal{H}$, vanishes on vectors in $\mathcal{V}$, and is G-invariant. Hence both sides push down to $\overline{\mathcal{H}}$.

Note that the original equations of motion are

$$X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{H}} = dH_{\mathcal{H}}$$

where $\mathcal{H}$ is a distribution in the constraint manifold $\mathcal{M}$. After the reduction of symmetry we obtain equations of the same type

$$\overline{X}_{\overline{\mathcal{H}}} \lrcorner\, \Omega_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{H}}},$$

where $\overline{\mathcal{H}}$ is a distribution in the reduced space $\overline{\mathcal{M}} = \mathcal{M}/G$.

3.3.3 Lagrangian Side

By using the Legendre transformation $\mathbb{F}L$, we can construct dual geometric structures on the tangent bundle TQ and formulate a similar Lagrangian reduction procedure. This allows us to better compare with the geometric constructions and analytic formulations on the manifold Q in [BKMM], and in the course of doing this, we realize that the requirement (see point (1) of last subsection) that the vector field $X_{\mathcal{H}}$ lies in the subdistribution $\mathcal{U}$ is equivalent to the extended Noether Theorem; that is, that any solution of the Lagrange-d'Alembert equations must satisfy the momentum equation.

Recall from Section 2. We consider $\mathcal{D}$ as a constraint submanifold of TQ and then construct the distribution

$$\mathcal{K} = \mathcal{C} \cap T\mathcal{D},$$

on TTQ, where

$$\mathcal{C} = (T\tau_Q)^{-1}(\mathcal{D}),$$

and $\tau_Q : TQ \to Q$. Clearly $\mathcal{D} = (\mathbb{F}L)^{-1}(\mathcal{M}), \mathcal{K} = (T\mathbb{F}L)^{-1}(\mathcal{H})$. The motion is then given by a vector field $X_{\mathcal{K}}$ on the manifold $\mathcal{D}$ which takes values in $\mathcal{K}$ and satisfies the equation

$$X_{\mathcal{K}} \lrcorner \, \Omega_{\mathcal{K}} = dE_{\mathcal{K}}, \tag{3.30}$$

where $dE_{\mathcal{K}}$ and $\Omega_{\mathcal{K}}$ are the restrictions of $dE_{\mathcal{D}}$ and $\Omega_{\mathcal{D}}$ respectively to the distribution $\mathcal{K}$.

Now let G be the symmetry group of this system and assume that the quotient space $\overline{\mathcal{D}} = \mathcal{D}/G$ of the G-orbit in $\mathcal{D}$ is a smooth quotient manifold with projection map $\lambda : \mathcal{D} \to \overline{\mathcal{D}}$. Since G is a symmetric group, all intrinsically defined vector fields and distributions push down to $\overline{\mathcal{D}}$. In particular, the vector field $X_{\mathcal{D}}$ on $\mathcal{D}$ pushes down to a vector field $\overline{X}_{\overline{\mathcal{D}}} = \lambda_* X_{\mathcal{D}}$, and the distribution $\mathcal{K}$ pushes down to a distribution $\lambda_* \mathcal{K}$ on $\overline{\mathcal{D}}$. Here we use the push forward symbol λ_* to mean that the vector fields are λ-related.

For the same reason as the Hamiltonian side, $\Omega_{\mathcal{K}}$ need not push down to a two-form defined on $\lambda_* \mathcal{K}$, despite the fact that $\Omega_{\mathcal{K}}$ is G-invariant. We can restrict $\Omega_{\mathcal{K}}$ to the subdistribution $\mathcal{W}$ of $\mathcal{K}$ defined by

$$\mathcal{W} = \{w \in \mathcal{K} \mid \Omega_{\mathcal{K}}(w, v) = 0 \text{ for all } v \in \mathcal{T} \cap \mathcal{K}\} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{D})^{\perp}, \tag{3.31}$$

where $\mathcal{T}$ is the distribution on $\mathcal{D}$ tangent to the orbits of G in $\mathcal{D}$ and is spanned by the infinitesimal symmetries. Clearly, $\mathcal{W}$ and $\mathcal{T}$ are both G-invariant, $\mathcal{W}$ projects down to $\overline{\mathcal{D}}$

and $\lambda_* \mathcal{T} = 0$. Define $\overline{\mathcal{K}}$ by

$$\overline{\mathcal{K}} = \lambda_* \mathcal{W}.$$

Since the above constructions are dual to those in the Hamiltonian side, we also have

1. The vector field $X_{\mathcal{K}}$ which satisfies the above equation (3.30) takes values in the distribution $\mathcal{W}$.

2. The restriction $\Omega_{\mathcal{W}}$ of Ω_L to the distribution $\mathcal{W}$, pushes down to a nondegenerate 2-form $\Omega_{\overline{\mathcal{K}}} = \lambda_* \Omega_{\mathcal{W}}$ on $\overline{\mathcal{K}}$, which is modeled by the symplectic space $(\mathcal{T} \cap \mathcal{K})^{\perp}/(\mathcal{T} \cap \mathcal{K}) \cap (\mathcal{T} \cap \mathcal{K})^{\perp}$.

3. The reduced equations of motion are given by

$$\overline{X}_{\overline{\mathcal{K}}} \lrcorner \, \Omega_{\overline{\mathcal{K}}} = d\overline{E}_{\overline{\mathcal{K}}}, \tag{3.32}$$

where $\overline{E}_{\overline{\mathcal{D}}} = \lambda_* E_{\mathcal{D}}$ is the pushdown of the restriction to $\mathcal{D}$ of the energy function E. This is because the equation $X_{\mathcal{K}} \lrcorner \, \Omega_{\mathcal{K}} = dE_{\mathcal{K}}$, restricted to $\mathcal{W} \subset \mathcal{K}$, vanishes on vectors in $\mathcal{T}$, and is G-invariant. Hence both sides push down to $\overline{\mathcal{K}}$. All these will become clearer in the subsequent computations.

3.3.4 The equivalence of Hamiltonian and Lagrangian Reductions

Theorem 8 *Consider a simple nonholonomic mechanical system with symmetry and assume that it is in the principal case. Then the reduction procedure on TQ described in the preceding section gives the same set of equations as in [BKMM].*

Proof The first difficulty is how to represent the constraint submanifold $\mathcal{D} \subset TQ$ in a way that is both intrinsic and ready for reduction. The comparison with the geometric constructions in [BKMM] and the desire to have the dynamics break up in a way that are ready for reconstruction give hints that we should use the tools like nonholonomic momentum p and the nonholomomic connection $\mathcal{A}$ in [BKMM] to describe the constraint submanifold $\mathcal{D}$

Recall that in [BKMM], the nonholonomic constraints together with the basic identity of the nonholonomic momentum map are used to synthesis a nonholonomic connection $\mathcal{A}$ and the nonholonomic constraints are then written in the form

$$g^{-1}\dot{g} = -A(r)\dot{r} + \Gamma(r)p,$$

where p is G-invariant. Hence, the constraint manifold is nothing but

$$\mathcal{D} = \{(g, r, \dot{g}, \dot{r}) \mid \dot{g} = g(-A(r)\dot{r} + \Gamma(r))p)\}.$$

It is a submanifold in TQ and we can use $(g, r, \dot{r}, p)$ as its induced local coordinates. Then, clearly, the corresponding coordinates for $\overline{\mathcal{D}} = \mathcal{D}/G$ are $(r, \dot{r}, p)$. From now on, we will use $A(r)$ to abbreviate $\mathcal{A}^{\rm nh}_{\rm loc}(r)$.

The next difficulty is to find the corresponding representations for the distribution $\mathcal{K}$, the subdistribution $\mathcal{T} \cap \mathcal{K}$ and its annihilator distribution $\mathcal{W}$ where

$$\mathcal{W} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{K})^{\perp}.$$

Recall that in [BKMM], a body fixed basis

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

has been constructed such that the infinitesimal generators $(e_i(g, r))_Q$ of its first m elements at a point q span $\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q))$. Assume that G is a matrix group and e_i^d is the component of $e_i(r)$ with respect to a fixed basis $\{b_a\}$ of the Lie algebra $\mathfrak{g}$ where $(b_a)_Q = \partial_{g^a}$, then

$$(e_i(g, r))_Q = g^a_d e^d_i \partial_{g^a}.$$

Since $\mathcal{K} = (T\tau)^{-1}(\mathcal{D})$ where $\mathcal{D}_q$ is the direct sum of $\mathcal{S}_q$ and the horizontal space of the nonholonomic connection $\mathcal{A}^{\rm nh}$, it can be represented in the induced coordinates by

$$\mathcal{K} = \mathrm{span}\{g^a_d e^d_i \partial_{g^a}, -g^a_b A^b_\alpha \partial_{g^a} + \partial_{r^\alpha}, \partial_{\dot{r}}, \partial_p\}. \tag{3.33}$$

Also, we have

$$\mathcal{T} \cap \mathcal{K} = \mathrm{span}\{g^a_d e^d_i \partial_{g^a}\}. \tag{3.34}$$

To find the distribution $\mathcal{W}$, we have to compute $g^a_d e^d_i \partial_{g^a} \,\lrcorner\, \Omega_{\mathcal{D}}$, for all $i = 1, \dots, m$. Since L is G-invariant, we have

$$\begin{aligned}
\Omega_{\mathcal{D}} &= dg^a \wedge d\left(\frac{\partial L}{\partial \dot{g}^a}\right) + dr^\alpha \wedge d\left(\frac{\partial L}{\partial \dot{r}^\alpha}\right) \\
&= dg^a \wedge d\left((g^{-1})^b_a \frac{\partial l}{\partial \xi^b}\right) + dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot{r}^\alpha}\right) \\
&= \frac{\partial (g^{-1})^b_a}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^a \wedge dg^c + (g^{-1})^b_a dg^a \wedge d\left(\frac{\partial l}{\partial \xi^b}\right) + dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot{r}^\alpha}\right)
\end{aligned}$$

Hence

$$\begin{aligned}
g_f^a e_i^f \partial_{g^a} \lrcorner \Omega_{\mathcal{D}} &= g_f^a e_i^f \frac{\partial (g^{-1})_a^b}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^c - g_f^c e_i^f \frac{\partial (g^{-1})_a^b}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^a + e_i^b d\left(\frac{\partial l}{\partial \xi^b}\right) \\
&= e_i^f \left(\left(g_f^c \frac{\partial (g^{-1})_c^b}{\partial g^a} - \frac{\partial (g^{-1})_a^b}{\partial g^c} g_f^c \right) \frac{\partial l}{\partial \xi^b} dg^a + d\left(\frac{\partial l}{\partial \xi^f}\right) \right) \\
&= e_i^f \left((g^{-1})_\sigma^b \left(-\frac{\partial g_f^\sigma}{\partial g^\tau} g_a^\tau + \frac{\partial g_a^\sigma}{\partial g^\tau} g_f^\tau \right) \frac{\partial l}{\partial \xi^b} (g^{-1})_e^a dg^e + d\left(\frac{\partial l}{\partial \xi^f}\right) \right) \\
&= e_i^f \left(-C_{af}^b \frac{\partial l}{\partial \xi^b} (g^{-1})_e^a dg^e + d\left(\frac{\partial l}{\partial \xi^f}\right) \right) \\
&= dp_i - \frac{\partial l}{\partial \xi^f} d(e_i^f) - C_{af}^b \frac{\partial l}{\partial \xi^b} e_i^f (g^{-1})_e^a dg^e.
\end{aligned}$$

Here, C_{af}^b is the structural constants for the Lie algebra $\mathfrak{g}$ and $p_i = \frac{\partial l}{\partial \xi^f} e_i^f$ as defined in (3.22). Therefore, the subdistribution $\mathcal{W} \subset \mathcal{K}$ is

$$\mathcal{W} = \ker \left\{ dp_i - \frac{\partial l}{\partial \xi^f} d(e_i^f) - C_{af}^b \frac{\partial l}{\partial \xi^b} e_i^f (g^{-1})_e^a dg^e \right\}. \tag{3.35}$$

Since the constraint manifold $\mathcal{D}$ has the induced local coordinates $(g, r, \dot{r}, p)$, any vector field $X_{\mathcal{D}}$ on the manifold $\mathcal{D}$ is of the form

$$X_{\mathcal{D}} = \dot{g}^a \partial_{g^a} + \dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial_{p_i}.$$

If $X_{\mathcal{D}}$ lies in the distribution $\mathcal{K}$, then we have $\dot{g} = g(-A\dot{r} + \Gamma p)$. Moreover, if $X_{\mathcal{D}}$ lies in the distribution $\mathcal{W}$, then for each j, we have

$$\dot{p}_j - \frac{\partial l}{\partial \xi^d} \frac{\partial e_j^d}{\partial r^\alpha} \dot{r}^\alpha - C_{ad}^b \frac{\partial l}{\partial \xi^b} \xi^a e_j^d = 0,$$

i.e.,

$$\dot{p}_j = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_j] + \dot{e}_j \right\rangle, \tag{3.36}$$

which gives exactly the momentum equation (3.23). Therefore, any vector field $X_{\mathcal{W}}$ taking values in $\mathcal{W}$ must be of the form

$$X_{\mathcal{W}} = g_b^a \xi^b \partial_{g^a} + \dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial_{p_i},$$

where

$$\xi = -A\dot{r} + \Gamma p \quad \dot{p}_j = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_j] + \dot{e}_j \right\rangle, \tag{3.37}$$

Now we are ready to do the reduction. But before that, we need to compute all the ingredients of the equation

$$X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{K}} = dE_{\mathcal{K}}.$$

Notice first that since E is G-invariant, we have

$$\begin{aligned} E &= \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L \\ &= \frac{\partial L}{\partial \dot{g}^a}\dot{g}^a + \frac{\partial L}{\partial \dot{r}^\alpha}\dot{r}^\alpha - L \\ &= \frac{\partial l}{\partial \xi^a}\xi^a + \frac{\partial l}{\partial \dot{r}^\alpha}\dot{r}^\alpha - l \end{aligned}$$

After restricting it to the submanifold $\mathcal{D}$, we have

$$\begin{aligned} E_{\mathcal{D}} &= \frac{\partial l}{\partial \xi^a}(-A^a_\alpha \dot{r}^\alpha + \Gamma^{ai}p_i) + \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} + A^a_\alpha \frac{\partial l}{\partial \xi^a}\right)\dot{r}^\alpha - l_c \\ &= \frac{\partial l}{\partial \xi^a}\Gamma^{ai}p_i + \frac{\partial l_c}{\partial \dot{r}^\alpha}\dot{r}^\alpha - l_c \end{aligned}$$

Therefore,

$$\begin{aligned} dE_{\mathcal{D}} &= \frac{\partial l}{\partial \xi^a}\left(\frac{\partial \Gamma^{ai}}{\partial r^\alpha}p_i dr^\alpha + \Gamma^{ai}dp_i\right) + \Gamma^{ai}p_i\left(\frac{\partial^2 l}{\partial r^\alpha \partial \xi^a}dr^\alpha + \frac{\partial^2 l}{\partial \dot{r}^\alpha \partial \xi^a}d\dot{r}^\alpha + \frac{\partial^2 l}{\partial p_j \partial \xi^a}dp_j\right) \\ &\quad + \dot{r}^\alpha\left(\frac{\partial^2 l_c}{\partial r^\beta \partial \dot{r}^\alpha}dr^\beta + \frac{\partial^2 l_c}{\partial \dot{r}^\beta \partial \dot{r}^\alpha}d\dot{r}^\beta + \frac{\partial^2 l_c}{\partial p_i \partial \dot{r}^\alpha}dp_i\right) - \frac{\partial l_c}{\partial r^\alpha}dr^\alpha - \frac{\partial l_c}{\partial p_i}dp_i. \end{aligned}$$

Furthermore,

$$\begin{aligned} X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{D}} &= g^a_f \xi^f \frac{\partial (g^{-1})^b_a}{\partial g^c}\frac{\partial l}{\partial \xi^b}dg^c - g^c_f \xi^f \frac{\partial (g^{-1})^b_a}{\partial g^c}\frac{\partial l}{\partial \xi^b}dg^a + g^a_f \xi^f (g^{-1})^b_a d\left(\frac{\partial l}{\partial \xi^b}\right) \\ &\quad - \left(\frac{\partial}{\partial r^\alpha}\left(\frac{\partial l}{\partial \xi^b}\right)\dot{r}^\alpha + \frac{\partial}{\partial \dot{r}^\alpha}\left(\frac{\partial l}{\partial \xi^b}\right)\ddot{r}^\alpha + \frac{\partial}{\partial p_i}\left(\frac{\partial l}{\partial \xi^b}\right)\dot{p}_i\right)(g^{-1})^b_a dg^a \\ &\quad + (\dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial p_i) \lrcorner \left(dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot{r}^\alpha}\right)\right) \\ &= \xi^f d\left(\frac{\partial l}{\partial \xi^f}\right) + \left(C^b_{fa}\frac{\partial l}{\partial \xi^b}\xi^f - \frac{d}{dt}\left(\frac{\partial l}{\partial \xi^a}\right)\right)(g^{-1})^a_e dg^e \\ &\quad + (\dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial p_i) \lrcorner \left(dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot{r}^\alpha}\right)\right). \end{aligned} \tag{3.38}$$

Clearly, both sides of the equation

$$X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{K}} = dE_{\mathcal{K}}$$

are G-invariant, and when restricted to subdistribution $\mathcal{W} \subset \mathcal{K}$, they vanish on the distribution $\mathcal{T} \cap \mathcal{K}$. This can be shown to be true either by invoking how $\mathcal{W}$ has been constructed or by direct calculation, noticing that when

$$\left(C^b_{fa} \frac{\partial l}{\partial \xi^b} \xi^f - \frac{d}{dt} \left(\frac{\partial l}{\partial \xi^a} \right) \right) (g^{-1})^a_e dg^e \tag{3.39}$$

is paired with $g^f_c e^c_i$ in $\mathcal{T} \cap \mathcal{K}$, it is equal to zero on $\mathcal{W}$. Hence both sides push down to $\overline{\mathcal{K}}$ where

$$\overline{X}_{\overline{\mathcal{K}}} = \dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial_{p_i},$$

with

$$\dot{p}_i = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_i] + \dot{e}_i \right\rangle.$$

To find the remaining reduced equations, notice that the restriction of (3.39) to the subdistribution spanned by $\{-g^a_b A^b_\alpha \partial_{g^a} + \partial_{r^\alpha}, \partial_{\dot{r}^\alpha}, \partial_{p_i}\}$ is equivalent to

$$-\left(C^b_{fa} \frac{\partial l}{\partial \xi^b} \xi^f - \frac{d}{dt} \left(\frac{\partial l}{\partial \xi^a} \right) \right) A^a_\alpha dr^\alpha. \tag{3.40}$$

If we compute

$$\begin{aligned} -\left(C^b_{fa} \frac{\partial l}{\partial \xi^b} \xi^f - \frac{d}{dt} \left(\frac{\partial l}{\partial \xi^a} \right) \right) A^a_\alpha dr^\alpha + \xi^a d \left(\frac{\partial l}{\partial \xi^a} \right) \\ + (\dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial p_i) \lrcorner \left(dr^\alpha \wedge d \left(\frac{\partial l}{\partial \dot{r}^\alpha} \right) \right) \end{aligned}$$

and equate its terms with the corresponding terms of $d\overline{E}_{\overline{\mathcal{K}}}$ which is the same as $dE_{\mathcal{K}}$, we have the following equations after some computations

$$\frac{d}{dt} \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} \right) - \frac{\partial l_c}{\partial r^\alpha} = -C^b_{da} \frac{\partial l}{\partial \xi^b} \xi^d A^a_\alpha - \frac{\partial l}{\partial \xi^a} \left(\dot{A}^a_\alpha - \frac{\partial A^a_\beta}{\partial r^\alpha} \dot{r}^\beta + \frac{\partial \Gamma^{ai} p_i}{\partial r^\alpha} \right).$$

After plugging in the constraint $\xi = -A\dot{r} + \Gamma p$ and simplify, we get the desired reduced equations

$$\frac{d}{dt} \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} \right) - \frac{\partial l_c}{\partial r^\alpha} = -\frac{\partial l}{\partial \xi^b} (B^b_{\alpha\beta} \dot{r}^\beta + F^{bi} p_i), \tag{3.41}$$

where

$$B^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac} A^a_\beta A^c_\alpha \tag{3.42}$$

$$F^{bi}_\alpha = \frac{\partial \Gamma^{bi}}{\partial r^\alpha} - C^b_{ad} A^a_\alpha \Gamma^{di}. \tag{3.43}$$

In an orthogonal body frame where we choose our moving basis $e_b(g,r)$ to be orthogonal, that is, the corresponding generators $[e_b(g,r)]_Q$ are orthogonal in the given kinetic energy metric (actually, all that is needed is that the vectors in the set of basis vectors corresponding to the subspace $\mathcal{S}_q$ be orthogonal to the remaining basis vectors), the momentum equation (3.36) can be written as (see [BKMM])

$$\frac{d}{dt}p_i = C^j_{hi} I^{hl} p_j p_l + \mathcal{D}^j_{i\alpha}\dot r^\alpha p_j + \mathcal{D}_{\alpha\beta i}\dot r^\alpha \dot r^\beta, \tag{3.44}$$

where

$$\begin{aligned} \mathcal{D}^j_{i\alpha} &= -C^j_{ai}A^a_\alpha + \gamma^j_{i\alpha} + \lambda_{a'\alpha}C^{a'}_{li}I^{lj} \\ \mathcal{D}_{\alpha\beta i} &= \lambda_{a'\alpha}(-C^{a'}_{ai}A^a_\beta + \gamma^{a'}_{i\beta}). \end{aligned}$$

Here $\gamma^c_{b\alpha}$ and $\lambda_{a'\alpha}$ are defined by

$$\begin{aligned} \frac{\partial e_b}{\partial r^\alpha} &= \gamma^c_{b\alpha}e_c \\ \lambda_{a'\alpha} &= \frac{\partial l}{\partial \xi^{a'}\partial \dot r^\alpha} - \frac{\partial l}{\partial \xi^{a'}\partial \xi^b}A^b_\alpha. \end{aligned}$$

Notice that while the summation range of $a, b, c, d...$ are over all Lie algebra element (1 to k). those over $i, j, l, ...$ are the restricted (constrained) range (1 to m) and those over $a', b', ...$ run from $m+1$ to k (which correspond to the symmetry directions not aligned with the constraints).

Similarly we can rewrite the above reduced Lagrange-d'Alembert equations (3.41) using the orthogonal body frame. Essentially, it is a change of basis. Instead of using the natural fixed basis $\{b_a\}$ where $(b_a)_Q = \partial_{g^a}$, we do all the computations in the orthogonal body frame $\{e_b\}$. With the abuse of notation, we shall still use $l(r,\dot r,\xi)$ and $l_c(r,\dot r,p)$ to denote the reduced Lagrangian and the constrained reduced Lagrangian (in the orthogonal body frame) respectively. But it should be clear that for the following computations, $\xi = \xi^b e_b$. Similar interpretation should apply to all other notations. Now let us compute the right hand side of the reduced equations in the new basis. Since

$$\begin{aligned} \frac{\partial}{\partial r^\beta}(A^a_\alpha e_a) &= \frac{\partial A^b_\alpha}{\partial r^\beta}e_b + A^a_\alpha \gamma^b_{a\beta}e_b \\ \frac{\partial}{\partial r^\alpha}(\Gamma^{bi}p_i e_b) &= \frac{\partial \Gamma^{bi}}{\partial r^\alpha}p_i e_b + \Gamma^{ai}\gamma^b_{a\alpha}p_i e_b, \end{aligned}$$

we have

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot r^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} &= -\frac{\partial l}{\partial \xi^b}\left(\frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha + A^c_\alpha \gamma^b_{c\beta} - A^c_\beta \gamma^b_{c\alpha}\right)\dot r^\beta \\ &\quad -\frac{\partial l}{\partial \xi^b}\left(\frac{\partial \Gamma^{bi} p_i}{\partial r^\alpha} - C^b_{ad}A^a_\alpha \Gamma^{di} p_i + \Gamma^{ci}\gamma^b_{c\alpha} p_i\right) \end{aligned}$$

Now applying Proposition 7.1 of [BKMM] to the above reduced equations and notice that in the orthogonal body basis, $\Gamma^{bi} = 0$ for any $b > m$ (recall $\Gamma^{ji} = I^{ji}$), we can rewrite the reduced equations in the following form

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot r^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -(\mathcal{K}^{jl}_\alpha p_j p_l + \mathcal{K}^j_{\alpha\beta}\dot r^\beta p_j + \mathcal{K}_{\alpha\beta\delta}\dot r^\beta \dot r^\delta) \tag{3.45}$$

where

$$\begin{aligned} \mathcal{K}^{jl}_\alpha &= \frac{\partial I^{jl}}{\partial r^\alpha} - C^j_{bh}A^b_\alpha I^{hl} + \gamma^j_{h\alpha} I^{hl} \\ \mathcal{K}^j_{\alpha\beta} &= \lambda_{a'\beta}(-C^{a'}_{bh}A^b_\alpha I^{hj} + \gamma^{a'}_{h\alpha} I^{hj}) + B^j_{\alpha\beta} \\ \mathcal{K}_{\alpha\beta\delta} &= \lambda_{a'\delta} B^{a'}_{\alpha\beta}. \end{aligned}$$

Here

$$B^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha + A^c_\alpha \gamma^b_{c\beta} - A^c_\beta \gamma^b_{c\alpha}. \quad \blacksquare$$

Remarks

1. A careful reading of the proof of Theorem 8 and the subsections 3.2 and 3.3 shows that the Hamiltonian reduction procedure still works as long as the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ is invertible. This is important because in some examples like the bicycle the Legendre transform $\mathbb{F}L$ is singular, but its restriction to the constraint submanifold $\mathcal{D}$ is invertible and the Hamiltonian reduction procedure is also applicable.

2. In many examples like the snakeboard and the bicycle, the constraints satisfy a special condition, namely, they involve only the velocities of the group variables $\dot g$ and are independent of the velocities of the shape variables $\dot r$ (see equations (3.12) and (3.13)). Under this special condition, the distribution $\mathcal{K}$ in equation (3.33) can be represented by

$$\mathcal{K} = \mathrm{span}\{g^a_d e^d_i \partial_{g^a}, \partial_r, \partial_{\dot r}, \partial_p\}. \tag{3.46}$$

This representation simplifies the computation for finding the reduced equations because the restriction of the one form (3.39) to the subdistribution $\overline{\mathcal{K}}$ spanned by $\{\partial_r, \partial_{\dot{r}}, \partial_p\}$ will equal to zero. Hence in pushing down $X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{D}}$ in (3.38) to $\overline{\mathcal{K}}$, we can simply omit the one form (3.39). In the following subsections, we will use this simplified procedure for the examples of the snakeboard and the bicycle. We will use a modified version of a nonholonomically constrained particle to illustrate the general procedure.

3. Since the momentum equation is central to the theory of nonholonomic mechanical systems with symmetry, we make a few additional remarks about it. Before that, we state the following proposition, the result of which is implicit in both [BKMM] and Ostrowski [1996].

Proposition 2 *For a nonholonomic mechanical system with symmetry, we have*

$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}\right)(\xi^q_Q)^i = \frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi^q_Q)^i\right) - \frac{\partial L}{\partial \dot{q}^i}\left(\frac{d}{dt}\xi^q\right)^i_Q$$

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

Proof: Choose a section of $\mathfrak{g}^{\mathcal{D}}$ and apply the chain rule to give

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(\xi^q_Q)^i\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi^q_Q)^i + \frac{\partial L}{\partial \dot{q}^i}\left((T\xi^q_Q \cdot \dot{q})^i + \left(\frac{d}{dt}\xi^q\right)^i_Q\right).$$

Invariance of the Lagrangian implies that

$$L(exp(s\xi^q)\cdot q, \exp(s\xi^q)\cdot \dot{q}) = L(q, \dot{q}).$$

Differentiating this expression and evaluating it at $s = 0$, we get

$$\frac{\partial L}{\partial \dot{q}^i}(\xi^q_Q)^i + \frac{\partial L}{\partial \dot{q}^i}(T\xi^q_Q \cdot \dot{q})^i = 0$$

After eliminating the term $\frac{\partial L}{\partial \dot{q}^i}(T\xi^q_Q \cdot \dot{q})^i$ from the above two equations, we arrive at the desired result. ■

The above equation can be rewritten as

$$\left\langle (dE - X \lrcorner \Omega_L)|_{\mathcal{D}}, (\xi^q_Q)' \right\rangle = \frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi^q_Q)^i\right) - \frac{\partial L}{\partial \dot{q}^i}\left(\frac{d}{dt}\xi^q\right)^i_Q,$$

where $(\xi_Q^q)' \in \mathcal{T} \cap \mathcal{K}$ and $T\tau_Q((\xi_Q^q)') = \xi_Q^q$. Since both the energy function E and the submanifold $\mathcal{D}$ are G-invariant, the left hand of the above equation reduces to $\Omega_{\mathcal{D}}(X_{\mathcal{D}}, (\xi_Q^q)')$ and hence any vector field $X_{\mathcal{D}}$ which takes values in $\mathcal{W} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{K})^{\perp}$ will make the left hand side zero and hence must satisfy the momentum equation (3.21)

$$\frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi_Q^q)^i\right) - \frac{\partial L}{\partial \dot{q}^i}\left(\frac{d}{dt}\xi^q\right)_Q^i = 0,$$

as we have already seen in the proof of Theorem 8.

In showing that the vector field $X_{\mathcal{H}}$ which satisfies the equation $X_{\mathcal{H}} \,\lrcorner\, \Omega_{\mathcal{H}} = dH_{\mathcal{H}}$ must lie in the subdistribution $\mathcal{U}$, one might think that any vector field $Y \in \mathcal{V} \cap \mathcal{H}$ can be expressed as a linear combination of infinitesimal generators (generated by fixed Lie algebra elements). But this is not the case, as we have pointed out earlier in the Lagrangian side, in general $(\xi_Q^q)'$ is the (vertical) lift of a section of the bundle $\mathcal{S}$ (generated by a section of the bundle $\mathfrak{g}^{\mathcal{D}}$). This is also true on the Hamiltonian side.

3.3.5 Example: The Snakeboard Revisited

Now we return to the snakeboard and discuss the role of the symmetry group $G = SE(2)$. Recall from our earlier discussion that the Lagrangian is

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

which is independent of the configuration of the board and hence it is invariant to all possible group actions.

The Constraint Submanifold. The condition of rolling without slipping gives rise to the constraint one forms

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

which are invariant under the $SE(2)$ action. The constraints determine the kinematic distribution $\mathcal{D}_q$:

$$\mathcal{D}_q = \text{span}\{\partial_\psi, \partial_\phi, a\partial_x + b\partial_y + c\partial_\theta\},$$

where $a = -2r\cos^2\phi\cos\theta, b = -2r\cos^2\phi\sin\theta, c = \sin 2\phi$. The tangent space to the orbits of the $SE(2)$ action is given by

$$T_q(\text{Orb}(q)) = \text{span}\{\partial_x, \partial_y, \partial_\theta\}$$

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

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \mathrm{span}\{a\partial_x + b\partial_y + c\partial_\theta\}.$$

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D} \cap T\mathrm{Orb}$ 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

$$\xi_Q^q = a\partial_x + b\partial_y + c\partial_\theta,$$

which is invariant under the action of $SE(2)$ on Q. The nonholonomic momentum is thus given by

$$\begin{aligned} p &= \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i \\ &= ma\dot{x} + mb\dot{y} + mr^2 c\dot{\theta} + J_0 c\dot{\psi}. \end{aligned}$$

The kinematic 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 the point $\phi = \pi/2$),

$$\begin{bmatrix} \cos\theta\dot{x} + \sin\theta\dot{y} \\ -\sin\theta\dot{x} + \cos\theta\dot{y} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} -\dfrac{J_0}{2mr}\sin 2\phi\dot{\psi} \\ 0 \\ \dfrac{J_0}{mr^2}\sin^2\phi\dot{\psi} \end{bmatrix} = \begin{bmatrix} \dfrac{-1}{2mr}p \\ 0 \\ \dfrac{\tan\phi}{2mr^2}p \end{bmatrix}. \tag{3.47}$$

These equations have the form

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

where

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

These are precisely the terms which appear in the nonholonomic connection relative to the (global) trivialization (r, g).

After applying the constrained Legendre transformation and its inverse to the constraint equations (3.47), we have

$$\begin{bmatrix} \cos\theta p_x + \sin\theta p_y \\ -\sin\theta p_x + \cos\theta p_y \\ p_\theta \end{bmatrix} + \begin{bmatrix} -\dfrac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi \\ 0 \\ -\dfrac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi \end{bmatrix} = \begin{bmatrix} \dfrac{-mr}{2(mr^2 - J_0\sin^2\phi)}p \\ 0 \\ \dfrac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}p \end{bmatrix},$$

where

$$p = -2r\cos^2\phi\cos\theta p_x - 2r\cos^2\phi\sin\theta p_y + \sin 2\phi p_\theta$$

and is $SE(2)$-invariant.

Therefore, the constraint submanifold $\mathcal{M} \subset T^*Q$ is defined by

$$\begin{aligned} p_x &= \frac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi\cos\theta - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}p\cos\theta \\ p_y &= \frac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi\sin\theta - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}p\sin\theta \\ p_\theta &= \frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}p \end{aligned}$$

It is a submanifold in T^*Q and we can use $(x, y, \theta, \psi, \phi, p_\psi, p_\phi, p)$ as its induced local coordinates.

The Distributions $\mathcal{H}, \mathcal{V} \cap \mathcal{H}$ and $\mathcal{U}$. With the induced coordinates, the distribution $\mathcal{H}$ on $\mathcal{M}$ is

$$\mathcal{H} = \operatorname{span}\{-2r\cos^2\phi\cos\theta\partial_x - 2r\cos^2\phi\sin\theta\partial_y + \sin 2\phi\partial_\theta, \partial_\psi, \partial_\phi, \partial_{p_\psi}, \partial_{p_\phi}, \partial_p\}$$

and the subdistribution $\mathcal{V} \cap \mathcal{H}$ is

$$\mathcal{V} \cap \mathcal{H} = \operatorname{span}\{-2r\cos^2\phi\cos\theta\partial_x - 2r\cos^2\phi\sin\theta\partial_y + \sin 2\phi\partial_\theta\}.$$

As for the subdistribution $\mathcal{U}$, we first calculate the two form $\Omega_\mathcal{M}$. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned}
\Omega_{\mathcal{M}} &= dx \wedge dp_x + dy \wedge dp_y + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \\
&= (\cos\theta dx + \sin\theta dy) \wedge \left(\frac{mr \sin 2\phi}{2(mr^2 - J_0 \sin^2\phi)} dp_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)} dp \right) \\
&\quad + (\cos\theta dx + \sin\theta dy) \wedge \frac{mr(mr^2 \cos 2\phi + J_0 \sin^2\phi)}{(mr^2 - J_0\sin^2\phi)^2} p_\psi d\phi \\
&\quad - (\cos\theta dx + \sin\theta dy) \wedge \frac{mr J_0 \sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2} p d\phi \\
&\quad + d\theta \wedge \left(\frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)} dp_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)} dp \right) \\
&\quad + d\theta \wedge \frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2} p_\psi d\phi \\
&\quad + d\theta \wedge + \frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2} p d\phi \\
&\quad + (-\sin\theta dx + \cos\theta dy) \wedge \left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)} p_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)} p \right) d\theta \\
&\quad + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi
\end{aligned}$$

Since $\mathcal{U} = (\mathcal{V} \cap \mathcal{H})^\perp = \ker\{(\mathcal{V}\cap\mathcal{H}) \lrcorner \Omega_{\mathcal{H}}\}$, we need to calculate $(\mathcal{V}\cap\mathcal{H}) \lrcorner \Omega_{\mathcal{M}}$, and restrict it to $\mathcal{H}$:

$$\begin{aligned}
&(\mathcal{V}\cap\mathcal{H}) \lrcorner \Omega_{\mathcal{H}} = \\
&\quad -2r\cos^2\phi \left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)} dp_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)} dp \right) \\
&\quad -2r\cos^2\phi \left(\frac{mr(mr^2\cos 2\phi + J_0\sin^2\phi)}{(mr^2 - J_0\sin^2\phi)^2} p_\psi d\phi - \frac{mr J_0 \sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2} p d\phi \right) \\
&\quad + \sin 2\phi \left(\frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)} dp_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)} dp \right) \\
&\quad + \sin 2\phi \frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2} p_\psi d\phi \\
&\quad + \sin 2\phi \frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2} p d\phi \\
&= dp - \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi} p_\psi d\phi + \frac{(mr^2 + J_0\cos 2\phi)\tan\phi}{mr^2 - J_0\sin^2\phi} p d\phi
\end{aligned}$$

Hence,

$$\mathcal{U} = \ker\left\{ dp - \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi} p_\psi d\phi + \frac{(mr^2 + J_0\cos 2\phi)\tan\phi}{mr^2 - J_0\sin^2\phi} p d\phi \right\}.$$

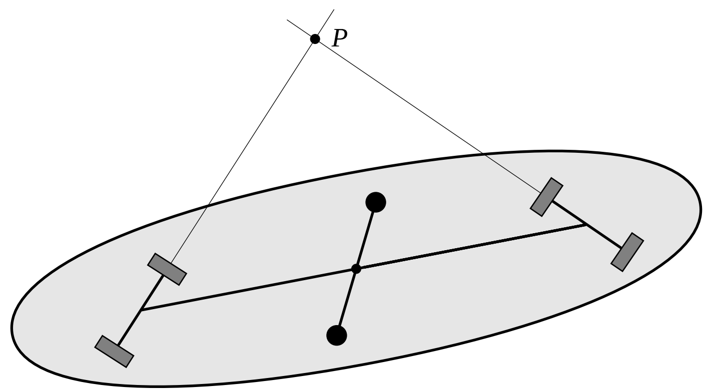

Figure 3.2: The momentum p is the angular momentum of the snakeboard system about the point P.

The Reconstruction and Momentum Equations A vector field $X_{\mathcal{U}}$ taking values in $\mathcal{U}$ must be of the form

$$X_{\mathcal{U}} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi} + \dot{p}\partial_p$$

where

$$\begin{aligned} \dot{x} &= \frac{J_0}{2mr}\sin 2\phi\dot{\psi}\cos\theta - \frac{1}{2mr}p\cos\theta \\ \dot{y} &= \frac{J_0}{2mr}\sin 2\phi\dot{\psi}\sin\theta - \frac{1}{2mr}p\sin\theta \\ \dot{\theta} &= -\frac{J_0}{mr^2}\sin^2\phi\dot{\psi} + \frac{\tan\phi}{2mr^2}p \end{aligned}$$

and

$$\dot{p} = \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}p_\psi\dot{\phi} - \frac{(mr^2 + J_0\cos 2\phi)\tan\phi}{mr^2 - J_0\sin^2\phi}p\dot{\phi} \tag{3.48}$$

The equations for $\dot{x}$, $\dot{y}$ and $\dot{\theta}$ are the same reconstruction equations as equations (3.47) and the last one for $\dot{p}$ is the momentum equation on the Hamiltonian side. As noted in [BKMM], the momentum p is the angular momentum of the system about the point P shown in figure 3.2.

It can be checked that the momentum equation (3.48) is equivalent to the equation (3.14) via a change of variables with

$$\begin{aligned} p &= -2r\cos^2\phi\cos\theta p_x - 2r\cos^2\phi\sin\theta p_y + \sin 2\phi p_\theta \\ &= \frac{2(mr^2 - J_0\sin^2\phi)\cot\phi}{mr^2 - J_0}p_\theta - \frac{2mr^2\cos^2\phi\cot\phi}{mr^2 - J_0}p_\psi \end{aligned}$$

as the key link. Similarly the two full sets of equations of motion in both section 3.2.5 and this section are also related in the same way.

The Reduced Hamilton Equations. To find the remaining reduced equations, we need to compute

$$X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} = dH_{\mathcal{M}},$$

restrict it to the subdistribution $\mathcal{U}$ and then push it down to the reduced constraint submanifold $\overline{\mathcal{M}}$. Let us first compute $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$

$$\begin{aligned}
X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} = &\\
&(\dot{x}\cos\theta + \dot{y}\sin\theta)\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}dp_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}dp\right)\\
&+(\dot{x}\cos\theta + \dot{y}\sin\theta)\left(\frac{mr(mr^2\cos 2\phi + J_0\sin^2\phi)}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi - \frac{mrJ_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right)\\
&+\dot{\theta}\left(\frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}dp_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}dp\right)\\
&+\dot{\theta}\left(\frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi + \frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right)\\
&+\dot{\psi}dp_\psi + \dot{\phi}dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi\\
&-\dot{\theta}\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}p_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}p\right)(-\sin\theta dx + \cos\theta dy)\\
&-mr\left(\frac{mr^2\cos 2\phi + J_0\sin^2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi\dot{\phi} - \frac{J_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}p\dot{\phi}\right)(\cos\theta dx + \sin\theta dy)\\
&-\left(\frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi\dot{\phi} + \frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2}p\dot{\phi}\right)d\theta\\
&-\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}\dot{p}_\psi(\cos\theta dx + \sin\theta dy) - \frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}\dot{p}_\psi d\theta\\
&+\frac{mr}{2(mr^2 - J_0\sin^2\phi)}(\cos\theta dx + \sin\theta dy) - \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}\dot{p}d\theta.
\end{aligned}$$

As for $dH_{\mathcal{H}}$, recall that the constrained Hamiltonian $H_{\mathcal{M}}$ is

$$H_{\mathcal{M}} = \frac{mr^2}{2(mr^2 - J_0)^2}\cot^2\phi(p_\theta - p_\psi)^2 + \frac{1}{2J_0}p_\psi^2 + \frac{1}{2(mr^2 - J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1}p_\phi^2.$$

Notice that $H_{\mathcal{M}}$ is $SE(2)$-invariant and hence $H_{\mathcal{M}} = h_{\overline{\mathcal{M}}}$ where

$$\begin{aligned}
h_{\overline{\mathcal{M}}} &= \frac{mr^2}{2}\left(\frac{1}{2(mr^2 - J_0\sin^2\phi)}p - \frac{\sin^2\phi}{2(mr^2 - J_0\sin^2\phi)}p_\psi\right)^2 + \frac{1}{2J_0}p_\psi^2\\
&\quad+\frac{mr^2 - J_0}{2}\left(\frac{\tan\phi}{2(mr^2 - J_0\sin^2\phi)}p - \frac{\sin^2\phi}{mr^2 - J_0\sin^2\phi}p_\psi\right)^2 + \frac{1}{4J_1}p_\phi^2.
\end{aligned}$$

Compute $dH_{\mathcal{M}} = dh_{\overline{\mathcal{M}}}$ and we have

$$\begin{aligned}
dh_{\overline{\mathcal{M}}} = &\\
&\frac{mr^2(p-\sin 2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}\left(\frac{1}{2(mr^2-J_0\sin^2\phi)}dp - \frac{\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}dp_\psi\right)\\
&+\frac{mr^2(p-\sin 2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}\left(pd\left(\frac{1}{2(mr^2-J_0\sin^2\phi)}\right) - p_\psi d\left(\frac{\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}\right)\right)\\
&+\frac{(mr^2-J_0)(\tan\phi p - 2\sin^2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}\left(\frac{\tan\phi}{2(mr^2-J_0\sin^2\phi)}dp - \frac{\sin^2\phi}{(mr^2-J_0\sin^2\phi)}dp_\psi\right)\\
&+\frac{(mr^2-J_0)(\tan\phi p - 2\sin^2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}pd\left(\frac{\tan\phi}{2(mr^2-J_0\sin^2\phi)}\right)\\
&-\frac{(mr^2-J_0)(\tan\phi p - 2\sin^2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}p_\psi d\left(\frac{\sin^2\phi}{(mr^2-J_0\sin^2\phi)}\right)\\
&+\frac{1}{J_0}p_\psi dp_\psi + +\frac{1}{2J_1}p_\phi dp_\phi.
\end{aligned}$$

It is easy to check that $X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}} = dH_{\mathcal{M}}$ is $SE(2)$-invariant, and vanishes on $\mathcal{V}\cap\mathcal{H}$ when restricted to $\mathcal{U}$. Hence both sides push down to $\overline{\mathcal{H}}$. The push down of $X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}}$ is given by

$$\begin{aligned}
X_{\overline{\mathcal{H}}} \lrcorner\, \Omega_{\overline{\mathcal{H}}} = &\\
&\left(\frac{J_0}{2mr}\sin(2\phi)\dot\psi - \frac{1}{2mr}p\right)\left(\frac{mr\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}dp_\psi - \frac{mr}{2(mr^2-J_0\sin^2\phi)}dp\right)\\
&+\left(\frac{J_0}{2mr}\sin(2\phi)\dot\psi - \frac{1}{2mr}p\right)\frac{mr(mr^2\cos 2\phi + J_0\sin^2\phi)}{(mr^2-J_0\sin^2\phi)^2}p_\psi d\phi\\
&-\left(\frac{J_0}{2mr}\sin(2\phi)\dot\psi - \frac{1}{2mr}p\right)\frac{mrJ_0\sin 2\phi}{2(mr^2-J_0\sin^2\phi)^2}pd\phi\\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot\psi + \frac{\tan\phi}{2mr^2}p\right)\left(\frac{mr^2\cos^2\phi}{(mr^2-J_0\sin^2\phi)}dp_\psi + \frac{(mr^2-J_0)\tan\phi}{2(mr^2-J_0\sin^2\phi)}dp\right)\\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot\psi + \frac{\tan\phi}{2mr^2}p\right)\frac{mr^2(J_0-mr^2)\sin 2\phi}{(mr^2-J_0\sin^2\phi)^2}p_\psi d\phi\\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot\psi + \frac{\tan\phi}{2mr^2}p\right)\frac{(mr^2-J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2-J_0\sin^2\phi)^2}pd\phi\\
&+\dot\psi dp_\psi + \dot\phi dp_\phi - \dot p_\psi d\psi - \dot p_\phi d\phi.
\end{aligned}$$

Equating the terms of $dh_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{M}}}$ with those of the push down of $X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}}$ gives the remaining reduced Hamilton equations:

$$\dot{\psi} = -\frac{\tan\phi}{2(mr^2 - J_0\sin^2\phi)}p + \frac{mr^2}{J_0(mr^2 - J_0\sin^2\phi)}p_\psi \tag{3.49}$$

$$\dot{\phi} = \frac{p_\phi}{2J_1} \tag{3.50}$$

$$\dot{p}_\psi = 0 \tag{3.51}$$

$$\dot{p}_\phi = 0. \tag{3.52}$$

Notice that both the momentum equation (3.48) and the above set of reduced equations are independent of the group elements of the symmetry group $SE(2)$. If we add in the set of reconstruction equations (3.47), we recover the full dynamics of the system, and in a form that is suitable for control theoretical purposes.

Finding the Reduced Equations on the Lagrangian Side As shown in the proof of Theorem 3.2, we can derive the reduced Lagrange-d'Alembert equations in two ways. Here we will first use the equations (3.41).

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -\frac{\partial l}{\partial \xi^b}(B^b_{\alpha\beta}\dot{r}^\beta + F^{bi}p_i), \tag{3.53}$$

where

$$\begin{aligned} B^b_{\alpha\beta} &= \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha \\ F^{bi}_\alpha &= \frac{\partial \Gamma^{bi}}{\partial r^\alpha} - C^b_{ad}A^a_\alpha \Gamma^{di}. \end{aligned}$$

From the Lagrangian L, we find the reduced Lagrangian

$$l(r, \dot{r}, \xi) = \frac{1}{2}m((\xi^1)^2 + (\xi^2)^2) + \frac{1}{2}mr^2(\xi^3)^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_0\dot{\psi}(\xi^3) + J_1\dot{\phi}^2,$$

where $\xi = g^{-1}\dot{g}$. After plugging in the constraints (3.47), we have the constrained reduced Lagrangian

$$l_c(r, \dot{r}, p) = -\frac{J_0^2}{2mr^2}\sin^2\phi\dot{\psi}^2 + \frac{1}{8mr^2}\sec^2\phi p^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_1\dot{\phi}^2.$$

Let us find all the ingredients of the above equations:

$$\begin{aligned} \frac{\partial l}{\partial \xi^1} &= m\xi^1 = m\left(\frac{J_0}{2mr}\sin 2\phi\dot{\psi} - \frac{1}{2mr}p\right) \\ \frac{\partial l}{\partial \xi^2} &= m\xi^2 = 0 \\ \frac{\partial l}{\partial \xi^3} &= mr^2\left(-\frac{J_0}{mr^2}\sin^2\phi\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right) + J_0\dot{\psi}; \end{aligned}$$

since $\frac{\partial l}{\partial \xi^2} = 0$, we do not need to compute $B^2_{\alpha\beta}$ and F^2_α (notice that $i = 1$). Also it is straightforward to find

$$\begin{aligned}
B^1_{12} &= \frac{\partial}{\partial\phi}\left(-\frac{J_0}{2mr}\sin 2\phi\right) = -\frac{J_0}{mr}\cos 2\phi \\
B^3_{12} &= \frac{\partial}{\partial\phi}\left(\frac{J_0}{mr}\sin^2\phi\right) = \frac{J_0}{mr}\sin 2\phi \\
F^3_2 &= \frac{\partial}{\partial\phi}\left(\frac{\tan\phi}{2mr^2}\right) = \frac{\sec^2\phi}{2mr^2},
\end{aligned}$$

and $F^1_1 = F^3_1 = F^1_2 = 0$. Substituting into (3.53), we get the reduced equations after some computations

$$\left(1 - \frac{J_0}{mr^2}\sin^2\phi\right)\ddot{\psi} = \frac{J_0}{2mr^2}\sin 2\phi\dot{\psi}\dot{\phi} - \frac{J_0}{2mr^2}\dot{\phi}p \tag{3.54}$$

$$J_1\ddot{\phi} = 0 \tag{3.55}$$

It is easy to check that these two equations are equivalent to the set of reduced equations (3.49)-(3.52) on the Hamiltonian side through the constrained Legendre transformation $\mathbb{F}L|\mathcal{D}$.

Next we will find the reduced equations use the equations (3.45)

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -(\mathcal{K}^{jl}_\alpha p_j p_l + \mathcal{K}^j_{\alpha\beta}\dot{r}^\beta p_j + \mathcal{K}_{\alpha\beta\delta}\dot{r}^\beta\dot{r}^\delta) \tag{3.56}$$

where

$$\begin{aligned}
\mathcal{K}^{jl}_\alpha &= \frac{\partial I^{jl}}{\partial r^\alpha} - C^j_{bh}A^b_\alpha I^{hl} + \gamma^j_{h\alpha}I^{hl} \\
\mathcal{K}^j_{\alpha\beta} &= \lambda_{a'\beta}(-C^{a'}_{bh}A^b_\alpha I^{hj} + \gamma^{a'}_{h\alpha}I^{hj}) + B^j_{\alpha\beta} \\
\mathcal{K}_{\alpha\beta\delta} &= \lambda_{a'\delta}B^{a'}_{\alpha\beta}.
\end{aligned}$$

Here

$$B^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha + A^c_\alpha\gamma^b_{c\beta} - A^c_\beta\gamma^b_{c\alpha}.$$

First we need to construct the orthogonal body frame. Recall that

$$(e_1(g,r))_Q = g^a_d e^d_1 \partial_{g^a} = -2r\cos^2\phi\cos\theta\partial_x - 2r\cos^2\phi\sin\theta\partial_y + \sin 2\phi\partial_\theta.$$

Hence

$$e_1 = -2r\cos^2\phi e_x + \sin 2\phi e_\theta,$$

where e_x, e_y, e_θ are the generators of $\partial_x, \partial_y, \partial_\theta$. Using the kinetic energy metric, we find

$$\begin{aligned} e_2 &= -\frac{1}{m}\sin\phi e_x + \frac{1}{m}\cos\phi e_y - \frac{1}{mr}\cos\phi e_\theta \\ e_3 &= \frac{1}{m}\sin\phi e_x + \frac{1}{m}\cos\phi e_y + \frac{1}{mr}\cos\phi e_\theta \end{aligned}$$

Recall that we only need e_1 to be orthogonal to e_2 and e_3.

Let η^b be the components of ξ in the new basis, i.e., $\xi = \xi^1 e_x + \xi^2 2_y + \xi^3 e_\theta = \eta^a e_a$, then

$$\begin{aligned} \xi^1 &= -2r\cos^2\phi\eta^1 - \frac{1}{m}\sin\phi\eta^2 + \frac{1}{m}\sin\phi\eta^3 \\ \xi^2 &= \frac{1}{m}\cos\phi\eta^2 + \frac{1}{m}\cos\phi\eta^3 \\ \xi^1 &= \sin 2\phi\eta^1 - \frac{1}{mr}\cos\phi\eta^2 + \frac{1}{mr}\cos\phi\eta^3, \end{aligned}$$

and $\bar{l}(r, \dot{r}, \eta^a) = l(r, \dot{r}, T_a^b \eta^a)$ where T_b^a is defined as above by $\xi^b = T_a^b \eta^a$.

Notice that in the new basis, the constraints (3.47) become

$$\begin{bmatrix} \eta^1 \\ \eta^2 \\ \eta^3 \end{bmatrix} = -\begin{bmatrix} \frac{J_0}{2mr^2}\tan\phi\dot{\psi} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{4mr^2}\sec^2\phi p \\ 0 \\ 0 \end{bmatrix}, \tag{3.57}$$

but the constrained reduced equation $\bar{l}_c(r, \dot{r}, p)$ remains the same and is equal to $l_c(r, \dot{r}, p)$.

Let us find all the ingredients of equations (3.56). After finding from (3.57) that

$$\begin{aligned} A_1^1 &= \frac{J_0}{2mr^2}\tan\phi \\ I^{11} &= \frac{1}{4mr^2}\sec^2\phi \end{aligned}$$

and the rest of A_α^b equal to zero (which is not true in general), it is straightforward to calculate

$$\begin{aligned}
\mathcal{K}_1^{11} &= 0 \\
\mathcal{K}_2^{11} &= \frac{1}{4mr^2}\sec^2\phi\tan\phi \\
\mathcal{K}_{11}^{1} &= 0 \\
\mathcal{K}_{12}^{1} &= \frac{J_0}{2mr^2} \\
\mathcal{K}_{21}^{1} &= 0 \\
\mathcal{K}_{22}^{1} &= 0 \\
\mathcal{K}_{121} &= \frac{J_0^2}{2mr^2}\sin 2\phi \\
\mathcal{K}_{122} &= 0
\end{aligned}$$

After substituting into (3.56) we get the same reduced equations as (3.54) and (3.55).

3.3.6 Example: The Bicycle

Control of the bicycle is a rich problem offering a number of considerable challenges of current research interest in the area of mechanical and robotic control. The bicycle is an underactuated system, subject to nonholonomic contact constraints associated with the rolling constraints on the front and rear wheels. It is unstable (except under certain combinations of fork geometry and speed) when not controlled. It is also, when considered to traverse flat ground, a system subject to symmetries; its Lagrangian and constraints are invariant with respect to translations and rotations in the ground plane.

Here a simplified bicycle model will be considered. The wheels of the bicycle are considered to have negligible inertia moments, mass, radii, and width, and roll without side or longitudinal slip. The vehicle is assumed to have a fixed steering axis that is perpendicular to the flat ground when the bicycle is upright. For simplicity we concern ourselves with a point mass bicycle. The rigid frame of the bicycle will be assumed to be symmetric about a plane containing the rear wheel.

Consider a ground fixed inertial reference frame with x and y axis in the ground plane and z-axis perpendicular to the ground plane in the direction opposite to gravity. The intersection of the vehicle's plane of symmetry with the ground plane forms a contact line. The contact line is rotated about the z-direction by a yaw angle θ. The contact line is considered directed, with its positive direction from the rear to the front of the vehicle. The

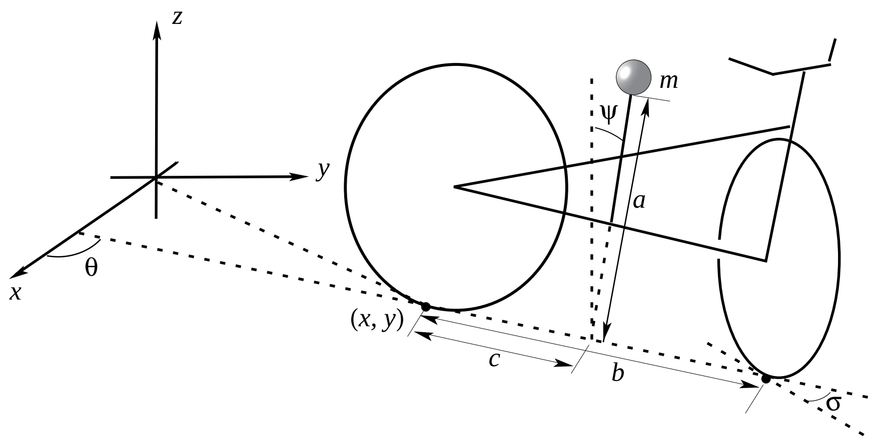

Figure 3.3: Notation for the bike.

yaw angle θ is zero when the contact line is in the x-direction. The angle that the bicycle's plane of symmetry makes with the vertical direction is the roll angle $\psi \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Front and rear wheel contacts are constrained to have velocities parallel to the lines of intersection of their respective wheel planes and the ground plane, but free to turn about an axis through the wheel/ground contact and parallel to the z-axis. Let $\sigma \in (-\frac{\pi}{2}, \frac{\pi}{2})$ be the steering angle between the front wheel plane/ground plane intersection and the contact line. With σ we associate a moment of inertia J which depends both on ψ and σ. We will parametrize the steering angle by $\phi := \tan\sigma/b$. For more details, see Getz and Marsden [1995] and Getz [1996]. See figure 3.3.

The configuration space is $Q = SE(2) \times S^1 \times S^1$ and the Lagrangian $L : TQ \to \mathbb{R}$ is the total kinetic energy minus potential energy of the system and is given by

$$
\begin{aligned}
L \;=\; & -mga\cos\psi + \frac{1}{2}J(\psi,\phi)\dot{\phi}^2 \\
& + \frac{m}{2}(\cos\theta\dot{x} + \sin\theta\dot{y} + a\sin\psi\dot{\theta})^2 \\
& + \frac{m}{2}\left((-\sin\theta\dot{x} + \cos\theta\dot{y} - a\cos\psi\dot{\psi} + c\dot{\theta})^2 + (-a\sin\psi\dot{\psi})^2\right)
\end{aligned}
$$

where m is the mass of the bicycle, considered for simplicity to be a point mass, and $J(\psi,\phi)$ is the moment of inertia associated with the steering action. The nonholonomic constraints associated with the front and rear wheels, assumed to roll without slipping, are expressed by

$$\begin{aligned} \dot{\theta} - \phi(\cos\theta\dot{x} + \sin\theta\dot{y}) &= 0 \\ -\sin\theta\dot{x} + \cos\theta\dot{y} &= 0. \end{aligned}$$

Clearly both the Lagrangian and the constraints are invariant under the $SE(2)$ action.

Notice that the Legendre transform $\mathbb{F}L$ is singular but by the remark following Theorem 8 the Hamiltonian procedure still works because the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ is invertible.

The Constraint Submanifold The constraints above give rise to the constraint one forms

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

which determine the kinematic distribution $\mathcal{D}_q$:

$$\mathcal{D}_q = \text{span}\{\partial_\psi, \partial_\phi, \cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}.$$

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

$$T_q(\text{Orb}(q)) = \text{span}\{\partial_x, \partial_y, \partial_\theta\},$$

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

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\text{Orb}(q)) = \text{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}.$$

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

$$\xi_Q^q = \cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta,$$

which is invariant under the action of $SE(2)$ on Q. The nonholonomic momentum map is thus given by

$$
\begin{aligned}
p &= \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i \\
&= m(\dot{x} + a\sin\psi\cos\theta\dot{\theta} + a\cos\psi\sin\theta\dot{\psi} - c\sin\theta\dot{\theta})\cos\theta \\
&\quad +m(\dot{y} + a\sin\psi\sin\theta\dot{\theta} - a\cos\psi\cos\theta\dot{\psi} + c\cos\theta\dot{\theta})\sin\theta \\
&\quad +m(\cos\theta\dot{x} + \sin\theta\dot{y} + a\sin\psi\dot{\theta})a\phi\sin\psi \\
&\quad +m(-\sin\theta\dot{x} + \cos\theta\dot{y} - a\cos\psi\dot{\psi} + c\dot{\theta})c\phi.
\end{aligned}
$$

The kinematic constraints plus the momentum are given by

$$
\begin{aligned}
0 &= \xi^3 - \phi\xi^1 \\
0 &= \xi^2 \\
p &= m(\xi^1 + a\sin\psi\xi^3) + ma\phi\sin\psi(\xi^1 + a\sin\psi\xi^3) \\
&\quad m\phi(c\xi^2 - ca\cos\psi\dot{\psi} + c^2\xi^3)
\end{aligned}
$$

where

$$
\begin{aligned}
\xi^1 &= \cos\theta\dot{x} + \sin\theta\dot{y} \\
\xi^2 &= -\sin\theta\dot{x} + \cos\theta\dot{y} \\
\xi^3 &= \dot{\theta}
\end{aligned}
$$

Adding, subtracting, and scaling these equations, we can write

$$
\begin{bmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{bmatrix} + \begin{bmatrix} -\dfrac{ca\phi\cos\psi}{K}\dot{\psi} \\ 0 \\ -\dfrac{ca\phi^2\cos\psi}{K}\dot{\psi} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{mK}p \\ 0 \\ \dfrac{\phi}{mK}p \end{bmatrix} \tag{3.58}
$$

where

$$
K = (1 + a\phi\sin\psi)^2 + c^2\phi^2. \tag{3.59}
$$

These equations have the form

$$
g^{-1}\dot{g} + A(r)\dot{r} = \Gamma(r)p.
$$

Next find the Legendre transform $\mathbb{F}L$ and restrict it to the constraint submanifold $\mathcal{D} \subset TQ$, we get

$$\begin{aligned} p_x &= m(1 + a\phi \sin\psi)\xi^1 \cos\theta - m(c\phi\xi^1 - a\cos\psi\dot{\psi})\sin\theta \\ p_y &= m(1 + a\phi \sin\psi)\xi^1 \sin\theta + m(c\phi\xi^1 - a\cos\psi\dot{\psi})\cos\theta \\ p_\theta &= ma\sin\psi(1 + a\phi\sin\psi)\xi^1 + m(c^2\phi\xi^1 - ca\cos\psi\dot{\psi}) \\ p_\psi &= ma^2\dot{\psi} - mac\cos\psi\phi\xi^1 \\ p_\phi &= J(\psi,\phi)\dot{\phi}. \end{aligned}$$

After applying the constrained Legendre transformation $\mathbb{F}L|\mathcal{D}$ and its inverse to the constraint equations (3.58), we have

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} -\dfrac{c\phi\cos\psi(1 + a\phi\sin\psi)}{F}\dfrac{p_\psi}{a} \\ \dfrac{(1 + a\phi\sin\psi)^2\cos\psi}{F}\dfrac{p_\psi}{a} \\ \dfrac{c\cos\psi(1 + a\phi\sin\psi)}{F}\dfrac{p_\psi}{a} \end{bmatrix} = \begin{bmatrix} \dfrac{1 + a\phi\sin\psi}{F}p \\ \dfrac{c\phi\sin^2\psi}{F}p \\ \dfrac{(1 + a\phi\sin\psi)a\sin\psi + c^2\phi\sin^2\psi}{F}p \end{bmatrix},$$

where

$$\begin{aligned} \mu_1 &= \cos\theta p_x + \sin\theta p_y \\ \mu_2 &= -\sin\theta p_x + \cos\theta p_y \\ \mu_3 &= p_\theta \end{aligned}$$

and

$$F = (1 + a\phi\sin\psi)^2 + c^2\phi^2\sin^2\psi \tag{3.60}$$

$$p = p_x\cos\theta + p_y\sin\theta + p_\theta\phi. \tag{3.61}$$

Therefore, the constraint submanifold $\mathcal{M} \subset T^*Q$ is defined by

$$\begin{aligned} p_x &= \mu_1\cos\theta - \mu_2\sin\theta \\ p_y &= \mu_1\sin\theta + \mu_2\cos\theta \\ p_\theta &= \mu_3. \end{aligned}$$

It is a submanifold in T^*Q and we can use $(x, y, \theta, \psi, \phi, p_\psi, p_\phi, p)$ as its induced local coordinates.

The Distributions $\mathcal{H}, \mathcal{V}\cap\mathcal{H}$ and $\mathcal{U}$. Using the induced coordinates, the distribution $\mathcal{H}$ on $\mathcal{M}$ is

$$\mathcal{H} = \mathrm{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta, \partial_\psi, \partial_\phi, \partial_{p_\psi}, \partial_{p_\phi}, \partial_p\}$$

and the subdistribution $\mathcal{V}\cap\mathcal{H}$ is

$$\mathcal{V}\cap\mathcal{H} = \mathrm{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}.$$

Notice that in the case of the bicycle, the constraints are independent of the velocities of the shape variables and hence the simplified procedure employed in the snakeboard is also used here.

As for the subdistribution $\mathcal{U}$, we first calculate the two form $\Omega_{\mathcal{M}}$. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned}\Omega_{\mathcal{M}} &= dx\wedge dp_x + dy\wedge dp_y + d\theta\wedge dp_\theta + d\psi\wedge dp_\psi + d\phi\wedge dp_\phi \\ &= (\cos\theta dx + \sin\theta dy)\wedge d\mu_1 + \mu_1(-\sin\theta dx + \cos\theta dy)\wedge d\theta \\ &\quad +(-\sin\theta dx + \cos\theta dy)\wedge d\mu_2 - \mu_2(\cos\theta dx + \sin\theta dy)\wedge d\theta \\ &\quad +d\theta\wedge d\mu_3 + d\psi\wedge dp_\psi + d\phi\wedge dp_\phi\end{aligned}$$

Since $\mathcal{U} = (\mathcal{V}\cap\mathcal{H})^\perp = \ker\{(\mathcal{V}\cap\mathcal{H})\,\lrcorner\,\Omega_{\mathcal{H}}\}$, we need to calculate $(\mathcal{V}\cap\mathcal{H})\,\lrcorner\,\Omega_{\mathcal{M}}$, and restrict it to $\mathcal{H}$:

$$\begin{aligned}(\mathcal{V}\cap\mathcal{H})\,\lrcorner\,\Omega_{\mathcal{H}} &= d\mu_1 - \mu_1\phi(-\sin\theta dx + \cos\theta dy) \\ &\quad -\mu_2 d\theta + \mu_2\phi(\cos\theta dx + \sin\theta dy) + \phi d\mu_3 \\ &= d\mu_1 + \phi d\mu_3 \\ &= dp + \frac{c\cos\psi(1 + a\phi\sin\psi)}{F}\frac{p_\psi}{a}d\phi - \frac{a\sin\psi(1 + a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}pd\phi.\end{aligned}$$

Hence,

$$\mathcal{U} = \ker\left\{dp + \frac{c\cos\psi(1 + a\phi\sin\psi)}{F}\frac{p_\psi}{a}d\phi - \frac{a\sin\psi(1 + a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}pd\phi\right\}. \tag{3.62}$$

The Reconstruction and Momentum Equations A vector field $X_{\mathcal{U}}$ taking values in $\mathcal{U}$ must be of the form

$$X_{\mathcal{U}} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi} + \dot{p}\partial_p$$

where

$$\begin{aligned}\dot{x} &= \xi^1\cos\theta - \xi^2\sin\theta = \left(\frac{ca\phi\cos\psi}{K}\dot{\psi} + \frac{1}{mK}p\right)\cos\theta \\ \dot{y} &= \xi^1\sin\theta + \xi^2\cos\theta = \left(\frac{ca\phi\cos\psi}{K}\dot{\psi} + \frac{1}{mK}p\right)\sin\theta \\ \dot{\theta} &= \phi\xi^1 = \left(\frac{ca\phi^2\cos\psi}{K}\dot{\psi} + \frac{\phi}{mK}p\right)\end{aligned}$$

and

$$\dot{p} = -\frac{c\cos\psi(1+a\phi\sin\psi)}{F}\frac{p_\psi}{a}\dot{\phi} + \frac{a\sin\psi(1+a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}p\dot{\phi}. \tag{3.63}$$

The equations for $\dot{x}$, $\dot{y}$ and $\dot{\theta}$ are the same reconstruction equations as equations (3.58) and the last one for $\dot{p}$ is the momentum equation on the Hamiltonian side. Similar to the example of the snakeboard, the momentum p equals the angular momentum of the system about a fixed point P that can be determined in the same way as in the case of the snakeboard. Notice also that the last equation can be written simply as $\dot{p} = \mu_3\dot{\phi}$.

The Reduced Hamilton Equations. To find the remaining reduced equations, we need to compute

$$X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}} = dH_{\mathcal{M}},$$

restrict it to the subdistribution $\mathcal{U}$ and then push it down to the reduced constraint submanifold $\overline{\mathcal{M}}$. Let us first compute $X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}}$

$$\begin{aligned}X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}} = \\ &(\cos\theta\dot{x} + \sin\theta\dot{y})d\mu_1 + \mu_1(-\sin\theta\dot{x} + \cos\theta\dot{y})d\theta - \mu_1\dot{\theta}(-\sin\theta dx + \cos\theta dy) \\ &+(-\sin\theta\dot{x} + \cos\theta\dot{y})d\mu_2 - \mu_2(\cos\theta\dot{x} + \sin\theta\dot{y})d\theta + \mu_2\dot{\theta}(\cos\theta dx + \sin\theta dy) \\ &+\dot{\theta}d\mu_3 + \dot{\psi}dp_\psi + \dot{\phi}dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi \\ &-((\dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi} + \dot{p}\partial_p) \lrcorner\, d\mu_1)(\cos\theta dx + \sin\theta dy) \\ &-((\dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi} + \dot{p}\partial_p) \lrcorner\, d\mu_2)(-\sin\theta dx + \cos\theta dy)\end{aligned}$$

As for $dH_{\mathcal{H}}$, we can find the constrained Hamiltonian $H_{\mathcal{M}}$ via the constrained Legendre transform and have

$$\begin{aligned}H_{\mathcal{M}} &= mga\cos\psi + \frac{1}{2J}p_\phi^2 + \\ &\frac{1}{2m}\left(\mu_1^2 + \mu_2^2 + \left(\frac{K\sin\psi}{F}\frac{p_\psi}{a} + \frac{c\phi\sin\psi\cos\psi}{F}p\right)^2\right).\end{aligned}$$

Notice that $H_{\mathcal{M}}$ is $SE(2)$-invariant and hence $H_{\mathcal{M}} = h_{\overline{\mathcal{M}}}$. Compute $dH_{\mathcal{M}} = dh_{\overline{\mathcal{M}}}$ and we have

$$\begin{aligned} dh_{\overline{\mathcal{M}}} = & \\ & -mga\sin\psi d\psi + \frac{1}{J}p_\phi dp_\phi - \frac{1}{2J^2}p_\phi^2\left(\frac{\partial J}{\partial\psi}d\psi + \frac{\partial J}{\partial\phi}d\phi\right) \\ & +\frac{1}{m}\left(\mu_1 d\mu_1 + \mu_2 d\mu_2\right) \\ & +\frac{1}{m}\left(\frac{K\sin\psi}{F}\frac{p_\psi}{a} + \frac{c\phi\sin\psi\cos\psi}{F}p\right) d\left(\frac{K\sin\psi}{F}\frac{p_\psi}{a} + \frac{c\phi\sin\psi\cos\psi}{F}p\right). \end{aligned}$$

It can be checked that $X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}} = dH_{\mathcal{M}}$ is $SE(2)$-invariant, and vanishes on $\mathcal{V}\cap\mathcal{H}$ when restricted to $\mathcal{U}$. Hence both sides push down to $\overline{\mathcal{H}}$. The push down of $X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}}$ is given by

$$\begin{aligned} X_{\overline{\mathcal{H}}} \lrcorner\, \Omega_{\overline{\mathcal{H}}} &= (\cos\theta\dot{x} + \sin\theta\dot{y})d\mu_1 + \dot{\theta}d\mu_3 + \dot{\psi}dp_\psi + \dot{\phi}dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi \\ &= \xi^1 d\mu_1 + \xi^3 d\mu_3 + \dot{\psi}dp_\psi + \dot{\phi}dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi \end{aligned}$$

Equating the terms of $dh_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{M}}}$ with those of the push down of $X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{M}}$ gives the remaining reduced Hamilton equations:

$$\begin{aligned} \dot{\psi} &= \frac{1}{ma}\left(\frac{K}{F}\frac{p_\psi}{a} + \frac{c\phi\cos\psi}{F}p\right) \\ \dot{\phi} &= \frac{p_\phi}{J} \\ \dot{p}_\psi &= mga\sin\psi + \frac{1}{2J^2}p_\phi^2\frac{\partial J}{\partial\psi} + m(1 + a\phi\sin\psi)a\phi\cos\psi(\xi^1)^2 + mca\phi\sin\psi\xi^1\dot{\psi} \\ \dot{p}_\phi &= \frac{1}{2J^2}\frac{\partial J}{\partial\phi}p_\phi^2, \end{aligned}$$

where

$$\xi^1 = \frac{c\phi\cos\psi}{K}\dot{\psi} + \frac{1}{mK}p = \frac{c\phi\cos\psi}{mF}\frac{p_\psi}{a} + \frac{1}{mF}p$$

as defined earlier in (3.58). The first two equations are nothing but the inverse of the constrained Legendre transform. Notice that both the momentum equation (3.63) and the above set of reduced equations are independent of the group elements of the symmetry group $SE(2)$. If we add in the set of reconstruction equations (3.58), we recover the full dynamics of the system, and in a form that is suitable for control theoretical purposes.

3.3.7 Example: A Nonholonomically Constrained Particle

In [BS], the example of a nonholonomically constrained particle has been used to illustrate its theory. Here, we would like to modify this example slightly in order to show concretely what need to be done to find the reduced equations of motion if the constraints involve also the velocities of the shape variables.

Consider a particle with the Lagrangian

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

and the nonholonomic constraint

$$\dot{z} = y\dot{x} + \dot{y}.$$

The constraint and the Lagrangian are invariant under the $\mathbb{R}^2$ action on $\mathbb{R}^3$ given by

$$(x, y, z) \mapsto (x + \lambda, y, z + \mu).$$

Notice that in the original example used in [BS], the constraint does not involve the $\dot{y}$-term and hence it also satisfies the special condition that the constraints are independent of the velocities of the shape variables. But the slight modification changes all these.

The Constraint Submanifold The constraint above gives rise to the constraint one form

$$\omega(q) = dz - ydx - dy.$$

The tangent space to the orbits of this group action is given by

$$T_q(\mathrm{Orb}(q)) = \mathrm{span}\{\partial_x, \partial_z\},$$

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

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \mathrm{span}\{\partial_x + y\partial_z\}.$$

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D} \cap T\mathrm{Orb}$ 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

$$\xi_Q^q = \partial_x + y\partial_z.$$

The nonholonomic momentum map is thus given by

$$p = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i = \dot{x} + y\dot{z}.$$

The kinematic constraint plus the momentum are

$$\begin{aligned} -y\dot{x} + \dot{z} &= \dot{y} \\ \dot{x} + y\dot{z} &= p. \end{aligned}$$

Solving for $\dot{x}$ and $\dot{z}$, we get

$$\begin{aligned} \dot{x} &= -\frac{y}{1+y^2}\dot{y} + \frac{1}{1+y^2}p \\ \dot{z} &= \frac{1}{1+y^2}\dot{y} + \frac{y}{1+y^2}p. \end{aligned}$$

After applying the constrained Legendre transform, we find that the constraint submanifold $\mathcal{M} \subset T^*Q$ is defined by

$$p_x = -\frac{y}{1+y^2}p_y + \frac{1}{1+y^2}p \tag{3.64}$$

$$p_z = \frac{1}{1+y^2}p_y + \frac{y}{1+y^2}p. \tag{3.65}$$

It is a submanifold in T^*Q and we can use (x, y, z, p_y, p) as its induced local coordinates.

The Distributions $\mathcal{H}, \mathcal{V} \cap \mathcal{H}$ and $\mathcal{U}$. With the induced coordinates, the distribution $\mathcal{H}$ on $\mathcal{M}$ is

$$\mathcal{H} = \operatorname{span}\left\{\partial_x + y\partial_z, -\frac{y}{1+y^2}\partial_x + \frac{1}{1+y^2}\partial_z + \partial_y, \partial_{p_y}, \partial_p\right\} \tag{3.66}$$

Notice that we are using $-g^a_b A^b_\alpha \partial_{g^a} + \partial_{r^\alpha}$, i.e, $-\frac{y}{1+y^2}\partial_x + \frac{1}{1+y^2}\partial_z + \partial_y$ instead of ∂_y. In fact, ∂_y does not even lie in the distribution $\mathcal{H}$.

The subdistribution $\mathcal{V} \cap \mathcal{H}$ is

$$\mathcal{V} \cap \mathcal{H} = \operatorname{span}\{\partial_x + y\partial_x\}. \tag{3.67}$$

As for the subdistribution $\mathcal{U}$, we first calculate the two form $\Omega_{\mathcal{M}}$. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned} \Omega_{\mathcal{M}} &= dx \wedge dp_x + dy \wedge dp_y + dz \wedge dp_z \\ &= dx \wedge d\left(-\frac{y}{1+y^2}p_y + \frac{1}{1+y^2}p\right) \\ &\quad + dy \wedge dp_y + dz \wedge d\left(\frac{1}{1+y^2}p_y + \frac{y}{1+y^2}p\right). \end{aligned}$$

Since $\mathcal{U} = (\mathcal{V} \cap \mathcal{H})^{\perp} = \ker\{(\mathcal{V} \cap \mathcal{H}) \lrcorner \Omega_{\mathcal{H}}\}$, we need to calculate $(\mathcal{V} \cap \mathcal{H}) \lrcorner \Omega_{\mathcal{M}}$, and restrict it to $\mathcal{H}$:

$$\begin{aligned}(\partial_x + y\partial_z) \lrcorner \Omega_{\mathcal{H}} &= -\frac{y}{1+y^2}dp_y - \frac{1-y^2}{(1+y^2)^2}p_y dy + \frac{1}{1+y^2}dp - \frac{2y}{(1+y^2)^2}pdy \\ &\quad + \frac{y}{1+y^2}dp_y - \frac{2y^2}{(1+y^2)^2}p_y dy + \frac{y^2}{1+y^2}dp + \frac{y(1-y^2)}{(1+y^2)^2}pdy \\ &= dp - \frac{1}{1+y^2}p_y dy - \frac{y}{1+y^2}pdy\end{aligned}$$

Hence,

$$\mathcal{U} = \ker\left\{dp - \frac{1}{1+y^2}p_y dy - \frac{y}{1+y^2}pdy\right\}. \tag{3.68}$$

The Reconstruction and Momentum Equations A vector field $X_{\mathcal{U}}$ taking values in $\mathcal{U}$ must be of the form

$$X_{\mathcal{U}} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{z}\partial_z + \dot{p}_y\partial_{p_y} + \dot{p}\partial_p$$

where

$$\dot{x} = -\frac{y}{1+y^2}\dot{y} + \frac{1}{1+y^2}p \tag{3.69}$$

$$\dot{z} = \frac{1}{1+y^2}\dot{y} + \frac{y}{1+y^2}p. \tag{3.70}$$

and

$$\dot{p} - \frac{1}{1+y^2}p_y\dot{y} - \frac{y}{1+y^2}p\dot{y} = 0 \tag{3.71}$$

The first set are the reconstruction equations and the last one is the momentum equation on the Hamiltonian side.

The Reduced Hamilton Equations. To find the remaining reduced equations, we need to compute

$$X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} = dH_{\mathcal{M}}, \tag{3.72}$$

restrict it to the subdistribution $\mathcal{U}$ and then push it down to the reduced constraint submanifold $\overline{\mathcal{M}}$. Let us first compute $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$

$$\begin{aligned} X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} &= \left(\frac{1-y^2}{(1+y^2)^2}\dot{y}p_y + \frac{2y(1-y^2)}{(1+y^2)^2}\dot{y}p + \frac{y}{1+y^2}\dot{p}_y - \frac{1}{1+y^2}\dot{p}\right) dx \\ &+ \left(\frac{2y}{(1+y^2)^2}\dot{y}p_y - \frac{1-y^2}{(1+y^2)^2}\dot{y}p - \frac{1}{1+y^2}\dot{p}_y - \frac{y}{1+y^2}\dot{p}\right) dz \\ &+ \left(-\frac{1-y^2}{(1+y^2)^2}\dot{x}p_y - \frac{2y}{(1+y^2)^2}\dot{x}p - \frac{2y}{(1+y^2)^2}\dot{z}p_y + \frac{1-y^2}{(1+y^2)^2}\dot{z}p - \dot{p}_y\right) dy \\ &+ \left(-\frac{y}{1+y^2}\dot{x} + \dot{y} + \frac{1}{1+y^2}\dot{z}\right) dp_y + \left(\frac{1}{1+y^2}\dot{x} + \frac{y}{1+y^2}\dot{z}\right) dp \end{aligned}$$

Notice that in pushing down $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$, we cannot simply just throw away the terms involving dx and dz, instead we have to replace them by $-A^a_\alpha dr^\alpha$, i.e.,by $-\frac{y}{1+y^2}dy$ and $\frac{1}{1+y^2}dy$ respectively, as it has been done in the proof of Theorem 8.

As for $dH_{\mathcal{H}}$, we first find the constrained Hamiltonian $H_{\mathcal{M}}$

$$\begin{aligned} H_{\mathcal{M}} &= \frac{1}{2}\left(\left(-\frac{y}{1+y^2}p_y + \frac{1}{1+y^2}p\right)^2 + p_y^2 + \left(\frac{1}{1+y^2}p_y + \frac{y}{1+y^2}p\right)^2\right) \\ &= \frac{1}{2}\left(\frac{p^2}{1+y^2} + p_y^2 + \frac{p_y^2}{1+y^2}\right) \end{aligned}$$

Clearly $H_{\mathcal{M}}$ is $\mathbb{R}^2$-invariant and hence $H_{\mathcal{M}} = h_{\overline{\mathcal{M}}}$. Compute $dH_{\mathcal{M}} = dh_{\overline{\mathcal{M}}}$ and we have

$$dh_{\overline{\mathcal{M}}} = \frac{1}{1+y^2}pdp - \frac{y}{(1+y^2)^2}p^2dy + p_ydp_y + \frac{1}{1+y^2}p_ydp_y - \frac{y}{(1+y^2)^2}p_y^2dy$$

Equating the terms of $dh_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{M}}}$ with those of the push down of $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$ gives the remaining reduced Hamilton equations:

$$\dot{y} = p_y \tag{3.73}$$
$$\dot{p}_y = -\frac{1}{2+y^2}p_xp_y, \tag{3.74}$$

where

$$p_x = -\frac{y}{1+y^2}p_y + \frac{1}{1+y^2}p,$$

as defined earlier in equation (3.64).

Conclusions.

In this chapter we have analyzed the relation between the Lagrangian and Hamiltonian approaches to problems in nonholonomic mechanics. In the course of doing this, we have clarified each of the pictures. For example, we have shown how the momentum equation first found on the Lagrangian side fits into the Hamiltonian approach. We have also explored the reduced Lagrange-d'Alembert equations in greater detail than was known previously. An example, a simplified model of the bicycle is used to illustrate the ideas.

This chapter concentrates in comparing different but equivalent formulations of mechanics with nonholonomic constraints from the *intrinsic* point of view. While a further comparison with the *extrinsic* point of view taken in Dezord [1994] and Marle [1995] would be interesting, we will leave it to the future.

Chapter 4

Poisson Geometry of Nonholonomic Systems

4.1 Introduction

The General Setting. On the Hamiltonian side, besides the symplectic point of view, one can also develop the Poisson point of view. Because of the momentum equation, it is natural to let the value of momentum be a variable and for this a Poisson rather than a symplectic viewpoint is more natural. Some of this theory has been started in van der Schaft and Maschke [1994], hereafter denoted [VM]. In this chapter, we build on their work and develop the Poisson reduction for the nonholonomic systems with symmetry. We use this Poisson reduction procedure to obtain specific formulas for the nonholonomic Hamiltonian dynamics. We also show that the equations given by the Poisson reduction are equivalent to those given by the Lagrangian reduction via a reduced constrained Legendre transform.

Two interesting complications make this effort especially interesting. First of all, it is well known that symmetry need not lead to conservation laws but rather to a momentum equation. Second, the natural Poisson bracket fails to satisfy the Jacobi identity. In fact, the so-called Jacobiizer (the cyclic sum that vanishes when the Jacobi identity holds) is an interesting expression involving the curvature of the underlying distribution describing the nonholonomic constraints. We shall explore these in detail in the forthcoming paper "Poisson Reduction of Nonholonomic Mechanical Systems with Symmetry".

These results are important for the future development of the stability theory for nonholonomic mechanical systems with symmetry. In particular, they will be required for the development of the powerful block diagonalization properties of the energy-momentum method developed by Simo, Lewis and Marsden [1991]. This technique is very important for the development of systematic methods for stability criteria.

Outline of the Chapter. In Section 2, we first consider general nonholonomic systems without symmetry assumptions. In this section,

1. we review the Poisson formulation of nonholonomic systems in [VM] which includes a procedure for finding the equations of motion for nonholonomic systems from the Poisson point of view.

2. With the help of the Ehresmann connection, we use the Poisson procedure to write a compact formula for the equations of motion of the nonholonomic Hamiltonian dynamics.

3. We prove the equivalence of the Poisson and Lagrange-d'Alembert formulations for the nonholonomic mechanics.

4. We apply the Poisson procedure to the example of the snakeboard.

In Section 3, we add the hypothesis of symmetry to the preceding development. In this section,

1. we build on the work of [VM] and develop the Poisson reduction, using the tools like the nonholonomic connection and nonholonomic momentum. We write the equations of motion for the reduced constrained Hamiltonian dynamics using a reduced Poisson bracket. This Poisson reduction procedure breaks the Hamiltonian nonholonomic dynamics into a reconstruction equation, a momentum equation and a set of reduced Hamilton equations.

2. We prove that the set of equations given by the Poisson reduction is equivalent to those given by the Lagrangian reduction via a reduced Legendre transform.

3. We apply the Poisson reduction procedure to the example of the snakeboard.

4.2 General Nonholonomic Mechanical Systems

Following the approaches of [BKMM], we first consider mechanics in the presence of homogeneous linear nonholonomic velocity constraints. For now, no symmetry assumptions are made; we add such assumptions in the following section.

4.2.1 Review of the Poisson Formulation

The approach of [VM] starts on the Lagrangian side with a configuration space Q and a Lagrangian L (possibly of the form kinetic energy minus potential energy, *i.e.*,

$$L(q, \dot{q}) = \frac{1}{2}\langle\langle \dot{q}, \dot{q} \rangle\rangle - V(q),$$

where $\langle\langle \, , \, \rangle\rangle$ is a metric on Q defining the kinetic energy and V is a potential energy function.)

As above, our nonholonomic constraints are given by a distribution $\mathcal{D} \subset TQ$. We also let $\mathcal{D}^{\mathrm{o}} \subset T^*Q$ denote the annihilator of this distribution. Using a basis ω^a of the annihilator $\mathcal{D}^{\mathrm{o}}$, we can write the constraints as

$$\omega^a(\dot{q}) = 0.$$

where $a = 1, \ldots, k$.

As above, the basic equations are given by the Lagrange d'Alembert principle and are written as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \lambda_a \omega^a_i,$$

where λ_a is a set of Lagrange multipliers.

The Legendre transformation $\mathbb{F}L : TQ \to T^*Q$, assuming that it is a diffeomorphism, is used to define the Hamiltonian $H : T^*Q \to \mathbb{R}$ in the standard fashion (ignoring the constraints for the moment):

$$H = \langle p, \dot{q} \rangle - L = p_i \dot{q}^i - L.$$

Here, the momentum is $p = \mathbb{F}L(v_q) = \partial L / \partial \dot{q}$. Under this change of variables, the equations of motion are written in the Hamiltonian form as

$$\dot{q}^i = \frac{\partial H}{\partial p_i}. \tag{4.1}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q^i} + \lambda_a \omega^a_i, \tag{4.2}$$

where $i = 1, \dots, n$, together with the constraint equations

$$\omega_i^a \dot{q}^i = \omega_i^a \frac{\partial H}{\partial p_i} = 0.$$

The preceding constrained Hamiltonian equations can be rewritten as

$$\begin{pmatrix} \dot{q}^i \\ \dot{p}_i \end{pmatrix} = J \begin{pmatrix} \frac{\partial H}{\partial q^j} \\ \frac{\partial H}{\partial p_j} \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_a \omega_i^a \end{pmatrix}, \qquad \omega_i^a \frac{\partial H}{\partial p_i} = 0. \tag{4.3}$$

Recall that the cotangent bundle T^*Q is equipped with a canonical Poisson bracket and is expressed in the canonical coordinates (q, p) as

$$\{F, G\}(q, p) = \frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i} = \left(\frac{\partial F^T}{\partial q}, \frac{\partial F^T}{\partial p}\right) J \begin{pmatrix} \frac{\partial G}{\partial q} \\ \frac{\partial G}{\partial p} \end{pmatrix}.$$

Here J is the canonical Poisson tensor

$$J = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix},$$

which is intrinsically determined by the Poisson bracket $\{,\}$ as

$$J = \begin{pmatrix} \{q^i, q^j\} & \{q^i, p_j\} \\ \{p_i, q^j\} & \{p_i, p_j\} \end{pmatrix}. \tag{4.4}$$

On Lagrangian side, we saw that one can get rid of the Lagrangian multipliers. On the Hamiltonian side, it is also desirable to model the Hamiltonian equations without the Lagrange multipliers by a vector field on a submanifold of T^*Q. In [VM], it is done through a clever change of coordinates.

First, a constraint phase space $\mathcal{M} = \mathbb{F}L(\mathcal{D}) \subset T^*Q$ is defined in the same way as in [BS] so that the constraints on the Hamiltonian side are given by $p \in \mathcal{M}$. In local co-ordinates,

$$\mathcal{M} = \left\{ (q, p) \in T^*Q \mid \omega_i^a \frac{\partial H}{\partial p_i} = 0 \right\}.$$

Let $\{X_\alpha\}$ be a local basis for the constraint distribution $\mathcal{D}$ and let $\{\omega^a\}$ be a local basis for the annihilator $\mathcal{D}^0$. Let $\{\omega_a\}$ span the complementary subspace to $\mathcal{D}$ such that $\langle \omega^a, \omega_b \rangle = \delta_b^a$ where δ_b^a is the usual Kronecker delta. Here $a = 1, \dots, k$ and $\alpha = 1, \dots, n-k$. Define a coordinate transformation $(q, p) \to (q, \tilde{p}_\alpha, \tilde{p}_a)$ by

$$\tilde{p}_\alpha = X_\alpha^i p_i, \qquad \tilde{p}_a = \omega_a^i p_i. \tag{4.5}$$

[VM] shows that in the new (generally not canonical) coordinates $(q, \tilde{p}_\alpha, \tilde{p}_a)$, the Poisson tensor becomes

$$\tilde{J}(q,\tilde{p}) = \begin{pmatrix} \{q^i, q^j\} & \{q^i, \tilde{p}_j\} \\ \{\tilde{p}_i, q^j\} & \{\tilde{p}_i, \tilde{p}_j\} \end{pmatrix}. \tag{4.6}$$

and the constrained Hamiltonian equations (4.3) transform into

$$\begin{pmatrix} \dot{q}^i \\ \dot{\tilde{p}}_\alpha \\ \dot{\tilde{p}}_a \end{pmatrix} = \tilde{J}(q,\tilde{p}) \begin{pmatrix} \frac{\partial \tilde{H}}{\partial q^j} \\ \frac{\partial \tilde{H}}{\partial \tilde{p}_\beta} \\ \frac{\partial \tilde{H}}{\partial \tilde{p}_b} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \lambda_b K_a^b(q) \end{pmatrix}, \qquad \frac{\partial \tilde{H}}{\partial \tilde{p}_a}(q,\tilde{p}) = 0. \tag{4.7}$$

with $K_a^b(q) := \omega_a^i \omega_i^b$ and where $\tilde{H}(q,\tilde{p})$ is the Hamiltonian $H(q,p)$ expressed in the new coordinates $(q,\tilde{p})$.

Then, let $(\tilde{p}_\alpha, \tilde{p}_a)$ satisfy the constraint equations $\frac{\partial \tilde{H}}{\partial \tilde{p}_a}(q,\tilde{p}) = 0$. Since

$$\mathcal{M} = \left\{ (q, \tilde{p}_\alpha, \tilde{p}_a) \mid \frac{\partial \tilde{H}}{\partial \tilde{p}_a}(q, \tilde{p}_\alpha, \tilde{p}_a) = 0 \right\},$$

[VM] uses $(q, \tilde{p}_\alpha)$ as an induced local coordinates for $\mathcal{M}$. It is easy to show that

$$\begin{aligned} \frac{\partial \tilde{H}}{\partial q^j}(q, \tilde{p}_\alpha, \tilde{p}_a) &= \frac{\partial H_{\mathcal{M}}}{\partial q^j}(q, \tilde{p}_\alpha) \\ \frac{\partial \tilde{H}}{\partial \tilde{p}_\beta}(q, \tilde{p}_\alpha, \tilde{p}_a) &= \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta}(q, \tilde{p}_\alpha) \end{aligned}$$

where $H_{\mathcal{M}}$ is the constrained Hamiltonian on $\mathcal{M}$ expressed in the induced coordinates.

Now we are ready to eliminate the Lagrange multipliers. Notice that $\frac{\partial \tilde{H}}{\partial \tilde{p}_b}(q,\tilde{p}) = 0$ on $\mathcal{M}$, and by restricting the dynamics on $\mathcal{M}$, we can disregard the last equations involving λ in equations (4.7). In fact, we can also truncate the Poisson tensor $\tilde{J}$ in (4.6) by leaving out its last k columns and last k rows and then describe the constrained dynamics on $\mathcal{M}$ expressed in the induced coordinates $(q^i, \tilde{p}_\alpha)$ as follows

$$\begin{pmatrix} \dot{q}^i \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = J_{\mathcal{M}}(q,\tilde{p}_\alpha) \begin{pmatrix} \frac{\partial H_{\mathcal{M}}}{\partial q^j}(q,\tilde{p}_\alpha) \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta}(q,\tilde{p}_\alpha) \end{pmatrix}, \qquad \begin{pmatrix} q^i \\ \tilde{p}_\alpha \end{pmatrix} \in \mathcal{M}. \tag{4.8}$$

Here $J_{\mathcal{M}}$ is the $(2n-k) \times (2n-k)$ truncated matrix of $\tilde{J}$ restricted to $\mathcal{M}$ and is expressed in the induced coordinates.

The matrix $J_{\mathcal{M}}$ defines a bracket $\{,\}_{\mathcal{M}}$ on the constraint submanifold $\mathcal{M}$ as follows

$$\{F_{\mathcal{M}}, G_{\mathcal{M}}\}_{\mathcal{M}}(q,\tilde{p}_\alpha) := \begin{pmatrix} \frac{\partial F_{\mathcal{M}}^T}{\partial q^i} & \frac{\partial F_{\mathcal{M}}^T}{\partial \tilde{p}_\alpha} \end{pmatrix} J_{\mathcal{M}}(q^i, \tilde{p}_\alpha) \begin{pmatrix} \frac{\partial G_{\mathcal{M}}}{\partial q^j} \\ \frac{\partial G_{\mathcal{M}}}{\partial \tilde{p}_\beta} \end{pmatrix}$$

for any two smooth functions $F_{\mathcal{M}}, G_{\mathcal{M}}$ on the constraint submanifold $\mathcal{M}$. Clearly this bracket satisfies the first two defining properties of a Poisson bracket, namely, skew symmetry and Leibniz rule, and it is shown in [VM] that it satisfies the Jacobi identity if and only if the constraints are holonomic.

In the coming paper, we will develop a general formula for the Jacobiizer (the cyclic sum that vanishes when the Jacobi identity holds) which is an interesting expression involving the curvature of the underlying distribution that describes the nonholonomic constraints. From this formula, we can see clearly that the Poisson bracket defined here satisfies the Jacobi identity if and only if the constraints are holonomic.

4.2.2 A Formula for the Constrained Hamilton Equations

In holonomic mechanics, it is well known that the Poisson and the Lagrangian formulations are equivalent via a Legendre transform. And it is natural to ask whether the same relation holds for the nonholonomic mechanics as developed in [VM] and [BKMM]. But before we answer this question in the next subsection, we would like to first use the general procedures of [VM] to write down a compact formula for the nonholonomic equations of motion.

Theorem 9 *Assume that we have same setup as in the preceding subsection. Let* $q^i = (r^\alpha, s^a)$ *be the local coordinates in which* ω^a *has the form*

$$\omega^a(q) = ds^a + A^a_\alpha(r,s)dr^\alpha \tag{4.9}$$

where $A^a_\alpha(r,s)$ *is the coordinate expression of the Ehresmann connection described in section 2.1 in Chapter 3. Then the nonholonomic constrained Hamilton equation of motion on* $\mathcal{M}$ *can be written as*

$$\dot{r}^\alpha = \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\alpha} \tag{4.10}$$

$$\dot{s}^a = -A^a_\beta \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta} \tag{4.11}$$

$$\dot{\tilde{p}}_\alpha = -\frac{\partial H_{\mathcal{M}}}{\partial r^\alpha} + A^b_\alpha \frac{\partial H_{\mathcal{M}}}{\partial s^b} - p_b B^b_{\alpha\beta} \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta} \tag{4.12}$$

where $B^b_{\alpha\beta}$ *are the coefficients of the curvature of the Ehresmann connection given in equation (3.5) in Chapter 3. Here,* p_b *should be understood as* p_b *restricted to* $\mathcal{M}$ *and more precisely should be denoted as* $(p_b)_{\mathcal{M}}$.

Proof As mentioned in subsection 2.1 in Chapter 3, no additional assumption is needed since one can always choose a local coordinates where

$$\omega^a(q) = ds^a + A^a_\alpha(r,s)dr^\alpha.$$

In this local coordinate system,

$$\mathcal{D} = \text{span}\{\partial_{r^\alpha} - A^a_\alpha \partial_{s^a}\}. \tag{4.13}$$

Then the new coordinates $(r^\alpha, s^a, \tilde{p}_\alpha, \tilde{p}_a)$ of [VM] are defined by

$$\tilde{p}_\alpha = p_\alpha - A^a_\alpha p_a, \qquad \tilde{p}_a = p_a + A^\alpha_a p_\alpha \tag{4.14}$$

and we can use $(r^\alpha, s^a, \tilde{p}_\alpha)$ as the induced coordinates on $\mathcal{M}$.

Moreover, we can find the constrained Poisson structure matrix $J_\mathcal{M}(r^\alpha, s^a, \tilde{p}_\alpha)$ by computing $\{q^i, q^j\}, \{q^i, \tilde{p}_\alpha\}, \{\tilde{p}_\alpha, \tilde{p}_\beta\}$ and then restrict them to $\mathcal{M}$. Recall that $J_\mathcal{M}$ is constructed out of the Poisson tensor $\tilde{J}$ in equation (4.6) by leaving out its last k columns and last k rows and restricting its remaining elements to $\mathcal{M}$.

Clearly

$$\{q^i, q^j\} = 0.$$

And we have

$$\begin{aligned}
\{r^\beta, \tilde{p}_\alpha\} &= \{r^\beta, p_\alpha - A^a_\alpha p_a\} = \{r^\beta, p_\alpha\} - \{r^\beta, A^a_\alpha p_a\} = \delta^\beta_\alpha \\
\{s^b, \tilde{p}_\alpha\} &= \{s^b, p_\alpha - A^a_\alpha p_a\} = \{s^b, p_\alpha\} - \{s^b, A^a_\alpha p_a\} = -A^b_\alpha,
\end{aligned}$$

where δ^β_α is the usual Kronecker delta. It is also straightforward to find

$$\begin{aligned}
\{\tilde{p}_\alpha, \tilde{p}_\beta\} &= \{p_\alpha - A^a_\alpha p_a, p_\beta - A^b_\beta p_b\} \\
&= -\{p_\alpha, A^b_\beta p_b\} - \{A^a_\alpha p_a, p_\beta\} + \{A^a_\alpha p_a, A^b_\beta p_b\} \\
&= \frac{\partial A^b_\beta}{\partial r^\alpha} p_b - \frac{\partial A^b_\alpha}{\partial r^\beta} p_b + \frac{\partial A^a_\alpha}{\partial s^b} p_a A^b_\beta - A^a_\alpha \frac{\partial A^b_\beta}{\partial s^a} p_b \\
&= \left(\frac{\partial A^b_\beta}{\partial r^\alpha} - \frac{\partial A^b_\alpha}{\partial r^\beta} + A^a_\beta \frac{\partial A^b_\alpha}{\partial s^a} - A^a_\alpha \frac{\partial A^b_\beta}{\partial s^a} \right) p_b \\
&= -B^b_{\alpha\beta} p_b.
\end{aligned}$$

After restricting the above results to $\mathcal{M}$, all other terms remain the same but the last line should be understood as $-B^b_{\alpha\beta}(p_b)_\mathcal{M}$. But for notational simplicity, we keep writing it as $-B^b_{\alpha\beta}p_b$. Putting the above computations together, we can write the nonholonomic equations of motion as follows

$$\begin{pmatrix} \dot{r}^\alpha \\ \dot{s}^a \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 & \delta^\alpha_\beta \\ 0 & 0 & -A^a_\beta \\ -\delta^\beta_\alpha & (A^b_\alpha)^T & -p_c B^c_{\alpha\beta} \end{pmatrix} \begin{pmatrix} \frac{\partial H_{\mathcal{M}}}{\partial r^\beta} \\ \frac{\partial H_{\mathcal{M}}}{\partial s^b} \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta} \end{pmatrix} \tag{4.15}$$

which is the desired result.

4.2.3 The Equivalence of the Poisson and the Lagrange-d'Alembert Formulations

Now we are ready to state and prove the equivalence of the Poisson and Lagrange-d'Alembert formulations.

Theorem 10 *The Lagrange-d'Alembert equations*

$$\dot{s}^a = -A^a_\alpha \dot{r}^\alpha \tag{4.16}$$

$$\frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot{s}^b} B^b_{\alpha\beta}\dot{r}^\beta \tag{4.17}$$

are equivalent to the constrained Hamilton equations

$$\dot{r}^\alpha = \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\alpha} \tag{4.18}$$

$$\dot{s}^a = -A^a_\beta \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta} \tag{4.19}$$

$$\dot{\tilde{p}}_\alpha = -\frac{\partial H_{\mathcal{M}}}{\partial r^\alpha} + A^b_\alpha \frac{\partial H_{\mathcal{M}}}{\partial s^b} - p_b B^b_{\alpha\beta}\frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta} \tag{4.20}$$

via a constrained Legendre transform which are given by

$$\tilde{p}_\alpha = \frac{\partial L_c}{\partial \dot{r}^\alpha} \qquad \dot{r}^\alpha = \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\alpha}. \tag{4.21}$$

Proof Recall that

$$\mathcal{D} = \{(r, s, \dot{r}, \dot{s}) \in TQ \mid \dot{s} + A^a_\alpha \dot{r}^\alpha = 0\}.$$

And we can use $(r, s, \dot{r})$ as the induced coordinates for the submanifold $\mathcal{D}$. Since the constrained Lagrangian is given by

$$L_c(r^\alpha, s^a, \dot{r}^\alpha) = L(r^\alpha, s^a, \dot{r}^\alpha, -A^a_\alpha(r, s)\dot{r}^\alpha),$$

We have

$$\frac{\partial L_c}{\partial \dot{r}^\alpha} = \frac{\partial L}{\partial \dot{r}^\alpha} - \frac{\partial L}{\partial s^a} A^a_\alpha = p_\alpha - p_a A^a_\alpha = \tilde{p}_\alpha. \tag{4.22}$$

Hence, $\frac{\partial L_c}{\partial \dot{r}^\alpha} = \tilde{p}_\alpha$ does define the right constrained Legendre transform between the submanifolds $\mathcal{D}$ and $\mathcal{M}$ with the corresponding induced coordinates $(r^\alpha, s^a, \dot{r}^\alpha)$ and $(r^\alpha, s^a, \tilde{p}_\alpha)$.

Now notice that if $E = \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L$, then restricting it to $\mathcal{D}$ we will get

$$\begin{aligned} E_{\mathcal{D}} &= \left.\left(\frac{\partial L}{\partial \dot{r}^\alpha}\dot{r}^\alpha + \frac{\partial L}{\partial \dot{s}^a}\dot{s}^a\right)\right|_{\mathcal{D}} - L_c \\ &= \frac{\partial L_c}{\partial \dot{r}^\alpha}\dot{r}^\alpha + A^a_\alpha \frac{\partial L}{\partial \dot{s}^a}\dot{r}^\alpha - A^a_\alpha \frac{\partial L}{\partial \dot{s}^a}\dot{r}^\alpha - L_c \\ &= \frac{\partial L_c}{\partial \dot{r}^\alpha}\dot{r}^\alpha - L_c. \end{aligned}$$

Hence, the constrained Hamiltonian is given by

$$H_{\mathcal{M}} = \tilde{p}_\alpha \dot{r}^\alpha - L_c. \tag{4.23}$$

And it is straightforward to show that

$$\frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\alpha} = \dot{r}^\alpha + \tilde{p}_\beta \frac{\partial \dot{r}^\beta}{\partial \tilde{p}_\alpha} - \frac{\partial L_c}{\partial \dot{r}^\beta}\frac{\partial \dot{r}^\beta}{\partial \tilde{p}_\alpha} = \dot{r}^\alpha$$

which gives the equation (4.18). Clearly, $\dot{s}^a = -A^a_\beta \dot{r}^\beta$ together with equation (4.18) gives equation (4.19).

Furthermore, we have

$$\frac{\partial H_{\mathcal{M}}}{\partial r^\beta} = \tilde{p}_\alpha \frac{\partial \dot{r}^\alpha}{\partial r^\beta} - \frac{\partial L_c}{\partial r^\beta} - \frac{\partial L_c}{\partial \dot{r}^\alpha}\frac{\partial \dot{r}^\alpha}{\partial r^\beta} = -\frac{\partial L_c}{\partial r^\beta}, \tag{4.24}$$

and

$$\frac{\partial H_{\mathcal{M}}}{\partial s^b} = \tilde{p}_\alpha \frac{\partial \dot{r}^\alpha}{\partial s^b} - \frac{\partial L_c}{\partial s^b} - \frac{\partial L_c}{\partial \dot{r}^\alpha}\frac{\partial \dot{r}^\alpha}{\partial s^b} = -\frac{\partial L_c}{\partial s^b}. \tag{4.25}$$

Substituting the results of (4.24) and (4.25) into equation (4.17), we get the remaining equation (4.20).

4.2.4 Example: 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 [BKMM] and the references listed there.

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. 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. See figure 3.1 in Chapter 3.

In [BKMM], 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 $J + J_0 + J_1 + J_2 = mr^2$, 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_1, J_2 are the inertia of the wheels. This simplification eliminates some terms in the derivation but does not affect the essential geometry of the problem. Setting $\phi = \phi_1 = -\phi_2$, then the configuration space becomes $Q = SE(2) \times S^1 \times S^1$ and the Lagrangian $L : TQ \to \mathbb{R}$ is the total kinetic energy of the system and is given by

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}J_0\dot{\psi}^2 + J_0\dot{\psi}\dot{\theta} + J_1\dot{\phi}^2.$$

The Constraints. 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. The constraints are defined by

$$\begin{aligned} -\sin(\theta + \phi)\dot{x} + \cos(\theta + \phi)\dot{y} - r\cos\phi\dot{\theta} &= 0 \\ -\sin(\theta - \phi)\dot{x} + \cos(\theta - \phi)\dot{y} + r\cos\phi\dot{\theta} &= 0 \end{aligned}$$

and can be simplified as

$$\begin{aligned} \dot{x} &= -r\cot\phi\cos\theta\dot{\theta} \\ \dot{y} &= -r\cot\phi\sin\theta\dot{\theta}. \end{aligned}$$

Since the coordinate expressions of the Ehresmann connection A^a_α are zeroes except

$$A^1_1 = r\cot\phi\cos\theta \qquad A^2_1 = r\cot\phi\sin\theta,$$

the coefficients of the curvature of this connection are zeroes except

$$\begin{aligned} B^1_{13} = -B^1_{31} &= -r\csc^2\phi\cos\theta \\ B^2_{13} = -B^2_{31} &= -r\csc^2\phi\sin\theta. \end{aligned}$$

Also, the constrained Lagrangian is given by

$$\begin{aligned} L_c &= L(r^\alpha, s^a, \dot r^\alpha, -A^a_\alpha \dot r^\alpha) \\ &= \frac{1}{2}mr^2\csc^2\phi\dot\theta^2 + \frac{1}{2}J_0\dot\psi^2 + J_0\dot\psi\dot\theta + J_1\dot\phi^2 \end{aligned}$$

The Constrained Hamiltonian. The constrained Legendre transform on the constraint $\mathcal{D}$ is given by

$$\begin{aligned} \tilde p_\theta = \frac{\partial L_c}{\partial\dot\theta} &= mr^2\csc^2\phi\dot\theta + J_0\dot\psi \\ \tilde p_\psi = \frac{\partial L_c}{\partial\dot\psi} &= J_0\dot\psi + J_0\dot\theta \\ \tilde p_\phi = \frac{\partial L_c}{\partial\dot\phi} &= 2J_1\dot\phi. \end{aligned}$$

Its inverse are

$$\begin{aligned} \dot\theta &= \frac{\sin^2\phi}{mr^2 - J_0\sin^2\phi}(\tilde p_\theta - \tilde p_\psi) \\ \dot\psi &= \frac{mr^2\tilde p_\psi - J_0\sin^2\phi\tilde p_\theta}{J_0(mr^2 - J_0\sin^2\phi)} \\ \dot\phi &= \frac{\tilde p_\phi}{2J_1}, \end{aligned}$$

And we can find the corresponding constrained Hamiltonian on the submanifold $\mathcal{M}$ via the inverse of the constrained Legendre transform in the following way

$$\begin{aligned} H_{\mathcal{M}} &= \tilde p_\theta\dot\theta + \tilde p_\psi\dot\psi + \tilde p_\phi\dot\phi - (\frac{1}{2}mr^2\csc^2\phi\dot\theta^2 + \frac{1}{2}J_0\dot\psi^2 + J_0\dot\psi\dot\theta + J_1\dot\phi^2) \\ &= \frac{\sin^2\phi}{2(mr^2 - J_0\sin^2\phi)^2}(\tilde p_\theta - \tilde p_\psi)^2 + \frac{1}{2J_0}\tilde p_\psi^2 + \frac{1}{4J_1}\tilde p_\phi^2. \end{aligned}$$

The Equations of Motion. Now we can write the constrained Hamilton equations of motion using the constrained Poisson matrix as follow

$$\begin{pmatrix} \dot{\theta} \\ \dot{\psi} \\ \dot{\phi} \\ \dot{x} \\ \dot{y} \\ \dot{\tilde{p}}_\theta \\ \dot{\tilde{p}}_\psi \\ \dot{\tilde{p}}_\phi \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -A_1^1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -A_1^2 & 0 & 0 \\ -1 & 0 & 0 & A_1^1 & A_1^2 & 0 & 0 & -p_b B_{13}^b \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -p_b B_{31}^b & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \frac{\partial H_{\mathcal{M}}}{\partial \phi} \\ 0 \\ 0 \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\theta} \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\psi} \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\phi} \end{pmatrix}$$

where $A_1^1 = r \cot\phi \cos\theta, A_1^2 = r\cot\phi \sin\theta$ and

$$\begin{aligned} \frac{\partial H_{\mathcal{M}}}{\partial \phi} &= \frac{mr^2 \sin\phi\cos\phi}{(mr^2 - J_0 \sin^2\phi)^2}(\tilde{p}_\theta - \tilde{p}_\psi)^2 \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\theta} &= \frac{\sin^2\phi}{mr^2 - J_0\sin^2\phi}(\tilde{p}_\theta - \tilde{p}_\psi) \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\psi} &= \frac{mr^2\tilde{p}_\psi - J_0\sin^2\phi\tilde{p}_\theta}{J_0(mr^2 - J_0\sin^2\phi)} \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\phi} &= \frac{1}{2J_1}\tilde{p}_\phi. \end{aligned}$$

As for $p_b B_{13}^b$, notice first that

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \qquad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y}.$$

Restricting them to $\mathcal{M}$ is the same as restricting them to $\mathcal{D}$ and then applying the constrained Legendre transform, i.e.,

$$\begin{aligned} p_x &= -mr\cot\phi\cos\theta\dot{\theta} = -\frac{mr\sin\phi\cos\phi}{mr^2 - J_0\sin^2\phi}(\tilde{p}_\theta - \tilde{p}_\psi)\cos\theta \\ p_y &= -mr\cot\phi\sin\theta\dot{\theta} = -\frac{mr\sin\phi\cos\phi}{mr^2 - J_0\sin^2\phi}(\tilde{p}_\theta - \tilde{p}_\psi)\sin\theta. \end{aligned}$$

Therefore,

$$\begin{aligned} p_b B_{13}^b &= p_x(-r\csc^2\phi\cos\theta) + p_y(-r\csc^2\phi\sin\theta) \\ &= \frac{mr^2\cot\phi}{mr^2 - J_0\sin^2\phi}(\tilde{p}_\theta - \tilde{p}_\psi). \end{aligned}$$

After simplification, we have the constrained Hamilton equations

$$\begin{aligned}
\dot{\theta} &= \frac{\sin^2\phi}{mr^2 - J_0\sin^2\phi}(\tilde{p}_\theta - \tilde{p}_\psi) \\
\dot{\psi} &= \frac{mr^2\tilde{p}_\psi - J_0\sin^2\phi\tilde{p}_\theta}{J_0(mr^2 - J_0\sin^2\phi)} \\
\dot{\phi} &= \frac{\tilde{p}_\phi}{2J_1} \\
\dot{\tilde{p}}_\theta &= -\frac{mr^2\cot\phi}{mr^2 - J_0\sin^2\phi}(\tilde{p}_\theta - \tilde{p}_\psi)\frac{\tilde{p}_\phi}{2J_1} \\
\dot{\tilde{p}}_\psi &= 0 \\
\dot{\tilde{p}}_\phi &= 0
\end{aligned}$$

together with the constrained equations

$$\begin{aligned}
\dot{x} &= -\frac{r\sin\phi\cos\phi}{mr^2 - J_0\sin^2\phi}(\tilde{p}_\theta - \tilde{p}_\psi)\cos\theta \\
\dot{y} &= -\frac{r\sin\phi\cos\phi}{mr^2 - J_0\sin^2\phi}(\tilde{p}_\theta - \tilde{p}_\psi)\sin\theta
\end{aligned}$$

4.3 Nonholonomic Mechanical Systems with Symmetry

Now we add the hypothesis of symmetry to the preceding development. Assume that we have a configuration manifold Q, a Lagrangian of the form kinetic minus potential, and a distribution $\mathcal{D}$ that describes the kinematic nonholonomic constraints. We also assume there is a symmetry group G (a Lie group) that leaves the Lagrangian invariant, and that acts on Q (by isometries) and also leaves the distribution invariant, *i.e.,* the tangent of the group action maps $\mathcal{D}_q$ to $\mathcal{D}_{gq}$ (for more details, see [BKMM].) Later, we shall refer this as a **simple nonholonomic mechanical system**. Furthermore, this section uses many results of the Lagrangian reduction developed in [BKMM]. For a brief review of Lagrangian reduction, see subsection 3.1 in Chapter 3 of this thesis.

4.3.1 Poisson Reduction

Now let G be the symmetry group of the system and assume that the quotient space $\bar{\mathcal{M}} = \mathcal{M}/G$ of the G-orbit in $\mathcal{M}$ is a quotient manifold with projection map $\rho : \mathcal{M} \longrightarrow \bar{\mathcal{M}}$. Since G is a symmetry group, all intrinsically defined vector fields push down to $\mathcal{M}$. In this subsection, we will write the equations of motion for the reduced constrained Hamiltonian dynamics using a reduced "Poisson" bracket on the reduced constraint phrase space $\bar{\mathcal{M}}$. Moreover, an explicit expression for this bracket will be provided.

The crucial step here is how to represent the constraint distribution $\mathcal{D}$ in a way that is both intrinsic and ready for reduction. The work in both [BKMM] and Koon and Marsden [1997] suggest that we should use the tools like nonholonomic momentum p and the nonholonomic connection $\mathcal{A}$ in [BKMM] to describe $\mathcal{D}$

Recall that in [BKMM], a body fixed basis

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

has been constructed such that the infinitesimal generators $(e_i(g,r))_Q$ of its first m elements at a point q span $\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q))$. Assume that G is a matrix group and e_i^d is the component of $e_i(r)$ with respect to a fixed basis $\{b_a\}$ of the Lie algebra $\mathfrak{g}$ where $(b_a)_Q = \partial_{g^a}$, then

$$(e_i(g,r))_Q = g_d^a e_i^d \partial_{g^a}.$$

Since $\mathcal{D}_q$ is the direct sum of $\mathcal{S}_q$ and the horizontal space of the nonholonomic connection $\mathcal{A}$, it can be represented by

$$\mathcal{D} = \mathrm{span}\{g_d^a e_i^d \partial_{g^a}, -g_b^a A_\alpha^b \partial_{g^a} + \partial_{r^\alpha}\}. \tag{4.26}$$

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

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).

Then the induced coordinates $(g^a, r^\alpha, \tilde{p}_i, \tilde{p}_\alpha)$ for the constraint submanifold $\mathcal{M}$ are defined by

$$\tilde{p}_i = g_d^a e_i^d p_a = \mu_d e_i^d \tag{4.27}$$

$$\tilde{p}_\alpha = p_\alpha - g_b^a A_\alpha^b p_a = p_\alpha - \mu_b A_\alpha^b. \tag{4.28}$$

Here μ is an element of the dual of the Lie algebra $\mathfrak{g}^*$ and μ_a is its coordinates with respect to a fixed dual basis. Notice that $\tilde{p}_i$ are nothing but the corresponding momentum functions on the Hamiltonian side.

We can find the constrained Poisson structure matrix $J_{\mathcal{M}}(g^a, r^\alpha, \tilde{p}_i, \tilde{p}_\alpha)$ by computing $\{g^a, g^b\}$, etc. and then restrict them to $\mathcal{M}$. Recall that $J_{\mathcal{M}}$ is constructed out of the Poisson tensor $\tilde{J}$ in (4.6) by leaving out its last k columns and last k rows and restricting its remaining elements to $\mathcal{M}$.

Clearly

$$\{g^a, g^b\} = 0, \qquad \{g^a, r^\alpha\} = 0, \qquad \{r^\alpha, r^\beta\} = 0.$$

And we also have

$$\begin{aligned}
\{g^a, \tilde{p}_i\} &= \{g^a, g^b_c e^c_i p_b\} = g^a_c e^c_i \\
\{g^a, \tilde{p}_\alpha\} &= \{g^a, p_\alpha - g^c_b A^b_\alpha p_c\} = -g^a_b A^b_\alpha \\
\{r^\alpha, \tilde{p}_i\} &= \{r^\beta, g^b_c e^c_i p_b\} = 0 \\
\{r^\alpha, \tilde{p}_\beta\} &= \{r^\alpha, p_\beta - g^c_b A^b_\beta p_c\} = \delta^\alpha_\beta
\end{aligned}$$

It is also straightforward to find

$$\begin{aligned}
\{\tilde{p}_i, \tilde{p}_j\} &= \{g^a_c e^c_i p_a, g^b_d e^d_j p_b\} \\
&= p_b \frac{\partial g^b_c}{\partial g^\sigma} e^c_i g^\sigma_d e^d_j - p_b \frac{\partial g^b_d}{\partial g^\tau} e^c_i g^\tau_c e^d_j \\
&= p_b \left(\frac{\partial g^b_c}{\partial g^\sigma} g^\sigma_d - \frac{\partial g^b_d}{\partial g^\tau} g^\tau_c \right) e^c_i e^d_j \\
&= -p_a g^a_b C^b_{cd} e^c_i e^d_j \\
&= -\mu_a C^a_{cd} e^c_i e^d_j,
\end{aligned}$$

where C^a_{cd} is the structure coefficients of the Lie algebra $\mathfrak{g}$. Similarly, we have

$$\begin{aligned}
\{\tilde{p}_i, \tilde{p}_\alpha\} &= \{g^a_c e^c_i p_a, p_\alpha - g^b_d A^b_\alpha p_d\} \\
&= \{g^a_c e^c_i p_a, p_\alpha\} - \{g^a_c e^c_i p_a, g^b_d A^b_\alpha p_d\} \\
&= \mu_a \frac{\partial e^a_i}{\partial r^\alpha} + \mu_a C^a_{bd} e^b_i A^d_\alpha
\end{aligned}$$

and

$$\begin{aligned}
\{\tilde{p}_\alpha, \tilde{p}_\beta\} &= \{p_\alpha - g^a_b A^b_\alpha p_a, p_\beta - g^c_d A^d_\beta p_c\} \\
&= -\{p_\alpha, g^c_d A^d_\beta p_c\} - \{g^a_b A^b_\alpha p_a, p_\beta\} + \{g^a_b A^b_\alpha p_a, g^c_d A^d_\beta p_c\} \\
&= \mu_b \frac{\partial A^b_\beta}{\partial r^\alpha} - \mu_b \frac{\partial A^b_\alpha}{\partial r^\beta} - \mu_b C^b_{ac} A^a_\alpha A^c_\beta \\
&= -\mu_b B^b_{\alpha\beta},
\end{aligned}$$

where $B^b_{\alpha\beta}$ are the coefficients of the curvature of the nonholonomic connection and are given by

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

Therefore the constrained Hamilton equations can be written as follows

$$\begin{pmatrix} \dot{g}^a \\ \dot{r}^\alpha \\ \dot{\tilde{p}}_i \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 & g^a_c e^c_j & -g^a_c A^c_\beta \\ 0 & 0 & 0 & \delta^\alpha_\beta \\ -(g^b_c e^c_i)^T & 0 & -\mu_a C^a_{bd} e^b_i e^d_j & \mu_a F^a_{i\beta} \\ (g^b_c A^c_\alpha)^T & -\delta^\beta_\alpha & -(\mu_a F^a_{j\alpha})^T & -\mu_a B^a_{\alpha\beta} \end{pmatrix} \begin{pmatrix} \frac{\partial H_{\mathcal{M}}}{\partial g^b} \\ \frac{\partial H_{\mathcal{M}}}{\partial r^\beta} \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_j} \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta} \end{pmatrix} \tag{4.29}$$

where $F^a_{i\beta}$ is defined by

$$F^a_{i\beta} = \frac{\partial e^a_i}{\partial r^\beta} + C^a_{bd} e^b_i A^d_\beta \tag{4.30}$$

Since G is the symmetry group of the system and the Hamiltonian H is G-invariant, we have $H_{\mathcal{M}} = h_{\bar{\mathcal{M}}}$. Hence

$$\begin{aligned}
\frac{\partial H_{\mathcal{M}}}{\partial g^b} &= 0 \\
\frac{\partial H_{\mathcal{M}}}{\partial r^\beta} &= \frac{\partial h_{\mathcal{M}}}{\partial r^\beta} \\
\frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_j} &= \frac{\partial h_{\mathcal{M}}}{\partial \tilde{p}_j} \\
\frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta} &= \frac{\partial h_{\mathcal{M}}}{\partial \tilde{p}_\beta}.
\end{aligned}$$

After the reduction by the symmetry group G, we have

$$\begin{pmatrix} \xi^b \\ \dot{r}^\alpha \\ \dot{\tilde{p}}_i \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 & e^b_j & -A^b_\beta \\ 0 & 0 & 0 & \delta^\alpha_\beta \\ -(e^c_i)^T & 0 & -\mu_a C^a_{bd} e^b_i e^d_j & \mu_a F^a_{i\beta} \\ (A^c_\alpha)^T & -\delta^\beta_\alpha & -\mu_a (F^a_{j\alpha})^T & -\mu_a B^a_{\alpha\beta} \end{pmatrix} \begin{pmatrix} 0 \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial r^\beta} \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} \end{pmatrix} \tag{4.31}$$

where $\xi^b = (g^{-1})^b_a \dot{g}^a$.

From the above computations, we have proved the following theorem

Theorem 11 *The momentum equation and the reduced Hamilton equations on the reduced constraint submanifold $\bar{\mathcal{M}}$ can be written as follows*

$$\dot{\tilde{p}}_i = -\mu_a C^a_{bd} e^b_i e^d_j \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} + \mu_a F^a_{i\beta} \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} \tag{4.32}$$

$$\dot{r}^\alpha = \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\alpha} \tag{4.33}$$

$$\dot{\tilde{p}}_\alpha = -\frac{\partial h_{\bar{\mathcal{M}}}}{\partial r^\alpha} - \mu_a F^a_{j\alpha} \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} - \mu_a B^a_{\alpha\beta} \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta}. \tag{4.34}$$

Adding in the following reconstruction equation

$$\dot{\xi}^b = -A^b_\beta \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} + e^b_j \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j}, \tag{4.35}$$

we recover the full dynamics of the system.

Notice that equation (4.32) can be considered as the momentum equation on the Hamiltonian side which corresponds to the momentum equation developed in [BKMM]. It generalizes the Lie-Poisson equation to the nonholonomic case.

Moreover, if we now truncate the reduced Poisson matrix $\tilde{J}$ in equation (4.31) by leaving out its first column and first row, the new matrix $J_{\bar{\mathcal{M}}}$ given by

$$J_{\bar{\mathcal{M}}} = \begin{pmatrix} 0 & 0 & \delta^\alpha_\beta \\ 0 & -\mu_a C^a_{bd} e^b_i e^d_j & \mu_a F^a_{i\beta} \\ -\delta^\beta_\alpha & -\mu_a (F^a_{j\alpha})^T & -\mu_a B^a_{\alpha\beta} \end{pmatrix} \tag{4.36}$$

defines a bracket $\{,\}_{\bar{\mathcal{M}}}$ on the reduced constraint submanifold $\bar{\mathcal{M}}$ as follows

$$\{F_{\bar{\mathcal{M}}}, G_{\bar{\mathcal{M}}}\}_{\bar{\mathcal{M}}}(r^\alpha, \tilde{p}_i, \tilde{p}_\alpha) := \begin{pmatrix} \frac{\partial F^T_{\bar{\mathcal{M}}}}{\partial r^\alpha} & \frac{\partial F^T_{\bar{\mathcal{M}}}}{\partial \tilde{p}_i} & \frac{\partial F^T_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\alpha} \end{pmatrix} J_{\bar{\mathcal{M}}}(r^\alpha, \tilde{p}_i, \tilde{p}_\alpha) \begin{pmatrix} \frac{\partial G_{\bar{\mathcal{M}}}}{\partial r^\beta} \\ \frac{\partial G_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} \\ \frac{\partial G_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} \end{pmatrix}$$

for any two smooth functions $F_{\bar{\mathcal{M}}}, G_{\bar{\mathcal{M}}}$ on the reduced constraint submanifold $\bar{\mathcal{M}}$ whose induced coordinates are $(r^\alpha, \tilde{p}_i, \tilde{p}_\alpha)$. Clearly this bracket satisfies the first two defining properties of a Poisson bracket, namely, skew-symmetry and Leibniz rule.

4.3.2 The equivalence of Poisson and Lagrangian Reduction

Theorem 12 *The equations (4.33) to (4.35) given by the Poisson reduction are equivalent to the equations given by the Lagrangian reduction*

$$\xi^b = -A^b_\beta \dot{r}^\beta + \Gamma^{bi} p_i = -A^b_\beta \dot{r}^\beta + e^b_j \Omega^j \tag{4.37}$$

$$\dot{p}_i = \frac{\partial l}{\partial \xi^b}\left(C^a_{bd}\xi^a e^d_i + \frac{\partial e^a_i}{\partial r^\beta}\dot{r}^\beta \right) \tag{4.38}$$

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -\frac{\partial l}{\partial \xi^b}(B^b_{\alpha\beta}\dot{r}^\beta + F^b_{\alpha i}\Omega^i), \tag{4.39}$$

via a reduced Legendre transform

$$\tilde{p}_\alpha = \frac{\partial l_c}{\partial \dot{r}^\alpha} \qquad \tilde{p}_i = \frac{\partial l_c}{\partial \Omega^i}.$$

Proof Define the reduced constrained Lagrangian

$$l_c(r, \dot{r}, \Omega) = l(r, \dot{r}, -A\dot{r} + \Omega e).$$

where Ω is the body angular velocity and $e(r)$ is the body fixed basis at the identity defined earlier. Notice first that

$$\frac{\partial l}{\partial \dot{r}^\alpha} = \frac{\partial L}{\partial \dot{r}^\alpha} = p_\alpha.$$

Since

$$p_b = \frac{\partial L}{\partial \dot{g}^b} = \frac{\partial l}{\partial \xi^a}\frac{\partial \xi^a}{\partial \dot{g}^b} = \frac{\partial l}{\partial \xi^a}(g^{-1})^a_b,$$

we have

$$\frac{\partial l}{\partial \xi^a} = \mu_a.$$

Hence,

$$\begin{aligned}
\frac{\partial l_c}{\partial \dot{r}^\alpha} &= \frac{\partial l}{\partial \dot{r}^\alpha} + \frac{\partial l}{\partial \xi^a}\frac{\partial \xi^a}{\partial \dot{r}^\alpha} \\
&= \frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial \xi^a}A^a_\alpha \\
&= p_\alpha - \mu_a A^a_\alpha \\
&= \tilde{p}_\alpha,
\end{aligned}$$

and

$$\frac{\partial l_c}{\partial \Omega^i} = \frac{\partial l}{\partial \xi^a}\frac{\partial \xi^a}{\partial \Omega^i} = \frac{\partial l}{\partial \xi^a}e^a_i = \tilde{p}_i.$$

That is, $\tilde{p}_\alpha = \frac{\partial l_c}{\partial \dot{r}^\alpha}$ and $\tilde{p}_i = \frac{\partial l_c}{\partial \Omega^i}$ do define the right reduced constrained Legendre transform between the reduced constraint submanifolds $\bar{\mathcal{D}}$ and $\bar{\mathcal{M}}$ with the corresponding reduced coordinates $(r^\alpha, \dot{r}^\alpha, \Omega^i)$ and $(r^\alpha, \tilde{p}_\alpha, \tilde{p}_i)$.

To find the reduced constrained Hamiltonian $h_{\mathcal{M}}$, notice first that since E is G-invariant, we have

$$\begin{aligned} E &= \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L \\ &= \frac{\partial L}{\partial \dot{g}^a}\dot{g}^a + \frac{\partial L}{\partial \dot{r}^\alpha}\dot{r}^\alpha - L \\ &= \frac{\partial l}{\partial \xi^a}\xi^a + \frac{\partial l}{\partial \dot{r}^\alpha}\dot{r}^\alpha - l \end{aligned}$$

After restricting it to the submanifold $\mathcal{D}$, we have

$$\begin{aligned} E_{\mathcal{D}} &= \frac{\partial l}{\partial \xi^a}(-A^a_\alpha \dot{r}^\alpha + \Omega^i e^a_i) + \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} + A^a_\alpha \frac{\partial l}{\partial \xi^a}\right)\dot{r}^\alpha - l_c \\ &= \frac{\partial l}{\partial \xi^a}\Omega^i e^a_i + \frac{\partial l_c}{\partial \dot{r}^\alpha}\dot{r}^\alpha - l_c \\ &= \frac{\partial l_c}{\partial \Omega^i}\Omega^i + \frac{\partial l_c}{\partial \dot{r}^\alpha}\dot{r}^\alpha - l_c \end{aligned}$$

Therefore, we have

$$h_{\bar{\mathcal{M}}} = \tilde{p}_i \Omega^i + \tilde{p}_\alpha \dot{r}^\alpha - l_c, \tag{4.40}$$

via the Legendre transform $(r^\alpha, \dot{r}^\alpha, \Omega^i) \longrightarrow (r^\alpha, \tilde{p}_\alpha, \tilde{p}_i)$. Differentiate $h_{\bar{\mathcal{M}}}$ with respect to $\tilde{p}_\alpha$ and $\tilde{p}_j$ and use the Legendre transform, we have

$$\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\alpha} = \tilde{p}_i \frac{\partial \Omega^i}{\partial \tilde{p}_\alpha} + \tilde{p}_\beta \frac{\partial \dot{r}^\beta}{\partial \tilde{p}_\alpha} + \dot{r}^\alpha - \frac{\partial l_c}{\partial \dot{r}_\beta}\frac{\partial \dot{r}^\beta}{\partial \tilde{p}_\alpha} - \frac{\partial l_c}{\partial \Omega^i}\frac{\partial \Omega^i}{\partial \tilde{p}_\alpha} = \dot{r}^\alpha$$

which is equation (4.33). Also, we have

$$\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} = \Omega^j + \tilde{p}_i \frac{\partial \Omega^i}{\partial \tilde{p}_j} + \tilde{p}_\alpha \frac{\partial \dot{r}^\alpha}{\partial \tilde{p}_j} - \frac{\partial l_c}{\partial \dot{r}_\alpha}\frac{\partial \dot{r}^\alpha}{\partial \tilde{p}_j} - \frac{\partial l_c}{\partial \Omega^i}\frac{\partial \Omega^i}{\partial \tilde{p}_j} = \Omega^j,$$

which, together with equation (4.37), gives equation (4.35). Moreover, since $\frac{\partial l}{\partial \xi^b} = g^a_b \frac{\partial L}{\partial \dot{g}^a} = \mu_b$ and $\tilde{p}_i = p_i$, we have

$$\begin{aligned} \dot{\tilde{p}}_i &= \frac{\partial l}{\partial \xi^b}\left(C^a_{bd}\xi^a e^d_i + \frac{\partial e^a_i}{\partial r^\beta}\dot{r}^\beta\right) \\ &= \mu_a \left(C^a_{bd} e^d_i \left(-A^b_\beta \dot{r}^\beta + e^b_j \Omega^j\right) + \frac{\partial e^a_i}{\partial r^\beta}\dot{r}^\beta\right) \\ &= \mu_a \left(C^a_{bd} e^d_i \left(-A^b_\beta \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} + e^b_j \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j}\right) + \frac{\partial e^a_i}{\partial r^\beta}\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta}\right) \\ &= \mu_a C^a_{bd} e^d_i e^d_j \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} + \mu_a \left(C^a_{bd} e^b_i A^d_\beta + \frac{\partial e^a_i}{\partial r^\beta}\right)\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta}, \end{aligned}$$

which is equation (4.32).

Finally, differentiate $h_{\bar{\mathcal{M}}}$ with respect to r^α, we have

$$\frac{\partial h_{\bar{\mathcal{M}}}}{\partial r^\alpha} = \tilde{p}_i \frac{\partial \Omega^i}{\partial \tilde{r}^\alpha} + \tilde{p}_\beta \frac{\partial \dot{r}^\beta}{\partial r^\alpha} - \frac{\partial l_c}{\partial r^\alpha} - \frac{\partial \tilde{l}_c}{\partial \dot{r}^\beta}\frac{\partial \dot{r}^\beta}{\partial r^\alpha} - \frac{\partial l_c}{\partial \Omega^i}\frac{\partial \Omega^i}{\partial r^\alpha} = -\frac{\partial l_c}{\partial r_\alpha}$$

which together with equation (4.39) gives

$$\dot{\tilde{p}}_\alpha = -\frac{\partial h_{\bar{\mathcal{M}}}}{\partial r^\alpha} - \mu_a F^a_{j\alpha}\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} - \mu_a B^a_{\alpha\beta}\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta}$$

which is equation (4.34).

Remark: Notice that equations (4.39) are the same as the reduced Lagrange-d'Alembert equations in [BKMM]. The only difference is that in this chapter, the reduced constrained Lagrangian l_c is a function of $r, \dot{r}, \Omega$ where in [BKMM] (and in Chapter 3) it is considered as a function of $r, \dot{r}, p$. Since it is more natural to use the body angular velocity as a variable on the Lagrangian side, the formulation here looks better.

4.3.3 Example: The Snakeboard Revisited

Now we return to the snakeboard and discuss the role of the symmetry group $G = SE(2)$. Recall from our earlier discussion that the Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mr^2\dot{\theta}^2 + +\frac{1}{2}J_0\dot{\psi}^2 + J_0\dot{\psi}\dot{\theta} + J_1\dot{\phi}_1^2, \tag{4.41}$$

which is independent of the configuration of the board and hence it is invariant to all possible group actions.

The Constraint Submanifold. The condition of rolling without slipping gives rise to the constraint one forms

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

which are invariant under the $SE(2)$ action. The constraints determine the kinematic distribution $\mathcal{D}_q$:

$$\mathcal{D}_q = \text{span}\{\partial_\psi, \partial_\phi, a\partial_x + b\partial_y + c\partial_\theta\},$$

where $a = -2r\cos^2\phi\cos\theta, b = -2r\cos^2\phi\sin\theta, c = \sin 2\phi$. The tangent space to the orbits of the $SE(2)$ action is given by

$$T_q(\mathrm{Orb}(q)) = \mathrm{span}\{\partial_x, \partial_y, \partial_\theta\}$$

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

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \mathrm{span}\{a\partial_x + b\partial_y + c\partial_\theta\}.$$

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D} \cap T\mathrm{Orb}$ 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

$$\xi_Q^q = a\partial_x + b\partial_y + c\partial_\theta,$$

which is invariant under the action of $SE(2)$ on Q. The nonholonomic momentum is thus given by

$$\begin{aligned} p &= \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i \\ &= ma\dot{x} + mb\dot{y} + mr^2c\dot{\theta} + J_0c\dot{\psi}. \end{aligned}$$

The kinematic 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 the point $\phi = \pi/2$),

$$\begin{bmatrix} \cos\theta\dot{x} + \sin\theta\dot{y} \\ -\sin\theta\dot{x} + \cos\theta\dot{y} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} -\dfrac{J_0}{2mr}\sin 2\phi\dot{\psi} \\ 0 \\ \dfrac{J_0}{mr^2}\sin^2\phi\dot{\psi} \end{bmatrix} = \begin{bmatrix} \dfrac{-1}{2mr}p \\ 0 \\ \dfrac{\tan\phi}{2mr^2}p \end{bmatrix}. \tag{4.42}$$

These equations have the form

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

where

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

These are precisely the terms which appear in the nonholonomic connection relative to the (global) trivialization (r, g).

Since $\Gamma p = \Omega e$, we can rewrite the constraints using the angular momentum Ω as follows

$$\begin{bmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{bmatrix} = \begin{bmatrix} \frac{J_0}{2mr}\sin 2\phi\dot{\psi} \\ 0 \\ -\frac{J_0}{mr^2}\sin^2\phi\dot{\psi} \end{bmatrix} + \begin{bmatrix} -2r\cos^2\phi\Omega \\ 0 \\ \sin 2\phi\Omega \end{bmatrix}. \tag{4.43}$$

The Reduced Constrained Hamiltonian From the Lagrangian L, we find the reduced Lagrangian

$$l(r, \dot{r}, \xi) = \frac{1}{2}m((\xi^1)^2 + (\xi^2)^2) + \frac{1}{2}mr^2(\xi^3)^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_0\dot{\psi}(\xi^3) + J_1\dot{\phi}^2,$$

where $\xi = g^{-1}\dot{g}$. After plugging in the constraints (4.43), we have the reduced constrained Lagrangian

$$l_c(r, \dot{r}, \Omega) = -\frac{J_0^2}{2mr^2}\sin^2\phi\dot{\psi}^2 + 2mr^2\cos^2\phi\Omega^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_1\dot{\phi}^2. \tag{4.44}$$

Then the reduced constrained Legendre transform is given by

$$\begin{aligned} p &= \frac{\partial l_c}{\partial \Omega} = 4mr^2\cos^2\phi\Omega \\ \tilde{p}_\psi &= \frac{\partial l_c}{\partial \dot{\psi}} = -\frac{J_0^2}{mr^2}\sin^2\phi\dot{\psi} + J_0\dot{\psi} \\ \tilde{p}_\phi &= \frac{\partial l_c}{\partial \dot{\phi}} = 2J_1\dot{\phi}. \end{aligned}$$

And its inverse is

$$\begin{aligned} \Omega &= \frac{p}{4mr^2\cos^2\phi} \\ \dot{\psi} &= \frac{mr^2\tilde{p}_\psi}{J_0(mr^2 - J_0\sin^2\phi)} \\ \dot{\phi} &= \frac{\tilde{p}_\phi}{2J_1} \end{aligned}$$

Therefore, the reduced constrained Hamiltonian $h_{\bar{\mathcal{M}}}$ is

$$\begin{aligned} h_{\bar{\mathcal{M}}} &= p\Omega + \tilde{p}_\psi \dot{\psi} + \tilde{p}_\phi \dot{\phi} - l_c \\ &= \frac{\sec^2\phi}{8mr^2}p^2 + \frac{mr^2}{2J_0(mr^2 - J_0\sin^2\phi)}p_\psi^2 + \frac{1}{4J_1}p_\phi^2 \end{aligned}$$

The Reduced Poisson Structure Matrix Recall that in computing the reduced structural matrix, we only need to calculate $\{\tilde{p}_\alpha, \tilde{p}_\beta\}$, etc. and then restrict them to $\bar{\mathcal{M}}$. Since

$$\begin{aligned} p &= -2r\cos^2\phi\cos\theta p_x - 2r\cos^2\phi\sin\theta p_y + \sin 2\phi p_\theta \\ \tilde{p}_\psi &= \frac{J_0}{2mr^2}\sin 2\phi\cos\theta p_x + \frac{J_0}{2mr^2}\sin 2\phi\sin\theta p_y - \frac{J_0}{mr^2}\sin^2\phi p_\theta + p_\psi \\ \tilde{p}_\phi &= p_\phi, \end{aligned}$$

we have

$$\{p, \tilde{p}_\phi\} = \{-2r\cos^2\phi\mu_1, p_\phi\} + \{\sin 2\phi\mu_3, p_\phi\} = 2r\sin 2\phi\mu_1 + 2\cos 2\phi\mu_3. \tag{4.45}$$

Similarly, we find

$$\{\tilde{p}_\psi, \tilde{p}_\phi\} = \frac{J_0}{mr}\cos 2\phi\mu_1 - \frac{J_0}{mr}\sin 2\phi\mu_3 \tag{4.46}$$

$$\{p, \tilde{p}_\psi\} = 0. \tag{4.47}$$

As for μ_1, μ_2, μ_3 (restricted to $\bar{\mathcal{M}}$), recall that

$$\begin{aligned} \mu_1 &= \cos\theta p_x + \sin\theta p_y \\ &= \cos\theta(m\dot{x}) + \sin\theta(m\dot{y}) \\ &= -mr\cot\phi\dot{\theta} \\ &= -mr\cot\phi\left(-\frac{J_0}{mr^2}\sin^2\phi\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right) \\ &= \frac{J_0}{r}\cos\phi\sin\phi\dot{\psi} - \frac{1}{2r}p \\ &= \frac{mr\sin\phi\cos\phi}{mr^2 - J_0\sin^2\phi}\tilde{p}_\psi - \frac{1}{2r}p. \end{aligned}$$

We can also find μ_2, μ_3 in a similar way. Therefore

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \dfrac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi \\ 0 \\ \dfrac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi \end{bmatrix} + \begin{bmatrix} \dfrac{-1}{2r}p \\ 0 \\ \dfrac{\tan\phi}{2}p \end{bmatrix}. \tag{4.48}$$

So after substituting the constraints (4.48) into equations (4.45) to (4.47), we have

$$\{p, \tilde{p}_\phi\}_{\bar{\mathcal{M}}} = -\tan\phi p + \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}\tilde{p}_\psi \tag{4.49}$$

$$\{\tilde{p}_\psi, \tilde{p}_\phi\}_{\bar{\mathcal{M}}} = -\frac{J_0}{2mr^2}p - \frac{J_0\sin\phi\cos\phi}{mr^2 - J_0\sin^2\tilde{p}_\phi}p_\psi \tag{4.50}$$

$$\{p, \tilde{p}_\psi\}_{\bar{\mathcal{M}}} = 0. \tag{4.51}$$

Therefore the reduced Poisson structure matrix is given by

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -2r\cos^2\phi & \frac{J_0}{2mr}\sin 2\phi & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sin 2\phi & -\frac{J_0}{mr^2}\sin^2\phi & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 2r\cos^2\phi & 0 & -\sin 2\phi & 0 & 0 & 0 & 0 & \{p, \tilde{p}_\phi\}_{\bar{\mathcal{M}}} \\ -\frac{J_0}{2mr}\sin 2\phi & 0 & \frac{J_0}{mr^2}\sin^2\phi & -1 & 0 & 0 & 0 & \{\tilde{p}_\psi, \tilde{p}_\phi\}_{\bar{\mathcal{M}}} \\ 0 & 0 & 0 & 0 & -1 & -\{p, \tilde{p}_\phi\}_{\bar{\mathcal{M}}} & -\{\tilde{p}_\psi, \tilde{p}_\phi\}_{\bar{\mathcal{M}}} & 0 \end{pmatrix}$$

where $\{p, \tilde{p}_\phi\}_{\bar{\mathcal{M}}}$ and $\{\tilde{p}_\psi, \tilde{p}_\phi\}_{\bar{\mathcal{M}}}$ are given as above by (4.49) and (4.50).

The Reduced Constrained Hamilton Equations It is straightforward to find that

$$\begin{aligned} \frac{\partial h_{\bar{\mathcal{M}}}}{\partial\psi} &= 0 \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial\phi} &= \frac{\sec^2\phi\tan\phi}{4mr^2}p^2 + \frac{mr^2\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}\tilde{p}_\psi^2 \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial p} &= \frac{\sec^2\phi}{4mr^2}p \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial\tilde{p}_\psi} &= \frac{mr^2}{J_0(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial\tilde{p}_\phi} &= \frac{1}{2J_1}\tilde{p}_\phi. \end{aligned}$$

Then by using the formula in (4.31) and after some computations, we obtain the momentum equation and the reduced constrained Hamilton equations as follows

$$\dot{p} = \left(-\tan\phi p + \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}\tilde{p}_\psi\right)\frac{1}{2J_1}\tilde{p}_\phi \tag{4.52}$$

$$\dot{\psi} = \frac{mr^2}{J_0(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi \tag{4.53}$$

$$\dot{\phi} = \frac{1}{2J_1}\tilde{p}_\phi \tag{4.54}$$

$$\dot{\tilde{p}}_\psi = -\left(\frac{J_0}{mr^2}p + \frac{J_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi\right)\frac{1}{2J_1}\tilde{p}_\phi \tag{4.55}$$

$$\dot{\tilde{p}}_\psi = 0. \tag{4.56}$$

Also, we can obtain the following reconstruction equations on the Hamiltonian side

$$\dot{x} = \xi^1\cos\theta - \xi^2\sin\theta = \left(-\frac{1}{2mr}p + \frac{r\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi\right)\cos\theta \tag{4.57}$$

$$\dot{y} = \xi^1\sin\theta - \xi^2\cos\theta = \left(-\frac{1}{2mr}p + \frac{r\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi\right)\sin\theta \tag{4.58}$$

$$\dot{\theta} = \xi^3 = \frac{\tan\phi}{2mr^2}p - \frac{\sin^2\phi}{mr^2 - J_0\sin^2\phi}\tilde{p}_\psi. \tag{4.59}$$

Together, these two sets of equations give us the dynamics of the full constrained systems but in a form that is suitable for control theoretical purposes.

Chapter 5

Conclusions and Future Work

This thesis builds on the recent advances made by Bates and Sniatycki [1993], van der Schaft and Maschke [1994], Bloch, Krishnaprasad, Marsden and Murray [1996] and others in the study of nonholonomic systems. It helps to lay a firm foundation for a gauge viewpoint of such systems.

[BKMM] has started this work on the Lagrangian side and generalized the use of connections and momentum maps associated with a given symmetry group to nonholonomic systems. It has shown how Ehresmann connections can be used to write the kinematic constraints as the condition of horizontality with respect to the connection and shown how the equations of motion can be written in terms of base variables and that these equations involve the curvature of the connection. It has also shown that the presence of symmetries in the nonholonomic case may or may not lead to conservation laws and has developed the momentum equation, which plays an important role in control problems of such systems. The process of reduction and reconstruction for these systems is worked out by making use of a nonholonomic connection which is obtained by synthesizing the mechanical connection and constraint connection. Moreover, using the tools of Lagrangian reduction, it developed the reduced Lagrange-d'Alembert equations.

The major part of this thesis extends this gauge viewpoint of such systems to the Hamiltonian side, building on the works of [BS] and [VM]. With the help of nonholonomic connections and momentum maps, we have developed the Poisson reduction of nonholonomic systems with symmetry. We have shown that the Lagrangian reduction for the nonholonomic mechanics is equivalent to both the symplectic reduction and the Poisson reduction via a reduced constrained Legendre transform. But most importantly, we have shown where the momentum equation is lurking on the Hamiltonian side and how this is related to the breaking up the Hamiltonian dynamics of such systems into a reconstruction equation, a momentum equation and the reduced Hamilton equations.

Furthermore, we have developed in the first part of this thesis a "reduced Lagrangian optimization" procedure which can be used to find the optimal controls for such systems. The techniques use the framework obtained in [BKMM] for studying nonholonomic mechanical control systems with symmetry that might have a nontrivial momentum equation. The snakeboard has been used to illustrate the method.

Future Work

In the immediate future, we will finish the work started in Chapter 4 in exploring the failure of the Jacobi identity when the constraints are nonholonomic. As pointed out in that chapter, the so-called Jacobiizer (the cyclic sum that vanishes when the Jacobi identity holds) is an interesting expression involving the curvature of the underlying distribution describing the nonholonomic constraints. From this formula, we can see clearly that the Poisson bracket satisfies the Jacobi identity if and only if the constraints are holonomic.

Besides this, some other interesting topics for future work are

Optimal Control and Numerical Method In our work, we have initiated the investigation of optimal control for nonholonomic systems like the snakeboard, using the Lagrangian framework developed in [BKMM] and coupling it with the method of Lagrange multipliers and Lagrangian reduction. Interestingly, Gregory and Lin [1992] has used the same method of Lagrange multipliers to devise a general, accurate and efficient numerical method to solve the constrained optimal control problem. Ostrowski, Desai and Kumar [1997] has built on these advances to study the optimal gait selection for nonholonomic locomotion systems. This kind of finite element method applied to the variational problem in integral form developed in Gregory and Lin [1992] fits well with the Lagrangian framework and gives good and interesting results in the case of a relatively complicated problem, namely the optimal control of a snakeboard. We would like to use this Lagrangian approach to study the optimal control of a simplified model of the bicycles, which is an underactuated balance system.

Geometric Phase for Nonholonomic Systems The geometric effect of holonomy plays an important role in the understanding of phase drifts and is a crucial ingredient in problems of stabilization and tracking. The basic idea of holonomy is that if the system undergoes cyclic motion in the shape space (this is sometimes the control space), then it need not undergo cyclic motion in the configuration space. The difference between the beginning and the end of the motion is given by a drift in the group variables and this is the geometric phase. But the basic theory for the holonomy is not as well developed in the case of nonholonomic systems as for holonomic ones.

The geometric tools to further develop the theory for systems with nonholonomic constraints are laid in Marsden, Montgomery and Ratiu [1990] and in [BKMM]. We aim to develop the theory by combining the approach in these two papers and also by making the calculations more concrete and accessible. In particular, in [BKMM] the notion of the nonholonomic connection is defined and this is what replaces the mechanical connection in the case of holonomic constraints. What makes this theory more interesting is the presence of of the constraint distribution as well as the fact that the momentum need not be conserved.

Stability and (eventually) Bifurcation theories for Nonholonomic Systems Because of the momentum equation, it is natural to let the value of the momentum be a variable and for this a Poisson rather than a symplectic viewpoint is more natural. This approach is also natural for understanding the block diagonalization procedure in the energy-momentum method developed by Simo, Lewis and Marsden [1991]. With the development of the Poisson geometry in this thesis, we hope that these results will lead to further progress on the stability issues started by Zenkov, Bloch and Marsden [1997].

Bibliography

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

Abraham, R., J.E. Marsden, and T.S. Ratiu [1988] *Manifolds, Tensor Analysis, and Applications.* 2nd Edition, Applied Mathematical Sciences **75**, Springer-Verlag.

Arnold, V.I. [1988] *Dynamical Systems III.* Encyclopedia of Mathematics **3**, Springer-Verlag.

Arnold, V. I. [1978] *Mathematical Methods of Classical Mechanics* Graduate Texts in Math. **60**, Springer Verlag.

Bates, L., H.Graumann, and C.MacDonnell [1996] Examples of gauge conservation laws in nonholonomic systems. *Rep. Math. Phys.*, **37**, 295–308.

Bates, L. and J. Sniatycki [1993] Nonholonomic reduction, *Reports on Math. Phys.* **32**, 99–115.

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.E. Crouch [1994] Reduction of Euler Lagrange problems for constrained variational problems and relation with optimal control problems. *Proc. CDC, IEEE* **33**, 2584–2590.

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.S. Krishnaprasad, J.E. Marsden, and R. Murray [1996] Nonholonomic mechanical systems with symmetry. *Arch. Rat. Mech. An.*, **136**, 21–99.

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 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.

Bloch, A.M., N. Leonard and J.E. Marsden [1997] Stabilization of Mechanical Systems Using Controlled Lagrangians, *preprint*.

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.

Cushman, R., J. Hermans, and D. Kemppainen [1995] The rolling disc. in *Nonlinear dynamical systems and chaos (Groningen, 1995)*, Progr. Nonlinear Differential Equations Appl., **19**, Birkhduser, Basel, 21–60.

Dazord, P [1994] Mecanique hamiltonienne en presence de constraintes. *Illinois J. of Math.***38**:148-175.

Getz, N.H. [1994] Control of balance for a nonlinear nonholonomic non-minimum phase model of a bicycle. *ACC Conf.*, Baltimore, June, 1994.

Getz, N.H. & J. E. Marsden [1995] Control for an autonomous bicycle, *International Conference on Robotics and Automation, IEEE*, Nagoya, Japan, May, 1995.

Gregory, J. and C. Lin [1992] *Constrained optimization in the calculus of variations and optimal control theory.* Van Nostrand Reinhold, NY.

Hermans, J. [1995] A symmetric sphere rolling on a surface. *Nonlinearity* **8**, 4:493-515.

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.

Koiller, J. [1992] Reduction of some classical nonholonomic systems with symmetry. *Arch. Rat. Mech. An.* **118**, 113–148.

Koon, W.S. and J.E. Marsden [1997a] Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction. *SIAM J. Control and Optim.* **35**, 901–929.

Koon, W.S. and J.E. Marsden [1997b] The Hamiltonian and Lagrangian Approaches to the Dynamics of Nonholonomic Systems, *Report on Mathematical Physics* (to appear).

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.

Leonard, N.E. and J.E. Marsden [1996] Stability and Drift of Underwater Vehicle Dynamics: Mechanical Systems with Rigid Motion Symmetry, *Physica D, to appear.*

Lewis, A. [1996] Affine connections and distributions, *preprint*.

Marle, C.-M. [1995] Reduction of constrained mechanical systems and stability of relative equilibria. *Comm. in Math. Phys.* **174**:295-318.

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

Marsden, J.E., R. Montgomery, and T.S. Ratiu [1990] *Reduction, symmetry, and phases in mechanics.* Memoirs AMS **436**.

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 T.S. Ratiu [1986] Reduction of Poisson manifolds, *Lett. in Math. Phys.* **11**, 161–170.

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.

Marsden, J.E. and J. Ostrowski [1997] Symmetries in Motion: Geometric Foundations of Motion Control (to appear).

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.

Murray, R., Z. Li, and S. Sastry [1994]. *A Mathematical Introduction to Robotic Manipulation.* CRC Press, Boca Raton, Fl.

Neimark, Ju. I. & N.A. Fufaev [1972] *Dynamics of Nonholonomic Systems.* Translations of Mathematical Monographs, AMS, vol. **33**.

Ostrowski, J. [1996] *The Mechanics and Control of Undulatory Robotic Locomotion*, PhD thesis, California Institute of Technology.

Ostrowski, J., J.P. Desai, and V. Kumar [1997] Optimal gait selection for nonholonomic locomotion systems, *Proc. CDC, to appear.*

Simo, J.C., D.R. Lewis, and J.E. Marsden [1991] Stability of relative equilibria I: The reduced energy momentum method, *Arch. Rat. Mech. Anal.* **115**, 15-59.

van der Schaft, A.J. & B.M. Maschke [1994] On the Hamiltonian formulation of nonholonomic mechanical systems, *Rep. on Math. Phys.* **34**, 225–233.

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.*

Weber, R.W. [1986] Hamiltonian systems with constraints and their meaning in mechanics. *Arch. Rat. Mech. An.* **91**, 309–335.

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

Zenkov, D.V., A.M. Bloch, and J.E. Marsden [1997] The Energy Momentum Method for the Stability of Nonholonomic Systems *In preparation.*