Poisson Reduction for Nonholonomic Mechanical Systems with Symmetry

Wang Sang Koon
Control and Dynamical Systems
California Institute of Technology 107-81
Pasadena, CA 91125
koon@cds.caltech.edu

Jerrold E. Marsden
Control and Dynamical Systems
California Institute of Technology 107-81
Pasadena, CA 91125
marsden@cds.caltech.edu

January, 1998

Abstract

This paper continues the work of Koon and Marsden [1997b] that began the comparison of the Hamiltonian and Lagrangian formulations of nonholonomic systems. Because of the necessary replacement of conservation laws with the momentum equation, it is natural to let the value of momentum be a variable and for this reason it is natural to take a Poisson viewpoint. Some of this theory has been started in van der Schaft and Maschke [1994]. We build on their work, further develop the theory of nonholonomic Poisson reduction, and tie this theory to other work in the area. We use this reduction procedure to organize nonholonomic dynamics into a reconstruction equation, a nonholonomic momentum equation and the reduced Lagrange d'Alembert equations in Hamiltonian form. We also show that these equations are equivalent to those given by the Lagrangian reduction methods of Bloch, Krishnaprasad, Marsden and Murray [1996]. Because of the results of Koon and Marsden [1997b], this is also equivalent to the results of Bates and Sniatycki [1993], obtained by nonholonomic symplectic reduction.

Two interesting complications make this effort especially interesting. First of all, as we have mentioned, 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), or equivalently, the Schouten bracket, is an interesting expression involving the curvature of the underlying distribution describing the nonholonomic constraints.

The Poisson reduction results in this paper are important for the future development of the stability theory for nonholonomic mechanical systems with

symmetry, as begun by Zenkov, Bloch and Marsden [1997]. In particular, they should be useful for the development of the powerful block diagonalization properties of the energy-momentum method developed by Simo, Lewis and Marsden [1991].

Contents

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 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.

Bloch, Krishnaprasad, Marsden and Murray [1996], hereafter denoted [BKMM], applied the methods of geometric mechanics to the Lagrange-d'Alembert formulation and extended the use of connections and momentum maps associated with a given symmetry group to this case. The resulting framework, including the nonholonomic momentum equation and nonholonomic mechanical connection, provides a setting for studying nonholonomic mechanical control systems that may have a nontrivial evolution of their nonholonomic momentum.

The setting is a configuration space Q with a (nonintegrable) distribution $\mathcal{D} \subset TQ$ describing the constraints. For simplicity, we consider only homogeneous velocity constraints. We are given a Lagrangian L on TQ and a Lie group G acting

2

on the configuration space that leaves the constraints and the Lagrangian invariant. In many example, the group encodes 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 *shape space*.

The dynamics of such a system is described by a set of equations of the following form:

$$
\begin{aligned}
g^{-1}\dot{g} &= -\mathcal{A}^{\mathrm{nh}}(r)\dot{r} + I^{-1}(r)p & (1.1)\\
\dot{p} &= \dot{r}^T H(r)\dot{r} + \dot{r}^T K(r)p + p^T D(r)p & (1.2)\\
M(r)\ddot{r} &= \delta(r,\dot{r},p) + \tau. & (1.3)
\end{aligned}
$$

The first equation is a reconstruction equation for a group element g, the second is an equation for the nonholonomic momentum p (strictly speaking p is the *body* representation of the nonholonomic momentum map, which is not conserved in general), and the third are equations of motion for the reduced variables r which describe the "shape" of the system. The momentum equation is bilinear in $(\dot{r}, p)$. 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.

This framework has been very useful for studying controllability, gait selection and locomotion for systems such as the snakeboard. It has also helped in the study of optimality of certain gaits, by using optimal control ideas in the context of nonholonomic mechanics (Koon and Marsden [1997a] and Ostrowski, Desai and Kumar [1997]). Hence, it is natural to explore ways to develop similar framework on the Hamiltonian side.

Bates and Sniatycki [1993], developed the symplectic geometry on the Hamiltonian side of nonholonomic systems, while [BKMM] explored the Lagrangian side. It was not obvious how these two approaches were equivalent, especially how the momentum equation, the reduced Lagrange-d'Alembert equations and the reconstruction equation correspond to the developments in Bates and Sniatycki [1993].

Koon and Marsden [1997b], established the specific links between these two sides and used 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 Lagrangian reduction and symplectic reduction, we have shown where the momentum equation lies on the Hamiltonian side and how this is related to the organization of the dynamics of nonholonomic systems with symmetry into the three parts displayed above: a reconstruction equation for the group element g, an equation for the nonholonomic momentum p and the reduced Hamilton equations for the shape variables r, p_r. Koon and Marsden [1997b] illustrate the basic theory with the snakeboard, as well as a simplified model of the bicycle (see Getz and Marsden [1995]). The latter is an important prototype control system because it is an underactuated balance system.

However, on the Hamiltonian side, besides the symplectic approach, 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]. We build on their work and develop the theory of 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 Poisson reduction are equivalent to those given by the Lagrangian reduction via a reduced constrained Legendre transform.

Two special features of nonholonomic systems make this study interesting. First, as we have mentioned, 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 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. As is well known (see, for example, Marsden and Ratiu [1994], §10.8), the failure of Jacobi's identity may be equivalently measured by the Schouten bracket of the Poisson tensor with itself.

These results are important for the future development of the stability theory for nonholonomic mechanical systems with symmetry. Important progress towards a stability theory for these systems has been made in Zenkov, Bloch and Marsden [1997]. However, additional insight 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 applicability of the energy-momentum method to complex systems.

Outline of the Paper. In Section 2, we first consider general nonholonomic systems without symmetry assumptions. 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 Poisson formulation of nonholonomic systems in [VM].

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

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

5. We develop a formula for the Jacobiizer that involves the curvature of the Ehresmann connection.

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

4

1. we recall the basic ideas and results of [BKMM] on simple nonholonomic mechanical systems, especially on how it extended 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 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.

3. 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.

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

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

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.

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 $\mathcal{D}_q \subset T_q Q$ of the tangent space to Q at each point $q \in Q$. Consistent with the fact that each $\mathcal{D}_q$ is a *linear* subspace, 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

$$-\delta L := \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right) \delta q^i = 0, \tag{2.1}$$

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 $(i = 1, \dots, n)$, 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 take 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 $\{\omega^a\}$, $a = 1, \dots, k$ 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 $\alpha = 1, \dots n - k$, in which ω^a has the form

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

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

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

The equations of motion, (2.1) may be rewritten by noting that the allowed variations $\delta q^i = (\delta r^\alpha, \delta s^a)$ satisfy $\delta s^a + A_\alpha^a \delta r^\alpha = 0$. Substitution into (2.1) gives

$$\left(\frac{d}{dt} \frac{\partial L}{\partial \dot{r}^\alpha} - \frac{\partial L}{\partial r^\alpha} \right) = A_\alpha^a \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a} \right). \tag{2.2}$$

Equation (2.2) combined with the constraint equations

$$\dot{s}^a = -A_\alpha^a \dot{r}^\alpha \tag{2.3}$$

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 (2.2) can be written in the following way:

$$\frac{d}{dt} \frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A_\alpha^a \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot{s}^b} B_{\alpha\beta}^b \dot{r}^\beta, \tag{2.4}$$

where

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

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

$$B_{\alpha\beta}^b = \left(\frac{\partial A_\alpha^b}{\partial r^\beta} - \frac{\partial A_\beta^b}{\partial r^\alpha} + A_\alpha^a \frac{\partial A_\beta^b}{\partial s^a} - A_\beta^a \frac{\partial A_\alpha^b}{\partial s^a} \right). \tag{2.5}$$

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

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

and hence the equations of motion have the form

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

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 $\pi_{Q,R} : Q \rightarrow R$ 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 vector valued one form on Q such that

1. $A_q : T_qQ \rightarrow V_q$ 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:

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

and the horizontal projection is the map

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

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

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

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:

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

so that the local expression for curvature is given by

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

where the coefficients $B_{\alpha\beta}^a$ are given by (2.5).

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

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

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 $T_q Q / \mathcal{D}_q$ 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.

2.2 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 $a = 1, \dots, k$.

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

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

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

$$\dot{q}^i = \frac{\partial H}{\partial p_i}. \tag{2.6}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q^i} + \lambda_a \omega_i^a, \tag{2.7}$$

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

$$\begin{pmatrix} \dot{q}^i \\ \dot{p}_i \end{pmatrix} = J \begin{pmatrix} \frac{\partial H}{\partial q^j} \\ \frac{\partial H}{\partial p_j} \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_a \omega_i^a \end{pmatrix}, \qquad \omega_i^a \frac{\partial H}{\partial p_i} = 0. \tag{2.8}$$

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

$$J = \begin{pmatrix} \{q^i, q^j\} & \{q^i, p_j\} \\ \{p_i, q^j\} & \{p_i, p_j\} \end{pmatrix}. \tag{2.9}$$

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. We now recall how they do this.

First, a constraint phase space $\mathcal{M} = \mathbb{F}L(\mathcal{D}) \subset T^*Q$ is defined in the same way as in Bates and Sniatycki [1993] so that the constraints on the Hamiltonian side are given by $p \in \mathcal{M}$. In local coordinates,

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

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 $a = 1, \ldots, k$ and $\alpha = 1, \ldots, n - k$. Define a coordinate transformation $(q, p) \to (q, \tilde{p}_\alpha, \tilde{p}_a)$ by

$$\tilde{p}_\alpha = X_\alpha^i p_i, \qquad \tilde{p}_a = \omega_a^i p_i. \tag{2.10}$$

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

$$\tilde{J}(q, \tilde{p}) = \begin{pmatrix} \{q^i, q^j\} & \{q^i, \tilde{p}_j\} \\ \{\tilde{p}_i, q^j\} & \{\tilde{p}_i, \tilde{p}_j\} \end{pmatrix}. \tag{2.11}$$

and the constrained Hamiltonian equations (2.8) transform into

$$\begin{pmatrix} \dot{q}^i \\ \dot{\tilde{p}}_\alpha \\ \dot{\tilde{p}}_a \end{pmatrix} = \tilde{J}(q, \tilde{p}) \begin{pmatrix} \frac{\partial \tilde{H}}{\partial q^j} \\ \frac{\partial \tilde{H}}{\partial \tilde{p}_\beta} \\ \frac{\partial \tilde{H}}{\partial \tilde{p}_b} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \lambda_a \end{pmatrix}, \qquad \frac{\partial \tilde{H}}{\partial \tilde{p}_a}(q, \tilde{p}) = 0. \tag{2.12}$$

where $\tilde{H}(q, \tilde{p})$ is the Hamiltonian $H(q, p)$ expressed in the new coordinates $(q, \tilde{p})$.

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

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

[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 (2.12). In fact, we can also truncate the Poisson tensor $\tilde{J}$ in (2.11) 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

$$
\begin{pmatrix} \dot{q}^i \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = J_\mathcal{M}(q, \tilde{p}_\alpha) \begin{pmatrix} \frac{\partial H_\mathcal{M}}{\partial q^j}(q, \tilde{p}_\alpha) \\ \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\beta}(q, \tilde{p}_\alpha) \end{pmatrix}, \qquad \begin{pmatrix} q^i \\ \tilde{p}_\alpha \end{pmatrix} \in \mathcal{M}. \qquad (2.13)
$$

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

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

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. Furthermore, the constrained Hamiltonian $H_\mathcal{M}$ is an integral of motion for the constrained dynamics on $\mathcal{M}$ due to the skew symmetry of the bracket.

In section (2.5), 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, one can see clearly that the Poisson bracket defined here satisfies the Jacobi identity if and only if the constraints are holonomic.

11

Remarks. Marle [1997] has been shown that the bracket $\{,\}_{\mathcal{M}}$ obtained in [VM] can be given an intrinsic interpretation as follows:

1. Given a constrained Hamiltonian system $\{T^*Q, \{,\}, H, \mathcal{M}, \mathcal{V}\}$ where the first four objects are defined as above and the last object $\mathcal{V}$ is a vector subbundle of $T_{\mathcal{M}}(T^*Q)$ defined by

$$\mathcal{V}_p = \{\mathrm{vert}_p(\eta) \mid \eta \in \mathcal{D}^0\}.$$

   Here, $T_{\mathcal{M}}(T^*Q)$ is the restriction of the tangent bundle of $T(T^*Q)$ to the constraint submanifold $\mathcal{M}$, and $\mathrm{vert}_p(\eta) \in T_p(T^*Q)$ is the vertical lift of $\eta \in T_q^*Q$ where $p \in T^*Q$; in coordinates, $\mathrm{vert}_{(q,p)}(q, \eta) = (q, p, 0, \eta)$. Marle [1997] uses this subbundle $\mathcal{V}$ to encode the fact that the *constraint forces* obey the Lagrange d'Alembert principle.

   It can be shown that the sum of the vector subbundles $T\mathcal{M}$ and $\mathcal{V}$ of $T_{\mathcal{M}}(T^*Q)$ is a direct sum, and

$$T\mathcal{M} \oplus V = T_{\mathcal{M}}(T^*Q).$$

   Moreover, the restriction $X_H|_{\mathcal{M}}$ of the Hamiltonian vector field to the constraint submanifold $\mathcal{M}$ splits into a sum

$$X_H|_{\mathcal{M}} = X_{\mathcal{M}} + X_{\mathcal{V}}$$

   where $X_{\mathcal{M}}$ is the constrained Hamiltonian vector field which is tangent to $\mathcal{M}$ and $X_{\mathcal{V}}$ is a smooth section of the subbundle $\mathcal{V}$ (whose opposite is the *constraint force* field).

2. The canonical Poisson tensor Λ (associated with the Poisson bracket $\{,\}$) of T^*Q can be projected on $\mathcal{M}$, and its projection $\Lambda_{\mathcal{M}}$ is a contravariant skew symmetric 2-tensor on $\mathcal{M}$. More precisely, let $p \in \mathcal{M}$, α and $\beta \in T_p^*\mathcal{M}$. There exists a unique pair $(\hat{\alpha}, \hat{\beta})$ of elements of $T_p^*(T^*Q)$ which vanish on the vector subspace $\mathcal{V}_p$ of $T_p(T^*Q)$, and whose restrictions to the vector subspace $T_p\mathcal{M}$ are α and β, respectively. We can therefore define $\Lambda_{\mathcal{M}}(p)$ by setting

$$\Lambda_{\mathcal{M}}(p)(\alpha, \beta) = \Lambda(p)(\hat{\alpha}, \hat{\beta}).$$

   Then the vector field $X_{\mathcal{M}}$, whose integral curves are the motions of the system, is

$$X_{\mathcal{M}} = \Lambda_{\mathcal{M}}^{\sharp}(dH_{\mathcal{M}}).$$

   Using the 2-tensor $\Lambda_{\mathcal{M}}$, Marle [1997] defines an intrinsic bracket $\{,\}_{\mathcal{M}}$ on the space of smooth functions on $\mathcal{M}$ by setting

$$\{f, g\}_{\mathcal{M}} = \Lambda_{\mathcal{M}}(df, dg).$$

   In the local coordinates $(q, \tilde{p})$, the truncated matrix $J_{\mathcal{M}}(q, \tilde{p}_\alpha)$ obtained in [VM] is exactly the matrix associated with the 2-tensor $\Lambda_{\mathcal{M}}$ and is nothing but the projection of the Poisson matrix $\tilde{J}(q, \tilde{p})$ onto the constraint submanifold $\mathcal{M}$.

   The above considerations show that the nonholonomic brackets constructed in the present paper by a quotient operation, are also intrinsic.

2.3 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 section, we would like to first use the general procedures of [VM] to write down a compact formula for the nonholonomic equations of motion.

Theorem 2.1 *Assume that we have same setup as in the preceding section. Let $q^i = (r^\alpha, s^a)$ be the local coordinates in which ω^a has the form*

$$\omega^a(q) = ds^a + A_\alpha^a(r,s)dr^\alpha \tag{2.14}$$

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

$$\dot{s}^a = -A_\beta^a \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\beta} \tag{2.15}$$

$$\dot{r}^\alpha = \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\alpha} \tag{2.16}$$

$$\dot{\tilde{p}}_\alpha = -\frac{\partial H_\mathcal{M}}{\partial r^\alpha} + A_\alpha^b \frac{\partial H_\mathcal{M}}{\partial s^b} - p_b B_{\alpha\beta}^b \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\beta} \tag{2.17}$$

where $B_{\alpha\beta}^b$ are the coefficients of the curvature of the Ehresmann connection given in equation (2.5). 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 section 2.1, no additional assumption is needed since one can always choose local coordinates in which

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

In this local coordinate system,

$$\mathcal{D} = \text{span}\{\partial_{r^\alpha} - A_\alpha^a \partial_{s^a}\}. \tag{2.18}$$

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

$$\tilde{p}_\alpha = p_\alpha - A_\alpha^a p_a, \qquad\qquad \tilde{p}_a = p_a + A_a^\alpha p_\alpha \tag{2.19}$$

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 (2.11) by leaving out its last k columns and last k rows and restricting its remaining elements to $\mathcal{M}$.

Clearly

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

In addition, we have

$$
\begin{aligned}
\{r^\beta, \tilde{p}_\alpha\} &= \{r^\beta, p_\alpha - A_\alpha^a p_a\} = \{r^\beta, p_\alpha\} - \{r^\beta, A_\alpha^a p_a\} = \delta_\alpha^\beta \\
\{s^b, \tilde{p}_\alpha\} &= \{s^b, p_\alpha - A_\alpha^a p_a\} = \{s^b, p_\alpha\} - \{s^b, A_\alpha^a p_a\} = -A_\alpha^b,
\end{aligned}
$$

where δ_α^β is the usual Kronecker delta. It is also straightforward to find

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

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

$$
\begin{pmatrix} \dot{s}^a \\ \dot{r}^\alpha \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 & -A_\beta^a \\ 0 & 0 & \delta_\beta^\alpha \\ (A_\alpha^b)^T & -\delta_\alpha^\beta & -p_c B_{\alpha\beta}^c \end{pmatrix} \begin{pmatrix} \frac{\partial H_\mathcal{M}}{\partial s^b} \\ \frac{\partial H_\mathcal{M}}{\partial r^\beta} \\ \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\beta} \end{pmatrix} \tag{2.20}
$$

which is the desired result. Notice that the order of the variables r^α and s^a have been switched to make the block diagonalization of the constrained Poisson tensor more apparent. Also, it may be important to point out here that $(p_b)_\mathcal{M}$ can be expressed in terms of the induced coordinates $(q, \tilde{p}_\beta)$ on the constraint submanifold $\mathcal{M}$ by the following formula

$$
\begin{aligned}
(p_b)_\mathcal{M} &= (g_{b\alpha} - g_{ba} A_\alpha^a) \dot{r}^\alpha \\
&= (g_{b\alpha} - g_{ba} A_\alpha^a) \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\alpha} = K_b^\beta \tilde{p}_\beta,
\end{aligned}
$$

where $g_{b\alpha}$ and g_{ba} are the components of the kinetic energy metric and K_b^β is defined by the last equality.

2.4 The Equivalence of Poisson and Lagrange-d'Alembert Formulations

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

Theorem 2.2 *The Lagrange-d'Alembert equations*

$$\dot{s}^a = -A_\alpha^a \dot{r}^\alpha \tag{2.21}$$

$$\frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A_\alpha^a \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot{s}^b}B_{\alpha\beta}^b\dot{r}^\beta \tag{2.22}$$

are equivalent to the constrained Hamilton equations

$$\dot{s}^a = -A_\beta^a \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\beta} \tag{2.23}$$

$$\dot{r}^\alpha = \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\alpha} \tag{2.24}$$

$$\dot{\tilde{p}}_\alpha = -\frac{\partial H_\mathcal{M}}{\partial r^\alpha} + A_\alpha^b \frac{\partial H_\mathcal{M}}{\partial s^b} - p_b B_{\alpha\beta}^b \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\beta} \tag{2.25}$$

via a constrained Legendre transform which are given by

$$\tilde{p}_\alpha = \frac{\partial L_c}{\partial \dot{r}^\alpha} \qquad \dot{r}^\alpha = \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\alpha}. \tag{2.26}$$

Proof Recall that

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

And we can use $(r,s,\dot{r})$ as the induced coordinates for the submanifold $\mathcal{D}$. Since the constrained Lagrangian is given by

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

we have

$$\frac{\partial L_c}{\partial \dot{r}^\alpha} = \frac{\partial L}{\partial \dot{r}^\alpha} - \frac{\partial L}{\partial \dot{s}^a}A_\alpha^a = p_\alpha - p_a A_\alpha^a = \tilde{p}_\alpha. \tag{2.27}$$

Hence, $\frac{\partial L_c}{\partial \dot{r}^\alpha} = \tilde{p}_\alpha$ does define the right constrained Legendre transform between the sub-manifolds $\mathcal{D}$ 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

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

Hence, the constrained Hamiltonian is given by

$$H_\mathcal{M} = \tilde{p}_\alpha \dot{r}^\alpha - L_c. \tag{2.28}$$

15

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 (2.24). Clearly, $\dot{s}^a = -A_\beta^a \dot{r}^\beta$ together with equation (2.24) gives equation (2.23).

Furthermore, we have

$$\frac{\partial H_{\mathcal{M}}}{\partial r^\beta} = \tilde{p}_\alpha \frac{\partial \dot{r}^\alpha}{\partial r^\beta} - \frac{\partial L_c}{\partial r^\beta} - \frac{\partial L_c}{\partial \dot{r}^\alpha}\frac{\partial \dot{r}^\alpha}{\partial r^\beta} = -\frac{\partial L_c}{\partial r^\beta}, \tag{2.29}$$

and

$$\frac{\partial H_{\mathcal{M}}}{\partial s^b} = \tilde{p}_\alpha \frac{\partial \dot{r}^\alpha}{\partial s^b} - \frac{\partial L_c}{\partial s^b} - \frac{\partial L_c}{\partial \dot{r}^\alpha}\frac{\partial \dot{r}^\alpha}{\partial s^b} = -\frac{\partial L_c}{\partial s^b}. \tag{2.30}$$

Substituting the results of (2.29) and (2.30) into equation (2.22), we get the remaining equation (2.25).

2.5   A Formula for the Jacobiizer

Recall that in the proof of Theorem 2.1, we have obtained

$$
\begin{aligned}
\{q^i, q^j\}_{\mathcal{M}} &= 0 \\
\{r^\beta, \tilde{p}_\alpha\}_{\mathcal{M}} &= \delta_\alpha^\beta \\
\{s^b, \tilde{p}_\alpha\}_{\mathcal{M}} &= -A_\alpha^b \\
\{\tilde{p}_\alpha, \tilde{p}_\beta\}_{\mathcal{M}} &= B_{\alpha\beta}^b(p_b)_{\mathcal{M}} = B_{\alpha\beta}^b K_b^\gamma \tilde{p}_\gamma.
\end{aligned}
$$

Clearly

$$
\begin{aligned}
\{\{q^i, q^j\}_{\mathcal{M}}, q_k\}_{\mathcal{M}} + \text{cyclic} &= 0 & (2.31) \\
\{\{q^i, q^j\}_{\mathcal{M}}, \tilde{p}_\alpha\}_{\mathcal{M}} + \text{cyclic} &= 0 & (2.32)
\end{aligned}
$$

It is also straightforward to obtain

$$
\begin{aligned}
\{\{r^\gamma, \tilde{p}_\alpha\}_{\mathcal{M}}, \tilde{p}_\beta\}_{\mathcal{M}} + \text{cyclic} &= K_b^\gamma B_{\alpha\beta}^b & (2.33) \\
\{\{s^a, \tilde{p}_\alpha\}_{\mathcal{M}}, \tilde{p}_\beta\}_{\mathcal{M}} + \text{cyclic} &= -B_{\alpha\beta}^a - A_\gamma^a K_b^\gamma B_{\alpha\beta}^b. & (2.34)
\end{aligned}
$$

As for $\{\{\tilde{p}_\alpha, \tilde{p}_\beta\}_{\mathcal{M}}\}_{\mathcal{M}} + \text{cyclic}$, it takes slightly more work to find

$$
\begin{aligned}
\{\{\tilde{p}_\alpha, \tilde{p}_\beta\}_{\mathcal{M}}, \tilde{p}_\gamma\}_{\mathcal{M}} + \text{cyclic} &= \tilde{p}_\tau K_a^\tau B_{\delta\gamma}^a K_b^\delta B_{\alpha\beta}^b - \tilde{p}_\tau \left(\frac{\partial K_b^\tau}{\partial r^\gamma} - A_\gamma^a \frac{\partial K_b^\tau}{\partial s^a} \right) B_{\alpha\beta}^b \\
&\quad - \tilde{p}_\tau K_b^\tau \left(\frac{\partial B_{\alpha\beta}^b}{\partial r^\gamma} - A_\gamma^a \frac{\partial B_{\alpha\beta}^b}{\partial s^a} \right) + \text{cyclic}.
\end{aligned}
$$

Notice that the right-hand side of the last equation involves the derivatives of the curvature. However, it can be shown that the following identity

$$\frac{\partial B_{\alpha\beta}^b}{\partial r^\gamma} - A_\gamma^c \frac{\partial B_{\alpha\beta}^b}{\partial s^c} + B_{\alpha\beta}^a \frac{\partial A_\gamma^b}{\partial s^a} + \text{cyclic} = 0,$$

16

holds, and we can apply it to rewrite the last equation using only the curvature but not its derivatives

$$
\begin{aligned}
\{\{\tilde{p}_\alpha, \tilde{p}_\beta\}_\mathcal{M}, \tilde{p}_\gamma\}_\mathcal{M} + \text{cyclic} \quad = \quad & \tilde{p}_\tau K_a^\tau B_{\delta\gamma}^a K_b^\delta B_{\alpha\beta}^b - \tilde{p}_\tau \left(\frac{\partial K_b^\tau}{\partial r^\gamma} - A_\gamma^a \frac{\partial K_b^\tau}{\partial s^a} \right) B_{\alpha\beta}^b \\
& + \tilde{p}_\tau K_a^\tau \frac{\partial A_\gamma^a}{\partial s^b} B_{\alpha\beta}^b + \text{cyclic.} \qquad (2.35)
\end{aligned}
$$

Equations (2.31) to (2.35) give the Jacobiizer of the Poisson bracket on $\mathcal{M}$. Of course, one can also use the formalism of the Schouten bracket to do the computations and obtain the same results.

Notice that from the formulas for the Jacobiizer, one can see clearly that if the constraints is holonomic and hence the Ehresmann connection has zero curvature, then the Jacobiizer is zero and the Jacobi identity holds. Conversely, if the Jacobi identity holds, then we have

$$
\begin{aligned}
0 &= K_b^\gamma B_{\alpha\beta}^b \\
0 &= -B_{\alpha\beta}^a - A_\gamma^a K_b^\gamma B_{\alpha\beta}^b.
\end{aligned}
$$

Therefore, $B_{\alpha\beta}^a = 0$ and the constraints are holonomic.

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**.

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 3.1 *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 $\mathcal{A}^{\mathrm{nh}}$ 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

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

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 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

$$\langle J^{\mathrm{nh}}(v_q), \xi \rangle = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q)^i, \tag{3.1}$$

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:

$$\frac{d}{dt}\left(J^{\mathrm{nh}}(\xi^{q(t)})\right) = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}(\xi^{q(t)})\right]^i_Q. \tag{3.2}$$

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

The Momentum Equation in 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

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

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

$$\left\langle J^{\mathrm{nh}}(r, g, \dot{r}, \dot{g}), e_b(r, g) \right\rangle = \left\langle \frac{\partial l}{\partial \xi}, e_b(r) \right\rangle := p_b, \tag{3.3}$$

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

$$\frac{d}{dt} p_i = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_i] + \frac{\partial e_i}{\partial r^\alpha} \dot{r}^\alpha \right\rangle, \tag{3.4}$$

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,

$$\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_qQ, \tag{3.5}$$

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 synthesize 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,

$$\mathcal{A}^{\mathrm{nh}}(v_q) = \mathbb{I}^{\mathrm{nh}}(q)^{-1}J^{\mathrm{nh}}(v_q). \tag{3.6}$$

In a body fixed basis, (3.6) can be written as

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

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

$$g^{-1}\dot{g} = \xi = -\mathcal{A}(r)\dot{r} + \Gamma(r)p, \tag{3.7}$$

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

Moreover, with the help of 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.

3.2 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 section, 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 [1997b] 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

$$\mathcal{D} = \mathrm{span}\{g_d^a e_i^d \partial_{g^a}, -g_b^a A_\alpha^b \partial_{g^a} + \partial_{r^\alpha}\}. \tag{3.8}$$

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).

Then the induced coordinates $(g^a, r^\alpha, \tilde{p}_i, \tilde{p}_\alpha)$ for the constraint submanifold $\mathcal{M}$ are defined by

$$\tilde{p}_i \;=\; g_d^a e_i^d p_a = \mu_d e_i^d \tag{3.9}$$

$$\tilde{p}_\alpha \;=\; p_\alpha - g_b^a A_\alpha^b p_a = p_\alpha - \mu_b A_\alpha^b. \tag{3.10}$$

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 (2.11) 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} are the structure constants 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

$$
\begin{pmatrix} \dot{g}^a \\ \dot{\tilde{p}}_i \\ \dot{r}^\alpha \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = \begin{pmatrix} 0 & g_c^a e_j^c & 0 & -g_c^a A_\beta^c \\ -(g_c^b e_i^c)^T & -\mu_a C_{bd}^a e_i^b e_j^d & 0 & \mu_a F_{i\beta}^a \\ 0 & 0 & 0 & \delta_\beta^\alpha \\ (g_c^b A_\alpha^c)^T & -(\mu_a F_{j\alpha}^a)^T & -\delta_\alpha^\beta & -\mu_a B_{\alpha\beta}^a \end{pmatrix} \begin{pmatrix} \frac{\partial H_\mathcal{M}}{\partial g^b} \\ \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_j} \\ \frac{\partial H_\mathcal{M}}{\partial r^\beta} \\ \frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\beta} \end{pmatrix} \qquad (3.11)
$$

where $F_{i\beta}^a$ is defined by

$$
F_{i\beta}^a = \frac{\partial e_i^a}{\partial r^\beta} + C_{bd}^a e_i^b A_\beta^d \qquad (3.12)
$$

Notice that the order of the variables r^α and $\tilde{p}_i$ have been switched to make the diagonalization of the constrained Poisson tensor more apparent.

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

$$
\begin{aligned}
\frac{\partial H_\mathcal{M}}{\partial g^b} &= 0 \\
\frac{\partial H_\mathcal{M}}{\partial \tilde{p}_j} &= \frac{\partial h_\mathcal{M}}{\partial \tilde{p}_j} \\
\frac{\partial H_\mathcal{M}}{\partial r^\beta} &= \frac{\partial h_\mathcal{M}}{\partial r^\beta} \\
\frac{\partial H_\mathcal{M}}{\partial \tilde{p}_\beta} &= \frac{\partial h_\mathcal{M}}{\partial \tilde{p}_\beta}.
\end{aligned}
$$

After the reduction by the symmetry group G, we have

$$
\begin{pmatrix} \xi^b \\ \dot{\tilde{p}}_i \\ \dot{r}^\alpha \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = \begin{pmatrix} 0 & e_j^b & 0 & -A_\beta^b \\ -(e_i^c)^T & -\mu_a C_{bd}^a e_i^b e_j^d & 0 & \mu_a F_{i\beta}^a \\ 0 & 0 & 0 & \delta_\beta^\alpha \\ (A_\alpha^c)^T & -\mu_a (F_{j\alpha}^a)^T & -\delta_\alpha^\beta & -\mu_a B_{\alpha\beta}^a \end{pmatrix} \begin{pmatrix} 0 \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial r^\beta} \\ \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} \end{pmatrix} \qquad (3.13)
$$

where $\xi^b = (g^{-1})_a^b \dot{g}^a$.

The above computations prove the following theorem

Theorem 3.2 *The momentum equation and the reduced Hamilton equations on the reduced constraint submanifold $\bar{\mathcal{M}}$ can be written as follows*

$$
\dot{\tilde{p}}_i = -\mu_a C_{bd}^a e_i^b e_j^d \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} + \mu_a F_{i\beta}^a \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} \qquad (3.14)
$$

$$
\dot{r}^\alpha = \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\alpha} \qquad (3.15)
$$

$$
\dot{\tilde{p}}_\alpha = -\frac{\partial h_{\bar{\mathcal{M}}}}{\partial r^\alpha} - \mu_a F_{j\alpha}^a \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} - \mu_a B_{\alpha\beta}^a \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta}. \qquad (3.16)
$$

23

Adding in the following reconstruction equation

$$\dot{\xi}^b = -A^b_\beta \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} + e^b_j \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j}, \tag{3.17}$$

we recover the full dynamics of the system.

Notice that equation (3.14) 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 (3.13) by leaving out its first column and first row, the new matrix $J_{\bar{\mathcal{M}}}$ given by

$$J_{\bar{\mathcal{M}}} = \begin{pmatrix} -\mu_a C^a_{bd} e^b_i e^d_j & 0 & \mu_a F^a_{i\beta} \\ 0 & 0 & \delta^\alpha_\beta \\ -\mu_a (F^a_{j\alpha})^T & -\delta^\beta_\alpha & -\mu_a B^a_{\alpha\beta} \end{pmatrix} \tag{3.18}$$

defines a bracket $\{,\}_{\bar{\mathcal{M}}}$ on the reduced constraint submanifold $\bar{\mathcal{M}}$ as follows

$$\{F_{\bar{\mathcal{M}}}, G_{\bar{\mathcal{M}}}\}_{\bar{\mathcal{M}}}(\tilde{p}_i, r^\alpha, \tilde{p}_\alpha) := \begin{pmatrix} \frac{\partial F^T_{\bar{\mathcal{M}}}}{\partial \tilde{p}_i} & \frac{\partial F^T_{\bar{\mathcal{M}}}}{\partial r^\alpha} & \frac{\partial F^T_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\alpha} \end{pmatrix} J_{\bar{\mathcal{M}}}(\tilde{p}_i, r^\alpha, \tilde{p}_\alpha) \begin{pmatrix} \frac{\partial G_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} \\ \frac{\partial G_{\bar{\mathcal{M}}}}{\partial r^\beta} \\ \frac{\partial G_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} \end{pmatrix}$$

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 $(\tilde{p}_i, r^\alpha, \tilde{p}_\alpha)$. Clearly this bracket satisfies the first two defining properties of a Poisson bracket, namely, skew-symmetry and Leibniz rule. Moreover, the reduced constrained Hamiltonian $h_{\bar{\mathcal{M}}}$ is an integral of motion for the reduced Hamiltonian dynamics on $\bar{\mathcal{M}}$ due to the skew symmetry of the reduced bracket.

3.3 The equivalence of Poisson and Lagrangian Reduction

Theorem 3.3 *The equations (3.14) to (3.17) given by the Poisson reduction are equivalent to the equations given by the Lagrangian reduction*

$$\xi^b = -A^b_\beta \dot{r}^\beta + \Gamma^{bi} p_i = -A^b_\beta \dot{r}^\beta + e^b_j \Omega^j \tag{3.19}$$

$$\dot{p}_i = \frac{\partial l}{\partial \xi^a} \left(C^a_{bd} \xi^b e^d_i + \frac{\partial e^a_i}{\partial r^\beta} \dot{r}^\beta \right) \tag{3.20}$$

$$\frac{d}{dt} \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} \right) - \frac{\partial l_c}{\partial r^\alpha} = -\frac{\partial l}{\partial \xi^b} (B^b_{\alpha\beta} \dot{r}^\beta + F^b_{\alpha i} \Omega^i), \tag{3.21}$$

via a reduced Legendre transform

$$\tilde{p}_\alpha = \frac{\partial l_c}{\partial \dot{r}^\alpha} \qquad \tilde{p}_i = \frac{\partial l_c}{\partial \Omega^i}.$$

24

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

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

25

Therefore, we have

$$h_{\bar{\mathcal{M}}} = \tilde{p}_i \Omega^i + \tilde{p}_\alpha \dot{r}^\alpha - l_c, \tag{3.22}$$

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 (3.15). 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 (3.19), gives equation (3.17). Moreover, since $\frac{\partial l}{\partial \xi^b} = g_b^a \frac{\partial L}{\partial \dot{g}^a} = \mu_b$ and $\tilde{p}_i = p_i$, we have

$$
\begin{aligned}
\dot{\tilde{p}}_i &= \frac{\partial l}{\partial \xi^a} \left(C_{bd}^a \xi^b e_i^d + \frac{\partial e_i^a}{\partial r^\beta} \dot{r}^\beta \right) \\
&= \mu_a \left(C_{bd}^a e_i^d \left(-A_\beta^b \dot{r}^\beta + e_j^b \Omega^j \right) + \frac{\partial e_i^a}{\partial r^\beta} \dot{r}^\beta \right) \\
&= \mu_a \left(C_{bd}^a e_i^d \left(-A_\beta^b \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} + e_j^b \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} \right) + \frac{\partial e_i^a}{\partial r^\beta} \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta} \right) \\
&= \mu_a C_{bd}^a e_i^b e_j^d \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} + \mu_a \left(C_{bd}^a e_i^b A_\beta^d + \frac{\partial e_i^a}{\partial r^\beta} \right) \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta},
\end{aligned}
$$

which is equation (3.14).

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 (3.21) gives

$$\dot{\tilde{p}}_\alpha = -\frac{\partial h_{\bar{\mathcal{M}}}}{\partial r^\alpha} - \mu_a F_{j\alpha}^a \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_j} - \mu_a B_{\alpha\beta}^a \frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\beta}$$

which is equation (3.16).

Remark. Notice that equations (3.21) are the same as the reduced Lagrange-d'Alembert equations in [BKMM]. The only difference is that in this paper, the reduced constrained Lagrangian l_c is a function of $r, \dot{r}, \Omega$ where in [BKMM] and Koon and Marsden [1997b] 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.

3.4 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.
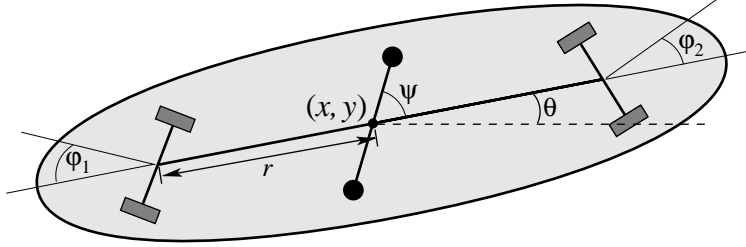

Figure 3.1: The geometry of the snakeboard.

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

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

Clearly L is independent of the configuration of the board and hence it is invariant to all possible group actions by the Euclidean group $SE(2)$.

The Constraint Submanifold. The rolling of the front and rear wheels of the snakeboard is modeled using nonholonomic constraints which allow the wheels to

27

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

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

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

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

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

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

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D} \cap T\operatorname{Orb}$ regarded as a bundle over Q. Since $\mathcal{D}_q \cap T_q\operatorname{Orb}(q)$ 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 $\phi = \pi/2$),

$$
\begin{bmatrix}
\cos\theta\dot{x} + \sin\theta\dot{y} \\
-\sin\theta\dot{x} + \cos\theta\dot{y} \\
\dot{\theta}
\end{bmatrix}
+
\begin{bmatrix}
-\dfrac{J_0}{2mr}\sin 2\phi\dot{\psi} \\
0 \\
\dfrac{J_0}{mr^2}\sin^2\phi\dot{\psi}
\end{bmatrix}
=
\begin{bmatrix}
\dfrac{-1}{2mr}p \\
0 \\
\dfrac{\tan\phi}{2mr^2}p
\end{bmatrix}.
\tag{3.23}
$$

28

These equations have the form

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

where

$$A(r) = -\frac{J_0}{2mr}\sin 2\phi\, e_x\, d\psi + \frac{J_0}{mr^2}\sin^2\phi\, e_\theta\, d\psi$$

$$\Gamma(r) = \frac{-1}{2mr}e_x + \frac{1}{2mr^2}\tan\phi\, e_\theta.$$

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

$$\begin{bmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{bmatrix} = \begin{bmatrix} \dfrac{J_0}{2mr}\sin 2\phi\dot{\psi} \\ 0 \\ -\dfrac{J_0}{mr^2}\sin^2\phi\dot{\psi} \end{bmatrix} + \begin{bmatrix} -2r\cos^2\phi\Omega \\ 0 \\ \sin 2\phi\Omega \end{bmatrix}. \tag{3.24}$$

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 (3.24), we have the reduced constrained Lagrangian

$$l_c(r, \dot{r}, \Omega) = -\frac{J_0^2}{2mr^2}\sin^2\phi\dot{\psi}^2 + 2mr^2\cos^2\phi\Omega^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_1\dot{\phi}^2. \tag{3.25}$$

Then the reduced constrained Legendre transform is given by

$$p = \frac{\partial l_c}{\partial \Omega} = 4mr^2\cos^2\phi\Omega$$

$$\tilde{p}_\psi = \frac{\partial l_c}{\partial \dot{\psi}} = -\frac{J_0^2}{mr^2}\sin^2\phi\dot{\psi} + J_0\dot{\psi}$$

$$\tilde{p}_\phi = \frac{\partial l_c}{\partial \dot{\phi}} = 2J_1\dot{\phi}.$$

And its inverse is

$$\Omega = \frac{p}{4mr^2\cos^2\phi}$$

$$\dot{\psi} = \frac{mr^2\tilde{p}_\psi}{J_0(mr^2 - J_0\sin^2\phi)}$$

$$\dot{\phi} = \frac{\tilde{p}_\phi}{2J_1}$$

29

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

$$
\{p, \tilde{p}_\phi\} = \{-2r \cos^2 \phi \mu_1, p_\phi\} + \{\sin 2\phi \mu_3, p_\phi\} = 2r \sin 2\phi \mu_1 + 2 \cos 2\phi \mu_3. \quad (3.26)
$$

Similarly, we find

$$
\begin{aligned}
\{\tilde{p}_\psi, \tilde{p}_\phi\} &= \frac{J_0}{mr} \cos 2\phi \mu_1 - \frac{J_0}{mr} \sin 2\phi \mu_3 & (3.27) \\
\{p, \tilde{p}_\psi\} &= 0. & (3.28)
\end{aligned}
$$

As for μ_1, μ_2, μ_3 (restricted to $\bar{\mathcal{M}}$), recall that

$$
\begin{aligned}
\mu_1 &= \cos \theta p_x + \sin \theta p_y \\
&= \cos \theta (m\dot{x}) + \sin \theta (m\dot{y}) \\
&= \frac{J_0}{r} \cos \phi \sin \phi \dot{\psi} - \frac{1}{2r} p \\
&= \frac{mr \sin \phi \cos \phi}{mr^2 - J_0 \sin^2 \phi} \tilde{p}_\psi - \frac{1}{2r} p.
\end{aligned}
$$

We can also find μ_2, μ_3 in a similar way. Therefore

$$
\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \dfrac{mr \sin \phi \cos \phi}{(mr^2 - J_0 \sin^2 \phi)} \tilde{p}_\psi \\ 0 \\ \dfrac{mr^2 \cos^2 \phi}{(mr^2 - J_0 \sin^2 \phi)} \tilde{p}_\psi \end{bmatrix} + \begin{bmatrix} \dfrac{-1}{2r} p \\ 0 \\ \dfrac{\tan \phi}{2} p \end{bmatrix}. \quad (3.29)
$$

So after substituting the constraints (3.29) into equations (3.26) to (3.28), we have

$$
\begin{aligned}
\{p, \tilde{p}_\phi\}_{\bar{\mathcal{M}}} &= -\tan \phi p + \frac{2mr^2 \cos^2 \phi}{mr^2 - J_0 \sin^2 \phi} \tilde{p}_\psi & (3.30) \\
\{\tilde{p}_\psi, \tilde{p}_\phi\}_{\bar{\mathcal{M}}} &= -\frac{J_0}{2mr^2} p - \frac{J_0 \sin \phi \cos \phi}{mr^2 - J_0 \sin^2 \phi} p_\psi & (3.31) \\
\{p, \tilde{p}_\psi\}_{\bar{\mathcal{M}}} &= 0. & (3.32)
\end{aligned}
$$

30

Therefore the reduced Poisson structure matrix is given by

$$
\begin{pmatrix}
0 & 0 & 0 & -2r\cos^2\phi & 0 & 0 & \frac{J_0}{2mr}\sin 2\phi & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \sin 2\phi & 0 & 0 & -\frac{J_0}{mr^2}\sin^2\phi & 0 \\
2r\cos^2\phi & 0 & -\sin 2\phi & 0 & 0 & 0 & 0 & \{p,\tilde{p}_\phi\}_{\bar{\mathcal{M}}} \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
-\frac{J_0}{2mr}\sin 2\phi & 0 & \frac{J_0}{mr^2}\sin^2\phi & 0 & -1 & 0 & 0 & \{\tilde{p}_\psi,\tilde{p}_\phi\}_{\bar{\mathcal{M}}} \\
0 & 0 & 0 & -\{p,\tilde{p}_\phi\}_{\bar{\mathcal{M}}} & 0 & -1 & -\{\tilde{p}_\psi,\tilde{p}_\phi\}_{\bar{\mathcal{M}}} & 0
\end{pmatrix}
$$

where $\{p,\tilde{p}_\phi\}_{\bar{\mathcal{M}}}$ and $\{\tilde{p}_\psi,\tilde{p}_\phi\}_{\bar{\mathcal{M}}}$ are given as above by (3.30) and (3.31).

The Reduced Constrained Hamilton Equations.   It is straightforward to find that

$$
\begin{aligned}
\frac{\partial h_{\bar{\mathcal{M}}}}{\partial p} &= \frac{\sec^2\phi}{4mr^2}p \\
\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \psi} &= 0 \\
\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \phi} &= \frac{\sec^2\phi\tan\phi}{4mr^2}p^2 + \frac{mr^2\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}\tilde{p}_\psi^2 \\
\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\psi} &= \frac{mr^2}{J_0(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi \\
\frac{\partial h_{\bar{\mathcal{M}}}}{\partial \tilde{p}_\phi} &= \frac{1}{2J_1}\tilde{p}_\phi.
\end{aligned}
$$

Then by using the formula in (3.13) and after some computations, we obtain the momentum equation and the reduced constrained Hamilton equations as follows

$$
\dot{p} = \left(-\tan\phi\, p + \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}\tilde{p}_\psi\right)\frac{1}{2J_1}\tilde{p}_\phi \tag{3.33}
$$

$$
\dot{\psi} = \frac{mr^2}{J_0(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi \tag{3.34}
$$

$$
\dot{\phi} = \frac{1}{2J_1}\tilde{p}_\phi \tag{3.35}
$$

$$
\dot{\tilde{p}}_\psi = -\left(\frac{J_0}{mr^2}p + \frac{J_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}\tilde{p}_\psi\right)\frac{1}{2J_1}\tilde{p}_\phi \tag{3.36}
$$

$$
\dot{\tilde{p}}_\phi = 0. \tag{3.37}
$$

Also, we can obtain the following reconstruction equations on the Hamiltonian side

$$\dot{x} = \xi^1 \cos\theta - \xi^2 \sin\theta = \left(-\frac{1}{2mr}p + \frac{r\sin 2\phi}{2(mr^2 - J_0 \sin^2\phi)}\tilde{p}_\psi \right)\cos\theta \quad (3.38)$$

$$\dot{y} = \xi^1 \sin\theta - \xi^2 \cos\theta = \left(-\frac{1}{2mr}p + \frac{r\sin 2\phi}{2(mr^2 - J_0 \sin^2\phi)}\tilde{p}_\psi \right)\sin\theta \quad (3.39)$$

$$\dot{\theta} = \xi^3 = \frac{\tan\phi}{2mr^2}p - \frac{\sin^2\phi}{mr^2 - J_0 \sin^2\phi}\tilde{p}_\psi. \quad (3.40)$$

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.

4   Conclusions and Future Work

This paper has continued the work of Koon and Marsden [1997b] in comparing the Hamiltonian and Lagrangian approaches to nonholonomic systems. This paper, together with Koon and Marsden [1997b], 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; that is, the viewpoint of principal bundles (described in, for example, Marsden [1992]) that has been so useful in stability theory, geometric phases and related matters.

[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 introduced to write the kinematic constraints as the condition of horizontality with respect to this 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 for 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.

This paper together with Koon and Marsden [1997b] extend this gauge viewpoint of such systems to the Hamiltonian side, building on the works of Bates and Sniatycki [1993] and [VM]. With the help of nonholonomic connections and momentum maps, the present paper develops the Poisson reduction of nonholonomic systems with symmetry. It shows that Lagrangian reduction for nonholonomic mechanics is equivalent to both the symplectic reduction and the Poisson reduction via a reduced constrained Legendre transform. But most importantly, it shows where the momentum equation lies on the Hamiltonian side and how this is related to the organization of the Hamiltonian dynamics of such systems into a reconstruction

equation, a momentum equation and the reduced Hamilton equations. We also explore the failure of the Jacobi identity when the constraints are nonholonomic and show that 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. Using this formula, one sees clearly that the Poisson bracket satisfies the Jacobi identity if and only if the constraints are holonomic.

Future work

Some interesting topics for future work include

A Deeper Understanding of the Geometry behind this Gauge Viewpoint of Nonholonomic Systems. In studying the Lagrangian reduction by stages of holonomic systems, Cendra, Marsden and Ratiu [1997] has developed an intrinsic approach to the Lagrangian reduction theory of Marsden and Scheurle [1993a,b]. Their work is based on a deeper understanding of the geometry of the bundle $(TQ)/G$ over the shape space Q/G as a Whitney sum $T(Q/G) \oplus Ad(Q)$ where $Ad(Q)$ is the adjoint bundle (i.e., the associated bundle with fiber $\mathfrak{g}$ where the action on $\mathfrak{g}$ is the adjoint action). Extending this theory to the nonholonomic case will clarify further the geometry of Lagrangian reduction in [BKMM]. It will also shed light on the geometry of the gauge Hamiltonian reductions for nonholonomic dynamics.

Stability, Stabilization and 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 may also allow one to extend to the nonholonomic systems 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 paper, we hope that these results will lead to further progress on the stability issues started by Zenkov, Bloch and Marsden [1997]. In this light, the stability and stabilization of a simplified model of the bicycle (see Koon and Marsden [1997b]) is an important problem to tackle.

Optimal Control and Numerical Methods. Closely related to the work mentioned above is the optimal control of the bicycle, which is an underactuated balance system. Koon and Marsden [1997a] 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,

Geometric Phases 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 approaches 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. A start on this problem is made in Koon and Marsden [1997c].

References

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

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., 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., N. Leonard and J.E. Marsden [1997] Stabilization of Mechanical Systems Using Controlled Lagrangians, *preprint*.

Cendra, H., J. E. Marsden and T.S. Ratiu [1997] Lagrangian reduction by stages. *preprint.*

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.

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).

Koon, W.S. and J.E. Marsden [1997c] The Geometric Structure of Nonholonomic Mechanics. *(preprint).*

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. [1997] Various approaches to conservative and nonconservative nonholonomic systems, *preprint*.

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 [1994] *Symmetry and Mechanics.* Texts in Appl. Math., **17**, Springer.

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).

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.

Zenkov, D.V., A.M. Bloch, and J.E. Marsden [1997] The Energy Momentum Method for the Stability of Nonholonomic Systems *preprint.*