The Hamiltonization of Nonholonomic Systems and its Applications

by

Oscar E. Fernandez

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied and Interdisciplinary Mathematics)
in The University of Michigan
2009

Doctoral Committee:

Professor Anthony M. Bloch, Chair
Associate Professor Alberto Rojo, Oakland University, Co-Chair
Professor Peter Smereka
Professor Alejandro Uribe-Ahumada

For Zoraida and my parents

ACKNOWLEDGEMENTS

The work culminating in this thesis would not have been possible without the support and encouragement of many people and they are far too many to list here. However, five groups in particular deserve recognition above all others. To my parents for their love and encouragement throughout life, to my wife for her support and understanding, to Professor Bloch for his willingness to work with me and for his continuing support of me and my research, to all the collaborators with whom I have had the pleasure of working, especially Tom Mestdag and Alejandro Uribe and finally to Alberto Rojo, Peter Smereka, Alejandro Uribe and Anthony Bloch for serving on the doctoral committee, providing valuable feedback on the dissertation, and supporting me throughout the process. I cannot thank you all enough but thank you.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

ABSTRACT

A *nonholonomic mechanical system* is a pair $(L, \mathcal{D})$, where $L : TQ \to \mathbb{R}$ is a mechanical Lagrangian and $\mathcal{D} \subset TQ$ is a distribution which is non-integrable (in the Frobenius sense). Although such mechanical systems are manifestly not Hamiltonian (their mechanics are described by the Lagrange-d'Alembert principle, not Hamilton's principle), one can nevertheless attempt to formulate the mechanics of certain classes of nonholonomic systems as almost-Hamiltonian. In this dissertation we study various methods of so-called *Hamiltonization* of nonholonomic systems and discuss their application to optimal control and the quantization of nonholonomic systems.

We begin by constructing *second-order associated systems* for a class of nonholonomic systems and solving the Inverse Problem of the Calculus of Variations to derive Hamiltonians whose canonical equations, when restricted to certain invariant submanifolds, reproduce the original nonholonomic mechanics.

We also introduce the idea of *conditionally variational nonholonomic systems*, which arise from a comparison with the variational nonholonomic equations, and show that these systems give a straightforward Hamiltonization for certain classes of systems.

Lastly, we extend a classical theorem of S.A. Chaplygin, which allows a larger class of nonholonomic systems to be Hamiltonized by reparameterizing time, to higher dimensions. Moreover, in some cases we show that the requirement that the original system possess an invariant measure can be removed.

The results are then applied to show that under certain conditions the equations of motion of nonholonomic systems can be derived by considering an associated first-order optimal control problem, similar to the situation in holonomic systems. Moreover, the methods are illustrated throughout by various well known examples of nonholonomic systems. Several future directions based on the research presented are also discussed, among them the relatively new problem of quantizing a nonholonomically constrained system. With the advent of nanomachines we expect the importance of subatomic motions in wheeled robots to raise interest in the classical-quantum equations of motion governing these nonholonomic vehicles. Although there is currently no accepted quantum mechanical treatment of nonholonomic mechanics, we discuss the application of the results of the Hamiltonizations obtained herein to the quantization of a well known nonholonomic mechanical system.

CHAPTER I

Introduction

The mechanics of systems with nonintegrable constraints, also called *nonholonomic mechanics*, has a long and at times confusing history [6, 19, 83]. Part of the difficulty in the development of its basic structure can be traced back to incorrect applications of Lagrange's equations for unconstrained systems, resulting in incorrect equations of motion for nonholonomic systems. These errors were eventually clarified by, and the resulting true equations of motion derived by, Hertz [62], Poincare [76] and Ferrers [42] at the turn of the 20th century. Subsequent research by Appell [3, 4], Chaplygin [24, 25], Hamel [60] and Maggi [66], to name a few, then resulted in several different yet equivalent forms of the equations of motion of nonholonomic systems. The mechanics of nonholonomic systems was finally put in a geometric context beginning with the work of [13]. However, it quickly became clear that nonholonomic systems are not variational [6], and therefore cannot be represented by canonical Hamiltonian equations. Unfortunately, even today some authors incorrectly claim this to be possible (see [43] for a discussion), but perhaps due in part to these errors several authors (dating back at least as early as S.A. Chaplygin [24, 25]) have attempted to express the mechanics of nonholonomic systems in Hamilton-like forms through several methods. We make no attempt here to give a complete overview of all such attempts, but instead will focus on those relevant to this thesis (for excellent reviews, see [6, 33, 71, 83]). Moreover, let us emphasize that the work contained in this thesis differs fundamentally from the above referenced works in that we will study the Hamiltonization of the full nonholonomic system (reduced mechanics plus kinematic constraints). Based on this idea, we will present three different Hamiltonization methods in the subsequent chapters, each building on and extending some of the aforementioned works.

Perhaps the most straightforward method one can use in attempting to "Hamiltonize" nonholonomic systems is to apply the so-called Helmholtz conditions [31] in an attempt to extract a Lagrangian whose Euler-Lagrange equations reproduce the mixed first- and second-order equations of nonholonomic systems. The search for such a Lagrangian is called the *inverse problem of the Calculus of Variations* [81]. Of course, this approach will never work for as we have mentioned nonholonomic systems are not Hamiltonian systems. Some authors [1] have nevertheless found ways to construct second-order systems whose mechanics, when restricted to certain invariant submanifolds, reproduce the mechanics of certain nonholonomic systems by starting with the explicit trajectories of the system. However, we will show in Chapter IV that we can associate certain second-order systems to the equations of nonholonomic systems for which a solution to the inverse problem can be found without need to solve the system explicitly. With the resulting Lagrangians, we will then show how the nonholonomic equations arise from the restriction of the derived Hamiltonian system to certain invariant submanifolds.

The methods of Chapter IV will have some drawbacks however, forcing us to examine other avenues for Hamiltonization. Perhaps the most logical next step is to consider the dynamics of *variational nonholonomic systems* [6]. The distinction between variational nonholonomic dynamics and nonholonomic mechanics is perhaps best described by Hertz's terminology, which describes the former as equations of "shortest" curves and the latter as equations of "straightest" curves (see [22] and references therein). Geometrically, this is because variational dynamics are equivalent to optimal control problems under certain regularity conditions [6, 72], and as such its equations of motion are geodesics of a Levi-Civita connection. The nonholonomic equations of motion, on the other hand, are geodesics of a projected connection which is in general not metrizable (hence eliminating the possibility of viewing them as curves of minimum length). In fact, it is known that the resulting equations of motion are independent of the method used (Lagrange-d'Alembert or variational dynamics) if and only if the constraints are integrable [6, 22], or *holonomic*. As a result, many authors have studied the similarities and differences between nonholonomic mechanics and variational nonholonomic dynamics (see [7, 22, 27, 34, 48, 80]) in an effort to gain insight into the nonintegrable case. Unfortunately, some authors have also caused confusion by incorrectly claiming that the two methods both give the physical equations of motion (see [65] and references therein). However, despite this and the works cited above, in this thesis we will study the similarities between nonholonomic mechanics and variational nonholonomic dynamics with an eye toward Hamiltonization. In Chapter V we introduce the notion of *conditionally variational* nonholonomic systems as another avenue for Hamiltonization. These are systems whose equations of motion can be derived from the variational nonholonomic equations provided the nonholonomic constraints are satisfied initially (as they must be anyway). There we show that we can use the Lagrangian associated with the variational nonholonomic dynamics to Hamiltonize the nonholonomic system. However, we will see that this special class of nonholonomic systems is rare and seek to generalize the idea by appealing to Chaplygin's Reducing Multiplier Theorem.

S.A. Chaplygin's own *Reducing Multiplier Theorem* [24, 25] states that for nonholonomic systems in two generalized coordinates (q_1, q_2) possessing an invariant measure with density $N(q_1, q_2)$, the equations of motion can be written in Hamiltonian form after the time reparameterization $d\tau = N dt$[1], where N is known the as *reducing multiplier*, or simply the *multiplier*. This is often referred to as the *Hamiltonization* of nonholonomic systems, although we shall refer to it here as *Chaplygin Hamiltonization* instead[2]. Although Chaplygin's Theorem allows for a Hamiltonization of two degree of freedom nonholonomic systems, it is bounded by this restriction as well as the requirement that the system possess an invariant measure. Subsequent research on the Theorem has resulted in, among other things, an extension to the quasicoordinate context [73], a study of the geometry behind the theorem [44, 45], discoveries of isomorphisms between nonholonomic systems through the use of the theorem [18], an example of a system in higher dimensions Hamiltonizable through a similar time reparameterization [37], an investigation of the necessary conditions for Hamiltonization for abelian Chaplygin systems [64] (see Section 3.6.1 for a definition) and an investigation of rank two Poisson structures in nonholonomic systems [78]. However, two important aspects yet to be resolved are the extension of the theorem to general nonholonomic systems with symmetry of arbitrary degrees of freedom, and the extension to nonholonomic systems not possessing an invariant measure. In Chapter VI we consider the aforementioned questions for two special cases of the Hamilton-Poincaré-d'Alembert equations[3] and present results in Sections 6.1 and 6.2

[1] The second part of his Theorem states that if a nonholonomic system can be transformed into Hamiltonian form after the time reparameterization $d\tau = f dt$, then the original system has an invariant measure with density f^{m-1}, where m is the number of degrees of freedom.

[2] We introduce this term because we wish to differentiate it from the other forms of Hamiltonization discussed in this thesis.

[3] These are the governing equations for nonholonomic systems with symmetry satisfying the dimension assumption, see Section 3.6.

which generalize Chaplygin's theorem to higher dimensional nonholonomic systems with symmetry, focusing mainly on extending the work of [64] to the necessary conditions for Chaplygin Hamiltonization for nonabelian Chaplygin systems, as well to nonholonomic systems on Lie groups. For the latter type of systems, we present results which allow Chaplygin Hamiltonization even when the system does not posses an invariant measure. Furthermore, in Section 6.2 we extend the idea of conditionally variational systems introduced in [38] and apply it to Chaplygin Hamiltonize the entire nonholonomic system (reduced constrained equations plus the nonholonomic constraints).

We begin in Chapter II with a brief review of the relevant Hamiltonian mechanics, generalizing to nonholonomic mechanics in Chapter III. After discussing Hamiltonization by associated second-order systems in Chapter IV, we proceed to compare the variational nonholonomic and Lagrange-d'Alembert equations in Chapter V and introduce the class of systems known as conditionally variational nonholonomic systems. We generalize Chaplygin's theorem in Chapter VI and finally apply the results of the thesis to recover the correspondence between the Pontryagin maximum principle and constrained mechanics in Chapter VII. There we also illustrate the various results by examining some well-known nonholonomic systems. In Chapter VIII we conclude with a discussion of future research directions which make use of the results of this thesis to address the quantization of nonholonomic systems, to further the development of a Hamilton-Jacobi theory for them and to perhaps construct more efficient numerical schemes for integrating the equations of motion of nonholonomic systems.

CHAPTER II

Hamiltonian Mechanics

In this chapter we summarize some basic concepts in the geometric mechanics of unconstrained Hamiltonian systems, assuming familiarity with basic differential geometry [47, 67]. We begin with the traditional variational formulation of mechanical system in terms of Hamilton's Principle on the Lagrangian side. Through the Legendre transform we then summarize the geometric structure behind the Hamiltonian side and then discuss the related Poisson structure. We then add in symmetry and introduce the Euler-Poincare and Lie-Poisson equations. Finally, we give a brief summary of the inverse problem of the calculus of variations in preparation for its use in Chapter IV.

Our exposition here is largely based on that found in [6, 77] and we wish to remark that the Einstein summation convention is enforced throughout this thesis unless otherwise noted. In addition, we will restrict our attention throughout the thesis to finite dimensional systems.

2.1 Unconstrained Mechanics and Hamilton's Principle

Let Q be a manifold, with TQ its tangent bundle. Denote by q^i the coordinates on Q and by $(q^i, \dot{q}^i)$ the induced coordinates on TQ. Define the *mechanical Lagrangian* $L : TQ \to \mathbb{R}$ given by $L = T - V$, where $K(v) = \frac{1}{2}\langle v, v\rangle$ is the *kinetic energy* associated with a given Riemannian metric and where $V : Q \to \mathbb{R}$ is the *potential energy*. The trajectories of an unconstrained mechanical system are then given by *Hamilton's Principle*, which states that among the set of possible motions $q(t)$ of our mechanical system in any time interval $[a, b]$, the actual trajectories are such that[1]

$$\delta \int_a^b L(q(t), \dot{q}(t))\, dt = 0. \tag{2.1}$$

We say that a mechanical system (unconstrained or constrained) is *variational* if its equations of motion can be derived from Hamilton's principle.

Basic results in the calculus of variations (see [6]) show that the condition (2.1) is equivalent to the requirement that $q(t)$ satisfies the *Euler-Lagrange* equations:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0. \tag{2.2}$$

Now, if we define the *fiber derivative* $\mathbb{F}L : TQ \to T^*Q$ in coordinates by the map $(q^i, \dot{q}^j) \mapsto (q^i, p_j)$, where $p_j = \partial L/\partial \dot{q}^j$ is called the *momentum conjugate to* q^j, then assuming that L is hyperregular[2] we can define the *Hamiltonian* H by $H(q, p) = p_i \dot{q}^i - L$. The coordinates (q^i, p_i) on the cotangent bundle T^*Q are called the *canonical cotangent coordinates* and the change of data from L on TQ to H on T^*Q is called the *Legendre transform.*

As we shall see below, the Hamiltonian H is related to the total energy of the mechanical sytem, and since the cotangent space T^*Q carries a natural symplectic structure, we will summarize the rich geometry of Hamiltonian mechanics below and present the analogue of (2.2), the Hamiltonian equations of motion.

[1] For a detailed discussion of the variations of a curve see [6].
[2] A Lagrangian L is *hyperregular* if $\mathbb{F}L$ is a diffeomorphism.

2.2 Hamiltonian Mechanics and Symplectic Geometry

We begin our discussion by recalling some basic definitions in symplectic geometry [77]. A *symplectic form* on a smooth manifold M is a nondegenerate closed 2-form ω on M and a *symplectic manifold* is a pair (M, ω) with M a smooth manifold and ω a symplectic form on it. If (M_1, ω_1) and (M_2, ω_2) are symplectic manifolds then a C^∞ mapping $h : M_1 \to M_2$ is called *symplectic* if $h^*\omega_2 = \omega_1$. Using this, we can now define a Hamiltonian system in general.

Definition II.1. [77] Let (M, ω) be a symplectic manifold and[3] $H \in C^\infty(M, \mathbb{R})$ a smooth real valued function on M. The vector field X_H determined by the condition

$$i_{X_H}\omega = dH \tag{2.3}$$

is called the *Hamiltonian vector field with energy function* H. We call (M, ω, H) a *Hamiltonian mechanical system*.

Let us now take the case when $M = T^*Q$. In this case there is a unique 1-form θ on T^*Q such that in any choice of canonical cotangent coordinates, $\theta = p_i dq^i$. Using this we can then define the *canonical 2-form* ω by $\omega = -d\theta = dq^i \wedge dp_i$. It is then clear that (T^*Q, ω) is a symplectic manifold. A simple computation then shows that $(q(t), p(t))$ is an integral curve of X_H iff *Hamilton's equations* hold:

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}. \tag{2.4}$$

Moreover, (2.2) and (2.4) are equivalent via the Legendre transform.

Embedded in the definition of Hamiltonian systems above are the following two facts.

[3]Hereafter, $C^\infty(X, \mathbb{R})$ is the set of infinitely differentiable real-valued functions on the space X.

Proposition II.2. *[77] (Conservation of Energy) Let (M, ω, H) be a Hamiltonian mechanical system and let $c(t)$ be an integral curve for X_H. Then $H(c(t))$ is a constant in t. Moreover, if ϕ_t is the flow of X_H then $H \circ \phi_t = H$ for each t.*

Proposition II.3. *[77] (Volume Preservation) Let (M, ω, H) be a Hamiltonian mechanical system and let ϕ_t be the flow of X_H. Then for each t, $\phi_t^* \omega = \omega$, that is, ϕ_t is symplectic and volume preserving.*

These two facts are hallmarks of Hamiltonian systems. Nonholonomic systems, on the other hand, conserve energy yet do not in general preserve volume (see Chapter III). Moreover, we will see in Chapters V and VI that nonholonomic systems which *do* preserve volume are in a quantifiable sense closer to Hamiltonian systems than their volume changing counterparts. We will exploit this in those chapters in order to create and extend Hamiltonization methods.

Now, more general mechanical systems can be defined by using the notion of Poisson structures, to which we now turn.

2.3 Poisson Structures in Hamiltonian Mechanics

Let P be a manifold and consider the bracket operation denoted by

$$\{\cdot, \cdot\} : C^\infty(P, \mathbb{R}) \times C^\infty(P, \mathbb{R}) \to C^\infty(P, \mathbb{R}). \tag{2.5}$$

The pair $(P, \{\cdot, \cdot\})$ is called a *Poisson manifold* and $\{\cdot, \cdot\}$ a *Poisson bracket* if [6] $\{\cdot, \cdot\}$ is (i) bilinear, (ii) anticommutative, (iii) satisfies Leibniz's rule, and (iv) satisfies *Jacobi's identity*: $\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$. If only conditions (i)-(iii) hold (i.e. the Jacobi identity is not satisfied) then the bracket is called an *almost-Poisson bracket*. Letting z^i denote local coordinates on P, the Poisson bracket is given by

$$\{f, g\} = B^{ij} \frac{\partial f}{\partial z^i} \frac{\partial g}{\partial z^j}, \tag{2.6}$$

where B^{ij} are the components of the antisymmetric tensor B on P (see [6] for details).

Now, it is well-known that every symplectic manifold is a Poisson manifold [6, 67], with Poisson bracket given by $\{f, g\} = \omega(X_f, X_g)$. However, the converse is not true and is perhaps most importantly illustrated by the Lie-Poisson structure associated with the rigid body, to which we now turn.

2.4 The Euler-Poincaré and Lie-Poisson Equations

Suppose that the configuration space for our mechanical system is a Lie Group G and let $L : TG \to \mathbb{R}$ be a left-invariant Lagrangian. Denote by $l : \mathfrak{g} \to \mathbb{R}$ the restriction of L to the tangent space of G at the identity. Moreover, for a curve $g(t) \in G$ let $\xi(t) = g(t)^{-1}\dot{g}(t)$. Then the following are equivalent (see [6]):

(i) $g(t)$ satisfies the Euler-Lagrange equations for L on G.

(ii) The *Euler-Poincaré equations* hold:

$$\frac{d}{dt} \frac{\partial l}{\partial \xi} = \mathrm{ad}^*_{\xi} \frac{\partial l}{\partial \xi}, \tag{2.7}$$

where $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ is defined by $\mathrm{ad}_\xi \nu = [\xi, \nu]$ and ad^*_ξ is its dual. By making the following Legendre transformation from $\mathfrak{g}$ to $\mathfrak{g}^*$:

$$\mu = \frac{\partial l}{\partial \xi}, \quad h(\mu) = \mu_i \xi^i - l(\xi), \tag{2.8}$$

it follows that the Euler-Poincaré equations are equivalent to the *Lie-Poisson equations*:

$$\frac{d\mu}{dt} = \mathrm{ad}^*_{\partial h/\partial \mu}\mu. \tag{2.9}$$

Assuming that $\mathfrak{g}$ is finite-dimensional and choosing coordinates $(\xi^1, \ldots, \xi^n)$ on $\mathfrak{g}$ and corresponding dual coordinates $(\mu_1, \ldots, \mu_m)$ on $\mathfrak{g}^*$, the (minus) the Lie-Poisson equations (2.9) can be written as:

$$\dot{\mu} = \{\mu, h\}_-, \quad \text{where } \{f, k\}_- = -\mu_a C^a_{bc} \frac{\partial f}{\partial \mu_b} \frac{\partial k}{\partial \mu_c}, \tag{2.10}$$

with C^a_{bc} the *structure constants* of $\mathfrak{g}$ defined by $[e_a, e_b] = C^c_{ab} e_c$, where $(e_1, \ldots, e_n)$ is the coordinate basis of $\mathfrak{g}$, and where for $\xi \in \mathfrak{g}$ we write $\xi = \xi^a e_a$ and for $\mu \in \mathfrak{g}^*$ we write $\mu = \mu_a e^a$, with (e^a) the dual basis.

The bracket $\{\cdot, \cdot\}_-$ is called the *(minus) Lie-Poisson bracket* and the dual space $\mathfrak{g}^*$ becomes a Poisson (but not symplectic) manifold with respect to the Lie-Poisson bracket. The classical Euler equations [6, 67] for a rigid body provide an example of this particular Poisson manifold.

2.5 The Inverse Problem of the Calculus of Variations

The previous sections began with the definition of Hamiltonian systems and Hamilton's principle and derived from that the second order equations of motion for unconstrained mechanical systems. However, there is also an inverse procedure known as the *inverse problem of the calculus of variations* which determines if a given set of second-order differential equations are in fact the Euler-Lagrange equations of some Lagrangian. This procedure has a long history (for a recent survey on this history see [59]). A solution to the problem indicates that the system of equations under study is in fact variational and hence a Hamiltonian can be defined. Below we provide a brief outline of the method in preparation for its use in Chapter IV.

Let Q be a manifold with local coordinates (q^i) and assume we are given a system of second-order ordinary differential equations $\ddot{q}^i = f^i(q,\dot{q})$ on Q. In order for a regular Lagrangian $L(q,\dot{q})$ to exist we must be able to find functions $g_{ij}(q,\dot{q})$, so-called *multipliers*, such that

$$g_{ij}(\ddot{q}^j - f^j) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}.$$

It can be shown [31, 81] that the multipliers must satisfy the so-called *Helmholtz conditions*:

$$\begin{aligned}
&\det(g_{ij}) \neq 0, \qquad g_{ji} = g_{ij}, \qquad \frac{\partial g_{ij}}{\partial \dot{q}^k} = \frac{\partial g_{ik}}{\partial \dot{q}^j};\\
&\Gamma(g_{ij}) - \nabla^k_j g_{ik} - \nabla^k_i g_{kj} = 0,\\
&g_{ik}\Phi^k_j = g_{jk}\Phi^k_i;
\end{aligned}$$

where $\nabla^i_j = -\frac{1}{2}\partial_{\dot{q}^j} f^i$ and

$$\Phi^k_j = \Gamma\left(\partial_{\dot{q}^j} f^k\right) - 2\partial_{q^j} f^k - \frac{1}{2}\partial_{\dot{q}^j} f^l \partial_{\dot{q}^l} f^k.$$

Here the symbol Γ stands for the vector field $\dot{q}^i\partial_{q^i} + f^i\partial_{\dot{q}^i}$ on TQ that can naturally be associated to the system $\ddot{q}^i = f^i(q,\dot{q})$ and for our purposes here, we will fix from the start $g_{ij} = g_{ji}$ for $j \leq i$ and simply write g_{ijk} for $\partial_{\dot{q}^k} g_{ij}$ and also assume the notation to be symmetric over all its indices.

Conversely, if one can find functions g_{ij} satisfying these the Helmholtz conditions then the equations $\ddot{q}^i = f^i$ are derivable from a regular Lagrangian. Moreover, if a regular Lagrangian L can be found, then its Hessian $\dfrac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}$ is a multiplier.

Now, the Helmholtz conditions are a mixed set of coupled algebraic and partial differential conditions in (g_{ij}). We will refer to the penultimate condition as the '∇-condition,' and to the last one as the 'Φ-condition.' These algebraic Φ-conditions are of course the easiest to start from. But in fact, we can easily derive more algebraic conditions (see e.g. [29]). For example, by taking a Γ-derivative of the Φ-condition, and by replacing $\Gamma(g_{ij})$ everywhere by means of the ∇-condition, we arrive at a new algebraic condition of the form

$$g_{ik}(\nabla\Phi)^k_j = g_{jk}(\nabla\Phi)^k_i,$$

where $(\nabla\Phi)^i_j = \Gamma(\Phi^i_j) - \nabla^i_m\Phi^m_j - \nabla^m_j\Phi^i_m$. As in [29], we will call this new condition the $(\nabla\Phi)$-condition. It will, of course, only give new information as long as it is independent from the Φ-condition (this will not be the case, for example, if the commutator of matrices $[\Phi, \nabla\Phi]$ vanishes). One can repeat the above process on the $(\nabla\Phi)$-condition, and so on to obtain possibly independent $(\nabla\ldots\nabla\Phi)$-conditions.

A second route to additional algebraic conditions arises from the derivatives of the Φ-equation in $\dot{q}$-directions. One can sum up those derived relations in such a way that the terms in g_{ijk} disappear on account of the symmetry in all their indices. The new algebraic relation in g_{ij} is then of the form

$$g_{ij}R^j_{kl} + g_{lj}R^j_{ik} + g_{kj}R^j_{li} = 0,$$

where $R^j_{kl} = \partial_{\dot{q}^j}(\Phi^k_i) - \partial_{\dot{q}^i}(\Phi^k_j)$. For future use, we will call this the *R-condition.*

As before, this process can be continued to obtain more algebraic conditions, with any mixture of the above mentioned two processes leading to possibly new and independent algebraic conditions. Once we have used up all the information that we can obtain from this infinite set of algebraic conditions, we can start looking at the partial differential equations in the ∇-conditions.

We will employ the Helmholtz conditions in the next chapter to begin our Hamiltonization of nonholonomic systems. However, due to the complexity of the partial differential equations encountered, even the simplest nonholonomic systems can present a formidable challenge to Hamiltonize by solving the Helmholtz conditions.

Despite this, we will be able to extract some rather general results on the Hamiltonization of a certain class of nonholonomic systems. In Chapter VII we will then apply some of the obtained results to optimal control problems associated to this class of nonholonomic systems.

CHAPTER III

Nonholonomic Mechanics

As we saw in Chapter II, Hamiltonian mechanics can be described in three equivalent ways: (i) through the variational principle of Hamilton (Section 2.1), (ii) by the existence of a Lagrangian that satisfies the Euler-Lagrange equations (Section 2.2), or (iii) through the use of Poisson brackets (Section 2.3). In contrast, nonholonomic mechanics fails to be expressible in any of these ways (as we will see below). This thesis is largely concerned with presenting Hamiltonization methods based on this failure and we will address (i), (ii) and (iii) in Chapters IV, III and V below. For now let us summarize the situation on the nonholonomic side.

3.1 Constraints in Mechanics

Consider a mechanical system subject to linear velocity constraints that can be expressed in generalized coordinates as $a_{ij}(q)\dot{q}^j = 0$ for $i = 1, \ldots, k < dim(Q)$. We call this constraint *holonomic* if (locally) there is a real-valued function $h(q)$ such that the constraint can be written as $h(q) = constant$. If no such function exists, the constraint is said to be *nonholonomic*. Equivalently, constraints are holonomic if their corresponding distribution[1] is integrable (in the sense of Frobenius) and nonholonomic otherwise.

[1] A distribution $\mathcal{D}$ is a collection of linear subspaces denoted by $\mathcal{D}_q \subset T_qQ$, one for each q.

3.2 Constrained Mechanics and the Lagrange-d'Alembert Principle

Suppose now that we have a mechanical system on Q with Lagrangian L and let $\mathcal{D}$ be the distribution describing the kinematic constraints of Section 3.1. The equations of motion are then given by the *Lagrange-d'Alembert principle* [6]:

$$\delta \int_a^b L(q^i, \dot{q}^i)\, dt = 0, \tag{3.1}$$

where the variations $\delta q(t)$ of the curve $q(t)$ satisfy $\delta q(t) \in \mathcal{D}_q(t)$ for each $t \in [a, b]$ and $\delta q(a) = \delta q(b) = 0$, along with the requirement that $\dot{q}(t) \in \mathcal{D}_{q(t)}$ for all t (i.e. that the curve $q(t) \in Q$ satisfy the constraints). In order to distinguish the holonomic and nonholonomic situations let us consider the constrained equations of motion resulting from (3.1).

3.3 The Equations of Motion for Constrained Mechanics

The equations of motion for constrained mechanics (both holonomic and nonholonomic) can be derived from a generalization of Hamilton's principle (as we shall do below). For our purposes, we shall consider a *constrained mechanical system* on a configuration manifold Q to be a pair $(L, \mathcal{D})$, where $L : TQ \to \mathbb{R}$ is a regular Lagrangian of mechanical type $L = T - V$, where $T : TQ \to \mathbb{R}$ is the kinetic energy corresponding to a Riemannian metric g on Q, and $V : Q \to \mathbb{R}$ is the potential energy, and $\mathcal{D}$ is the vector subbundle of TQ defined by the null space of k independent constraint one-forms ω^a [6, 26]. Moreover, in a neighborhood of each point, one can choose a local coordinate chart such that ω^a and $\mathcal{D}$ take the form:

$$\omega^a = ds^a + A^a_\alpha(r,s)dr^\alpha, \tag{3.2}$$

$$\mathcal{D} = \text{span}\,\{\partial_{r^\alpha} - A^a_\alpha(r,s)\partial_{s^a}\}, \tag{3.3}$$

respectively, where $q = (r,s) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$ and hereafter we make the index conventions: $a, b, c = 1, \ldots, k$, $\alpha, \beta, \gamma = 1, \ldots, m$, where $n = \dim\, Q$ and $m := n - k$ is the number of *degrees of freedom* of the constrained system.

Now define the vector bundle with coordinates (r^α, s^a) and projection map $\pi : (r^\alpha, s^a) \to r^\alpha$. Introducing the vertical space $V_q := \ker\, T_q\pi$, we can define an Ehresmann connection $A_q : T_qQ \to V_q$ represented locally by the vector-valued differential form ω^a:

$$A = (ds^a + A^a_\alpha(r,s)dr^\alpha)\partial_{s^a}. \tag{3.4}$$

It follows that the horizontal space $H_q := \ker\, A_q = \mathcal{D}$, and that $TQ = V_q \oplus H_q$, so that we can project a tangent vector onto its vertical and horizontal parts using the connection[2]. Defining the *constrained Lagrangian* $L_c(q, \dot{r}) = L(q, \text{hor}\ \dot{q})$, the corresponding equations of motion for a constrained mechanical system are obtained through the Lagrange-d'Alembert principle (3.1), and are given by:

$$\begin{aligned} \delta L_c &= \langle \mathbb{F}L, B(\dot{q}, \delta q)\rangle, \\ \text{where } \delta L_c &= \langle \delta q^\alpha, \frac{\partial L_c}{\partial q^\alpha} - \frac{d}{dt}\frac{\partial L_c}{\partial \dot{q}^\alpha}\rangle, \end{aligned} \tag{3.5}$$

and where $\langle \cdot, \cdot \rangle$ denotes the pairing between a vector and a dual vector. Here, $\mathbb{F}L : TQ \to T^*Q$ is the fiber derivative of Section 2.1, δq is a horizontal variation,

[2] In coordinates, the horizontal projection hor X_q of a vector $X_q \in T_qQ$ is the map $(\dot{r}^\alpha, \dot{s}^a) \mapsto (\dot{r}^\alpha, -A^a_\alpha(r,s)\dot{r}^\alpha)$.

and B is the curvature of the Ehresmann connection A regarded as a vertical-valued two-form (see below). These equations are to be supplemented by the constraint equations (given by (3.7) in local coordinates).

The curvature of A is the vertical-vector-valued-two-form B on Q defined by $B(X,Y) = -A([\mathrm{hor}\, X, \mathrm{hor}\, Y])$, where the Jacobi-Lie bracket of vector fields on the right hand side is obtained by extending the vectors X and Y on Q to vector fields[3]. In this form it becomes apparent that the curvature exactly measures the failure of the horizontal distribution $\mathcal{D}$ to be integrable (in the Frobenius sense). Hence, for holonomic systems this curvature vanishes, meaning from (3.5) that the equations of motion are given by $\delta L_c = 0$ along with the (integrable) constraint equations (3.7). This is then equivalent to the statement that for holonomic systems, one can "plug in the constraints" and compute the Euler-Lagrange equations of L_c to obtain the mechanics of the unconstrained variables. In nonholonomic systems, on the other hand, even after substituting in the constraints and arriving at L_c the constraint forces given by the right hand side of the first line in (3.5) must be taken into account. An important result is then that [6] *a nonholonomic system is variational iff it is holonomic*. We will see this more directly in Section 3.3 below.

Locally, we therefore have the constrained equations of motion along with the constraint equations given as:

$$
\begin{aligned}
&(3.6) \qquad & \frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} &= -\left(\frac{\partial L}{\partial \dot{s}^b}\right)_c B^b_{\alpha\beta}\dot{r}^\beta - A^a_\alpha \frac{\partial L_c}{\partial s^a}, \\
&(3.7) \qquad & \dot{s}^a &= -A^a_\alpha(r,s)\dot{r}^\alpha,
\end{aligned}
$$

where the local expression for the curvature is given by:

[3] Moreover, the curvature can be shown to be independent of the extension of the vector fields [6].

$$B^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial A^b_\beta}{\partial s^a} - A^a_\beta \frac{\partial A^b_\alpha}{\partial s^a}. \tag{3.8}$$

Note that the first term on the right hand side of (3.6) is the only instance where the $\dot{s}^b$ occur. We can eliminate this through the constraints (3.7), and we shall henceforth denote with a subscript c any expression for which we have eliminated the fiber dependency by using (3.7).

Equations (3.6) and (3.7) then describe the mechanics of the nonholonomic system $(L, \mathcal{D})$, and for ease of use later, we henceforth define $\phi^a := \dot{s}^a + A^a_\alpha(r, s)\dot{r}^\alpha$.

3.4 Variational Constrained Mechanics

Another way to derive equations of motion for a constrained mechanical system is to use Lagrange's Multiplier theorem [43]. In brief, define the space $\tilde{Q} = Q \times M$, where dim Q=n as before, and dim M=k, and where locally we denote the extra coordinates of $\tilde{Q}$ by $\mu_1(t), \ldots, \mu_k(t)$. We shall call the $\mu_a(t)$ the *multipliers* and form the augmented Lagrangian $L_V : T\tilde{Q} \to \mathbb{R}$:

$$L_V = L - \mu_a \phi^a. \tag{3.9}$$

Note that L_V is automatically singular, due to the absence of $\dot{\mu}$.

Applying Hamilton's principle to the augmented Lagrangian then yields the unconstrained equations of motion

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \dot{\mu}_a \frac{\partial \phi^a}{\partial \dot{q}^i} + \mu_a \left(\frac{d}{dt}\frac{\partial \phi^a}{\partial \dot{q}^i} - \frac{\partial \phi^a}{\partial q^i} \right), \tag{3.10}$$

where $i = 1, \ldots, n$, as well as the equations of constraint (3.7) which arise as the Euler-Lagrange equations of the μ_a coordinates. This set of equations is sometimes called the *vakonomic equations*, after Arnold, Kozlov and Neishtadt [2] who introduced nonholonomic dynamics under the Lagrange variational point of view. We shall thus prefer to call these equations the *variational constrained equations*.

Now, if the constraints are holonomic then one can show [43] that the variational constrained equations reproduce the equations of motion for a holonomic system, showing more directly that holonomically constrained systems are variational. However, if the constraints are nonholonomic then the variational constrained equations do *not* reproduce the nonholonomic equations[4]. When the constraints are nonholonomic, we will call the variational constrained equations the *variational nonholonomic equations*.

3.5 Almost Poisson Structures in Nonholonomic Mechanics

Returning to the equations of motion (3.6), one can define the constrained momenta $p_\alpha = \partial L_c / \partial \dot{r}^\alpha$ and, assuming L_c is hypperregular, the constrained Hamiltonian $H_{\mathcal{M}}$ on the constraint phase space $\mathcal{M} := \mathbb{F}L(\mathcal{D}) \subset T^*Q$. The equations of motion (3.6) and constraint (3.7) then become [6, 84]:

$$\dot{r}^\alpha = \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\alpha}, \tag{3.11}$$

$$\dot{p}_\alpha = -\frac{\partial H_{\mathcal{M}}}{\partial r^\alpha} + A^k_\alpha \frac{\partial H_{\mathcal{M}}}{\partial s^k} - p_\gamma R^\gamma_l B^l_{\alpha\beta} \frac{\partial H_{\mathcal{M}}}{\partial p_\beta}, \tag{3.12}$$

$$\dot{s}^k = -A^k_\beta \frac{\partial H_{\mathcal{M}}}{\partial p_\beta}, \tag{3.13}$$

respectively, where R^γ_l is defined by $M_{a\alpha} \partial H_{\mathcal{M}} / \partial p_\alpha = R^\gamma_a p_\gamma$, where $M_{a\alpha} = g_{a\alpha} - g_{ak} A^k_\alpha$, with g_{ij} the components of the kinetic energy metric of the unconstrained Lagrangian, $B^l_{\alpha\beta}$ is the curvature of the connection, $H_{\mathcal{M}} = p_\alpha \dot{r}^\alpha - L_c$ is the constrained Hamiltonian, and $p_\alpha = (\partial L / \partial \dot{r}^\alpha) - (\partial L / \partial \dot{s}^k) A^k_\alpha$. These are a set of $2n - m$ equations

[4]Sadly, several authors and even textbooks have incorrectly claimed otherwise, causing much confusion in the area (see the references in [43] for more details).

on the submanifold $\mathcal{M}$ with induced coordinates $(r^\alpha, s^k, p_\alpha)$, and are manifestly non-Hamiltonian, a reflection of the fact that the presence of nonholonomic constraints induces additional forces that enforce those constraints. Furthermore, the constrained equations (3.11)-(3.12) can be written in terms of the almost-Poisson bracket $\{\cdot,\cdot\}_{AP}$ as:

$$\dot{r}^\alpha = \{r^\alpha, H_{\mathcal{M}}\}_{AP}, \quad \dot{p}_\alpha = \{p_\alpha, H_{\mathcal{M}}\}_{AP}, \tag{3.14}$$

where the almost-Poisson bracket is given by[5]:

$$\{f,g\}_{AP} = \{f,g\}_{can} - A^b_\alpha \left(\frac{\partial f}{\partial s^b} \frac{\partial g}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha} \frac{\partial g}{\partial s^b} \right) - p_\gamma K^\gamma_l B^l_{\alpha\beta} \frac{\partial f}{\partial p_\alpha} \frac{\partial g}{\partial p_\beta}, \tag{3.15}$$

for any two real-valued functions f, g on $\mathcal{M}$. Now, since (3.15) is a bracket for only the constrained mechanics, there are rare nonholonomic systems for which it is in fact a Poisson bracket (see [38]). However, when one considers the almost-Poisson bracket for the entire system (3.11)-(3.13), one can then show [6] that this bracket is a Poisson bracket (i.e. satisfies the Jacobi identity) iff the constraints are holonomic. This is the third and final manifestation of the difference between Hamiltonian mechanics and nonholonomic mechanics: the latter is not variational, not the Euler-Lagrange mechanics of any Lagrangian and cannot be expressed in terms of a Poisson bracket.

3.6 Nonholonomic Systems with Symmetry

Consider, as before, a nonholonomic system with an n dimensional configuration manifold Q and mechanical Lagrangian L which is subject to k linear nonholonomic constraints described by the distribution $\mathcal{D}$. Suppose that we have a Lie group G

[5] Here $\{f,g\}_{can} = \frac{\partial f}{\partial r^\alpha}\frac{\partial g}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial r^\alpha}$.

which acts freely and properly on the configuration space Q, with the Lagrangian L and constraints $\mathcal{D}$ invariant with respect to the induced action of G on TQ. For simplicity, assume also that the constraints and the orbit directions span the entire tangent space to the configuration space:

$$\mathcal{D}_q + T_q\mathrm{Orb}(q) = T_qQ, \tag{3.16}$$

sometimes known as the *dimension assumption* [6]. Under this setup, the resulting nonholonomic equations of motion are known as the *Hamilton-Poincaré-d'Alembert* equations and split into a coupled set of second-order equations on the *shape space* $M := Q/G$ and first-order nonholonomic momentum equations on $\mathfrak{g}^*$ [6, 55], whose number equals *dim* $\mathcal{S}_q$, where $\mathcal{S}_q := \mathcal{D}_q \cap T_q\mathrm{Orb}(q)$.

These equations can also be collectively written in bracket form with respect to an almost-Poisson bracket $\{\cdot,\cdot\}_{\overline{\mathcal{M}}}$, where $\overline{\mathcal{M}} = \mathcal{M}/G$, with $\mathcal{M} = \mathbb{F}L(\mathcal{D}) \subset T^*Q$ as before. In general, this bracket does not satisfy the Jacobi identity, preventing these reduced, constrained equations of motion from being expressed as the equations of a Hamiltonian system.

For future reference, we will now describe two special cases: (1) where $\mathcal{S}_q = \{0\}$, known as the *purely kinematic* or *nonabelian Chaplygin* case, and (2) the case where $Q = G$, where the resulting equations represent a generalization of the *Euler-Poincaré-Suslov* equations [6].

3.6.1 Nonholonomic Chaplygin Systems

Consider the subclass of nonholonomic systems with symmetry corresponding to $\mathcal{S}_q = \{0\}$, known as the purely kinematic case [6], where the group orbits exactly complement the constraints. These are nonholonomic systems on an n dimensional configuration manifold Q characterized by a principal G-bundle $\pi : Q \to Q/G$, where $Q \neq G$, associated with a free and proper action Φ of G on Q. In addition they carry a mechanical Lagrangian $L : TQ \to \mathbb{R}$ and non-integrable distribution $\mathcal{D}$ describing the nonholonomic constraints which are both G-invariant with respect to the lifted action on TQ.

These systems are also known as *nonabelian Chaplygin systems* [6, 26, 54], since in the special case where $Q = \mathbb{R}^s \times S^r$ and G is either a torus action T^m or acts by translations $\mathbb{R}^{2m}$, they are called *abelian Chaplygin systems* and correspond to the classical exposition of Chaplygin systems [73] where there exist local coordinates (r^α, s^a), $\alpha = 1, \ldots, n-2m$, $a = n-2m+1, \ldots, n$ such that the Lagrangian L does not depend on the s^a coordinates, and where the constraints can be written as $\dot{s}^a = -A^a_\alpha(r)\dot{r}^\alpha$.

For nonabelian Chaplygin systems we can form the reduced velocity phase space TQ/G, and the Lagrangian L induces the reduced Lagrangian $l : TQ/G \to \mathbb{R}$ as well as the reduced constrained Lagrangian $l_c : T(Q/G) \to \mathbb{R}$. The nonholonomic Lagrange-d'Alembert equations of motion then induce well-defined reduced Lagrange-d'Alembert equations on $T(Q/G)$ [6, 26]:

$$\begin{aligned} \delta l_c &= \langle \frac{\partial l}{\partial \xi}, \mathcal{B}(\dot{r}, \delta r) \rangle, \\ \xi &= -\mathcal{A}\dot{r}, \end{aligned} \tag{3.17}$$

where $\delta r \in T(Q/G)$, $\mathcal{B}$ is the local expression for the curvature of $\mathcal{A}$, $\xi = g^{-1}\dot{g}$, and where we have chosen the bundle coordinates $r \in M := Q/G$, $g \in G$.

In a local trivialization $M \times G$ of π, the equations of motion are given as follows [26]. Let (r^α, g^a) be local coordinates for Q, where $\alpha = 1, \ldots, n-k$ and $a =$

$1, \ldots, k = \dim G$. Choosing a basis e_a $(a = 1, \ldots, k)$ of the Lie algebra $\mathfrak{g}$ and using the left trivialization $TG \cong G \times \mathfrak{g}$, we can write $\xi = \xi^a e_a \in \mathfrak{g}$ as $\xi = g^{-1}\dot{g}$. In terms of the coordinates $(r^\alpha, g^a, \dot{r}^\alpha, \xi^a)$ on TQ, the G-invariant reduced Lagrangian is $L = l(r^\alpha, \dot{r}^\alpha, \xi^a)$. Similarly, the constraints take the form $\xi^a = -\mathcal{A}^a_\alpha \dot{r}^\alpha$, where $\mathcal{A}^a_\alpha$ are the connection coefficients of the given principal connection. The reduced constrained Lagrangian is then given by $l_c = l(r^\alpha, \dot{r}^\alpha, -\mathcal{A}^a_\beta \dot{r}^\beta)$, and finally, the local equations corresponding to the system (3.17) are:

$$\frac{d}{dt}\frac{\partial l_c}{\partial \dot{r}^\alpha} - \frac{\partial l_c}{\partial r^\alpha} = -\left(\frac{\partial l}{\partial \xi^a}\right)_c \mathcal{B}^a_{\alpha\beta}\dot{r}^\beta, \tag{3.18}$$

$$\xi^a = -\mathcal{A}^a_\alpha(r)\dot{r}^\alpha, \tag{3.19}$$

where $\mathcal{B}^a_{\alpha\beta}$ are components of the curvature of $\mathcal{A}$ in local form, and where the subscript c denotes that we have substituted in the constraints. With the structure constants of $\mathfrak{g}$ with respect to our basis given by $[e_b, e_c] = C^a_{bc}e_a$, the curvature components are given by:

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

For easy reference later on, we also define the semi-basic two-form [20] Λ on TM with components

$$\Lambda_{\alpha\beta}(r, \dot{r}) := \left(\frac{\partial l}{\partial \xi^a}\right)_c \mathcal{B}^a_{\beta\alpha}, \tag{3.20}$$

so that the right hand side of (3.18) can also be expressed as $\Lambda_{\alpha\beta}\dot{r}^\beta$.

Defining the cojugate momenta $p_\alpha = \partial l_c / \partial \dot{r}^\alpha$ and assuming that l_c is hyperregular we can use the Legendre transform to define the reduced constrained Hamiltonian

$h_c : T^*M \to \mathbb{R}$, $h_c(r,p) = p_\alpha \dot{r}^\alpha - l_c|_{\dot{r}\to p}$, the nonholonomic equations (3.18) can be written with respect to an almost-Poisson bracket:

$$\dot{r}^\alpha = \{r^\alpha, h_c\}_{AP}, \quad \dot{p}_\alpha = \{p_\alpha, h_c\}_{AP}, \tag{3.21}$$

where the almost-Poisson bracket is given by [21]:

$$\{g,k\}_{AP}(r,p) = \{g,k\}_{can}(r,p) - \left[\left(\frac{\partial l}{\partial \xi^a}\right)_c \mathcal{B}^a_{\alpha\beta}\right]_{\dot{r}\to p} \frac{\partial g}{\partial p_\alpha}\frac{\partial k}{\partial p_\beta}, \tag{3.22}$$

for any two functions $g, k : T^*M \to \mathbb{R}$, and where $\{g,k\}_{can}$ was defined in footnote 5.

3.6.2 Nonholonomic Systems on Lie Groups

Consider now another subclass of nonholonomic systems with symmetry corresponding to the setting where the configuration space is a Lie group G, and the system is characterized by a left-invariant Lagrangian $l = \frac{1}{2}\langle I\xi, \xi\rangle$, where $\xi = g^{-1}\dot{g} \in \mathfrak{g}$, and $I : \mathfrak{g} \mapsto \mathfrak{g}^*$ is the inertia tensor. Suppose also that the system is subject to the left-invariant constraint

$$\langle a, \xi\rangle = a_I \xi^I = 0, \quad I = 1, \ldots, n, \tag{3.23}$$

where $a \in \mathfrak{g}^*$, and $n = dim(G)$. For simplicity, we shall restrict ourselves here to constraints for which the a_I are constant.

The equations of motion in this case are the *Euler-Poincaré-Suslov* equations [6, 55], and are given by:

$$\dot{\mu} = \mathrm{ad}^*_\xi \mu - \frac{\langle \mu, \mathrm{ad}_\xi I^{-1} a\rangle}{\langle a, I^{-1}a\rangle} a, \tag{3.24}$$

where $\mu = I\xi \in \mathfrak{g}^*$ is the body momentum.

We now discuss a more general form of (3.24) for the case of multiple constraints that emerges as a special case of the Hamilton-Poincaré-d'Alembert equations discussed above from the work of Section 5.8 of [6].

Suppose now that there are $k < n$ constraints of the form (3.23) such that k of the ξ's are dependent, so that $\xi^\alpha = -b_i^\alpha \xi^i$, where[6] $\alpha = 1, \ldots, k$ and $i = 1, \ldots, n-k$ and $\xi^I = (\xi^i, \xi^\alpha)$. Define the constrained Lagrangian $l_c(\xi^i) = l(\xi^i, \xi^\alpha = -b_i^\alpha \xi^i)$ as well as the constrained subspace $\mathfrak{g}^c$ of $\mathfrak{g}$ by $\mathfrak{g}^c = \{\xi \in \mathfrak{g} | \xi^\alpha = -b_i^\alpha \xi^i\}$. The basis vectors of this subspace are $u_i := e_i - b_i^\alpha e_\alpha$, and writing an element $\Omega \in \mathfrak{g}^c$ as $\Omega^i u_i$, we have the transformations $\xi^i = \Omega^i$, $\xi^\alpha = -b_j^\alpha \xi^j = -b_j^n \Omega^j$. Now, assuming enough regularity we can define the constrained reduced Hamiltonian $h_c(\Omega, \tilde{p}) = \tilde{p}_i \Omega^i - l_c$, where $\tilde{p}_i = \partial l_c / \partial \Omega^i$. The constrained Euler-Poincaré-Suslov equations are then given by [6]:

$$\dot{\tilde{p}}_i = -\mu_J C_{KL}^J e_i^K e_j^L \frac{\partial h_c}{\partial \tilde{p}_j}, \tag{3.25}$$

where the e_j^I are introduced through $\xi^I = e_j^I \Omega^j$ and in our particular case, as shown above,

$$e_i^K = (e_i^k, e_i^\alpha) = (\delta_i^k, -b_i^\alpha), \tag{3.26}$$

and $\tilde{p}_i = g_D^A p_A e_i^D =: \mu_D e_i^D$, with $p_A = \partial L / \partial \dot{g}^A = (\partial l / \partial \xi^B)(g^{-1})_A^B$. Here the g_J^I denote the lifted action of the group, C_{KL}^J are the structure constants of $\mathfrak{g}$ and δ_i^k is the Kronecker delta. Equation (3.25) represents the extension of the Lie-Poisson equations (2.10) to the nonholonomic context.

[6] When discussing Euler-Poincaré-Suslov systems we will use the index conventions that all uppercase indices $I, J, K, \ldots$ will range from 1 to n, all Greek indices from 1 to k and all lowercase indices $i, j, k, \ldots$ from 1 to $n-k$

We can also write the equations of motion (3.25) as:

$$\dot{\tilde{p}}_i = \{\tilde{p}_i, h_c\}_{AP}, \tag{3.27}$$

with $\{\cdot,\cdot\}_{AP} = \{\cdot,\cdot\}_-|_{\mathfrak{g}^c}$, where $\{\cdot,\cdot\}_-$ is the (minus) Lie-Poisson bracket on $\mathfrak{g}^*$. Although (3.25) gives the explicit form, we note here that the almost-Poisson bracket is equivalently given by:

$$\{g, k\}_{AP} = \{\tilde{p}_i, \tilde{p}_j\}\frac{\partial g}{\partial \tilde{p}_i}\frac{\partial k}{\partial \tilde{p}_j}, \tag{3.28}$$

for any two functions $g, k : (\mathfrak{g}^c)^* \to \mathbb{R}$, and where the bracket on the right hand side of (3.28) is computed by using the canonical bracket on T^*G and then restricting to $\mathfrak{g}^c$ (see [6], section 5.8 for more details). We will make use of this general form in Chapter VI when we consider the Hamiltonization of (3.25).

3.7 Invariant Measures of Nonholonomic Systems

Unlike the Hamiltonian systems of Chapter II, nonholonomic systems do not automatically preserve measure. In this section we summarize the conditions under which they do and extract special cases relevant to work presented in subsequent chapters.

3.7.1 Invariant Measures for Chaplygin Systems

Consider now a nonholonomic (nonabelian) Chaplygin system $(L, G, \mathcal{D})$ from Section 3.6.1 and assume that the constrained reduced Lagrangian l_c has an invertible kinetic energy matrix. Then we can express the right hand side of (3.18) in terms of the $\Lambda_{\alpha\beta}$ of (3.20) as

$$\Lambda_{\alpha\beta} = -K^{\epsilon}_{\alpha\beta} p_\epsilon := -M_{a\gamma} G^{\gamma\epsilon} \mathcal{B}^a_{\alpha\beta} p_\epsilon, \tag{3.29}$$

where $G^{\gamma\epsilon}$ is the inverse of the kinetic energy metric of l_c and $M_{a\gamma}$ was defined in Section 3.5. Although Λ in (3.20) was defined on TM, we will continue to use Λ to denote the Legendre transformed form in (3.29), an admitted abuse of notation.

The conditions for the existence of an invariant measure $N(r)$ for the system (3.18) are well studied [26, 57, 85], and we shall briefly review them here. If we denote by X_{nh} the vector field which solves (3.18)-(3.19) and further assume that the system has an invariant measure $N(r)\ dr^\alpha \wedge dp_\alpha = N(r)\omega^\alpha$, where ω^α denotes the standard measure on $\mathcal{M}$, then by definition $\mathcal{L}_{X_{nh}}(N\omega^k) = 0$, where $\mathcal{L}$ denotes the Lie-derivative. From this, we have:

$$0 = \mathrm{div}_{N\omega^k}(X_{nh}) = \mathrm{div}_{\omega^k}(X_{nh}) + \frac{1}{N} X_{nh}(N) = \frac{\partial \dot{r}^\alpha}{\partial r^\alpha} + \frac{\partial \dot{p}_\alpha}{\partial p_\alpha} + \frac{\dot{r}^\alpha}{N}\frac{\partial N}{\partial r^\alpha}. \tag{3.30}$$

Then using (3.18) this becomes:

$$\left(\frac{1}{N}\frac{\partial N}{\partial r^\beta} + \frac{\partial \Lambda_{\alpha\beta}}{\partial p_\alpha} \right) \dot{r}^\beta = 0, \tag{3.31}$$

and since the Λ's depend linearly on the momenta from above, we see that the quantities in parentheses in (3.31) depend only on the coordinates. Thus, the only way (3.31) vanishes is if the parenthetical terms vanish identically:

$$\frac{1}{N}\frac{\partial N}{\partial r^\beta} + \frac{\partial \Lambda_{\alpha\beta}}{\partial p_\alpha} = 0, \tag{3.32}$$

$$\implies \frac{1}{N}\frac{\partial N}{\partial r^\beta} - K^{\alpha}_{\alpha\beta} = 0. \tag{3.33}$$

Equation (3.32) locally describes explicitly how the measure density, metric, and curvature arising from the nonholonomic constraints interact. Moreover, since it is a system of first-order partial differential equations it can be easily solved by any of the popular mathematical software packages to yield the invariant measure density (if it exists).

3.7.2 Invariant Measures of Euler-Poincaré-Suslov Systems

In the unconstrained case, one can show [6] that the Euler-Poincaré equations have an integral invariant iff the group G is unimodular[7]. In the constrained case, for Euler-Poincaré-Suslov systems the existence of an invariant measure is given by a general result of Jovanović [50]:

Theorem III.1. *([50]) The Euler-Poincaré-Suslov equations (3.25) have an integral invariant with positive C^1 density iff*

$$K\, ad^*_{I^{-1}a}a + T = \gamma a, \tag{3.34}$$

for some $\gamma \in \mathbb{R}$, where $K = 1/\langle a, I^{-1}a\rangle$ and $T \in \mathfrak{g}^*$ is defined by $\langle T, \xi\rangle = Trace(ad_\xi)$.

The measure preservation conditions for more general nonholonomic systems are discussed in [16, 85].

The discussion of invariant measures in this section will allows us to understand how to generalize Chaplygin's theorem to higher dimensions. Recall from the Introduction that Chaplygin's reducing multiplier theorem rested on the assumption that the underlying system have an invariant measure. However, as we shall see in Chapter VII with the example of the Chaplygin sleigh, the lack of an invariant

[7] G being unimodular is equivalent to the requirement that the structure constants satisfy $C^c_{ac} = 0$.

measure in that case will not be an obstacle to its Hamiltonization using the results we develop in Chapter VI.

CHAPTER IV

Hamiltonization through the Inverse Problem of the Calculus of Variations

Recall from Section 2.5 that the inverse problem of the calculus of variations can be used to determine if a set of second-order differential equations are the Euler-Lagrange equations of some Lagrangian L. Now, although we already know that nonholonomic systems are not variational, because ultimately the mechanics of nonholonomic systems reduce to a set of mixed first- and second-order differential equations, it seems logical to begin the study of their Hamiltonization by considering the issue within the framework of the inverse problem. To that end, we will present various methods that associate to the nonholonomic equations of motion a family of systems of second-order ordinary differential equations. These methods will then allow us to apply the inverse problem of the calculus of variations [31, 81] to these associated systems instead. We shall then show that if an unconstrained (or free) regular Lagrangian exists for one of the associated systems, we will always be able to find an associated Hamiltonian H by means of the Legendre transformation. The canonical Hamiltonian equations resulting from H, when restricted to certain invariant submanifolds, will then reproduce the nonholonomic mechanics, hence accomplishing the Hamiltonization.

We will also show that our method only makes use of the equations of motion of the system and thus depends only on the Lagrangian and constraints of the nonholonomic system and not on the knowledge of the exact solutions of the system[1]. The application of the methods of the inverse problem will also provide us with families of regular Lagrangians which we expect to be useful in considering future applications of this work (see Chapter VIII).

4.1 Second-Order Dynamics Associated to a class of Nonholonomic Systems

Rather than abstractly describing the various ways of associating a second-order system to a given nonholonomic system, we will instead illustrate the method by means of one of the most interesting examples of a nonholonomic system.

The vertical rolling disk is a homogeneous disk rolling without slipping on a horizontal plane with configuration space $Q = \mathbb{R}^2 \times S^1 \times S^1$ and parameterized by the coordinates (x, y, θ, φ), where (x, y) is the position of the center of mass of the disk, θ is the angle that a point fixed on the disk makes with respect to the vertical and φ is measured from the positive x-axis (see Figure 7.1).

The system has the Lagrangian and constraints given by:

$$\begin{aligned} L &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2, \\ \dot{x} &= R\cos(\varphi)\dot{\theta}, \\ \dot{y} &= R\sin(\varphi)\dot{\theta}, \end{aligned} \tag{4.1}$$

where m is the mass of the disk, R is its radius, and I, J are the moments of inertia about the axis perpendicular to the plane of the disk and about the axis in the plane

[1] In contrast to some of the methods discussed in the Introduction.

of the disk, respectively. The constrained equations of motion are simply:

$$\ddot{\theta} = 0, \quad \ddot{\varphi} = 0, \quad \dot{x} = R\cos(\varphi)\dot{\theta}, \quad \dot{y} = R\sin(\varphi)\dot{\theta}. \tag{4.2}$$

The solutions of the first two equations are of course

$$\theta(t) = u_\theta t + \theta_0, \qquad \varphi(t) = u_\varphi t + \varphi_0,$$

and in the case where $u_\varphi \neq 0$, we get that the x- and y-solution is of the form

$$\begin{aligned} x(t) &= \left(\frac{u_\theta}{u_\phi}\right) R\sin(\varphi(t)) + x_0, \\ y(t) &= -\left(\frac{u_\theta}{u_\phi}\right) R\cos(\varphi(t)) + y_0, \end{aligned} \tag{4.3}$$

from which we can conclude that the disk follows a circular path. If $u_\varphi = 0$, we simply get the linear solutions

$$x(t) = R\cos(\varphi_0)u_\theta t + x_0, \quad y(t) = R\sin(\varphi_0)u_\theta t + y_0. \tag{4.4}$$

The situation in (4.4) corresponds to the case when φ remains constant, i.e. when the disk is rolling along a straight line. For much of what we will discuss in the next sections we will exclude these type of solutions from our framework, since our results will depend on $\dot{q}(0) \neq 0$.

Having introduced the vertical disk, let us take a closer look at the nonholonomic equations of motion (4.2). As a system of ordinary differential equations, these equations form a mixed set of coupled first- and second-order equations and as mentioned in the Introduction, it is well-known that these equations are never variational on their own [6, 26] in the sense that we can never find a regular Lagrangian whose (unconstrained) Euler-Lagrange equations are equivalent to the nonholonomic equations of motion. There are, however, systems of second-order equations (only) whose solution set contains the solutions of the nonholonomic equations (3.6). We shall call these second-order systems *associated second-order systems*, and in the next section will determine whether or not we can find a regular Lagrangian for at least one of those associated second-order systems. If so, we can then use the Legendre transformation to get a full Hamiltonian system on the associated phase space. The Legendre transformation will also map the constraint distribution onto a constraint submanifold in phase space and the nonholonomic solutions, considered as particular solutions of the Hamiltonian system, will then all lie on that submanifold.

Although there are infinitely many ways to arrive at an associated second-order system for a given nonholonomic system, we shall focus on three particular methods of constructing them in this chapter and illustrate these three choices below using the vertical rolling disk as an example.

4.1.1 Associated Second-Order Systems by Example

Firstly, consider taking the time derivative of the constraint equations so that a solution of the nonholonomic system (4.2) also satisfies the following complete set of second-order differential equations in all variables (θ, φ, x, y):

$$\ddot{\theta} = 0, \quad \ddot{\varphi} = 0, \quad \ddot{x} = -R\sin(\varphi)\dot{\theta}\dot{\varphi}, \quad \ddot{y} = R\cos(\varphi)\dot{\theta}\dot{\varphi}. \tag{4.5}$$

We shall call this system the *associated second-order system of type I*. Excluding for a moment the case where $u_\varphi = 0$, the solutions of equations (4.5) can be written as

$$\begin{aligned}
\theta(t) &= u_\theta t + \theta_0 \\
\varphi(t) &= u_\varphi t + \varphi_0 \\
x(t) &= \left(\frac{u_\theta}{u_\phi}\right) R\sin(\varphi(t)) + u_x t + x_0, \\
y(t) &= -\left(\frac{u_\theta}{u_\phi}\right) R\cos(\varphi(t)) + u_y t + y_0.
\end{aligned}$$

By restricting the above solution set to those that also satisfy the constraints $\dot{x} = \cos(\varphi)\dot{\theta}$ and $\dot{y} = \sin(\varphi)\dot{\theta}$ (i.e. to those solutions above with $u_x = u_y = 0$), we get back the solutions (4.3) of the non-holonomic equations (4.2). Similar reasoning holds for the solutions (4.4).

Now, taking note of the special structure of equations (4.5), we may alternately use the constraints (4.2) to eliminate the $\dot{\theta}$ dependency. This yields another plausible choice for an associated system:

$$\ddot{\theta} = 0, \quad \ddot{\varphi} = 0, \quad \ddot{x} = -\frac{\sin(\varphi)}{\cos(\varphi)}\dot{x}\dot{\varphi}, \quad \ddot{y} = \frac{\cos(\varphi)}{\sin(\varphi)}\dot{y}\dot{\varphi}. \tag{4.6}$$

We shall refer to this choice later as the *associated second-order system of type II.*

Lastly, we note that since the relation $\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y} = 0$ is satisfied on the constraint manifold we can easily add a multiple of this relation to some of the equations above. One way of doing so leads to the system

$$\begin{aligned}
\ddot{\varphi} &= -\frac{mR}{J}(\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y})\dot{\theta}, \\
\ddot{\theta} &= \frac{mR}{I + mR^2}(\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y})\dot{\varphi}, \\
\ddot{x} &= -R\sin(\varphi)\dot{\theta}\dot{\varphi} + \frac{mR^2}{I + mR^2}\cos(\varphi)(\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y})\dot{\varphi}, \\
\ddot{y} &= R\cos(\varphi)\dot{\theta}\dot{\varphi} + \frac{mR^2}{I + mR^2}\sin(\varphi)(\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y})\dot{\varphi}.
\end{aligned} \tag{4.7}$$

For later discussion we shall refer to this system as the *associated second-order system of type III.* We mention this particular second-order system here because, as we show in Chapter V (using techniques that are different than those we develop in this chapter), this complicated looking system is indeed variational! The Euler-Lagrange equations for the regular Lagrangian

$$L = -\frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2 + mR\dot{\theta}(\cos(\varphi)\dot{x} + \sin(\varphi)\dot{y}), \tag{4.8}$$

are indeed equivalent to equations (4.7) and when restricted to the constraint distribution reproduce the nonholonomic equations (4.2) exactly.

In Section 4.2 we will address the question of whether the second-order associated systems introduced above are equivalent to the Euler-Lagrange equations of some regular Lagrangian or not. However, let us now consider a more general setup.

4.1.2 Associated Second-Order Systems in General

We will, of course, not only be interested in the vertically rolling disk. It should be clear by now that there is no systematic way to categorize the second-order systems that are associated to a nonholonomic system. If no regular Lagrangian exists for one associated system, it may still exist for one of the infinitely many other associated systems. For many nonholonomic systems, the search for a Lagrangian using this method may therefore remain inconclusive. On the other hand, the solution of the inverse problem of any given associated second-order system is also too hard and too technical to tackle in full generality. Instead, we aim here to concisely formulate our results for a well-chosen class of nonholonomic systems which include the aforementioned example and for only a few choices of associated second-order systems.

To be more precise, let us assume from now on that the configuration space $Q = \mathbb{R}^n$ and that the base space of the fibre bundle is two dimensional, writing (r_1, r_2, s_α) for the coordinates. We will consider the class of nonholonomic systems where the Lagrangian is given by

$$L = \frac{1}{2}(I_1 \dot{r}_1^2 + I_2 \dot{r}_2^2 + \sum_\alpha I_\alpha \dot{s}_\alpha^2), \tag{4.9}$$

with all I_α positive constants and where the constraints take the following special form

$$\dot{s}_\alpha = -A_\alpha(r_1)\dot{r}_2. \tag{4.10}$$

Although this may seem to be a very thorough simplification, this interesting class of systems does include, for example, many of the classical examples of nonholonomic systems. We also remark that the above class of nonholonomic systems falls in the category of abelian Chaplygin systems from Section 3.6.1, which we will return to in Chapters V and VI.

In what follows, we will assume that none of the A_α are constant (in that case the constraints are, of course, holonomic). The nonholonomic equations of motion (3.6) are now

$$\ddot{r}_1 = 0, \quad \ddot{r}_2 = -N^2\Big(\sum_\beta I_\beta A_\beta A'_\beta\Big)\dot{r}_1\dot{r}_2, \quad \dot{s}_\alpha = -A_\alpha\dot{r}_2, \tag{4.11}$$

where N is shorthand for the function

$$N(r_1) = \frac{1}{\sqrt{I_2 + \sum_\alpha I_\alpha A_\alpha^2}}. \tag{4.12}$$

The function N has, in fact, a familiar interpretation in this case: it is the density of the invariant measure for the system (4.11). To see this, recall from Section 3.7 that we can obtain the invariant measure density by solving a set of partial differential equations. In the present case, these two equations read:

$$\frac{1}{N}\frac{\partial N}{\partial r_1} + \frac{\sum_\beta I_\beta A_\beta A'_\beta}{I_2 + \sum_\alpha I_\alpha A_\alpha^2} = 0, \qquad \frac{1}{N}\frac{\partial N}{\partial r_2} = 0, \tag{4.13}$$

whose solution is precisely the expression for N in (4.12) up to an irrelevant multiplicative constant. In the case of the vertically rolling disk, for example, it is a constant. Moreover, we shall see later in Proposition IV.2 that systems with a constant invariant measure always play a special role.

We are now in a position to generalize the associated second-order systems presented in Section 4.1.1 to the more general class of nonholonomic systems above. In the set-up above, the associated second-order system of type I for the more general systems (4.11) becomes the system

$$\ddot{r}_1 = 0, \quad \ddot{r}_2 = -N^2\big(\sum_\beta I_\beta A_\beta A'_\beta\big)\dot{r}_1\dot{r}_2, \quad \ddot{s}_\alpha = -(A'_\alpha \dot{r}_1\dot{r}_2 + A_\alpha \ddot{r}_2),$$

or equivalently, in normal form,

$$\begin{aligned} \ddot{r}_1 &= 0, \qquad\qquad \ddot{r}_2 = -N^2\big(\sum_\beta I_\beta A_\beta A'_\beta\big)\dot{r}_1\dot{r}_2, \\ \ddot{s}_\alpha &= -\Big(A'_\alpha - N^2 A_\alpha\big(\sum_\beta A_\beta A'_\beta\big)\Big)\dot{r}_1\dot{r}_2. \end{aligned} \tag{4.14}$$

For convenience, we will often simply write

$$\ddot{r}_1 = 0, \quad \ddot{r}_2 = \Gamma_2(r_1)\dot{r}_1\dot{r}_2, \quad \ddot{s}_\alpha = \Gamma_\alpha(r_1)\dot{r}_1\dot{r}_2,$$

for these types of second-order systems.

The associated second-order system of type II we encountered for the vertically rolling disk also translates to the more general setting. We get

$$\begin{aligned} \ddot{r}_1 &= 0, \qquad\qquad \ddot{r}_2 = -N^2\big(\sum_\beta I_\beta A_\beta A'_\beta\big)\dot{r}_1\dot{r}_2, \\ \ddot{s}_\alpha &= \Big(A'_\alpha - N^2 A_\alpha\big(\sum_\beta I_\beta A_\beta A'_\beta\big)\Big)\dot{r}_1\left(\frac{\dot{s}_\alpha}{A_\alpha}\right), \end{aligned} \tag{4.15}$$

where in the right-hand side of the last equation, there is no sum over α. A convenient byproduct of this way of associating a second-order system to (4.11) is that now all equations decouple except for the coupling with the r_1-equation. To highlight this, we will write this system as

$$\ddot{r}_1 = 0, \quad \ddot{q}_a = \Xi_a(r_1)\dot{q}_a\dot{r}_1$$

(no sum over a) where, from now on, $(q_a) = (r_2, s_\alpha)$ and $(q_i) = (r_1, q_a)$.

We postpone the discussion of the generalization of the associated second-order system of type III to our class of nonholonomic systems until Section 4.2.3.

4.2 Lagrangians for Associated Second-Order Systems

Using the Helmholtz conditions from Section 2.5, we are now in a position to investigate whether a Lagrangian exists for the two choices of associated systems, (4.14) and (4.15).

4.2.1 Lagrangians for Associated Second-Order Systems of Type I

The first second-order system of interest is of the form

$$\ddot{r}_1 = 0, \quad \ddot{r}_2 = \Gamma_2(r_1)\dot{r}_1\dot{r}_2, \quad \ddot{s}_\alpha = \Gamma_\alpha(r_1)\dot{r}_1\dot{r}_2. \tag{4.16}$$

The only non-zero components of (Φ^i_j) are

$$\begin{aligned} \Phi^2_1 &= (\frac{1}{2}\Gamma_2^2 - \Gamma_2')\dot{r}_1\dot{r}_2, \qquad \Phi^2_2 = -(\frac{1}{2}\Gamma_2^2 - \Gamma_2')\dot{r}_1^2, \\ \Phi^\alpha_1 &= (\frac{1}{2}\Gamma_\alpha\Gamma_2 - \Gamma_\alpha')\dot{r}_1\dot{r}_2, \qquad \Phi^\alpha_2 = -(\frac{1}{2}\Gamma_\alpha\Gamma_2 - \Gamma_\alpha')\dot{r}_1^2. \end{aligned}$$

For $\nabla\Phi$ and $\nabla\nabla\Phi$ we get

$$\begin{aligned} (\nabla\Phi)^2_1 &= (\Gamma_2\Gamma_2' - \Gamma_2'')\dot{r}_1^2\dot{r}_2, \qquad (\nabla\Phi)^2_2 = -(\Gamma_2\Gamma_2' - \Gamma_2'')\dot{r}_1^3, \\ (\nabla\Phi)^\alpha_1 &= (\Gamma_\alpha\Gamma_2 - \Gamma_\alpha'')\dot{r}_1^2\dot{r}_2, \qquad (\nabla\Phi)^\alpha_2 = -(\Gamma_\alpha\Gamma_2 - \Gamma_\alpha'')\dot{r}_1^3, \end{aligned}$$

and

$$\begin{aligned} (\nabla\nabla\Phi)^2_1 &= ((\Gamma_2')^2 + \Gamma_2\Gamma_2'' - \Gamma_2''')\dot{r}_1^3\dot{r}_2, \\ (\nabla\nabla\Phi)^2_2 &= -((\Gamma_2')^2 + \Gamma_2\Gamma_2'' - \Gamma_2''')\dot{r}_1^4 \\ (\nabla\nabla\Phi)^\alpha_1 &= (\Gamma_\alpha'\Gamma_2' + \frac{3}{2}\Gamma_\alpha\Gamma_2'' - \frac{1}{2}\Gamma_\alpha''\Gamma_2 - \Gamma_\alpha''')\dot{r}_1^3\dot{r}_2, \\ (\nabla\nabla\Phi)^\alpha_2 &= -(\Gamma_\alpha'\Gamma_2' + \frac{3}{2}\Gamma_\alpha\Gamma_2'' - \frac{1}{2}\Gamma_\alpha''\Gamma_2 - \Gamma_\alpha''')\dot{r}_1^4, \end{aligned}$$

and so on.

We can already draw some immediate consequences just by inspection, but to simplify the exposition let us illustrate the details using the four dimensional case. Then, the Φ-equations of the system (4.16) and their derivatives are all of the form

$$
\begin{aligned}
&g_{12}\Psi_2^2 + g_{13}\Psi_2^3 + g_{14}\Psi_2^4 = g_{22}\Psi_1^2 + g_{23}\Psi_1^3 + g_{24}\Psi_1^4,\\
&g_{23}\Psi_1^2 + g_{33}\Psi_1^3 + g_{34}\Psi_1^4 = 0,\\
&g_{23}\Psi_2^2 + g_{33}\Psi_2^3 + g_{34}\Psi_2^4 = 0,\\
&g_{24}\Psi_1^2 + g_{34}\Psi_1^3 + g_{44}\Psi_1^4 = 0,\\
&g_{24}\Psi_2^2 + g_{34}\Psi_2^3 + g_{44}\Psi_2^4 = 0,
\end{aligned}
\tag{4.17}
$$

where, within the same equation, Ψ stands for either Φ, $\nabla\Phi$, $\nabla\nabla\Phi$, $\nabla\nabla\nabla\Phi$, ... We will refer to the equations in the first line of (4.17) as *equations of the first type*, and to equations of the next four lines as *equations of the second type*. The first three equations of the first type, namely those for Φ, $\nabla\Phi$ and $\nabla\nabla\Phi$ are explicitly:

$$
\begin{aligned}
g_{12}\Phi_2^2 + g_{13}\Phi_2^3 + g_{14}\Phi_2^4 &= g_{22}\Phi_1^2 + g_{23}\Phi_1^3 + g_{24}\Phi_1^4, \qquad (4.18)\\
g_{12}(\nabla\Phi)_2^2 + g_{13}(\nabla\Phi)_2^3 + g_{14}(\nabla\Phi)_2^4 &= g_{22}(\nabla\Phi)_1^2 + g_{23}(\nabla\Phi)_1^3 + g_{24}(\nabla\Phi)_1^4,\\
g_{12}(\nabla\nabla\Phi)_2^2 + g_{13}(\nabla\nabla\Phi)_2^3 + g_{14}(\nabla\nabla\Phi)_2^4 &= g_{22}(\nabla\nabla\Phi)_1^2 + g_{23}(\nabla\nabla\Phi)_1^3 + g_{24}(\nabla\nabla\Phi)_1^4.
\end{aligned}
$$

For the systems at hand, the particular expression of Φ and its derivatives are such that

$$
\begin{aligned}
&\Phi_2^2(\nabla\Phi)_1^2 - \Phi_1^2(\nabla\Phi)_2^2 = 0,\\
&(\nabla\Phi)_2^2(\nabla\nabla\Phi)_1^2 - (\nabla\Phi)_1^2(\nabla\nabla\Phi)_2^2 = 0,
\end{aligned}
$$

and so on. By taking the appropriate linear combination of the first and the second, and of the second and the third equation in (4.18), we can therefore obtain two equations in which the unknowns g_{12} and g_{22} are eliminated. Moreover, under certain regularity conditions, these two equations can be solved for g_{13} and g_{14} in terms of g_{23} and g_{24} (we will deal with exceptions later on). So, if we can show that g_{23} and g_{24} both vanish, then so will also g_{13} and g_{14}. Then, in that case $g_{12}\Psi_2^2 = g_{22}\Psi_1^2$, but no further relation between g_{12} and g_{22} can be derived from this type of algebraic conditions.

The infinite set of equations given by those of the second type in (4.17) are all equations in the five unknowns g_{23}, g_{33}, g_{34}, g_{24} and g_{44}. Not all of these equations are linearly independent, however. In fact, given that the system (4.16) exhibits the property

$$\Psi_1^a\Psi_2^b - \Psi_1^b\Psi_2^a = 0,$$

(where Ψ is one of $\Phi, \nabla\Phi, \nabla\nabla\Phi, ...$), one can easily deduce that the last four lines of equations in (4.17) actually reduce to only two kinds of equations. If we assume that we can find among this infinite set five linearly independent equations, there will only be the zero solution

$$g_{23} = g_{33} = g_{34} = g_{24} = g_{44} = 0,$$

and from the above we then also know that $g_{13} = g_{14} = 0$.

To conclude, under the above mentioned assumptions, the matrix of multipliers

$$(g_{ij}) = \begin{pmatrix} g_{11} & g_{12} & 0 & 0 \\ g_{12} & g_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

is singular and we conclude that there is no regular Lagrangian for the system[2]. However, we note that the assumptions made above are not always satisfied and

[2]The above reasoning can, of course, be generalized to lower and higher dimensions.

need to be checked for every particular example. We will do so in Chapter VII where we discuss a wide variety of examples, but for now let us return to our discussion of the vertical disk.

The vertically rolling disk is in fact a special case, and so is any system (4.11) with the property that $\sum_\alpha I_\alpha A_\alpha^2$ is a constant (which is in fact equivalent with the geometric assumption that the density of the invariant measure N is constant[3]). In that case, we get $\Gamma_2 = 0$. Not only does Γ_2 vanish, but so do all Ψ_1^2 and Ψ_2^2 for $\Psi = \Phi, \nabla\Phi, \ldots$. We also have that $\Gamma_3 = -R\sin(\varphi)$ and $\Gamma_4 = R\cos(\varphi)$. Moreover, one can easily show that for the vertically rolling disk the three expressions (4.18), and any of the equations that follow in that set, are all linearly depending on the following two equations

$$\cos(\varphi)\dot{\varphi}g_{13} + \sin(\varphi)\dot{\varphi}g_{14} + \cos(\varphi)\dot{\theta}g_{23} + \sin(\varphi)\dot{\theta}g_{24} = 0,$$

$$\sin(\varphi)\dot{\varphi}g_{13} - \cos(\varphi)\dot{\varphi}g_{14} + \sin(\varphi)\dot{\theta}g_{23} - \cos(\varphi)\dot{\theta}g_{24} = 0.$$

Although these equations are already in a form where g_{12} and g_{22} do not show up, it is quite inconvenient that there is no way to relate these two unknowns to any of the other unknowns. However, as we did in the general case above, we can deduce from this an expression for g_{13} and g_{14} as a function of g_{23} and g_{24}. We get

$$g_{13} = -\frac{\dot{\theta}}{\dot{\varphi}}g_{23}, \qquad g_{14} = -\frac{\dot{\theta}}{\dot{\varphi}}g_{24}. \tag{4.19}$$

The infinite set of equations of the second type (i.e. the last four lines in (4.17)) are all linearly dependent to either one of the following four equations

$$\cos(\varphi)g_{33} + \sin(\varphi)g_{34} = 0, \quad \cos(\varphi)g_{34} + \sin(\varphi)g_{44} = 0$$

$$\sin(\varphi)g_{33} - \cos(\varphi)g_{34} = 0, \quad \sin(\varphi)g_{34} - \cos(\varphi)g_{44} = 0,$$

[3] This can be seen directly from (4.12).

from which $g_{33} = g_{34} = g_{44} = 0$ follows immediately. In comparison to the general case above, however, we can no longer conclude from the above that also g_{23} and g_{24} vanish, and therefore, we also cannot conclude from (4.19) that g_{13} and g_{14} vanish. This concludes, in fact, the information we can extract from the Φ-condition and the algebraic conditions that follow from taking its derivatives. Also, any attempt to create new algebraic conditions by means of the tensor R is fruitless since an easy calculation shows that when the above conclusions are taken into account all equations that can be derived from R are satisfied vacuously. However, since the determinant of the multiplier matrix

$$(g_{ij}) = \begin{pmatrix} g_{11} & g_{12} & \lambda g_{23} & \lambda g_{24} \\ g_{12} & g_{22} & g_{23} & g_{24} \\ \lambda g_{23} & g_{23} & 0 & 0 \\ \lambda g_{24} & g_{24} & 0 & 0 \end{pmatrix},$$

(with $\lambda = -\dot{\theta}/\dot{\varphi}$) clearly vanishes, this is a violation of one of the first Helmholtz conditions. Thus we can conclude that there does not exist a regular Lagrangian for the associated second-order system of type I of the vertically rolling disk.

4.2.2 Lagrangians for Associated Second-Order Systems of Type II

In this section we investigate the inverse problem for associated systems of type II,

$$\ddot{r}_1 = 0, \qquad \ddot{q}_a = \Xi_a(r_1)\dot{q}_a\dot{r}_1, \quad a = 2, \ldots, n, \tag{4.20}$$

where in the q_a-equations there is no sum over a and where $n = dim(Q)$. With respect to the formulation of the inverse problem in Section 2.5, we have $f_1 = 0$ and $f_a = \Xi_a \dot{q}_a \dot{r}_1$. Moreover, one can easily compute that the only non-vanishing components of Φ are now

$$\Phi_1^a = -\frac{1}{2}\dot r_1 \dot q_a (2\Xi_a' - \Xi_a^2), \quad \Phi_a^a = \frac{1}{2}\dot r_1^2 (2\Xi_a' - \Xi_a^2).$$

The Φ-conditions turn out to be quite simple: if $\Phi_a^a \neq 0$, then

$$\dot q_a g_{aa} = -\dot r_1 g_{1a}, \tag{4.21}$$

and if $\Phi_a^a \neq \Phi_b^b$ for $a \neq b$, then

$$g_{ab} = 0. \tag{4.22}$$

These restrictions on Φ lead to the assumptions that first $\Xi_a \neq 0$ and $\Xi_a \neq 2/(C - r_1)$, where C is any constant, second that $\Xi_a \neq \Xi_b$ and, formally, $\Xi_a - \Xi_b \neq E_b/(C - \int E_b dr_1)$, where $E_b(r_1) = \exp(\int 2\Xi_b dr_1)$.

Suppose for now that we are dealing with nonholonomic systems (4.15) where this is the case. Then one can easily show that all the other $\nabla \ldots \nabla\Phi$-conditions do not contribute any new information, as well as that the R-condition is satisfied vacuously. Thus we should therefore turn our attention to the ∇-condition, which is a partial differential equation. To simplify the subsequent analysis though, we note that although the multipliers g_{ij} can in general be functions of all variables $(r_1, q_a, \dot r_1, \dot q_a)$, in view of the symmetry of the system we shall assume them to be, without loss of generality, functions of $(r_1, \dot r_1, \dot q_a)$ only.

Now, by differentiating the algebraic conditions by r_1, $\dot r_1$ and $\dot q_a$, we get the additional conditions

$$\begin{aligned}
&\dot q_a g_{aa}' = -\dot r_1 g_{1a}' \\
&g_{aa} + \dot q_a g_{aaa} = -\dot r_1 g_{1aa}, \quad \dot q_a g_{1aa} = -g_{1a} - \dot r_1 g_{11a} \\
&g_{aab} = 0 = g_{1ab}, \quad \text{for } a \neq b.
\end{aligned}$$

Finally, the ∇-Helmholtz conditions are, with the above already incorporated,

$$g'_{11} + \sum_b \Xi_b (g_{11b}\dot q_b - g_{bb}\frac{\dot q_b^2}{\dot r_1^2}) = 0,$$
$$g'_{aa} + \Xi_a (g_{aaa}\dot q_a + g_{aa}) = 0.$$

In what follows we will implicitly assume everywhere that $\dot r_1 \neq 0$. As a consequence, the multipliers (g_{ij}) (and the Lagrangians we may derive from it) will only be defined for $\dot r_1 \neq 0$.

It is quite impossible to find the most general solution for (g_{ij}) though. We will show that there is an interesting class of solutions if we make the anszatz that $g_{bbb} = 0$ for all b. With that and with the above $g_{aab} = 0$ in mind, we conclude that all such g_{bb} will depend only on possibly r_1 and $\dot r_1$. Moreover, from the last ∇-conditions we can determine their dependency on the variable r_1. Since now

$$g'_{bb} + g_{bb}\Xi_b = 0,$$

it follows that $g_{bb}(r_1, \dot r_1) = F_b(\dot r_1)\exp(-\xi_b(r_1))$, where ξ_b is such that $\xi'_b = \Xi_b$ and where $F_b(\dot r_1)$ is still to be determined from the remaining conditions. From one of the above conditions we get $g_{1bb} = -g_{bb}/\dot r_1$ (since $g_{bbb} = 0$), so

$$\frac{dF_b}{d\dot r_1} = -\frac{F_b}{\dot r_1},$$

from which $F_b = C_b/\dot r_1$, with C_b a constant, and thus $g_{bb} = C_b \exp(-\xi_b)/\dot r_1$. Therefore, from the algebraic conditions, $g_{1b} = -(g_{bb}/\dot r_1)\dot q_b = -C_b\exp(-\xi_b)\dot q_b/\dot r_1^2$, and thus $g_{11b} = 2C_b\dot q_b \exp(-\xi_b)/\dot r_1^3$. With this, the first ∇-condition becomes

$$g'_{11} + \sum_b C_b \exp(-\xi_b)\xi'_b \frac{\dot q_b^2}{\dot r_1^3} = 0,$$

and thus

$$g_{11} = \sum_b C_b \exp(-\xi_b)\frac{\dot q_b^2}{\dot r_1^3} + C(\dot r_1, \dot q_b).$$

Given that $g_{11b} = 2C_b\dot{q}_b \exp(-\xi_b)/\dot{r}_1^3$, we can now determine the $\dot{q}_b$-dependence of C. We simply get

$$g_{11} = \sum_b C_b \exp(-\xi_b)\frac{\dot{q}_b^2}{\dot{r}_1^3} + F_1(\dot{r}_1).$$

Notice that g_{111} does not show up explicitly in the conditions or in the derived conditions. Therefore, there will always be some freedom in the g_{11}-part of the Hessian, represented here by the undetermined function $F_1(\dot{r}_1)$.

Up to a total time derivative, the most general Lagrangian whose Hessian $g_{ij} = \dfrac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}$ is the above multiplier, is then:

$$L = \rho(\dot{r}_1) + \frac{1}{2}\sum_b C_b \exp(-\xi_b)\frac{\dot{q}_b^2}{\dot{r}_1}, \tag{4.23}$$

where $d^2\rho/d\dot{r}_1^2 = F_1$. One can easily check that the Lagrangian is regular as long as $d^2\rho/d\dot{r}_1^2$ is not zero, and as long as none of the C_b are zero. We also note that the Lagrangian is only defined on the whole tangent space if $C_b = 0$ (and ρ is at least C^2 everywhere). We can therefore only conclude that there is a regular Lagrangian (with the ansatz $g_{bbb} = 0$) on that part of the tangent manifold where $\dot{r}_1 \neq 0$. As a consequence, the solution set of the Euler-Lagrange equations of the Lagrangian (4.23) will not include those solutions of the second-order system (4.16) where $\dot{r}_1 = 0$, which is why we will exclude them from our formalism. In the case of the vertically rolling disk, for example, these solutions are the special straight line solutions.

Recall that at the beginning of this section, we made the assumptions that $\Phi_a^a \neq 0$ and $\Phi_a^a \neq \Phi_b^b$. Suppose now that one of these assumptions is not valid, say $\Xi_2 = 0$ and therefore $\Phi_2^2 = 0$. Then, among the algebraic Helmholtz conditions there will no longer be a relation in (4.21) that relates g_{22} to g_{12}. In fact, since the g_{ij} now need to satisfy only a smaller number of algebraic conditions, the set of possible Lagrangians may be larger. We can, of course, still take the relation

$$\dot{q}_2 g_{22} = -\dot{r}_1 g_{12} \tag{4.24}$$

as an extra ansatz (rather than as a condition) and see whether there exist Lagrangians with that property. By following the same reasoning as before, we easily conclude that the function (4.23) is also a Lagrangian for systems with $\Phi_2^2 = 0$. In fact, it will be a Lagrangian if any of the assumptions is not valid.

Apart from (4.24), we are, of course, free to take any other ansatz on g_{12} and g_{22}. If we simply set

$$g_{12} = 0,$$

it can easily be verified that

$$L = \rho(\dot{r}_1) + \sigma(\dot{r}_2) + \frac{1}{2}\left(\sum_\alpha a_\alpha \exp(-\xi_\alpha)\frac{\dot{s}_\alpha^2}{\dot{r}_1}\right) \tag{4.25}$$

is also a Lagrangian for a system (4.20) with $\Xi_2 = 0$ (where, as usual, $(q_a) = (r_2, s_\alpha)$). Moreover, it is regular as long as both $d^2\rho/d\dot{r}_1^2$ and $d^2\sigma/d\dot{r}_2^2$ do not vanish. We can thus summarize the results of this section as the following Proposition.

Proposition IV.1. *The function*

$$L = \rho(\dot{r}_1) + \frac{1}{2N}\left(C_2\frac{\dot{r}_2^2}{\dot{r}_1} + \sum_\beta C_\beta \frac{\dot{s}_\beta^2}{A_\beta \dot{r}_1}\right), \tag{4.26}$$

with $d^2\rho/d\dot{r}_1^2 \neq 0$ and all $C_\alpha \neq 0$ is a regular Lagrangian for the associated systems of type II given by (4.15). Moreover, if the invariant measure density N is a constant, then

$$L = \rho(\dot{r}_1) + \sigma(\dot{r}_2) + \frac{1}{2N}\sum_\beta a_\beta \frac{\dot{s}_\beta^2}{A_\beta \dot{r}_1}, \tag{4.27}$$

where $d^2\rho/d\dot{r}_1^2 \neq 0$, $d^2\sigma/d\dot{r}_1^2 \neq 0$ and all $C_\alpha \neq 0$, is also a regular Lagrangian for the associated systems of type II given by (4.15).

Proof. For associated systems of type II, the second-order equations (4.20) are of the form (4.15). One easily verifies that in that case

$$\xi_2 = \ln N \quad \text{and} \quad \xi_\alpha = \ln(NA_\alpha) \tag{4.28}$$

are such that $\xi_a' = \Xi_a$. The Lagrangian (4.26) is then equal to the one given by (4.23). For a system with constant invariant measure N we get that $\Xi_2 = 0$ and therefore the function (4.25) is a valid Lagrangian. □

4.2.3 Lagrangians for Associated Second-Order Systems of Type III

In Section 4.1.1 we described the associated second-order system of type III (4.7) for the example of the vertically rolling disk. That complicated system actually results from a comparison of the variational nonholonomic and the Lagrange-d'Alembert nonholonomic equations of motion which will be the subject of Chapter V. Thus, we will defer the details of how the system was obtained until Chapter V, but note here that we will see there that, adapted to our current needs, the Lagrangian

$$\begin{aligned} L_V &:= L - \sum_\alpha \frac{\partial L}{\partial \dot{s}_\alpha}(\dot{s}_\alpha + A_\alpha \dot{r}_2) \\ &= \frac{1}{2}(I_1\dot{r}_1^2 + I_1\dot{r}_1^2 - \sum_\alpha I_\alpha \dot{s}_\alpha^2) - \sum_\alpha A_\alpha I_\alpha \dot{s}_\alpha \dot{r}_2, \end{aligned} \tag{4.29}$$

produces the Euler-Lagrange equations in normal form given by

$$\begin{aligned} \ddot{r}_1 &= -\big(\sum_\beta I_\beta A_\beta' \dot{s}_\beta\big)\dot{r}_2, \\ \ddot{r}_2 &= -N^2\big(\sum_\beta I_\beta A_\beta A_\beta'\big)\dot{r}_1\dot{r}_2 + \big(\sum_\beta I_\beta A_\beta' \dot{s}_\beta\big)\dot{r}_1, \\ \ddot{s}_\alpha &= -\Big(A_\alpha' - N^2 A_\alpha\big(\sum_\beta A_\beta A_\beta'\big)\Big)\dot{r}_1\dot{r}_2 - A_\alpha\big(\sum_\beta I_\beta A_\beta' \dot{s}_\beta\big)\dot{r}_1. \end{aligned} \tag{4.30}$$

Although in general these systems are not associated to our class of nonholonomic systems, that is, the restriction of their solutions to the constraint manifold $\dot{s}_\alpha = -A_\alpha \dot{r}_2$ are not necessarily solutions of the nonholonomic equations (4.11), in case the invariant measure density N is a constant, we have that $\sum_\beta I_\beta A_\beta A'_\beta = 0$. As a consequence, all the terms in the equations (4.30) that contain $\sum_\beta I_\beta A'_\beta \dot{s}_\beta$ vanish when we restrict those equations to the constraint manifold and the equations in $\ddot{s}_\alpha$ integrate to the equations of constraint (4.10). The restriction of the equations (4.30) is therefore equivalent with the nonholonomic equations (4.11). This proves the following.

Proposition IV.2. *If N is constant, the equations (4.30) form an associated second-order system and, by construction, they are equivalent to the Euler-Lagrange equations of the variational nonholonomic Lagrangian L_V.*

4.3 Hamiltonian formulation and the Constraints in Phase Space

In the situations where we have found a regular Lagrangian, the Legendre transformation leads to an associated Hamiltonian system. Since the base solutions of the Euler-Lagrange equations of a regular Lagrangian are also base solutions of Hamilton's equations of the corresponding Hamiltonian, the Legendre transformation $\mathbb{F}L$ will map those solutions of the Euler-Lagrange equations that lie in the constraint distribution $\mathcal{D}$ to solutions of the Hamilton equations that belong to the constraint manifold $\mathcal{M} = \mathbb{F}L(\mathcal{D})$ in phase space. Recall however that the Lagrangians for the associated second-order systems of type II (and their Legendre transformation) were not defined on $\dot{r}_1 = 0$, and thus will also be the case with the corresponding Hamiltonians.

Now, for convenience let us choose $\rho(\dot{r}_1) = \frac{1}{2} I_1 \dot{r}_1^2$ and $\sigma(\dot{r}_2) = \frac{1}{2} I_2 \dot{r}_2^2$ in the Lagrangians of Proposition IV.1. Then we have the following.

Proposition IV.3. *Given the associated second-order system of type II (4.15), the regular Lagrangian (4.26) (away from $\dot{r}_1 = 0$) and constraints (4.10) on TQ are mapped by the Legendre transform to the Hamiltonian and constraints in T^*Q given by:*

$$H = \frac{1}{2I_1}\left(p_1 + \frac{1}{2}N\left(\frac{p_2^2}{C_2} + \sum_\beta A_\beta \frac{p_\beta^2}{C_\beta}\right)\right)^2, \qquad C_2 p_\alpha = -C_\alpha p_2. \tag{4.31}$$

In case N is constant, the second Lagrangian (4.27) and constraints (4.10) are transformed into

$$H = \frac{1}{2I_2}p_2^2 + \frac{1}{2I_1}\left(p_1 + \frac{1}{2}N\left(\sum_\beta \frac{A_\beta}{a_\beta}p_\beta^2\right)\right)^2, \quad I_2 N\dot{r}_1 p_\alpha + a_\alpha p_2 = 0, \tag{4.32}$$

where $\dot{r}_1(r_1, p_1, p_\alpha) = (p_1 + \frac{1}{2}N\sum_\alpha A_\alpha p_\alpha^2/a_\alpha)/I_1$.

Proof. The Legendre transformation for the Lagrangian (4.23) gives

$$p_1 = I_1\dot{r}_1 - \frac{1}{2}\sum_b C_b \exp(-\xi_b)\frac{\dot{q}_b^2}{\dot{r}_1^2}, \qquad p_b = C_b\exp(-\xi_b)\frac{\dot{q}_b}{\dot{r}_1}, \tag{4.33}$$

from which one can easily verify that the corresponding Hamiltonian is

$$H = \frac{1}{2I_1}\left(p_1 + \frac{1}{2}\sum_b \exp(\xi_b)\frac{p_b^2}{C_b}\right)^2. \tag{4.34}$$

In the case of the associated second-order systems of type II in the form (4.15), the ξ_a take the form (4.28) and we obtain the Hamiltonian in expression (4.31). From (4.33) we can then compute the constraint manifold $\mathcal{M}$ in phase space. Since now

$$p_2 = C_2\frac{\dot{r}_2}{N\dot{r}_1} \quad \text{and} \quad p_\alpha = C_\alpha\frac{\dot{q}_\alpha}{N\dot{r}_1},$$

the constraints (4.10) can be rewritten as

$$\dot{r}_1\left(\frac{p_\alpha}{C_\alpha} + \frac{p_2}{C_2}\right) = 0,$$

where $\dot r_1 = \frac{1}{I_1}(p_1 + \frac{1}{2}N(p_2\dot r_2^2/C_2 + \sum_\beta A_\beta p_\beta^2/C_\beta))$. Assuming as always that $\dot r_1 \neq 0$ yields the constraint manifold $C_2 p_\alpha = -C_\alpha p_2$ for all α in phase space.

An analogous calculation with the Lagrangian (4.27) gives the Hamiltonian and the constraints in (4.32), in the case where N is constant. □

As mentioned in the Introduction, the authors in [1] Hamiltonized two nonholonomic systems in their paper by using the explicit solutions to the nonholonomic equations of motion. A quick application of the above shows that we can now derive their Hamiltonians from Proposition IV.3. As perhaps the simplest example, note that with $(r_1, r_2, s_\alpha) = (x, y, z)$, by taking C_2 and C_3 both to be 1 and $A(r_1) = x$, we recover the Hamiltonian and the constraint that appears in [1] for the nonholonomic free particle.

For the rolling disk the first Hamiltonian (4.31) is

$$H = \frac{1}{2J}\left(p_\varphi + \frac{1}{2\sqrt{I+mR^2}}\left(\frac{p_\theta^2}{C_2} - \frac{\cos(\varphi)p_x^2}{C_3} - \frac{\sin(\varphi)p_y^2}{C_4}\right)\right)^2,$$

and $C_2 p_x = -C_3 p_\theta$ and $C_2 p_y = -C_4 p_\theta$ for the constraints. However, these are not the Hamiltonian and the constraints that appear in [1] though. Instead the Hamiltonian and the constraints in [1] are in fact those that are associated to the second Hamiltonian (4.32), with, for example, $a_3 = a_4 = -J/\sqrt{I+mR^2}$:

$$H = \frac{1}{2I}p_\theta^2 + \frac{1}{2}\left(p_\varphi + \frac{1}{2}p_x^2\cos(\varphi) + \frac{1}{2}p_y^2\sin(\varphi)\right)^2,$$

and the constraints are

$$\dot\varphi p_x = p_\theta, \qquad \dot\varphi p_y = p_\theta,$$

where $\dot\varphi = p_\varphi + \frac{1}{2}\cos(\varphi)p_x^2 + \frac{1}{2}\sin(\varphi)p_y^2$ or, equivalently,

$$p_x - p_y = 0, \qquad \dot\varphi p_x - p_\theta = 0,$$

as they appear in [1].

Proposition IV.3 yields the main results of our application of the methods of the inverse problem to our goal of Hamiltonization for the class of nonholonomic systems given by (4.11). As we shall see in Chapter VII, we can apply these results to many of the well-known nonholonomic systems and obtain a Hamiltonization which can prove useful for a number of applications (see Chapter VIII). Now, the $\dot{r}_1 \neq 0$ restriction of the Proposition allows a degree of freedom which manifests itself as the family of Lagrangians obtained in Proposition IV.1. Although this restriction may not be relevant depending on the application of the results, it stands in contrast to the results of Proposition IV.2 where Hamiltonization is accomplished without any restrictions on the velocity of the system. Having seen the computational difficulties in obtaining explicit results by applying the Helmholtz conditions to even our simplified class of nonholonomic systems (4.11), the investigation of the content of Proposition IV.2 in a more general context will occupy the body of the next chapter.

CHAPTER V

Variational Nonholonomic Systems and Hamiltonization

In Chapter IV we investigated the Hamiltonization of the class of nonholonomic systems (4.11) by introducing the idea of associated systems and applying the inverse problem to them. For the first two types of associated systems we provided rules for their construction but did not discuss how to arrive at a general associated system of type III. The reason is that these systems arise from a comparison of the variational nonholonomic equations (3.10) and the nonholonomic Lagrange-d'Alembert equations (3.6)-(3.7) of Chapter III. Unlike Chapter IV, where we associated second-order systems to nonholonomic equations and applied the techniques of the inverse problem to derive the Lagrangian (and the Hamiltonian), here we start from a specific Lagrangian (the variational nonholonomic Lagrangian L_V in (3.9) of Section 3.4) and investigate the conditions under which its variational equations match the nonholonomic equations. Before we begin, we recall the index conventions of Section 2.4, where $a, b, c = 1, \ldots, k$, $\alpha, \beta, \gamma = 1, \ldots, m$, where $m = n - k$ and $n = dim(Q)$, with k the number of nonholonomic constraints. We also introduce the uppercase Roman indices I, J, K which range from 1 through n.

5.1 The Constrained Variational Nonholonomic Equations

Let us set $q = (r, s)$ as in Chapter III, considering the r and s equations separately and recalling the fundamentals of variational nonholonomic systems from Section 3.4. Writing E_I for the Euler-Lagrange operator in the I-th coordinate, we can rewrite (3.10) as

$$E_\alpha(L) = \dot{\mu}_a A^a_\alpha + \mu_a \left(\frac{d}{dt}(A^a_\alpha) - \frac{\partial A^a_\beta}{\partial r^\alpha} \dot{r}^\beta \right), \tag{5.1}$$

$$E_a(L) = \dot{\mu}_a - \mu_b \frac{\partial A^b_\alpha}{\partial s^a} \dot{r}^\alpha, \tag{5.2}$$

and substituting $\dot{\mu}_a$ from (5.2) into (5.1), we get

$$E_\alpha(L) = A^a_\alpha E_a(L) + \mu_a B^a_{\alpha\beta} \dot{r}^\beta. \tag{5.3}$$

One can then show [6] that by the definition of the constrained Lagrangian, (5.3) can be re-written in the more suggestive form:

$$E_\alpha(L_c) = \left(\mu_a - \frac{\partial L}{\partial \dot{s}^a} \right) B^a_{\alpha\beta} \dot{r}^\beta - A^a_\alpha \frac{\partial L_c}{\partial s^a}. \tag{5.4}$$

These along with the constraint equations (3.7) form the equations of motion for the constrained variational nonholonomic system, and a simple comparison of (5.4) and (3.6) reveals the extra term $\langle \mu, B(\dot{q}, \delta q) \rangle$, from which our analysis of the conditions under which the two equations of motion coincide will be based.

We should also point out that one can equivalently express the dynamics of variational nonholonomic systems directly in terms of the constrained Lagrangian L_c, without considering the Lagrangian equations on the full space $\tilde{Q}$ as we have done in (5.1)-(5.2) (see [27]). However, as we shall see in Section 5.2.2, equation (5.2) will lead to an *a priori* determination of the multipliers $\mu_a(t)$.

5.2 Conditionally Variational Nonholonomic Systems

In order to specify a unique solution to the variational nonholonomic problem, one must not only specify the initial values $(q_0, \dot{q}_0)$, but also the initial values of the multipliers μ_0. However, as we shall see, for certain initial values of the multipliers the trajectories obtained by solving (3.6) and (5.4) will coincide. In some cases, only some of the nonholonomic trajectories will coincide with variational trajectories, and in other cases all the nonholonomic trajectories will coincide with variational ones. To make these ideas more precise, we make the following definition.

Definition V.1. Consider the nonholonomic system (3.6)-(3.7) with initial conditions $(q_0, \dot{q}_0)$ and the associated variational nonholonomic system (5.4), (3.7) with initial conditions $(q_0, \dot{q}_0, \mu_0)$, where $(q_0, \dot{q}_0)$ satisfies the constraints (3.7). We make the following definitions:

(1) We shall say that the nonholonomic system is *conditionally variational* if for every initial condition $(q_0, \dot{q}_0)$ there exists an initial condition μ_0 such that the solution to (3.6)-(3.7) with initial condition $(q_0, \dot{q}_0)$ is the same as the solution to (5.4), (3.7) with initial condition $(q_0, \dot{q}_0, \mu_0)$.

(2) We will call a nonholonomic system *partially conditionally variational* if there exist some trajectories of the nonholonomic system which are also trajectories of the variational nonholonomic system, i.e. if there exist some initial data $(q_0, \dot{q}_0)$ for which there exist μ_0 such that the solution to (3.6)-(3.7) with initial condition $(q_0, \dot{q}_0)$ is the same as the solution to (5.4), (3.7) with initial condition $(q_0, \dot{q}_0, \mu_0)$.

Remark V.2. A nonholonomic system which is conditionally variational can thus be seen as a variational nonholonomic system with Lagrangian (3.9) and initial condition μ_0. We shall say more about the specifics of how to choose μ_0 in Proposition V.5 below.

5.2.1 The Equivalence Conditions

The actual conditions under which the nonholonomic system would be (in our nomenclature) conditionally variational were originally stated in [80]. There the author shows that the necessary and sufficient conditions for the equivalence between the two formalisms, in the notation adopted here, are:

$$\mu_a E_J(\phi^a)\delta q^J = 0. \tag{5.5}$$

These conditions are stated more in the language of analytical mechanics, and for our purposes we wish to have a more geometric and global view of them. The equivalent geometric condition we shall come to has already been hinted at near the end of Section 5.1, where we observed the difference in the two formalisms to depend on the multipliers and the curvature.

To that end, we can geometrize these conditions by first using the constraints (3.7) to relate (5.5) to the curvature (3.8) of the Ehresmann connection A:

$$E_J(\phi^a)\, dq^J \;\; = \;\; B^a_{\alpha\beta}\dot{r}^\beta dr^\alpha = d\omega^a(\dot{q}, \cdot). \tag{5.6}$$

Moreover, contracting with the vector $\delta q^J \partial_{q^J}$ and pre-multiplying by μ_a gives:

$$\mu_a E_J(\phi^a)\delta q^J = \mu_a d\omega^a(\dot{q}, \delta q). \tag{5.7}$$

Comparing the right hand side of the preceding with the vertical-vector-valued-two-form definition of curvature given in (3.8), we arrive at the geometric necessary and sufficient conditions for a nonholonomic system to be conditionally variational (see also [34]):

Proposition V.3. *The nonholonomic system (3.6)-(3.7) is conditionally variational with Lagrangian (3.9) if and only if*

$$\langle \mu, B(\dot{q}, \delta q) \rangle = 0. \tag{5.8}$$

Remark V.4. Since the curvature is vertical-vector-valued, Proposition V.3 intuitively says that a nonholonomic system is conditionally variational whenever the one-form $\mu = \mu_a ds^a$ is annihilated by $B(\dot{q}, \delta q)$. In the two degree of freedom case ($m = n - k = 2$) with k constraints, (5.8) reads $\mu_a B^a_{12} \dot{q}^2 = -\mu_a B^a_{12} \dot{q}^1 = 0$, and in this special case these two conditions are satisfied if and only if $\Lambda := \mu_a B^a_{12} = 0$. We shall call the conditions (5.8) the *equivalence conditions*, as originally named in [80].

Verifying (5.8) is often impractical though, since the conditions depend on the variational nonholonomic multipliers which are *a priori* unknown, thus requiring one to solve the variational dynamics explicitly. Below we will show how to remove this need in special cases, in addition to showing how to determine equivalence in some cases by simple inspection of the nonholonomic system.

5.2.2 Abelian Chaplygin Systems

Suppose now that we are considering an abelian Chaplygin nonholonomic system from Section 3.6.1. The curvature (3.8) then simplifies to just the difference of the first two terms, and the last term on the right hand side of (3.6) also vanishes. Thus, the equations of motion (3.6) reduce to the simpler form

$$E_\alpha(L_c) = F_{\gamma\alpha\beta}\dot{r}^\gamma \dot{r}^\beta, \tag{5.9}$$

where the $F_{\gamma\beta\alpha}$ are the components of $\Lambda_{\alpha\beta}$ from (3.20), i.e. $\Lambda_{\alpha\beta} = (g_{b\gamma} - g_{ba}A^a_\gamma)B^b_{\beta\alpha}\dot{r}^\gamma =: F_{\gamma\beta\alpha}\dot{r}^\gamma$. In this modified form, it becomes easier to state the first main result:

Proposition V.5. *Suppose the nonholonomic system (3.6)-(3.7) is (abelian) Chaplygin. Then we have the following:*

(1) The variational nonholonomic multipliers are given by:

$$\mu_a = \left(\frac{\partial L}{\partial \dot{s}^a}\right)_c + C_a, \qquad a = 1, \ldots, k, \tag{5.10}$$

where the subscript denotes that we have used the constraints to eliminate the $\dot{s}^b$, and where the C_a are integration constants.

(2) The system is conditionally variational if and only if the constrained nonholonomic equations (3.6) are Lagrangian.

(3) The property of being conditionally variational is unaffected by the addition of a potential function dependent on only the coordinates.

(4) The equivalence condition (5.8) reduces to the following conditions:

$$\begin{aligned} F_{\alpha\alpha\beta} &= 0 \quad \forall \alpha \neq \beta, && (5.11)\\ F_{\alpha\beta\gamma} + F_{\beta\alpha\gamma} &= 0 \quad \textit{for each}\ \gamma,\ \forall \alpha < \beta, \quad \alpha, \beta \neq \gamma. && (5.12) \end{aligned}$$

Consider now a general nonholonomic system (not necessarily Chaplygin). Then:

(5) Such a system with three generalized coordinates and one constraint is not conditionally variational.

Proof. For (1), consider the vakonomic momentum $p_a^V := \partial L_V/\partial \dot{s}^a = (\partial L/\partial \dot{s}^a) - \mu_a$, and note that $\partial L/\partial s^a = 0$ by the abelian Chaplygin assumption. Then by (5.2) we have $\dot{p}_a^V = 0$, from which the claim follows by integration and substitution of the constraints (see also Remark 1 below).

For (2), the reduced equations (5.9) are Lagrangian when the right hand side vanishes, which happens when $\Lambda_{\alpha\beta}\dot{r}^\beta = 0 \ \forall \alpha$. These are precisely the conditions (5.8) after taking into account (5.10). Conversely, suppose that the system is conditionally variational. Then by (5.8) the constrained equations (5.9) are Lagrangian.

For (3), simply note that if the added potential V is independent of $\dot{q}$, then the multipliers μ_a from Part (1) of the Proposition are unchanged, and so is condition (5.8). Thus, the system remains conditionally variational provided it was originally conditionally variational.

For (4), using (2) the system is conditionally variational when the right hand side of (5.9) vanishes. Since the $F_{\gamma\beta\alpha}$ are only functions of r^α, and $F_{\alpha\beta\beta} = 0 \ \forall \alpha, \beta$, these together imply that the coefficients of the products $\dot{r}^\gamma \dot{r}^\beta$ must vanish as stated.

For (5), by way of contradiction suppose it is conditionally variational. Then this means that (5.8) reads $\mu B_{12} = 0$ (taking into account Remark V.4). By assumption, the system is nonholonomic, meaning that it is not variational, so that μ is nonzero. Thus, the condition reduces to $B_{12} = 0$, which means that the system must actually be holonomic, in contradiction. □

Remarks:

1. Part (5) applies in the Chaplygin case as well. As such, it prevents some well-known nonholonomic systems from being conditionally variational simply by inspection: the knife edge, the Chaplygin sleigh, the Heisenberg free particle, and the

Euler-Poincaré-Suslov system on $SO(3)$ (for details see [6]), to name a few.

2. Part (1) of the Proposition allows one to determine the variational nonholonomic multipliers explicitly without having to solve the variational problem first. However as we noted in our definition of conditionally variational, these multipliers require initial conditions to be specified uniquely. To obtain the most general choice of the initial conditions on the μ_a that maintain the conditionally variational property we substitute (5.10) into (5.8). We see at once that one needs $\mu_a(0) = (\partial L/\partial \dot{s}^a)_c(0)$ $\forall a$ such that $B^a_{\alpha\beta} \neq 0$, and $\mu_a(0)$ may be chosen arbitrarily for each a such that $B^a_{\alpha\beta} = 0 \ \forall \alpha, \beta$.

3. The previous remark, along with conditions (5.8), now allows us to characterize conditionally variational systems in terms of the $\Lambda_{\alpha\beta}$ from (3.20). Namely, the two statements imply that to be conditionally variational requires $\Lambda_{\alpha\beta} = 0 \ \forall \alpha, \beta$.

4. Part (2) gives perhaps the simplest way to identify conditionally variational systems in the abelian Chaplygin case if we already know the system's constrained equations of motion.

5. Although verifying the conditions of Part (4) might be complex, the two degree of freedom case falls under Remark V.4, which is simpler to handle. Moreover, for a three degree of freedom system it leaves 9 conditions in (5.12) to be checked.

6. Although not part of Proposition V.5, it should be clear that given a conditionally variational nonholonomic system, if we add extra (constrained) coordinates s^b to its Lagrangian L and the extra constraints satisfy $B^b_{\alpha\beta} = 0 \ \forall \ \alpha, \beta$, then the new system will still be conditionally variational. This happens to be the case for the the two-wheeled carriage (see [38]).

5.2.3 Non-Abelian Chaplygin Systems

To derive the analogue of Proposition V.5 for nonabelian Chaplygin systems, we need the reduced constrained variational equations. As in the abelian Chaplygin case, we define the variational nonholonomic reduced Lagrangian $l_V = l - \mu_a(\xi - \mathcal{A}\dot{q})^a$. Then the Euler-Lagrange equations are [23, 68]:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial l}{\partial \xi} - \mu\right) &= \mathrm{ad}^*_\xi\left(\frac{\partial l}{\partial \xi} - \mu\right), \\ E_\alpha(l_c) &= \langle \mu - \frac{\partial l}{\partial \xi}, \mathcal{B}\rangle_\alpha, \end{aligned} \tag{5.13}$$

along with (3.19).

We can now compare these equations with (3.18) and write the analogue of Proposition V.5 in this context. However, due to the non-abelian character of the system, some aspects of Proposition V.5 no longer hold.

Corollary V.6. *Suppose that we have a non-abelian Chaplygin system given by (3.18)-(3.19) for which the right hand side of (3.18) vanishes. Then a solution to (3.18)-(3.19) with initial condition (q_0, ξ_0) is also a solution to the variational nonholonomic system (5.13) with*

$$\mu_b = \frac{\partial l}{\partial \xi} + C_b, \tag{5.14}$$

and initial condition (q_0, ξ_0, μ_0), where μ_0 and C_b are subject to Remark 2 of Section 5.2.2. Moreover, under this condition Part (3) of Proposition V.5 holds when the added potential is independent of ξ.

Proof. Suppose that the right hand side of (3.18) vanishes and consider a solution to the non-abelian system (3.18)-(3.19) with initial condition (q_0, ξ_0). Then the condition (5.8) with μ_b chosen as above, taking into account Remark 2 of Section 5.2.2, is satisfied. Thus, the solution to (3.18) is also a solution to (5.13) with μ as in (5.14) and μ_0 and C_b subject to Remark 2 of Section 5.2.2. Moreover, the last part of the Corollary follows by again observing that (5.8) is independent of the potential so long as the potential is independent of ξ. $\square$

Clearly nonabelian Chaplygin systems are more complicated than their abelian counterparts. For example, Part (1) of Proposition V.5 no longer applies, and here the variational multipliers are not given by the fiber derivative *a priori*. However, Corollary V.6 provides one with an alternative to solving the variational nonholonomic nonabelian problem to check for equivalence. The Corollary shows that if one can find a nonabelian Chaplygin system whose constrained reduced Euler-Lagrange equations are Lagrangian, then we can view its nonholonomic solutions with initial conditions (q_0, ξ_0) as variational solutions to (5.13) with initial conditions (q_0, ξ_0, μ_0), taking into account Remark 1 of Section 5.2.2 and (5.14).

5.2.4 Eliminating the Multipliers

We have defined and explored the idea of a nonholonomic system being conditionally variational in the preceding sections, but in the process have sacrificed the regularity of the new Lagrangian L_V, as we pointed out in Section 2.5. Thus, so far the variational system with Lagrangian L_V fails to describe the nonholonomic system via a regular Lagrangian, even though it describes it in terms of the dynamics both formalisms produce for certain initial data. However, we shall see below that this regularity may be regained in the abelian Chaplygin case by using a Lagrangian which is only a function of $(q, \dot{q})$ by eliminating the multipliers through Part (1) of Proposition V.5.

Suppose then that we again have an abelian Chaplygin nonholonomic system with regular mechanical Lagrangian and consider the following modification of the variational Lagrangian (3.9):

$$L_V = L - \frac{\partial L}{\partial \dot{s}^a}\phi^a. \tag{5.15}$$

Computing the Euler-Lagrange equations for the r and s variables gives:

$$0 = E_\alpha(L_V) = E_\alpha(L) + \frac{\partial L}{\partial \dot{s}^a}B^a_{\beta\alpha}\dot{r}^\beta - A^a_\alpha\left(\frac{\partial L}{\partial \dot{s}^a}\right)^{\cdot} + \left[\frac{\partial^2 L}{\partial r^\alpha \partial \dot{s}^a}\phi^a - (g_{\alpha a}\phi^a)^{\cdot}\right],$$

$$\implies E_\alpha(L_c) = (g_{\alpha a}\phi^a)^{\cdot} - \frac{\partial^2 L}{\partial r^\alpha \partial \dot{s}^a}\phi^a, \quad \text{and} \quad 0 = E_a(L_V) = \left(g_{ab}\phi^b\right)^{\cdot}, \tag{5.16}$$

$$\implies \left(g_{ab}\phi^b\right)(t) = g_{ab}(0)\phi^b(0), \tag{5.17}$$

We now see that if the constraints are satisfied initially (as they must be), and g_{ab} is invertible as a sub-matrix of g, then the constraints ϕ^a are satisfied for all subsequent times. If in addition we know that the the right hand side of the α equations in (5.16) vanishes, then the constrained Euler-Lagrange equations are Lagrangian and by Proposition V.5, Part (2) we would then know that the abelian nonholonomic Chaplygin system under consideration is actually conditionally variational with Lagrangian given by (5.15). Moreover, we can easily compute the Hessian of L_V to be the matrix

$$Hess(L_V) = g^V_{IJ} = \begin{pmatrix} g_{\alpha\beta} - g_{\alpha a}A^a_\beta - g_{\beta a}A^a_\alpha & -A^a_\alpha g_{ab} \\ -g_{ac}A^c_\alpha & g_{ab} \end{pmatrix}, \tag{5.18}$$

which is not automatically singular, as was the case for the $L(q, \dot{q}, \mu; t)$ Lagrangian (3.9). The preceding computations prove the following.

Proposition V.7. *Suppose that an abelian Chaplygin nonholonomic system with regular Lagrangian is known to be conditionally variational, and that in addition the sub-matrix g_{ab} is invertible. Then the nonholonomic mechanics can be derived from Hamilton's principle by using the Lagrangian (5.15) and with the initial data satisfying the constraints (3.7).*

Before discussing Proposition V.7, we should mention that its origins are rooted in momentum conservation. Assuming the Hessian (5.18) is invertible we can effect the Legendre transform and define $H : T^*Q \to \mathbb{R}$ as usual:

$$H_V(p_I, q^I) = p_I \dot{q}^I - L_V. \tag{5.19}$$

We can then examine the momenta conjugate to the s variables:

$$p_a := (\mathbb{F}L_V)_a = \frac{\partial L_V}{\partial \dot{s}^a} = g_{ab}\phi^b, \tag{5.20}$$

and from here it is clear that (5.17) is nothing but the a statement of conservation of momenta. However, Proposition V.7 shows us that additional conditions are required to translate this conservation of momentum into a statement about the preservation of the constraints throughout the motion. Moreover, it enables the description of conditionally variational nonholonomic systems without the use of the variational multipliers, hence possibly regaining regularity in the instances in which (5.18) is invertible. In fact, one such instance where (5.18) is nonsingular is the vertical rolling disk (see Chapter VII).

The Proposition also allows one to re-interpret conditionally variational systems as Hamiltonian systems restricted to certain subsets of phase space. This is because whenever the hypotheses of the Proposition are satisfied, (5.20) shows that enforcing the constraints is the same as setting p_a to zero. Thus, in these cases we can compute the Hamiltonian mechanics based on (5.19) and restrict to the submanifolds of T^*Q defined by $p_a = 0$ to recover the nonholonomic mechanics. In fact, this exactly turns out to be the case for the vertical disk (see Chapter VII). This interpretation also allows us to compare with the results of Proposition IV.3 of Section 4.4.

The final main idea to emphasize is that Proposition V.7 Hamiltonizes the entire nonholonomic system (constrained mechanics plus the kinematic constraints), and does so in some cases (the vertical disk being one) with a regular Lagrangian on TQ (i.e. without the use of Lagrange multipliers in an extended space).

5.3 Conditionally Variational Systems, Hamiltonization and Invariant Measures

As we saw in Section 3.4, variational nonholonomic dynamics can be derived from Hamilton's principle using the augmented Lagrangian (3.9). Hence, since the resulting system is Hamiltonian, it naturally preserves any non-zero constant multiple of the associated standard measure. However, intuitively nonholonomic systems which *do* possess invariant measures are in some sense closer to Hamiltonian systems and thus we expect them to be closer in structure to variational systems. Indeed we shall see below that in certain cases, having an invariant measure will render a nonholonomic system conditionally variational, and also Hamiltonize the nonholonomic system.

5.3.1 Conditionally Variational systems and Invariant Measures

Recall that Remark 3 from Section 5.2.2 tells us when an abelian Chaplygin system is conditionally variational, and it is clear from that Remark and (3.32) that nonholonomic systems with nonconstant measure densities cannot hope to be conditionally variational. Although this restricts the set of possible systems, the vertical rolling disk possesses a constant density invariant measure, and we shall make use of this in Chapter VII. However, as we shall see below having a constant invariant measure density is not sufficient for an abelian Chaplygin system to be conditionally variational.

Proposition V.8. *Suppose that the nonholonomic system (3.6)-(3.7) is an abelian Chaplygin system with an invariant measure with constant density N. Then the system satisfies:*

$$K^{\alpha}_{\alpha\beta} = 0. \tag{5.21}$$

In addition, if the system has two constraints and:

(a) Two degrees of freedom, it is also conditionally variational.

(b) Three degrees of freedom[1]*, if $R^3_2, R^3_1, R^2_1 \neq 0$ and the system satisfies the conditions (5.11), which in this case become:*

$$F_{112} = F_{113} = F_{221} = F_{223} = F_{331} = F_{332} = 0, \tag{5.22}$$

then it is also conditionally variational. For each of the aforementioned R^{α}_{γ} which are zero we must add to (5.22) the condition $F_{\gamma\alpha\beta} + F_{\alpha\gamma\beta} = 0$, for $\gamma \neq \alpha \neq \beta$. Also, for each R^{α}_{α} which is zero, we may omit the conditions $F_{\gamma\gamma\beta}$, $\gamma \neq \beta$ from (5.22).

Proof. Firstly, we have from (3.32) that having an invariant measure of constant density gives $K^{\alpha}_{\alpha\beta} = 0$. Now, for the two degree of freedom case of Part (a), (3.32) becomes (writing $\Lambda := \Lambda_{12}$):

[1]Recall the definition of the R^{α}_{β} from Section 3.5.

$$
\begin{aligned}
\frac{1}{N}\frac{\partial N}{\partial r^3} - \frac{\partial \Lambda}{\partial \tilde{p}_4} &= 0, \\
\frac{1}{N}\frac{\partial N}{\partial r^4} + \frac{\partial \Lambda}{\partial \tilde{p}_3} &= 0.
\end{aligned} \tag{5.23}
$$

However, from (3.29) we see that in general $\Lambda_{\alpha\beta} = \frac{\partial \Lambda_{\alpha\beta}}{\partial \tilde{p}_\gamma}\tilde{p}_\gamma$. Applying this here, we see at once that

$$
\begin{aligned}
\Lambda &= \frac{\partial \Lambda}{\partial \tilde{p}_3}\tilde{p}_3 + \frac{\partial \Lambda}{\partial \tilde{p}_4}\tilde{p}_4, \\
\Longrightarrow \Lambda &= \frac{1}{N}\left(\tilde{p}_4 \frac{\partial N}{\partial r^3} - \tilde{p}_3 \frac{\partial N}{\partial r^4}\right),
\end{aligned} \tag{5.24}
$$

where the last line follows by (5.23). By assumption of constant N (5.24) then shows that Λ vanishes, which by Remark 3 of Section 5.1.2 implies that the system is conditionally variational.

For Part (b), note that for the three degree of freedom case we can write out (5.12) by using the symmetry of $\tilde{g}$:

$$R_2^3 \left(F_{231} + F_{321}\right) = -\left[R_1^2 F_{121} + R_1^3 F_{131} + R_2^2 F_{221} + R_3^3 F_{331}\right], \tag{5.25}$$

$$R_1^3 \left(F_{132} + F_{312}\right) = -\left[R_1^1 F_{112} + R_2^1 F_{212} + R_2^3 F_{232} + R_3^3 F_{332}\right], \tag{5.26}$$

$$R_1^2 \left(F_{123} + F_{213}\right) = -\left[R_1^1 F_{113} + R_2^2 F_{223} + R_3^1 F_{313} + R_3^2 F_{323}\right]. \tag{5.27}$$

The parenthetical quantities are precisely (5.12) for the three degree of freedom case, and vanish precisely under the assumptions (5.22), by using $F_{\gamma\alpha\beta} = -F_{\gamma\beta\alpha}$. Moreover, it is clear that for each R_α^γ in (5.25)-(5.27) which is zero we must add the extra conditions as stated in Part (b). $\square$

Remark V.9. A similar result to Part (b) holds for the greater than three degree of freedom case, and also for the $k > 2$ case but isn't very computationally useful.

5.3.2 Hamiltonization and Conditionally Variational Systems

As discussed in the Introduction, another method which has been used with great success to Hamiltonize nonholonomic systems is Chaplygin's Reducing Multiplier Theorem. Having quantified the exact relationship between conditionally variational systems and their invariant measures in the previous section, we are now in a position to relate conditionally variational systems to the process of Chaplygin Hamiltonization.

Proposition V.10. *Suppose that the constrained abelian Chaplygin nonholonomic system (3.6)-(3.7) has two degrees of freedom. Then if the system is conditionally variational, it is also Chaplygin Hamiltonizable. Also, if the system is Chaplygin Hamiltonizable with constant measure density, then it is also conditionally variational.*

Proof. This follows directly from Chaplygin's reducibility theorem (see Theorem 1.2 in [35]) and Part (a) of Proposition V.8. □

Proposition V.10 gives the first relationship between two as yet unrelated avenues that both attempt to recover nonholonomic mechanics in Hamiltonian form. The main advantage is contained in its second statement, for there are numerous examples of abelian Chaplygin systems which are Hamiltonizable [18] and whose invariant measure densities are explicitly known. By considering particular values of the system parameters (i.e. moment of inertia, mass, etc.) we can then extract the constant measure density cases, and in the two degree of freedom case then express the nonholonomic system as a conditionally variational one thanks to Proposition V.10. Moreover, within the context of Chapter IV, we now see that Prop IV.2 of Section 4.2.3 shows that the system (4.11) is conditionally variational whenever

$\sum_\beta A_\beta A'_\beta = 0$, with L_V given by (4.29).

The restriction of constant invariant measure leaves little freedom to consider the Hamiltonization some of the other well-known nonholonomic systems through the methods in this chapter though. For this reason, we shall now turn our attention to extending Chaplygin's Theorem to arbitrary degrees of freedom (in addition to, in some cases, removing the restriction that the system possess an invariant measure) and thereby arrive at a more general formulation of conditionally variational systems.

CHAPTER VI

Hamiltonization through a Generalization of a Theorem of Chaplygin

6.1 Chaplygin Hamiltonization

To begin the generalization of Chaplygin's Theorem, we note that one can view Chaplygin's time reparametrization $d\tau = N(r)dt$ from the Introduction in a different way as follows: we have $\dot{r} = dr/dt = N(r)(dr/d\tau) =: N(r)\omega$, which defines the *quasivelocities* ω on $M = Q/G$. Thus, instead of considering which time reparametrization Hamiltonizes our system, we can rephrase the problem as one of finding a set of quasivelocities which enables one to rewrite the nonabelian Chaplygin equations (3.18) in terms of a Poisson bracket, or which enables one to write the Euler-Poincaré-Suslov equations (3.25) in terms of the (minus) Lie-Poisson bracket. To that end, we need to express the almost-Poisson brackets (3.22), (3.28) in terms of the quasivelocities ω, to which we now turn.

6.1.1 Chaplygin Hamiltonization of Chaplygin Systems

Proposition VI.1. *Consider a Chaplygin nonholonomic system $(L, G, \mathcal{D})$ with mechanical Lagrangian $L = T - V$, and define the quasivelocities ω through $\dot{r}^\alpha = f(r)\omega^\alpha$, where f is a never zero, smooth function, and let $\widetilde{j}$ be the map $\widetilde{j} : (r, \omega) \mapsto (r, \dot{r})$. Further, suppose that the constrained reduced Lagrangian l_c has an invertible kinetic* *energy metric. The constrained reduced nonholonomic equations (3.6) in the quasivelocities then become:*

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}_c}{\partial \omega}\right) - f\frac{\partial \mathcal{L}_c}{\partial r} = \left[\omega, \frac{\partial \mathcal{L}_c}{\partial \omega}\right]^* - \langle (\mathbb{F}\mathcal{L})_c, \mathcal{B}(\omega, f\partial_r)\rangle, \tag{6.1}$$

where $\mathcal{L}_c(r,\omega) = \widetilde{j}^* l_c$, $(\mathbb{F}\mathcal{L})_c = \widetilde{j}^*(\mathbb{F}l)_c$, *and where* $[\cdot,\cdot]^*$ *is the dual of the Jacobi-Lie bracket. Moreover, defining the conjugate momenta* $\mathcal{P}_\alpha = \partial\mathcal{L}_c/\partial\omega^\alpha$ *and the map* $j : (r,\mathcal{P}) \mapsto (r,p)$ *we can define the Hamiltonian* $\mathcal{H}_c := j^* h_c = \omega^\alpha \mathcal{P}_\alpha - \mathcal{L}_c|_{\omega\to\mathcal{P}}$ *and use it to obtain the quasi-Hamiltonian version of (6.1):*

$$\dot{r}^\alpha = f\{r^\alpha, \mathcal{H}_c\}'_{AP}, \quad \dot{\mathcal{P}}_\beta = f\{\mathcal{P}_\beta, \mathcal{H}_c\}'_{AP}, \tag{6.2}$$

where $\{\cdot,\cdot\}'_{AP}(r,\mathcal{P}) := (1/f)j^*\{\cdot,\cdot\}_{AP}(r,p)$ *(recall (3.22)) is the almost-Poisson bracket given by:*

$$\{G,K\}'_{AP}(r,\mathcal{P}) = \{G,K\}_{can}(r,\mathcal{P}) + \frac{1}{f}\left(fK^\epsilon_{\beta\alpha} - C^\epsilon_{\alpha\beta}\right)\mathcal{P}_\epsilon \frac{\partial G}{\partial \mathcal{P}_\alpha}\frac{\partial K}{\partial \mathcal{P}_\beta} \tag{6.3}$$

for any two functions $G, K : T^*\mathcal{Q} \to \mathbb{R}$, *where the* $C^\epsilon_{\alpha\beta}$ *are defined in (6.9) and the* $K^\epsilon_{\beta\alpha}$ *were defined in (3.29). Here* $T^*\mathcal{Q} = j^*T^*Q$ *and* $\{G,K\}_{can}(r,\mathcal{P})$ *is defined as*

$$\{G,K\}_{can}(r,\mathcal{P}) := \frac{\partial G}{\partial r^\alpha}\frac{\partial K}{\partial \mathcal{P}_\alpha} - \frac{\partial G}{\partial \mathcal{P}_\alpha}\frac{\partial K}{\partial r^\alpha}. \tag{6.4}$$

Proof. Let us begin with equations (3.18), expressed as:

$$\left(\frac{d}{dt}\frac{\partial l_c}{\partial \dot{r}^\alpha} - \frac{\partial l_c}{\partial r^\alpha} + \left(\frac{\partial l}{\partial \xi^a}\right)_c \mathcal{B}^a_{\alpha\beta}\dot{r}^\beta\right)\delta r^\alpha = 0. \tag{6.5}$$

Now, with $\delta r^\alpha = f\delta\omega^\alpha$, and the following transformations:

$$
\begin{aligned}
\frac{\partial \mathcal{L}_c}{\partial \omega^\alpha} &= f\frac{\partial l_c}{\partial \dot{r}^\alpha}, \\
\frac{\partial \mathcal{L}_c}{\partial r^\alpha} &= \frac{\partial l_c}{\partial r^\alpha} + \frac{\partial l_c}{\partial \dot{r}^\beta}\frac{\partial f}{\partial r^\alpha}\omega^\beta, \\
\dot{r}^\beta &= f\omega^\beta,
\end{aligned} \tag{6.6}
$$

the two parts of (6.5) transform as follows:

$$
\begin{aligned}
\left(\frac{d}{dt}\frac{\partial l_c}{\partial \dot{r}^\alpha} - \frac{\partial l_c}{\partial r^\alpha}\right)\delta r^\alpha &= \left(\frac{d}{dt}\frac{\partial \mathcal{L}_c}{\partial \omega^\alpha} - f\frac{\partial \mathcal{L}_c}{\partial r^\alpha} + C^\epsilon_{\alpha\beta}\frac{\partial \mathcal{L}_c}{\partial \omega^\epsilon}\omega^\beta\right)\delta\omega^\alpha, \\
\left(\left(\frac{\partial l}{\partial \xi^a}\right)_c \mathcal{B}^a_{\alpha\beta}\dot{r}^\beta\right)\delta r^\alpha &= \left(\left[\left(\frac{1}{f}\right) M_{a\gamma}G^{\gamma\epsilon}\frac{\partial \mathcal{L}_c}{\partial \omega^\epsilon}\right]\left(\mathcal{B}^a_{\alpha\beta}f\omega^\beta\right)f\right)\delta\omega^\alpha, \qquad (6.7)\\
&= fK^\epsilon_{\alpha\beta}\frac{\partial \mathcal{L}_c}{\partial \omega^\epsilon}\omega^\beta\delta\omega^\alpha, \qquad (6.8)
\end{aligned}
$$

where we have written l as $l(r,\dot{r},\xi) = (1/2)g_{IJ}(r)\dot{q}^I\dot{q}^J - V(r)$, with $\dot{q} = (\dot{r},\xi)$, and we remind the reader of the definition of $M_{a\gamma}$ from Section 3.5 and $K^\epsilon_{\alpha\beta}$ from (3.29). Here the constrained Lagrangian is $l_c = (1/2)G_{\alpha\beta}\dot{r}^\alpha\dot{r}^\beta - V(r)$, where $G_{\alpha\beta} = g_{\alpha\beta} - 2g_{a\alpha}\mathcal{A}^a_\beta + g_{ab}\mathcal{A}^a_\alpha\mathcal{A}^b_\beta$ and the $C^\epsilon_{\alpha\beta}$ are defined by:

$$
C^\epsilon_{\alpha\beta} := \left[\delta^\epsilon_\beta\frac{\partial f}{\partial r^\alpha} - \delta^\epsilon_\alpha\frac{\partial f}{\partial r^\beta}\right]. \tag{6.9}
$$

Finally, a simple renaming of indices along with the definition of the curvature 2-form yields:

$$
\frac{d}{dt}\left(\frac{\partial \mathcal{L}_c}{\partial \omega^\alpha}\right) - f\frac{\partial \mathcal{L}_c}{\partial r^\alpha} = \left[\omega, \frac{\partial \mathcal{L}_c}{\partial \omega}\right]^*_\alpha - \langle(\mathbb{F}\mathcal{L})_c, \mathcal{B}(\omega, f\partial_{r^\alpha})\rangle, \tag{6.10}
$$

which is the component version of (6.1).

Now, using the definition of $\mathcal{P}$ and $\mathcal{H}_c$ we can write (6.10) in the form:

$$\dot{\mathcal{P}}_\alpha = -f\frac{\partial \mathcal{H}_c}{\partial r^\alpha} + \left(fK^\epsilon_{\beta\alpha} - C^\epsilon_{\alpha\beta}\right)\mathcal{P}_\epsilon\frac{\partial \mathcal{H}_c}{\partial \mathcal{P}_\beta}, \tag{6.11}$$

which, along with $\dot{r}^\alpha = f\omega^\alpha = f(\partial\mathcal{H}_c/\partial\mathcal{P}_\alpha)$, allows us to construct the almost-Poisson bracket (6.3) and thus arrive at (6.2). □

Now, returning to (6.3), it is clear that the choice of multiplier f affects whether or not the system (6.2) is Chaplygin Hamiltonizable. More precisely, we have our first main result.

Theorem VI.2. *Suppose we have a Chaplygin nonholonomic system $(L, G, \mathcal{D})$ satisfying the assumptions of Proposition VI.1 and let $f(r) \in \mathcal{C}^1$ be a function which is nonzero everywhere on its domain. Then the necessary and sufficient conditions for Chaplygin Hamiltonization (using $d\tau = fdt$) on $M = Q/G$ are that f satisfy*

$$j^*\{g, k\}_{AP}(r, p) = f\{G, K\}_{can}(r, \mathcal{P}). \tag{6.12}$$

Moreover, the Hamiltonization of equations (3.21) are then:

$$\dot{r}^\alpha = f\{r^\alpha, \mathcal{H}_c\}_{can}, \quad \dot{\mathcal{P}}_\beta = f\{\mathcal{P}_\beta, \mathcal{H}_c\}_{can}. \tag{6.13}$$

Proof. By definition, a Chaplygin Hamiltonizable nonholonomic system can be written in Hamiltonian form with respect to the time reparameterization ($d\tau = fdt$). This is equivalent to the requirement that (6.3) be a Poisson bracket, which is equivalent to the requirement that it satisfy the Jacobi identity. A short calculation shows that (6.3) satisfies the Jacobi identity iff the second term in (6.3) vanishes $\forall\ \alpha, \beta, \epsilon$, which is equivalent to the requirement (6.12). Under this condition the Hamiltonized form of equations (3.21) then become (6.13). □

Remark VI.3. : In the context of Chaplygin's work, as well as to compare directly with [18], we note that the quasi-Hamiltonian form (6.13) is the "t-time" analogue of the Hamiltonian form stated for two degrees of freedom in the classical Chaplygin Reducibility Theorem in "τ-time," and the two are related through $\dot{r} = fr'$, $\dot{\mathcal{P}} = f\mathcal{P}'$, where $r' = dr/d\tau$ and $\mathcal{P}' = d\mathcal{P}/d\tau$.

The condition (6.12) is more (computationally) useful in its local form:

$$\frac{\partial f}{\partial r^\delta} G_{\alpha\nu} + \frac{\partial f}{\partial r^\nu} G_{\alpha\delta} - 2\frac{\partial f}{\partial r^\alpha} G_{\delta\nu} = f\left(K^\mu_{\alpha\delta} G_{\mu\nu} + K^\mu_{\alpha\nu} G_{\mu\delta}\right), \tag{6.14}$$

for all $\alpha, \nu, \delta = 1, \dots, m$, since this set of partial differential equations can then be solved (using any of the popular mathematics software programs) for a given nonholonomic system to determine the multiplier (if it exists).

Now, to extract Chaplygin's Reducibility Theorem as a special case[1] ($m = 2$), let us investigate the two degree of freedom case of Theorem VI.2.

Corollary VI.4. *The necessary and sufficient condition for a Chaplygin nonholonomic system $(L, G, \mathcal{D})$ for $m = 2$ to be Chaplygin Hamiltonizable is that*

$$\frac{\partial K^1_{12}}{\partial r^1} = -\frac{\partial K^2_{12}}{\partial r^2}. \tag{6.15}$$

The multiplier is then given by $f(r) = e^{\int K^1_{12} dr^2}$. Moreover, in case the system (3.18) is known to possesses an invariant measure with nonzero density $N(r) \in \mathcal{C}^1$, then the system is Chaplygin Hamiltonizable with $f = N$.

[1]Recall from Chapter III that $m := n - k$ is the number of degrees of freedom.

Proof. From Theorem VI.2, the only independent conditions in (6.12) in the two degree of freedom case are:

$$\left(\frac{\partial f}{\partial r^2} - fK^1_{12}\right) G_{11} - \left(\frac{\partial f}{\partial r^1} + fK^2_{12}\right) G_{12} = 0, \tag{6.16}$$

$$\left(\frac{\partial f}{\partial r^2} - fK^1_{12}\right) G_{21} - \left(\frac{\partial f}{\partial r^1} + fK^2_{12}\right) G_{22} = 0. \tag{6.17}$$

Since we have previously assumed that $G_{\alpha\beta}$ is invertible, (6.16)-(6.17) then implies that the satisfaction of these equations is equivalent to the vanishing of the parenthetical terms. The resulting set of equations is soluble only when (6.15) is satisfied, in which case f is given in explicit form as in the Corollary. Moreover, from (3.33) we see that the invariant measure density N of the reduced, constrained system (3.18) for $m = 2$ satisfies:

$$K^1_{12} = \frac{1}{N}\frac{\partial N}{\partial r^2}, \quad K^2_{12} = -\frac{1}{N}\frac{\partial N}{\partial r^1}. \tag{6.18}$$

One sees immediately that this satisfies (6.15), and hence $f = N$. □

The second part of Chaplygin's Reducibility Theorem[2] is that if such a multiplier f is found, the original nonholonomic system (3.18) has an invariant measure with density f^{m-1}. This too is the case in our situation, as shown below.

Proposition VI.5. *Suppose f satisfies the conditions of Theorem VI.2. Then the original system (3.6) has an invariant measure with density f^{m-1}.*

Proof. Suppose f satisfies (6.14). Multiplying by $\dot{r}^\delta$ and $\dot{r}^\nu$ and adding results in:

$$\frac{1}{f}\left(\frac{\partial f}{\partial r^\beta}p_\alpha - \frac{\partial f}{\partial r^\alpha}p_\beta\right)\dot{r}^\beta = \left(\frac{\partial l}{\partial \xi^a}\right)_c \mathcal{B}^a_{\alpha\beta}\dot{r}^\beta. \tag{6.19}$$

[2]See [35], or the discussion in the Introduction.

Comparing the $\dot{r}^\beta$ coefficients yields[3]

$$\Lambda_{\beta\alpha} = \frac{1}{f}\left(\frac{\partial f}{\partial r^\beta}p_\alpha - \frac{\partial f}{\partial r^\alpha}p_\beta\right), \tag{6.20}$$

where we remind the reader of the definition of Λ from (3.20), and have used $p_\alpha = \partial l_c/\partial \dot{r}^\alpha$. Thus, for a Chaplygin Hamiltonizable system, the right hand side of (3.18) can be written in terms on f as in (6.20). Now, suppose $X_{nh} = \dot{r}^\alpha \partial_{r^\alpha} + \dot{p}_\alpha \partial_{p^\alpha}$ is the nonholonomic vector field solution to the system (3.18)-(3.19). We will show that f^{m-1} is an invariant measure density by showing that the vector field $f^{m-1}X_{nh}$ has zero divergence. A straightforward calculation yields

$$\text{div}\left(f^{m-1}X_{nh}\right) = \frac{\partial(f^{m-1}\dot{r}^\alpha)}{\partial r^\alpha} + \frac{\partial(f^{m-1}\dot{p}_\alpha)}{\partial p_\alpha} = f^{m-2}\dot{r}^\alpha\left((m-1)\frac{\partial f}{\partial r^\alpha} + f\frac{\partial \Lambda_{\beta\alpha}}{\partial p^\beta}\right), \tag{6.21}$$

and a simple calculation of the last term in (6.21) using (6.20) then shows that the divergence does indeed vanish and completes the proof. □

On the one hand, we see from Corollary VI.4 that Chaplygin's original Theorem represented sufficiency conditions for Hamiltonization for $m = 2$, and its use was restricted to systems known to have invariant measures.

On the other hand, Theorem VI.2 yields the necessary conditions for the Hamiltonization of the Chaplygin system (3.18)-(3.19), which locally are first-order partial differential equations (6.14) in r. The theorem generalizes the classical Chaplygin reducibility theorem to systems with nonabelian symmetry groups in higher degrees of freedom. Proposition VI.5 then completes the generalization by providing us with the invariant measure density given a solution to Theorem VI.2. Thus, in light

[3] We note that this relationship was also presented as a sufficient condition for the existence of an invariant measure by [82] (see also [26]), but here is derived from the conditions of Theorem VI.2.

of this, whereas previously the hopes for Chaplygin Hamiltonization rested on the knowledge of a two degree of freedom system's invariant measure density, it now makes most sense to tackle the problem of Chaplygin Hamiltonization by solving (6.14), since then we get the invariant measure density for free if a multiplier exists. This path to Hamiltonization stands in contrast to the usual attempt to apply the ideas behind Chaplygin's theorem to higher dimensions (namely the guessing of a multiplier based on m and a known invariant measure and then a computation to verify that the equations in the reparameterized time are Hamiltonian), for (6.14) requires no guesswork on our part.

It may be the case, however, that (6.14) does not have a solution. This does not mean that the system is not Chaplygin Hamiltonizable though, since it may still possess more symmetries which reduce the degrees of freedom, and which allow one to seek such a solution on the second reduced phase space. We illustrate such a situation in the next section, making use of the classical Routhian [6] to explore the effect of additional simple symmetries in (3.18) on its Chaplygin Hamiltonizability.

6.1.2 Momentum Conservation and Chaplygin Hamiltonization

Suppose that (6.12) has no solutions, but that the nonholonomic system possesses momentum conservation laws that we have yet to account for. With the aid of these conservation laws, we can apply the reduction process to further reduce the degrees of freedom of the system and re-attempt a Hamiltonization on the reduced space.

In order to illustrate this in a simple manner, we restrict ourselves in this section to Chaplygin systems which we will call *nonholonomic cylic*. By this we mean that we have an abelian Lie group H acting on $M = Q/G$ by $M \ni r^\alpha = (w^{\alpha'}, v^i) \mapsto (w^{\alpha'}, v^i + h^i)$, $h \in H$, where $i = 1, \ldots l = dim(H)$ and $\alpha' = 1, \ldots, m - l$ and such that $\Lambda_{\alpha' i} = 0 \ \forall i, \alpha'$ (we shall hereafter denote the nonconserved conjugate variable indices w with a prime). Under these assumptions the v^i equations in (3.18) lead to the momentum conservation laws[4] and we can thus set $p_i = \mu_i =$ constant and perform a partial Legendre transform in the v^i variables to form the classical (constrained) Routhian [67, 68] $R_c(w, \dot{w})$ defined by

$$R_c(w, \dot{w}) := \left[l_c(w, \dot{w}, \dot{v}) - \mu_i \dot{v}^i \right]_{p_i = \mu_i}. \tag{6.22}$$

Since $R_c : T(M/H) \to \mathbb{R}$ we can now attempt to Hamiltonize on the reduced space. This gives the second main result.

Theorem VI.6. *Suppose that the Chaplygin nonholonomic system given by (3.18)-(3.19) is not Hamiltonizable by Theorem VI.2 but is nonholonomic cyclic. Further, suppose that the kinetic energy matrix of l_c and the sub-matrix $(\partial^2 l_c/\partial v^i \partial v^j)$ are invertible and define the maps $\overline{j} : (w, \omega) \mapsto (w, \dot{w})$ and $\overline{j}_p : (r, \mu_i) \mapsto (r, p_i)$. Then if there exists a multiplier $f(w)$, nonzero everywhere on its domain with $f(w) \in \mathcal{C}^1$, satisfying*

$$\left[\omega, \frac{\partial \mathcal{R}_c}{\partial \omega} \right]^* = \langle (\mathbb{F}\mathcal{R}_c)', \mathcal{B}(\omega, f\partial_w) \rangle, \tag{6.23}$$

where $\mathcal{R}_c(w, \omega) = \overline{j}^* R_c$ and[5] *$(\mathbb{F}\mathcal{R}_c)' = \overline{j}_p^*(\mathbb{F}R_c)'$, the reduced system is Chaplygin Hamiltonizable on $M' = M/H$ under the choice of quasivelocity $\dot{w} = f\omega$. Furthermore, its mechanics on $T(M/H)$ can be written in the quasi-Hamiltonian form:*

$$\dot{w}^{\alpha'} = f\{w^{\alpha'}, \mathcal{H}'\}_{AP}, \quad \dot{\mathcal{P}'}_{\beta'} = f\{\mathcal{P}'_{\beta'}, \mathcal{H}'\}_{AP}, \tag{6.24}$$

[4] Since the H-invariance implies that l_c does not depend explicitly on the v^i, these variables are *cyclic* and produce momentum conservation laws in unconstrained systems. However, due to the presence of the $\Lambda_{\alpha\beta}$, cyclic variables are not enough to produce the conservation laws, hence the introduction of nonholonomic cyclic.

[5] Where $(\mathbb{F}R_c)'$ is defined in the Proof.

where $\mathcal{P}'_{\alpha'} = \partial\mathcal{R}_c/\partial\omega^{\alpha'}$ *and* $\mathcal{H}' = \omega^{\alpha'}\mathcal{P}'_{\alpha'} - \mathcal{R}'_c|_{\omega\to\mathcal{P}'}$ *is the Hamiltonian and the almost-Poisson bracket is defined by:*

(6.25)
$$\{G,K\}_{AP}(w',\mathcal{P}') = \{G,K\}_{can}(w',\mathcal{P}') - \sum_{\alpha'<\beta'} fK^i_{\alpha'\beta'}\mu_i\left(\frac{\partial G}{\partial\mathcal{P}'_{\alpha'}}\frac{\partial K}{\partial\mathcal{P}'_{\beta'}} - \frac{\partial G}{\partial\mathcal{P}'_{\beta'}}\frac{\partial K}{\partial\mathcal{P}'_{\alpha'}}\right).$$

Moreover, the bracket satisfies the Jacobi identity for $dim(M') = 2$.

Proof. Locally, the equations of motion in the w variables are

$$\frac{d}{dt}\frac{\partial R_c}{\partial\dot{w}^{\alpha'}} - \frac{\partial R_c}{\partial w^{\alpha'}} = -\left(\frac{\partial l}{\partial\xi^a}\right)_c\mathcal{B}^a_{\alpha'\beta'}\dot{w}^{\beta'}. \tag{6.26}$$

However, the last term on the right hand side of (6.26) can be rewritten in terms of the Routhian:

$$\begin{aligned}\left(\frac{\partial l}{\partial\xi^a}\right)_c\mathcal{B}^a_{\alpha'\beta'} &= M_{a\alpha}G^{\alpha\beta}\frac{\partial l_c}{\partial\dot{r}^\beta}\mathcal{B}^a_{\alpha'\beta'} = \left[M_{a\alpha}G^{\alpha\epsilon'}\frac{\partial l_c}{\partial\dot{w}^{\epsilon'}} + M_{a\alpha}G^{\alpha i}\mu_i\right]\mathcal{B}^a_{\alpha'\beta'},\\ &= \left[\left(M_{a\alpha}G^{\alpha\epsilon'}\frac{\partial R_c}{\partial\dot{w}^{\epsilon'}}\right) + M_{a\alpha}G'^{\alpha i}\mu_i\right]\mathcal{B}^a_{\alpha'\beta'}, &(6.27)\\ &= K^{\epsilon'}_{\alpha'\beta'}\frac{\partial R_c}{\partial\dot{w}^{\epsilon'}} + K^i_{\alpha'\beta'}\mu_i, &(6.28)\end{aligned}$$

with the parenthetical term in (6.27) locally defining[6] $(\mathbb{F}R_c)'$ from (6.23) and where $G'^{\alpha i} = G^{\alpha i} + G^{\alpha\epsilon'}G^{ij}G_{j\epsilon'}$. Now, under the quasivelocity transformation $\dot{w} = f(w)\omega$ the reduced equations (6.26) become (taking into account (6.28)):

$$\frac{d}{dt}\frac{\partial\mathcal{R}_c}{\partial\omega^{\alpha'}} - f\frac{\partial\mathcal{R}_c}{\partial w^{\alpha'}} = \left(W^{\epsilon'}_{\alpha'\beta'}\frac{\partial\mathcal{R}_c}{\partial\omega^{\epsilon'}}\omega^{\beta'}\right) - f^2K^i_{\alpha'\beta'}\mu_i\omega^{\beta'}, \tag{6.29}$$

where $W^{\epsilon'}_{\alpha'\beta'} := fK^{\epsilon'}_{\beta'\alpha'} - C^{\epsilon'}_{\alpha'\beta'}$. Now, if f is chosen to satisfy (6.23), then the parenthetical term in (6.29) vanishes. By defining the Hamiltonian $\mathcal{H}'$ as in the

[6]We also remind the reader of the definition of the $K^\epsilon_{\alpha\beta}$ in (3.29).

statement of the Theorem, the equations of motion can then be written as in (6.24) with the almost-Poisson bracket (6.25). Lastly, a straightforward computation shows that the Jacobi identity is automatically satisfied for $\dim(M') = 2$, owing to the fact that non-canonical part in (6.25), $\{\mathcal{P}_{\alpha'}, \mathcal{P}_{\beta'}\}_{AP}$, is independent of the momenta. □

Theorem VI.6 will be used in Chapter VII when discussing the Chaplygin sphere, a classic example of how the failure of Hamiltonizability can be reversed in the presence of momentum conservation laws.

6.1.3 Chaplygin Hamiltonization of Euler-Poincaré-Suslov Systems

Keeping with the general idea of Section 6.1 and recalling the setup and index conventions of Section 3.6.2, we now consider the quasivelocity transformation $\xi^I = f(g)\omega^I$, where $g \in G$, of the system (3.27), assuming again that f is smooth and nowhere vanishing on its domain. Such a transformation is equivalent to the change of basis $e_I \mapsto v_I := f e_I$ of the Lie algebra $\mathfrak{g}$. We then have the analogue of Proposition VI.1 in this case.

Proposition VI.7. *Define the map $j_\omega : (\omega^I) \to (\xi^I)$, along with $\mathcal{P} = j_\omega^* \tilde{p}$ and $\mathcal{H}_c = j_\omega^* h_c$ and consider an Euler-Poincaré-Suslov system with Lagrangian $l(\xi) = \frac{1}{2} I_{IJ} \xi^I \xi^J$. Further, suppose that the constrained Lagrangian l_c has an invertible kinetic energy metric. Then the mechanics of the system (3.27) becomes:*

$$\dot{\mathcal{P}}_i = f\{\mathcal{P}_i, \mathcal{H}_c\}'_{AP}, \tag{6.30}$$

where $\{\cdot,\cdot\}'_{AP} := (1/f) j_\omega^* \{\cdot,\cdot\}_{AP}$ (recall (3.28)) is the almost-Poisson bracket given by:

$$\{G,K\}'_{AP} = -\frac{1}{f}\left(fK^m_{ji} - \overline{C}^m_{ij}\right)\mathcal{P}_m \frac{\partial G}{\partial \mathcal{P}_i}\frac{\partial K}{\partial \mathcal{P}_j}, \tag{6.31}$$

for any two functions $G, K : j^*_\omega(\mathfrak{g}^c)^* \to \mathbb{R}$, where K^m_{ji} and $\overline{C}^m_{ij}$ are defined, in analogy with (6.8) and (6.9), by (6.33) and (6.34) below, respectively.

Proof. We compute $j^*_\omega\{\tilde{p}_i, \tilde{p}_j\}$ as:

$$f\{\mathcal{P}_i, \mathcal{P}_j\}'_{AP} = -\left(fM^m_J C^J_{KL} e^K_i e^L_j - \frac{\partial f}{\partial g^J} g^J_K \left(e^K_j \delta^m_i - e^K_i \delta^m_j\right)\right)\mathcal{P}_m, \tag{6.32}$$

from which we define the K^m_{ji} and $\overline{C}^m_{ij}$ by:

$$K^m_{ji} := M^m_J C^J_{KL} e^K_i e^L_j, \tag{6.33}$$

$$\text{and } \overline{C}^m_{ij} := \frac{\partial f}{\partial g^J} g^J_K \left(e^K_j \delta^m_i - e^K_i \delta^m_j\right), \tag{6.34}$$

where M^m_J is defined from the constrained μ_I as follows. From $\xi^I = e^I_j \Omega^j$ we have $l_c(\Omega) = (1/2)I_{IJ} e^I_i e^J_j \Omega^i \Omega^j =: (1/2)G_{ij}\Omega^i\Omega^j$. Assuming G_{ij} is invertible, we can write $\Omega^j = G^{ji}\tilde{p}_i$ and so

$$(\mu_I)_c = \left(\frac{\partial l}{\partial \xi^I}\right)_c = (I_{IJ}\xi^J)_c = I_{IJ} e^J_j \Omega^j = I_{IJ} e^J_j G^{jm}\tilde{p}_m =: M^m_I \tilde{p}_m. \tag{6.35}$$

Using these in (3.28) then transforms the equations of motion (3.27) into (6.30) with the new bracket (6.31). □

Theorem VI.8 below now gives our last main result of this chapter, the analogue of Theorem VI.2 for Euler-Poincaré-Suslov nonholonomic systems.

Theorem VI.8. *Consider an Euler-Poincaré-Suslov system satisfying the conditions of Proposition VI.7 and let $f(g) \in \mathcal{C}^1$ be a function which is nonzero everywhere on its domain. Then the necessary conditions for Chaplygin Hamiltonization (using $d\tau = f dt$) on $(\mathfrak{g}^c)^*$ are that f satisfy*

$$j_\omega^*\{g, k\}_{AP} = f\{G, K\}_-^{j_\omega^*(\mathfrak{g}^c)^*}, \tag{6.36}$$

where $\{\cdot, \cdot\}_-^{j_\omega^*(\mathfrak{g}^c)^*}$ is the (minus) Lie-Poisson bracket on $j_\omega^*(\mathfrak{g}^c)^*$. Moreover, the Hamiltonized equations (3.27) are then:

$$\dot{\mathcal{P}}_i = f\{\mathcal{P}_i, \mathcal{H}_c\}_-^{j_\omega^*(\mathfrak{g}^c)^*}. \tag{6.37}$$

Proof. Similar to the proof of Theorem VI.2, we again recall that a Chaplygin Hamiltonizable nonholonomic system can be written in Hamiltonian form after the time reparameterization ($d\tau = f dt$). Now, since the natural Hamiltonian form for an Euler-Poincaré-Suslov system carries a (minus) Lie-Poisson bracket, this is equivalent to the requirement that f be chosen so as to make (6.31) a Poisson bracket, which is equivalent to the requirement that it satisfy the Jacobi identity. A short calculation shows that (6.31) satisfies the Jacobi identity iff the following relation is satisfied:

$$S_{km}^l S_{ij}^m + S_{jm}^l S_{ki}^m + S_{im}^l S_{jk}^m = 0, \quad \forall i, j, k, l = 1, \ldots, n-k, \tag{6.38}$$
$$\text{where } S_{km}^l := -\left(K_{mk}^l - \overline{C}_{km}^l\right).$$

If such an f exists, then $j_\omega^*(\mathfrak{g}^c)^*$ becomes a Lie algebra under the Poisson bracket (6.31), and this bracket then gives the Hamiltonization of equations (3.27) which then becomes (6.37). □

Taken together, Theorems VI.2 and VI.8 provide the necessary conditions for Hamiltonization of both nonabelian Chaplygin systems and Euler-Poincaré-Suslov systems in higher degrees of freedom. Through the local expressions of the theorems, equations (6.14) and (6.38), it is now possible to determine the reducing multiplier (if it exists) for a nonholonomic system in arbitrary degrees of freedom satisfying the conditions of the theorems directly from its given data $(L, G, \mathcal{D})$ by using any popular software package to solve the associated partial differential equations. Moreover, if a reducing multiplier does not exist we may still employ Theorem VI.6 in the presence of momentum conservation laws and re-attempt a Hamiltonization on the reduced space.

As a preliminary application of these theorems we shall now use Theorem VI.2 to extend the class of conditionally variational nonholonomic systems we introduced in Chapter V.

6.2 Conditionally Variational Systems in the Quasivelocity Context

Recall from Proposition V.10 that only nonholonomic systems with an invariant measure with constant density can hope to be conditionally variational. However, using Theorem VI.2 we can now extend this result to a more general setting if we instead focus on the Chaplygin Hamiltonized system. To that end we have the following result:

Theorem VI.9. *Suppose that for a given abelian Chaplygin nonholonomic system* $(L, G, \mathcal{D})$ *with constraints given by*

$$\phi^a(q, \dot{q}) = \dot{s}^a + A^a_\alpha(r)\dot{r}^\alpha, \quad a = 1, \ldots, k < n, \tag{6.39}$$

where $q = (r, s)$, *we have found an* f *as in Theorem VI.2 above and let* $\mathcal{L}(q, \omega) :=$

$L(q, \dot{r} = f\omega, \dot{s} = f\omega)$. Then if the matrix $\widetilde{g}_{ab} := (\partial^2 \mathcal{L}/\partial\omega^a \partial\omega^b)$ is invertible, the nonholonomic mechanics of the original system can be derived from the (almost) Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial \mathcal{L}_V}{\partial \omega^I} - f\frac{\partial \mathcal{L}_V}{\partial q^I} = 0, \qquad I = 1, \ldots, n, \tag{6.40}$$

by using the Lagrangian $\mathcal{L}_V(q, \omega)$ defined by

$$\mathcal{L}_V(q, \omega) = \mathcal{L}(q, \omega) - \frac{1}{f}\frac{\partial \mathcal{L}}{\partial \omega^a}\phi^a(q, \omega). \tag{6.41}$$

Proof. The existence of an f which Chaplygin Hamiltonizes the system guarantees, by Part (2) of Proposition V.5, that the system $(\mathcal{L}(q, \omega), \phi(q, \omega))$ is conditionally variational after the substitution $d\tau = f(r)dt$. Then, the theorem follows from Proposition V.7. $\square$

CHAPTER VII

Applications of the Theory

In the past three chapters we have developed three main methods of Hamiltonizing nonholonomic systems: (1) through associated second-order systems (Chapter IV), (2) through conditionally variational criteria (Chapter V), and (3) through Chaplygin Hamiltonization (Chapter VI). Moreover, we have also seen that each method has its advantages, disadvantages and area of applicability for a given nonholonomic system.

In this chapter we will discuss the application of all three methods. We begin with a theoretical application which closes the gap between nonholonomic mechanics and the optimal control dynamics of mechanical systems [6]. We then proceed to illustrate the three methods above through various examples of well-known nonholonomic systems.

7.1 The Pontryagin Maximum Principle and Nonholonomic Systems

It is well known that for holonomic systems the Lagrange-d'Alembert principle, variational constrained equations, and *Pontryagin's Maximum Principle* (see below) applied to the associated first-order control system all produce the same mechanics[1]. For nonholonomic systems, however, this equivalence breaks down in general. Having investigated the Hamiltonization of nonholonomic systems in the previous chapters

[1]Certain regularity conditions are needed, see Section 7.3 in [6].

though, as a theoretical application of the methods in Chapter IV we will show in this section that the Hamiltonians (4.31) and (4.32) can also be derived by associating a first-order controlled system to the (associated) nonholonomic systems of type II and applying Pontryagin's maximum principle.

To begin, consider the optimal control problem of finding the controls u that minimize a given cost function $G(x,u)$ under the constraint of a first order controlled system $\dot{x} = f(x,u)$. One of the hallmarks of continuous optimal control problems is that, under certain regularity assumptions, the optimal Hamiltonian can be found by applying the *Pontryagin Maximum Principle* [6]. Moreover, in most cases of physical interest the problem can be rephrased so as to be solved by using Lagrange multipliers p as follows. Form the Hamiltonian $H^P(x,p,u) = \langle p, f(x,u)\rangle - p_0 G(x,u)$ and calculate, if possible, the function $u^*(x,p)$ that satisfies the optimality conditions

$$\frac{\partial H^P}{\partial u}(x,p,u^*(x,p)) \equiv 0.$$

Then, an extremal $x(t)$ of the optimal control problem is also a base solution of Hamilton's equations for the optimal Hamiltonian given by $H^*(x,p) = H^P(x,p,u^*(x,p))$. The optimal controls $u^*(t)$ then follow from substituting the solutions $(x(t),p(t))$ of Hamilton's equations for H^* into $u^*(x,q)$.

We will show here that for the associated systems of type II

$$\ddot{r}_1 = 0, \qquad \ddot{q}_a = \Xi_a(r_1)\dot{q}_a\dot{r}_1,$$

we can also find the Hamiltonians of Section 4.4 via an application of Pontryagin's Maximum Principle. Hereto, let us put $\Xi_a = \xi_a'$ as before and observe that the above second-order system can easily be solved for $(\dot{r}_1(t), \dot{q}_a(t))$. Indeed, obviously $\dot{r}_1$ is constant along solutions, say u_1. We will suppose as before that $u_1 \neq 0$. From the q_a-equations it also follows that $\dot{q}_a/\exp(\xi_a)$ is constant, and we will denote this constant by u_a. To conclude,

$$\dot{r}_1(t) = u_1, \quad \dot{q}_a(t) = u_a \exp(\xi_a(r_1(t))).$$

Keeping that in mind, we can consider the following *associated controlled first-order system*:

$$\dot{r}_1 = u_1, \quad \dot{q}_a = u_a \exp(\xi_a(r_1)) \tag{7.1}$$

(no sum over a), where (u_1, u_a) are now interpreted as *controls*.

The next proposition relates the Hamiltonians of Proposition IV.3 to the optimal Hamiltonians for the optimal control problem of certain cost functions, subject to the constraints given by the controlled system (7.1).

Proposition VII.1. *The optimal Hamiltonian H^* of the optimal control problem of minimizing the cost function*

$$G_1(r_1, q_a, u_1, u_a) = \frac{1}{2}\left(I_1 u_1^2 + \sum_a C_a \exp(\xi_a(r_1))\frac{u_a^2}{u_1}\right)$$

subject to the dynamics (7.1) is given by:

$$H^*(q,p) = \frac{1}{2I_1}\left(p_{r_1} + \frac{1}{2}\sum_b \exp(\xi_b)\frac{p_b^2}{C_b}\right)^2. \tag{7.2}$$

If Ξ_2 is zero, the optimal Hamiltonian for the optimal control problem of minimizing the cost function

$$G_2(r_1, q_a, u_1, u_a) = \frac{1}{2}\left(I_1 u_1^2 + I_2 u_2^2 + \sum_\alpha a_\alpha \exp(\xi_\alpha(r_1))\frac{u_\alpha^2}{u_1}\right),$$

subject to the dynamics (7.1) is given by:

$$H^*(q,p) = \frac{1}{2I_2}p_2^2 + \frac{1}{2I_1}\left(p_1 + \frac{1}{2}\sum_\beta \exp(\xi_\beta)\frac{p_\beta^2}{a_\beta}\right)^2. \tag{7.3}$$

In case the controlled system is associated to a nonholonomic system (that is, in case the ξ_a take the form (4.28)), the above Hamiltonians are respectively the Hamiltonians (4.31) and (4.32) of Proposition IV.3.

Proof. The Hamiltonian H^P is

$$H^P(r_1, q_a, p_1, p_a, u_1, u_a) = p_1 u_1 + \sum_a p_a u_a \exp(\xi_a) - G_1. \tag{7.4}$$

The optimality conditions $\partial H^P/\partial u_1 = 0$, $\partial H^P/\partial u_a = 0$, together with the assumption that $u_1 \neq 0$, yield the following optimal controls as functions of (q,p):

$$\begin{aligned} I_1 u_1^* &= p_1 + \frac{1}{2}\sum_a \exp(\xi_a)\frac{p_a^2}{C_a}, \\ \frac{u_a^*}{u_1^*} &= \frac{p_a}{C_a}. \end{aligned}$$

For the Hamiltonian $H^*(q,p) = H^P(q,p,u^*(q,p))$, we get

$$\begin{aligned} H^*(q,p) &= \left(p_1 - \frac{1}{2}I_1 u_1^*\right)u_1^* + \sum_a \exp(\xi_a) u_a^* \left(p_a - \frac{1}{2}C_a \frac{u_a^*}{u_1^*}\right) \\ &= \frac{1}{I_x}\Bigg[\left(\frac{1}{2}p_1 - \frac{1}{4}\sum_a \exp(\xi_a)\frac{p_a^2}{C_a}\right)\left(p_1 + \frac{1}{2}\sum_b \exp(\xi_b)\frac{p_b^2}{C_b}\right) \\ &\qquad\qquad + \frac{1}{2}\sum_a \exp(\xi_a)\frac{p_a^2}{C_a}\left(p_1 + \frac{1}{2}\sum_b \exp(\xi_b)\frac{p_b^2}{C_b}\right)\Bigg] \\ &= \frac{1}{2I_1}\left(p_1 + \frac{1}{2}\sum_b \exp(\xi_b)\frac{p_b^2}{C_b}\right)^2, \end{aligned}$$

which is exactly the Hamiltonian (4.34).

For the second cost function, with $\Xi_2 = 0$, we get for Pontryagin's Hamiltonian

$$H^P = p_{r_1}u_{r_1} + p_{r_2}u_{r_2} + \sum_\alpha p_\alpha u_\alpha \exp(\xi_\alpha) - G_2.$$

The optimal controls as functions of (q,p) are now

$$\begin{aligned} I_1 u_{r_1}^* &= p_{r_1} + \frac{1}{2}\sum_\alpha \exp(\xi_\alpha)\frac{p_\alpha^2}{a_\alpha}, \\ I_2 u_{r_2}^* &= p_{r_2}, \\ \frac{u_\alpha^*}{u_{r_1}^*} &= \frac{p_\alpha}{C_\alpha}. \end{aligned}$$

With this the Hamiltonian becomes

$$H^*(q,p) = \frac{1}{2I_2}p_{r_2}^2 + \frac{1}{2I_1}\left(p_{r_1} + \frac{1}{2}\sum_\beta \exp(\xi_\beta)\frac{p_\beta^2}{a_\beta}\right)^2,$$

which is exactly (4.32) after the substitution (4.28). □

7.2 Examples of Hamiltonization

In this section we illustrate the three main avenues for Hamiltonization developed in this thesis and outlined above. Although we do not provide all of the details for each Hamiltonization method, it should be clear from the exposition of the previous chapters that not all of Hamiltonization methods developed earlier will apply for any given system. However, in view of the fact that at least one of our methods can be applied to most of the well-known nonholonomic systems, in an effort to conserve space we have limited our examples to those that best illustrate the results thus far.

7.2.1 The Vertical Rolling Disk

Consider the nonholonomic vertical rolling disk (see [6]) pictured in Figure 7.1 below[2] with configuration space $Q = \mathbb{R}^2 \times S^1 \times S^1$ and parameterized by the coordinates (x, y, θ, φ), where (x, y) is the position of the center of mass of the disk, θ is the angle that a point fixed on the disk makes with respect to the vertical and φ is measured from the positive x-axis. This system has Lagrangian and constraints given by:

$$\begin{aligned} L &= \frac{1}{2}m(\dot{x}^2+\dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2, \\ \phi^1 &= \dot{x} - R\cos\varphi\dot{\theta} = 0, \\ \phi^2 &= \dot{y} - R\sin\varphi\dot{\theta} = 0, \end{aligned} \tag{7.5}$$

[2] Used with permission from [6].

where m is the mass of the disk, R is its radius, and I, J are the moments of inertia about the axis perpendicular to the plane of the disk, and about the axis in the plane of the disk, respectively.

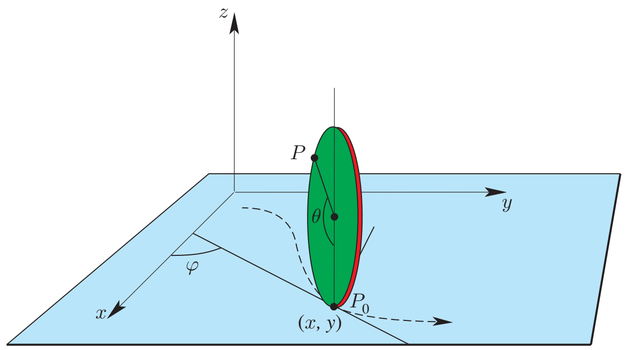

Figure 7.1: The Vertically Rolling Disk.

Viewing this system as an abelian Chaplygin system with $(s^1, s^2, r_1^\alpha, r_2^\alpha) = (x, y, \theta, \varphi)$ we compute (3.29) as:

$$\begin{aligned}\Lambda_{34} &= (mR\dot{x}\sin\varphi - mR\dot{y}\cos\varphi) \\ &= mR\Big((R\cos\varphi\dot{\theta})\sin\varphi - (R\sin\varphi\dot{\theta})\cos\varphi\Big) = 0,\end{aligned} \tag{7.6}$$

which by Remark V.4 shows that this system is conditionally variational. Moreover, the variational nonholonomic Lagrangian (5.15) is given by:

$$L_V = -\frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2 + mR\dot{\theta}(\dot{x}\cos\varphi + \dot{y}\sin\varphi). \tag{7.7}$$

The results of Proposition V.7 also apply here, and thus one could apply the Euler-Lagrange equations to L_V, and by imposing the constraints in (7.5) only initially, recover the full nonholonomic dynamics. It is also worth noting that a straightforward computation of (5.18) shows that L_V is regular, unlike in the variational approach where the Lagrangian is automatically singular.

Thus, we have regained regularity and may now pass to the Hamiltonian picture with momenta defined through the Legendre transform as in (5.20):

$$\begin{aligned} p_x &= -m\dot{x} + mR\cos\varphi\dot{\theta}, \\ p_y &= -m\dot{y} + mR\sin\varphi\dot{\theta}, \\ p_\theta &= I\dot{\theta} + mR\left(\dot{x}\cos\varphi + \dot{y}\sin\varphi\right), \\ p_\varphi &= J\dot{\varphi}. \end{aligned} \tag{7.8}$$

With these momenta, the constraints (4.1) are simply $p_x = p_y = 0$, and the Hamiltonian becomes:

$$\begin{aligned} H_V = \frac{p_\varphi^2}{2J} &+ \frac{1}{2m\beta}\left[m^2p_\theta^2 - \left(a^2\sin^2\varphi + Im\right)p_x^2 - \left(a^2\cos^2\varphi + Im\right)p_y^2\right] \\ &+ \frac{1}{2m\beta}\left[(a^2\sin 2\varphi)p_xp_y + \left(\cos\varphi p_x + \sin\varphi p_y\right)p_\theta\right], \end{aligned} \tag{7.9}$$

where $\beta = a^2 + Im$, and $a = mR$. Indeed, we see at once that since H is independent of x,y, and θ, the corresponding momenta are conserved. Moreover, after computing the associated canonical Hamilton equations and imposing on these the constraints $p_x = p_y = 0$, a straightforward verification shows that the resulting equations of motion reproduce the nonholonomic second-order equations of motion.

The Lagrangian (7.7) has already been encountered in (4.8). In that chapter we also discussed the Hamiltonization of the vertical disk through associated systems. By viewing the vertical disk within the framework of an associated system of type II, the first Lagrangian (4.26) is:

$$L = \rho(\dot{\varphi}) + \frac{\sqrt{I + mR^2}}{2}\left(C_2\frac{\dot{\theta}^2}{\dot{\varphi}} + C_3\frac{\dot{x}^2}{\cos(\varphi)\dot{\varphi}} + C_4\frac{\dot{y}^2}{\sin(\varphi)\dot{\varphi}}\right), \tag{7.10}$$

and the second Lagrangian (4.27) is:

$$L = \rho(\dot{\varphi}) + \sigma(\dot{\theta}) - \frac{\sqrt{I + mR^2}}{2}\left(a_3 \frac{\dot{x}^2}{\cos(\varphi)\dot{\varphi}} + a_4 \frac{\dot{y}^2}{\sin(\varphi)\dot{\varphi}}\right). \tag{7.11}$$

Moreover, as we saw in Section 4.4, we obtained a family of Hamiltonians that did not reproduce all of the nonholonomic solutions (only those for which $\dot{x} \neq 0$). The single Hamiltonian (7.9), in contrast, does reproduce all solutions on the submanifold defined by $p_x = p_y = 0$ but is also more complicated and allows no freedom. Compare this as well to the associated systems of type I of Section 4.3.1 for which there are no Hamiltonians at all.

Lastly, we mention that Corollary VI.4 adds no new information since it gives $f = N = const.$

7.2.2 The Nonholonomic Free Particle

The nonholonomic free particle consists of a free particle of mass m in $\mathbb{R}^3$ with position (x, y, z) subject to a nonholonomic constraint. The Lagrangian and constraint are given by [6]:

$$\begin{aligned} L &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \\ \dot{z} &= -x\dot{y}. \end{aligned} \tag{7.12}$$

We may view the system as an abelian Chaplygin system, and it is easy to show that it has an invariant measure with density $N = \frac{1}{\sqrt{1+x^2}}$, as can be verified directly through (5.23). Since this is nonconstant, we know that the system cannot be conditionally variational (this also follows immediately from Part (5) of Proposition V.5).

Since the system is not conditionally variational, not all nonholonomic solutions can be seen as variational ones. However, the solutions $(\dot{x}_0 t + x_0, y_0, z_0)$ for which $\dot{z} = 0$ can be seen as variational ones (for in this case $\Lambda = 0$ from Remark V.4).

Now, considering the associated system of type I from Chapter IV, we note that $\Gamma_2 = -x/(1+x^2)$ and $\Gamma_3 = -1/(1+x^2)$. The equations for $\Psi = \Phi, \nabla\Phi$ of the second type give the following two linearly independent equations

$$(\dot{x}^2 - 2)g_{23} + 3xg_{33} = 0, \quad (x^3 - 5x)g_{23} + (5x^2 - 1)g_{33} = 0.$$

We can easily conclude that $g_{23} = g_{33} = 0$ and hence the first two equations of the first type are

$$(x^2 - 2)\dot{x}g_{12} + 3x\dot{x}g_{13} + (x^2 - 2)\dot{y}g_{22} = 0,$$
$$(x^3 - 5x)\dot{x}g_{12} + (5x^2 - 1)\dot{x}g_{13} + (x^3 - 5x)\dot{y}g_{22}.$$

It follows from this $g_{13} = 0$ and $\dot{x}g_{12} = -\dot{y}g_{22}$, and that there is therefore no regular Lagrangian for the associated systems of type I. However, considering the associated system of type II we have that the Lagrangian (4.26) is:

$$L = \rho(\dot{x}) + \frac{1}{2}\sqrt{1+x^2}\left(C_2\frac{\dot{y}^2}{\dot{x}} + C_3\frac{\dot{z}^2}{x\dot{x}}\right). \tag{7.13}$$

Now, within the framework of Chapter VI, we note that by Corollary VI.4 $f(x) = (1+x^2)^{-1/2}$ and the quasivelocities are then defined by $\omega = \sqrt{1+x^2}\dot{r}$, where $r = (x, y)$. Corollary VI.4 then shows that the system is Chaplygin Hamiltonizable. To illustrate Theorem VI.9, note that $\mathcal{L}(q, \omega) = (1/2)f^2(\omega_x^2 + \omega_y^2 + \omega_z^2)$ and that since $\widetilde{g}_{zz} = (1/2)f^2$ equations (6.40) give:

$$\begin{aligned} \dot{\omega}_x &= \frac{x\omega_x^2}{(1+x^2)^{3/2}}, \\ \dot{\omega}_y &= 0, \\ \frac{d}{dt}(f^2(\omega_z + x\omega_y)) &= 0. \end{aligned} \tag{7.14}$$

Recalling that $\dot{q} = f\omega$, (7.14) expresses the conservation in time of the original constraint equation (7.12). Thus, enforcing the constraints initially is equivalent to choosing the integration constant to be zero in (7.14). In doing so, a short computation using the quasivelocity relations then shows that the dynamics in the system (7.14) are equivalent to the original nonholonomic mechanics. Theorem VI.9 then finally allows us to express the dynamics of the system (7.14) as the result of (6.40), where $\mathcal{L}_V$ is given by:

$$\mathcal{L}_V(q,\omega) = \frac{1}{2(1+x^2)}\left(\omega_x^2 + \omega_y^2 - \omega_z^2 - 2x\omega_x\omega_y\right), \tag{7.15}$$

and although (6.40) is not Hamiltonian, it is after Chaplygin's time reparameterization. Thus the nonholonomic free particle, like the vertical disk, is Hamiltonizable but since $f \neq \mathit{const}$ it is only conditionally variational after a reparameterization of time.

7.2.3 The Chaplygin Sphere

The Chaplygin sphere is a sphere rolling without slipping on a horizontal plane (see [6]) whose center of mass is at the geometric center, but the principal moments of inertia are distinct. In Euler angles (θ, ψ, φ) the Lagrangian and constraints are:

$$\begin{aligned} L = \frac{I_1}{2}\left(\dot\theta\cos\psi + \dot\varphi\sin\psi\sin\theta\right)^2 \; &+ \; \frac{I_2}{2}\left(-\dot\theta\sin\psi + \dot\varphi\cos\psi\sin\theta\right)^2 \\ &+ \; \frac{I_3}{2}\left(\dot\psi + \dot\varphi\cos\theta\right)^2 + \frac{m}{2}\left(\dot x^2 + \dot y^2\right), \end{aligned}$$

$$\begin{aligned} \phi^1 &= \dot x - \dot\theta\sin\varphi + \dot\psi\cos\varphi\sin\theta = 0, \\ \phi^2 &= \dot y + \dot\theta\cos\varphi + \dot\psi\sin\varphi\sin\theta = 0. \end{aligned} \tag{7.16}$$

where I_i are the moments of inertia about the q^i axes, and the ball is assumed to have mass m.

Now, since $q = (x, y, \theta, \psi, \varphi)$, and the constraints and Lagrangian are independent of x, y, we can view this system as an abelian Chaplygin system. The system also has an invariant measure whose density is in general non-constant (see [18]). However, the density is constant for the homogeneous sphere case (in which $I_1 = I_2 = I_3 = I$). Thus, the system might be conditionally variational if it satisfies (5.22). However, a quick computation shows that $F_{\psi\theta\psi} = \sin\theta\cos\theta$, which fails to satisfy (5.22).

There are still nonholonomic solutions that can be seen as variational ones though. The condition (5.8) becomes the matrix

$$\begin{pmatrix} m\dot{\psi}\sin\theta(\dot{\varphi} + \dot{\psi}\cos\theta) \\ -m\dot{\theta}\sin\theta(\dot{\varphi} + \dot{\psi}\cos\theta) \\ -m\dot{\psi}\,\dot{\theta}\sin\theta + m\dot{\psi}\,\dot{\theta}\sin\theta \end{pmatrix} = 0. \tag{7.17}$$

Clearly the last entry of (7.17) vanishes, and a straightforward computation shows that the variational nonholonomic Euler-Lagrange equation for φ reduces to the momentum conservation law:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}} = 0. \tag{7.18}$$

Now, for $I_1 = I_2 = I_3$ (7.18) expresses the conservation law $\dot{\varphi} + \dot{\psi}\cos\theta = C$, and so for the homogeneous sphere all the nonholonomic trajectories chosen such that $C = 0$ initially will annihilate (7.17). Thus all nonholonomic trajectories which satisfy $C = 0$ initially can be seen as variational ones. Moreover, since in this case we cannot view all of the nonholonomic trajectories as variational ones (only those which have $C = 0$), then system is not conditionally variational.

Returning to the general case, applying Theorem VI.2 shows that there does not exist an f which Hamiltonizes the three degree of freedom mechanics. However, it is easily seen that φ is a nonholonomic cyclic variable and leads to the momentum conservation law $p_\varphi = \mu_\varphi$. Thus we can form the constrained Routhian as in (6.22) and further reduce the dynamics to $M' = S^1 \times S^1$. We can then Hamiltonize on M' through Theorem VI.6, from which (6.23) shows that $f = N(\theta, \psi)$, where N is the invariant measure density for the resulting system on M' (see [18]). The non-canonical part of the almost-Poisson bracket (6.25) is then computed to be

$$\{\mathcal{P}'_1, \mathcal{P}'_2\} = -\mu_\psi (I_3 + 1) f^3 \sin\theta (I_1 \cos^2\varphi + I_2 \sin^2\varphi + 1), \tag{7.19}$$

and by the same Theorem (or simple computation) we know that this bracket satisfies the Jacobi identity and hence is indeed a Poisson bracket. This verifies the result obtained in [18] and is an example of a system that although is not Hamiltonizable at first is in fact so on the reduced space.

7.2.4 The Snakeboard

Another example of Theorem VI.6 is the Snakeboard [6, 55]. This system is modeled as a rigid body (the board) with two sets of independent 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, see Figure 7.2.

The configuration space is $Q = SE(2) \times S^1 \times S^1$ and the Lagrangian $L : TQ \to \mathbb{R}$ and constraints are given by:

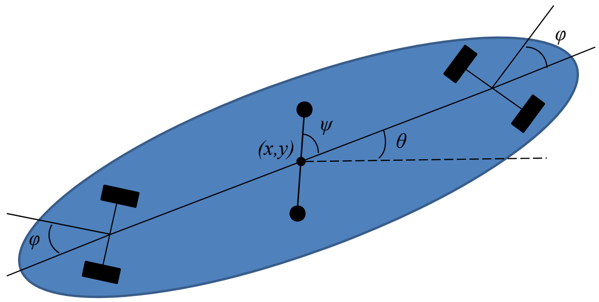

Figure 7.2: The Snakeboard.

$$L = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{\theta}^2 + \dot{\psi}^2 + 2\dot{\psi}\dot{\theta} + 2\dot{\phi}^2\right), \tag{7.20}$$

$$\dot{x} = -\cot\phi\cos\theta\dot{\theta}, \tag{7.21}$$

$$\dot{y} = -\cot\phi\sin\theta\dot{\theta}, \tag{7.22}$$

where we have set the mass m, moments of inertia, and the distance r from the center of the board to its wheels equal to unity. Here (x, y, θ) represents 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 relative to the board. Here we've made the simplification that $\phi_1 = -\phi_2$, as in [6, 55].

As stated, we can view this system as an abelian Chaplygin nonholonomic system with three degrees of freedom. Its equations of motion are given in [55] as:

$$\begin{aligned}
\dot{p}_\theta &= -\frac{1}{2}\sec\phi\csc\phi(p_\theta - p_\psi)p_\phi, & \dot{\theta} &= \tan^2\phi(p_\theta - p_\psi), \\
\dot{p}_\phi &= 0, & \dot{\phi} &= \frac{1}{2}p_\phi, \\
\dot{p}_\psi &= 0, & \dot{\psi} &= \frac{p_\psi - \sin^2\phi p_\theta}{\cos^2\phi}.
\end{aligned}$$

Since this system satisfies the conditions of Theorem VI.6 we can set $p_\psi = \mu_\psi = const.$ and focus on Hamiltonizing the reduced system. Given that this reduced system has the invariant measure $N(\phi) = \tan\phi$, by Corollary VI.4 $f = N$. The non-canonical part of the almost-Poisson bracket (6.25) is then computed to be:

$$\{\mathcal{P}_1', \mathcal{P}_2'\} = \sec^2\phi \mu_\psi, \tag{7.23}$$

and by the same Theorem we know that this bracket satisfies the Jacobi identity (since the reduced system has two degrees of freedom) and is thus a Poisson bracket.

7.2.5 The Chaplygin Sleigh

The Chaplygin Sleigh [6, 24, 25, 73] consists of a rigid body in the plane which is supported at three points, two of which slide freely without friction while the third is a knife edge, a constraint that allows no motion perpendicular to its edge. The configuration manifold $Q = \mathbb{R}^2 \times S^1$, where (x, y) are the coordinates of the contact point while θ is the angle the knife edge makes with the x-axis, see Figure 7.3 below. Moreover, we suppose here that the center of mass of the system C is not on top of the knife edge[3].

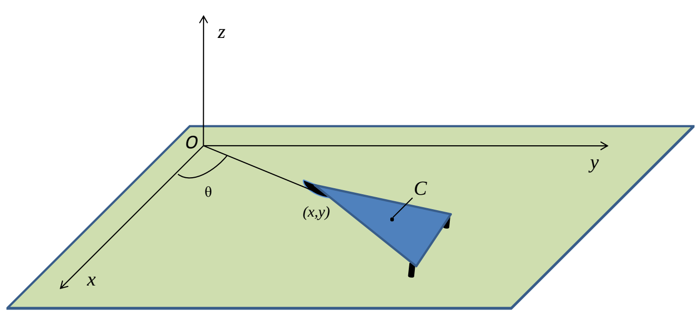

Figure 7.3: The Chaplygin Sleigh.

The Lagrangian L and constraints are given by:

[3] If it is, then one can show [6] that the sleigh reduces to another nonholonomic system known as the *knife edge*, which possesses an invariant measure.

$$L = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + 2\dot{\theta}^2 - 2(\dot{x}\sin\theta + \dot{y}\cos\theta)\dot{\theta}\right), \tag{7.24}$$

$$\dot{y}\cos\theta - \dot{x}\sin\theta = 0, \tag{7.25}$$

where for simplicity we have set all parameters to unity. Recall from Remark 1 in Section 5.2.2 that the Chaplygin sleigh is not conditionally variational. Moreover, due to its complicated Lagrangian, it does not fit into the class of systems (4.11) considered in Chapter V. However, it is Hamiltonizable within the framework of Chapter VI. Interestingly though, considered as an abelian Chaplygin system with symmetry group $G_1 = \mathbb{R}$ it is not Hamiltonizable, but using its $G_2 = SE(2)$ symmetry, considered as an Euler-Poincaré-Suslov system it is.

Since the Lagrangian and constraint are left invariant on the Lie group $G_2 = SE(2)$ we can treat the problem within the Euler-Poincaré-Suslov framework. Defining $\xi = g^{-1}\dot{g}$, where $g = (x, y, \theta)$, we can write the Lagrangian L in terms of ξ as $l(\xi) = \xi_3^2 + (1/2)(\xi_1^2 + \xi_2^2) + \xi_2\xi_3$, and the constraint as $\xi^2 = 0$. Thus from Section 3.6.2 we have that in this case $b_i^n = 0\ \forall i$.

With the structure constants given by $C_{13}^2 = -1 = -C_{23}^1$ and all other zero we see that $f = const.$ satisfies Theorem VI.8, which agrees with the recent result of [45]. Moreover, as is well-known the Chaplygin sleigh does not possess an invariant measure [6, 45], yet as we've seen above this system is Chaplygin Hamiltonizable.

This system is of critical importance in the study of Hamiltonization since unlike Proposition VI.5, the Chaplygin sleigh shows that just because a system is Hamiltonizable does not imply that it possesses an invariant measure. Thus, unlike in the nonabelian Chaplygin case, the Hamiltonizability of Euler-Poincaré-Suslov systems does not automatically imply that the system possesses an invariant measure. The traditional relationship between the existence of the invariant measure and the reducing multiplier f thus breaks down. However, as Theorem VI.8 and the above example shows, may still be able to Hamiltonize, or more properly, "Poissonize" (since the Hamiltonization doesn't always result in an invariant measure).

7.2.6 A Mathematical Example

Consider the following mathematical example due to Iliyev [64]. The Lagrangian and constraints are given by:

$$\begin{aligned} L &= \frac{1}{2}\left((\dot{q}^1)^2 + (\dot{q}^2)^2 + (\dot{q}^3)^2 + (\dot{q}^4)^2 + (\dot{q}^5)^2\right), \\ \dot{q}^4 &= \dot{q}^2 \tan(q^1), \\ \dot{q}^5 &= \dot{q}^3 \tan(q^1). \end{aligned} \tag{7.26}$$

This is a nonholonomic system with three degrees of freedom ($m = 3$) and can most easily be treated as an abelian Chaplygin system. Solving the conditions in (6.14) yields $f = \cos(q^1)$. Moreover, as a check of Proposition VI.5, we can use (3.33) to compute the system's invariant measure density to be $N = \cos^2(q^1)$, which indeed is equal to f^{m-1}, as the Proposition suggests.

CHAPTER VIII

Conclusion and Future Directions

8.1 Conclusion

The results of Chapters IV - VI show that although nonholonomic systems are not variational, in many cases they can be cast in Hamiltonian form through a variety of methods, not all of which are simultaneously applicable to a particular system. In this regard Hamiltonization is somewhat of an art form. This is perhaps best illustrated by the examples of the Chaplygin sleigh and vertical disk of Sections 7.2.5 and 7.2.1, respectively. Ignoring the symmetry of the Chaplygin sleigh leads to an inability to Hamiltonize it using any of the methods presented here, forcing one to consider its $SE(2)$ symmetry in order to Hamiltonize it. In the case of the vertical disk in Section 7.2.1, although Chaplygin Hamiltonization and Hamiltonization through the identification of the vertical disk as a conditionally variational system are both applicable, the family of Lagrangians for the associated second-order systems in some cases does not exist (for type I) and in other cases does not reproduce all of the solutions to the original system (7.5).

Given the art form involved in Hamiltonizing nonholonomic systems, one must ask why do so at all? An immediate partial answer to this question comes from Section 7.1, where we were able to preserve the equivalence present between the equations of motion of holonomic systems and their associated first-order optimal control problems. From a theoretical standpoint, the optimal control of nonholonomic systems is itself a large area of research, and Proposition VII.1 now makes possible a direct comparison between, in Hertz's terminology, the "straightest" curves and "shortest" curves which represent the physical trajectories of two different objectives (user defined optimal control on the one hand and actual motion on the other). However, there are other main areas to which we are currently applying the results of this thesis.

8.2 Future Directions

8.2.1 Quantization of Nonholonomic Systems

The quantization of nonholonomic systems has recently attracted the attention of many researchers [1, 17, 32, 52, 53, 75]. Although some would consider it a theoretical exercise, from a practical point of view the advent of nanomachines may influence this viewpoint, and from a theoretical point of view developing a consistent quantization scheme which can handle even simple holonomic constraints or even configuration spaces of non-zero curvature is still a challenge (see [52]). However, the variational Lagrangian (7.7) looks promising as a starting point for a quantization of a nonholonomic system based on the ideas developed in Chapter V. We are currently investigating the results of its quantization [12]. Moreover, we expect that the overall results obtained in the previous chapters (specifically the overarching idea of Hamiltonizing the entire nonholonomic system) will provide the groundwork for quantizing more general nonholonomic systems.

8.2.2 Numerical Schemes

Recently, the development of variational integrators (see [69] for a recent summary of the basic theory) has allowed substantially more accurate simulations at lower cost for conservative (or weakly dissipative) systems. Although we have emphasized the non-variational nature of nonholonomic systems, the various Hamiltonization methods presented in this thesis circumvent this to a large degree and allow the possibility of applying variational integrators to nonholonomic systems. We are currently exploring this byproduct of Hamiltonization [11].

8.2.3 Development of a Nonholonomic Hamilton-Jacobi Equation

The Hamilton-Jacobi equation of classical mechanics is well developed and highly useful in tackling the integrability of mechanical systems [67]. However, since it relies on a variational principle it is again inapplicable to nonholonomic systems. Recently, the authors in [63] proposed a Hamilton-Jacobi theory for nonholonomic systems and discussed its application to Chaplygin systems. We should note that these ideas are not new, and were implicitly used in [73], albeit in "quasicoordinates," to integrate the Chaplygin sleigh, among other systems. Given the success of the results of the previous chapters, especially with Chaplygin systems, it would be interesting to use the existing Hamilton-Jacobi theory in conjunction with our Hamiltonization mehods to develop a Hamilton-Jacobi theory for the classes of systems studied in this thesis. We are currently collaborating with the authors in [74] on this issue.

8.2.4 The Effect of Symmetry on the Hamiltonizability of a Nonholonomic System

We have commented on the fact that in some cases it is to our benefit (as far as Hamiltonization goes) to ignore the symmetry present in a nonholonomic system, whereas in other cases it is not. Moreover, most symmetries discussed in the literature [6] are *physical symmetries*, meaning that they arise based on consideration of the physical properties of the system. It also seems important though, to ask the following question: given a nonholonomic system, what is the set of symmetry groups under which a the system is (1) conditionally variational or (2) Chaplygin Hamiltonizable? For example for the Chaplygin sleigh of Section 7.2.5, $G = \mathbb{R}$ does not belong to this set, whereas $G = SE(2)$ does. We are presently exploring this in [39].

8.2.5 Hamiltonization by Stages

Finally, we note that the multi-dimensional Veselova system and multi-dimensional Chaplygin sphere have recently been Hamiltonized in [37] and [51], respectively. However, the methods and conditions for Hamiltonization presented in this thesis are inapplicable to those Hamiltonizations due to the particular Hamiltonization methods used by the authors. In the former, the authors constructed redundant coordinates and showed that the solutions of the multi-dimensional Veselova system can be mapped isomorphically into the solutions of an associated *different* Hamiltonian system known as the Neumann system. Within the framework of the methods presented here, this would be equivalent to the statement that after an appropriate time reparameterization, applying the inverse problem of the calculus of variations to the resulting system would yield the Neumann Lagrangian as a solution. In the latter case, the author Hamiltonizes the multi-dimensional Chaplygin sphere by constructing redundant coordinates and effecting a time-reparameterization. It is then shown that the reduced mechanics of the higher dimensional nonholonomic Chaplygin sphere emerge as the restriction to the invariant submanifolds of the Hamiltonian system resulting from the time reparameterization. Within our framework, this is equivalent to the statement that the author in [51] has succeeded in constructing an associated second-order system for the time-reparameterized mechanics. Thus, the main difference with our work is that in Chapter IV we constructed associated second-order systems for the original nonholonomic system (not the time reparameterized one).

Given the above discussion, we therefore expect that the aforementioned multi-dimensional Hamiltonizations can be realized as special cases of a synthesis of the general (yet mostly disjoint) methods presented in this thesis. We are currently exploring this [41].

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Abud Filho, M., Gomes, L.C., Simao F.R.A. and Coutinho, F.A.B. The Quantization of Classical Non-holonomic Systems, *Revista Brasileira de Fisica* **13** (1983) 384-406.

[2] Arnold, V.I., Kozlov, V.V. and Neishtadt, A.I. *Mathematical aspects of classical and celestial mechanics*, Springer-Verlag, Berlin, 1997.

[3] Appell, P. Sur l'intégration des équations du mouvement d'un corps pesant de révolution roulant par une arête circulaire sur un plan horizontal; cas parficulier du cerceau, *Rendiconti del circolo matematico di Palermo*, 14 (1900), 1-6.

[4] Appell, P. Sur les liasions experimées par des relations non linéaires entre les vitesses, C.R. Acad. Sc. Paris, 152 (1911), 1197-1199.

[5] Bates, L. Examples of Singular Nonholonomic Reduction, *Rep. Math. Phys.*, 42 (1998), 231-247.

[6] Bloch, A.M. with the collaboration of Baillieul, J., Crouch, P. and Marsden, J.E. *Nonholonomic Mechanics and Control*, Springer NY (2003).

[7] Bloch, A.M. and Crouch, P.E. Nonholonomic and vakonomic control systems on Riemannian manifolds, *Fields Institute Communications*, 1 (1993), 25-52.

[8] — Optimal Control and Geodesic flows, *Systems Control Lett.*, 28(2) (1996), 65-72.

[9] — Newton's law and integrability of nonholonomic systems, *SIAM J. Control Optim.*, 36 (1998), 2020-2039.

[10] Bloch, A.M., Fernandez, O.E. and Mestdag, T. Hamiltonization of Nonholonomic Systems and the Inverse Problem of the Calculus of Variations, *to appear, Rep. Math. Phys. 2009.*

[11] — Applying Variational Integrators to Nonholonomic Systems through Hamiltonization, *in preparation.*

[12] Bloch, A.M., Fernandez, O.E. and Uribe, A. Quantization of the Nonholonomic Vertically Rolling Disk, *in preparation.*

[13] Bloch, A.M., Krishnaprasad, P.S., Marsden, J.E. and Murray, R. Nonholonomic mechanical systems with symmetry, *Arch Rat. Mech. An.*, 136 (1996), 21-99.

[14] Bloch, A.M., Krishnaprasad, P.S., Marsden, J.E. and Ratiu, T.S. The Euler-Poincare equations and double bracket dissipation, *Comm. Math. Phys.*, 175 (1996), 1-42.

[15] — Nonholonomic mechanical systems with symmetry, *Arch. Rat. Mech. An.*, 136 (1996), 21-99.

[16] Bloch, A.M., Marsden, J.E. and Zenkov, D.V. Quasivelocities and Symmetries in Nonholonomic Systems, *Dyn. Sys.*, (2009) 1-36.

[17] Bloch, A.M. and Rojo, A. Quantization of a nonholonomic system, *Phys. Rev. Lett.*, 101 (2008) 030402.

[18] Borisov, A.V. and Mamaev, I.S. Hamiltonization of nonholonomic systems, *arXiv Pre-print: nlin/0509036*.

[19] — On the history of the development of the nonholonomic dynamics, *Regul. Chaotic Dyn.*, 7(1) (2002), 43-47.

[20] Cantrijn, F., Cortes, J., de Leon, M. and Martin de Diego, D. On the geometry of generalized Chaplygin systems, *Math. Proc. Cambridge Philos. Soc.*, 132(2) (2002), 323-351.

[21] Cantrijn, F., de Leon, M. and Martin de Diego, D. On almost-Poisson structures in nonholonomic mechanics, *Nonlinearity*, 12 (1999), 721-737.

[22] Cardin, F. and Favretti, M. On nonnholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints, *J. Geometry and Physics*, 18 (1996), 295-325.

[23] Cendra, H., Marsden, J.E and Ratiu, T.S. Geometric Mechanics, Lagrangian Reduction, and Nonholonomic Systems, *Mathematics Unlimited-2001 and Beyond*, pgs. 221-273.

[24] Chaplygin, S.A. On a rolling sphere on a horizontal plane (in Russian), *Mat. Sbornik*, 24 (1903), 139-168.

[25] Chaplygin, S.A. On the theory of motion of nonholonomic systems. Theorem on the reducing multiplier (in Russian), *Mat. Sbornik*, 28(2) (1911), 303-314.

[26] Cortes, J. *Geometric, control and numerical aspects of nonholonomic systems*, Springer (2002).

[27] Cortes, J., de Leon, M., de Diego, D.M. and Martinez, S. Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions, *SIAM Journal on Control and Optimization*, 41(5) (2003), 1389-1412.

[28] Crampin, M. and Mestdag, T. The inverse problem for invariant Lagrangians on a Lie group, *J. of Lie Theory* **18** (2008) 471–502.

[29] Crampin, M., Sarlet, W., Byrnes G.B. and Prince, G.E. Towards a geometrical understanding of Douglas's solution of the inverse problem of the calculus of variations, *Inverse Problems* **10** (1994) 245-260.

[30] Dirac, P.A.M., *Lectures on Quantum Mechanics* (Academic, New York, 1964).

[31] Douglas, J. Solution of the inverse problem of the calculus of variations, *Trans. Amer. Math. Soc.* **50** (1941) 71–128.

[32] Eden, R.J. The quantum mechanics of non-holonomic systems, *Proc. Royal Soc. London*, **205** (1951) 583-595.

[33] Ehlers, K, Koiller, J., Montgomery, R. and Rice, P.M. Nonholonomic systems via moving frames: Cartan's equivalence and Chaplygin Hamiltonization, in *The Breadth of Symplectic and Poisson Goemetry, Festschrift in Honor of Alan Weinstein*, Progress in Mathematics 232, 75-120, Birkhauser, Boston, 2004.

[34] Favretti, M. Equivalence of dynamics for nonholonomic systems with transverse constraints, *J. Dynamics and Differential Equations*, 10(4) (1998), 511-536.

[35] Federov, Y.N. and Jovanovic, B. Quasi-Chaplygin Systems and Nonholonomic Rigid Body Dynamics, *Lett. Math. Phys.*, 76(2-3) (2006), 215-230.

[36] — Integrable Nonholonomic Geodesic Flows on Compact Lie Groups, *Top. Methods in the theory of Integrable Systems*, (2005), 115-152.

[37] — Nonholonomic LR Systems as Generalized Chaplygin systems with an Invariant Measure and Geodesic Flows on Homogeneous Spaces, *J. Non. Sci.* 14 (2004), 341-381.

[38] Fernandez, O.E. and Bloch, A.M. Equivalence of the Dynamics of Nonholonomic and Variational Nonholonomic Systems for certain Initial Data, *J. Phys. A: Math. Theor.* 41 (2008), 344005.

[39] — The Set of Symmetry Groups under which a Nonholonomic System is Hamiltonizable, *in preparation.*

[40] Fernandez, O. E., Bloch, A.M. and Mestdag, T. The Pontryagin Maximum Principle applied to Nonholonomic Mechanics, *47th IEEE Proc. on Decision and Control* (2008), 4306-4311.

[41] Fernandez, O.E., Mestdag, T. and Bloch, A.M. Hamiltonization by Stages, *in preparation.*

[42] Ferrers, N.M. Extension of Lagrange's Equations, *Quart. J. of Pure and Appl. Math.*, 12(45) (1872), 1-5.

[43] Flannery, M.R. The enigma of nonholonomic constraints, *Am. J. Phys.*, 73(3) (2005), 265-272.

[44] Garcia-Naranjo, L.C. Reduction of Almost Poisson brackets and Hamiltonization of the Chaplygin Sphere, *eprint arXiv:0808.0854.*

[45] — Reduction of Almost Poisson Brackets for Nonholonomic Systems on Lie Groups, *Reg. Chaotic Dyn.*, 12(4) (2007), pp. 365-388.

[46] Gotay, M.J., Constraints, reduction, and quantization, *J. of Math. Phys.*, 27(8) (1986), pp.2051-2066.

[47] Guillemin, V. and Pollack, A. *Differential Topology*, Prentice Hall (1974).

[48] Guo, Y.X., Wang, Y., Chee, G.Y. and Mei, F.X. Nonholonomic versus vakonomic dynamics on a Riemann-Cartan manifold, *J. Mathematical Phys.*, 46 (2005), 062902.

[49] Guo, Y.X., Luo, S.K., Shang, M. and Mei, F.X. Birkhoffian formulations of nonholonomic constrained systems, *Rep. Math. Phys.*, 47(3) (2001), 313-332.

[50] Jovanović, B. Nonholonomic geodesic flows on lie groups and the integrable Suslov problem on $SO(4)$, *J. Phys. A: Math. Gen.* , 31 (1998), 1415-1422.

[51] Jovanović, B. Hamiltonization and integrability of the Chaplygin spehere in $\mathbb{R}^n$, *eprint arXiv:0902.4397.*

[52] Kaplan, K., Maitra, N.T. and Helleret, E.J. Quantizing constrained systems, *Phys. Rev. A* 56 (1997) 2592-2599.

[53] Klauder, J. *Beyond Conventional Quantization*, Cam. Univ. Press (1999).

[54] Koiller, J. Reduction of some classical nonholonomic systems with symmetry, *Arch. Rational Mech. Anal.*, 118 (1992), 113-148.

[55] Koon, W. S. and Marsden, J. E. The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems, *Reports on Math Phys.* 40 (1997), 21–62 .

[56] — The Poisson reduction of nonholonomic mechanical systems, *Reports on Math. Phys.*, 42 (1998), 101-134.

[57] Kozlov, V. V. *Dynamical Systems X: General Theory of Vortices*, Springer NY (2003).

[58] — Invariant measures of the Euler-Poincare equations on Lie algebras, *Functional Anal. Appl.*, 22, 69-70.

[59] Krupková, O. and Prince, G.E. Second-order ordinary differential equations in jet bundles and the inverse problem of the calculus of variations, Chapter 16 of D. Krupka and D. J. Saunders (eds.), *Handbook of Global Analysis*, Elsevier (2007).

[60] Hamel, G. *Theoretische Mechankik. Eine Einheitliche Einfuehrung in die gesamt Mechanik*, volume 57 of *Die Grundlehren der math. Wissenschaften.* Springer (1949).

[61] Henneaux, M. and Teitelboim, C. *Quantization of Gauge Systems*, Princeton Univ. Press (1994).

[62] Hertz, H. *Gessamelte Werke*, Band III. *Der Prinzipien der Mechanik in neum Zusammenhange dargestellt*, Barth, Leipzig, (1894), English translation MacMillan, London, 1899, reprint Dover, NY (1956).

[63] Iglesis-Ponte, David., et. al. Towards a Hamilton-Jacobi theory for nonholonomic mechanical systems, *J. Phys. A: Math. Theor.*, 41 (2008), 015205.

[64] Iliyev, IL. On the conditions for the existence of the reducing chaplygin factor, *PMM USSR*, 49(2), 295-301.

[65] Lewis, A. and Murray, R. Variational principles for constrained systems: theory and experiment, *The Intl. J. Nonlinear Mechanics*, 30(6) (1995), 793-815.

[66] Maggi, G.A. Da alcune nuove forma della equazioni della dinamica applicabile ai sistemi anolonomi, *Atti della Reale Acad. Naz. dei Lincei. Rend. Cl. fis. e math, Ser. 5*, 10(2) (1901), 287-291.

[67] Marsden, J.E. and Ratiu, T.S. *Introduction to Mechanics and Symmetry*, Second Ed., Springer NY (1999).

[68] Marsden, J.E and Scheurle, J. The reduced Euler-Lagrange equations, *Fields Inst. Com.*, 1 (1993), 139-164.

[69] Marsden, J.E. and West, M. Discrete mechanics and variational integrators, *Acta Num.* 10 (2001), 357-514.

[70] Martinez, S., Cortes, J. and de Leon, M. The geometrical theory of constraints applied to the dynamics of vakonomic mechanical systems: The vakonomic bracket, *Journal of Mathematical Physics*, 43(4) (2000), 2090-2120.

[71] Massa, E. and Pagani, E. Classical Dynamics of non-holonomic systems: a geometric approach, *Ann. Inst. Henri Poicar*, 55 (1991), 511-544.

[72] Montgomery, R. *A Tour of Subriemannian Geometries, Their Geodesics, and Applications*, American Mathematical Society (2001).

[73] Neimark, J.I. and Fufaev, N.A. *Dynamics of Nonholonomic Systems*, Amer. Math. Soc. (1972).

[74] Ohasawa, T. and Bloch, A.M. The nonholonomic Hamilton-Jacobi equation and integrability, *in preparation.*

[75] Pitanga, P. Quantization of a non-holonomic system with symmetry, *Il Nuovo Cimento B*, 109 (1994) 583-594.

[76] Poincaré, H. Hertz's ideas in mechanics, in *The last papers by H. Poincaré*, translation from French. Izhevsk: *Regul. Chaotic Dyn.* (2001) (in Russian).

[77] Puta, M. *Hamiltonian Mechanical Systems and Geometric Quantization*, Springer-Verlag NY (1993).

[78] Ramos, A. Poisson structures for reduced non-holonomic systems, *J. Phys. A*, 37(17) (2004), 4821-4842.

[79] Rosenberg, R.M.*Analytical Dynamics of Discrete Systems*, Plenum Press, NY (1977).

[80] Rumiantsev, V.V. On Hamilton's principle for nonholonomic systems, *P.M.M. USSR*, 42 (1978), 387-399.

[81] Santilli, R.M. *Foundations of Theoretical Mechanics: The Inverse Problem in Newtonian Mechanics*, Spinger-Verlag (1978).

[82] Stanchenko, S.V. Nonholonomic Chaplygin Systems (in Russian), *Prikl. Mat. Mehk,* 53(1) (1989), 16-23.

[83] Sumbatov, A.S. Nonholonomic Systems, *Regul. Chaotic Dyn.* 7(2) (2002), 221-238.

[84] van der Schaft, A. J. and Maschke, B. M. On the Hamiltonian formulation of nonholonomic mechanical systems, *Rep. on Math. Phys.* 34 (1994), 225–233.

[85] Zenkov, D.V. and Bloch, A.M. Invariant measures of nonholonomic flows with internal degrees of freedom, *Nonlinearity*, 16(5) 2003, 1793-1807.