

Equivalence of the dynamics of nonholonomic and variational nonholonomic systems for certain initial data

Oscar E Fernandez and Anthony M Bloch

Department of Mathematics, University of Michigan, 2074 East Hall, 530 Church Street,
Ann Arbor, MI 48109-1043, USA

E-mail: oscarum@umich.edu and abloch@umich.edu

Received 21 November 2007, in final form 26 April 2008

Published 11 August 2008

Online at stacks.iop.org/JPhysA/41/344005

Abstract

In this paper, we discuss the necessary and sufficient conditions for the equivalence of the dynamics of nonholonomic mechanics and variational nonholonomic (vakonomic) dynamics for certain initial conditions. We derive *a priori* results for identifying equivalence and, specializing to Abelian Chaplygin systems, prove that equivalence results if and only if the constrained nonholonomic equations are Lagrangian. We eliminate the need to solve the variational nonholonomic problem when checking equivalence by obtaining explicit formulae for the system's multipliers, and then derive conditions under which the multiplier free Lagrangian gives equivalence of the dynamics. We consider nonholonomic systems possessing invariant measures, showing when equivalence and Hamiltonization are the same. We also derive conditions under which measure-preserving systems exhibit equivalence. We apply the results to many of the known nonholonomic systems.

PACS numbers: 45.10.-b, 02.30.Yy, 45.20.-d

Introduction

The dynamics of nonholonomic systems is derivable from the Lagrange–D'Alembert (LD) principle, and is in general not *Lagrangian*, meaning that the resulting equations of motion are not the Euler–Lagrange equations of any Lagrangian [3]. That this is so can be attributed to the effects of the nonintegrable constraints, and how they are treated under the LD principle. However, instead of applying the LD principle, one may use Lagrange's multiplier theorem to incorporate the nonholonomic constraints. Such an application generally leads to different dynamics, resulting in what is called variational nonholonomic (also vakonomic) dynamics [16]. Although the LD principle gives the correct (physical) equations of motion, some

authors have incorrectly claimed that the variational nonholonomic treatment gives the physical dynamics (see [23] and references therein).

In a sense, the theory of variational nonholonomic dynamics may be viewed as an attempt to ‘Hamiltonize’ nonholonomic mechanics, so that the equations of motion can be derived from Hamilton’s principle. There have been many such attempts, approaching the question from different angles, and we only discuss a few of them here.

The most direct attempt is through comparison of the two methods themselves. 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 [7] and references therein). Geometrically, this is because variational dynamics are equivalent to optimal control problems under certain regularity conditions [3, 26], 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 [3, 7]. As a result, many authors have studied the similarities and differences between nonholonomic mechanics and variational nonholonomic dynamics (see [4, 11, 13, 17, 28]) in an effort to gain insight into the nonintegrable case.

In a slightly different direction, there has been much work done which casts nonholonomic systems into Hamiltonian form via classical Hamiltonization (in the sense of Chaplygin’s reducing multiplier), among which we mention [6] and [15]. The idea here is to reparametrize time with the aid of an invariant measure for the nonholonomic system such that in the new time variable the system is Hamiltonian.

Yet another direction is the almost Poisson route, where the authors, (see, e.g. [2, 20, 27, 29]) have written the nonholonomic equations of motion in Hamiltonian form with respect to an almost Poisson bracket.

Last, we mention the work of [18], where the authors construct canonical equations for nonholonomic systems by identification with Birkhoff mechanics. Unfortunately, as the authors themselves point out, the construction no longer preserves the original physical meaning of the nonholonomic system.

Clearly, these different directions have a common aim: to ‘Hamiltonize’ nonholonomic systems. As such, they will also have common points of intersection and can each benefit from results in the other. It is the aim of the paper to explore some of these interrelationships and draw some connections amongst the different approaches.

We begin in sections 1 and 2 with a review of the relevant dynamics of nonholonomic and variational nonholonomic systems, respectively. In section 3 we discuss the notions of *conditionally variational* nonholonomic systems (in short, the notion that there may exist certain initial conditions for the Lagrange multipliers of the variational problem for which all of the nonholonomic trajectories are solutions to the variational nonholonomic problem, and vice versa), and of *partially conditionally variational* (systems for which only some but not all nonholonomic trajectories can be viewed as variational nonholonomic trajectories). We geometrize the necessary and sufficient conditions originally presented in [28] and derive results that show *a priori* that certain systems cannot be conditionally variational. Then, because many nonholonomic systems are of Chaplygin type, we specialize to these and prove that Abelian Chaplygin systems are conditionally variational if and only if the reduced nonholonomic equations are Lagrangian, in addition to providing an explicit formula for the variational nonholonomic multipliers (which eliminates the need to first completely solve the

variational nonholonomic problem in order to check the conditions of section 3). We also show how to eliminate the multipliers in certain cases and provide the conditions under which the multiplier-free Lagrangian makes the nonholonomic system conditionally variational. Section 4 then considers nonholonomic systems which possess an invariant measure. There we prove a theorem relating Hamiltonization [6] to the idea of conditionally variational systems, and elaborate on some computational simplifications in searching for conditionally variational systems in proposition 6.

Finally, we conclude in sections 5 and 6 by applying the results to many of the common nonholonomic systems, showing which systems are conditionally variational, and which are only partially conditionally variational.

1. Nonholonomic mechanics

For our purposes, we shall consider a *nonholonomic system* on a configuration manifold Q to be a pair (L, \mathcal{D}) , where $L : TQ \rightarrow \mathbb{R}$ is a regular Lagrangian of mechanical type $L = T - V$, where $T : TQ \rightarrow \mathbb{R}$ is the kinetic energy corresponding to a Riemannian metric g on Q , and $V : Q \rightarrow \mathbb{R}$ 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 [3, 10]. In a neighborhood of each point, one can choose a local coordinate chart such that ω^a and \mathcal{D} take the form

$$\omega^a = ds^a + A_\alpha^a(r, s) dr^\alpha, \tag{1.1}$$

$$\mathcal{D} = \text{span}\{\partial_{r^\alpha} - A_\alpha^a(r, s)\partial_{s^a}\}, \tag{1.2}$$

respectively, where $q = (r, s) \in Q$, and for the remainder of the paper $a, b = 1, \dots, k, \alpha, \beta = k + 1, \dots, n$, where $n = \dim Q$. We will also call $n - k$ the *degrees of freedom* of the nonholonomic system.

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

$$A = (ds^a + A_\alpha^a(r, s) dr^\alpha)\partial_{s^a}. \tag{1.3}$$

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

The curvature of A is the vertical-vector-valued-two-form B on Q defined by $B(X, Y) = -A([\text{hor } X, \text{hor } Y])$, 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. Moreover, the curvature can be shown to be independent of the extension of the vector fields [3], and 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).

Now, using an identity¹ for the exterior derivative of a one-form, we can express the curvature as $B(X, Y) = d\omega^a(\text{hor } X, \text{hor } Y)\partial_{s^a}$, so that the local expression for the curvature is given by:

$$B(X, Y)^a = B_{\alpha\beta}^a X^\alpha Y^\beta, \quad \text{where} \quad B_{\alpha\beta}^a = \left(\frac{\partial A_\alpha^b}{\partial r^\beta} - \frac{\partial A_\beta^b}{\partial r^\alpha} + A_\alpha^a \frac{\partial A_\beta^b}{\partial s^a} - A_\beta^a \frac{\partial A_\alpha^b}{\partial s^a} \right). \tag{1.4}$$

¹ For two vector fields X, Y acting on Q , the exterior derivative of the one-form α is $(d\alpha)(X, Y) = X[\alpha(Y)] - Y[\alpha(X)] - \alpha([X, Y])$.

Defining the *constrained Lagrangian* $L_c(q, \dot{q}) = L(q, \text{hor } \dot{q})$, the corresponding equations of motion for the nonholonomic system are obtained through the *Lagrange–D’Alembert Principle* [3], and are given by

$$\delta L_c = \langle \mathbb{F}L, B(\dot{q}, \delta q) \rangle, \quad \text{where} \quad \delta L_c = \left\langle \delta q^\alpha, \frac{\partial L_c}{\partial q^\alpha} - \frac{d}{dt} \frac{\partial L_c}{\partial \dot{q}^\alpha} \right\rangle, \quad (1.5)$$

and where $\langle \cdot, \cdot \rangle$ denotes the pairing between a vector and a dual vector. Here, $\mathbb{F}L : TQ \rightarrow T^*Q$ is the fiber derivative, δq is a horizontal variation and B is the curvature regarded as a vertical-valued two-form. These equations are to be supplemented by the constraint equations $A(q) \cdot \dot{q} = 0$.

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

$$\frac{d}{dt} \frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} = - \left(\frac{\partial L}{\partial \dot{s}^b} \right)_c B_{\alpha\beta}^b \dot{r}^\beta - A_\alpha^a \frac{\partial L_c}{\partial s^a}, \quad (1.6)$$

$$\dot{s}^a = -A_\alpha^a(r, s) \dot{r}^\alpha, \quad (1.7)$$

Note that the first term on the right-hand side of (1.6) is the only instance where the \dot{s}^b occur. We can eliminate this through the constraints (1.7), and we shall henceforth denote with a subscript c any expression for which we have eliminated the fiber dependency by using (1.7).

Equations (1.6) and (1.7) then describe the mechanics of the nonholonomic system (L, \mathcal{D}) , and for ease of use later, we henceforth define $\phi^a := \dot{s}^a + A_\alpha^a(r, s) \dot{r}^\alpha$, and

$$\Lambda_{\alpha\beta} := \langle \mathbb{F}L, B(\partial_{r^\alpha}, \partial_{r^\beta}) \rangle. \quad (1.8)$$

2. Variational nonholonomic dynamics

Consider now 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 $\mu_1(t), \dots, \mu_k(t)$. We shall call the $\mu_a(t)$ the *multipliers*, and form the augmented Lagrangian $L_V : T\tilde{Q} \rightarrow \mathbb{R}$:

$$L_V = L - \mu_a \phi^a. \quad (2.1)$$

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

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^J} - \frac{\partial L}{\partial q^J} = \dot{\mu}_a \frac{\partial \phi^a}{\partial \dot{q}^J} + \mu_a \left(\frac{d}{dt} \frac{\partial \phi^a}{\partial \dot{q}^J} - \frac{\partial \phi^a}{\partial q^J} \right), \quad (2.2)$$

where $J = 1, \dots, n$, as well as the equations of constraint (1.7), which arise as the Euler–Lagrange equations of the μ_a coordinates. This set of equations is sometimes called the *vakonomic equations*, after [1] Arnold, Kozlov and Neishtadt who introduced nonholonomic mechanics under the Lagrange variational point of view. We shall thus prefer to call these equations the *variational nonholonomic equations*.

Now, by considering the r and s variables separately, and writing E_I for the Euler–Lagrange operator in the I th coordinate, we can rewrite (2.2) as

$$E_\alpha(L) = \dot{\mu}_a A_\alpha^a + \mu_a \left(\frac{d}{dt} (A_\alpha^a) - \frac{\partial A_\beta^a}{\partial r^\alpha} \dot{r}^\beta \right), \quad (2.3)$$

$$E_a(L) = \dot{\mu}_a - \mu_b \frac{\partial A_\alpha^b}{\partial s^a} \dot{r}^\alpha, \quad (2.4)$$

and substituting μ_a from (2.4) into (2.3), we get

$$E_\alpha(L) = A_\alpha^a E_a(L) + \mu_a B_{\alpha\beta}^a \dot{r}^\beta. \tag{2.5}$$

One can then show [3] that by the definition of the constrained Lagrangian, (2.5) can be re-written in the more suggestive form

$$E_\alpha(L_c) = \left(\mu_a - \frac{\partial L}{\partial \dot{s}^a} \right) B_{\alpha\beta}^a \dot{r}^\beta - A_\alpha^a \frac{\partial L_c}{\partial s^a}. \tag{2.6}$$

These along with the constraint equations (1.7) form the equations of motion for the variational nonholonomic system, and a simple comparison of (2.6) and (1.6) 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 (2.3)–(2.4) (see [11]). However, we shall see in section 3.3 that equation (2.4) will lead to an *a priori* determination of the multipliers $\mu_a(t)$.

3. 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 (1.6) and (2.6) 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 1. Consider the nonholonomic system (1.6), (1.7) with initial conditions (q_0, \dot{q}_0) , and the associated variational nonholonomic system (2.6), (1.7) with initial conditions (q_0, \dot{q}_0, μ_0) , where (q_0, \dot{q}_0) satisfies the constraints (1.7). 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 (1.6), (1.7) with initial condition (q_0, \dot{q}_0) is the same as the solution to (2.6), (1.7) with initial condition (q_0, \dot{q}_0, μ_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 (1.6), (1.7) with initial condition (q_0, \dot{q}_0) is the same as the solution to (2.6), (1.7) with initial condition (q_0, \dot{q}_0, μ_0) .

Remark 3.0.1: A nonholonomic system which is conditionally variational can thus be seen as a variational nonholonomic system with Lagrangian (2.1) and initial condition μ_0 . We shall say more about the specifics of how to choose μ_0 in proposition 3 below.

3.1. The equivalence conditions

As we have seen, variational nonholonomic systems arise from Hamilton’s principle of stationary action. Moreover, we have also noted that nonholonomic mechanics cannot be derived from such a principle, which has led us to introduce the idea of the system being

conditionally variational. The actual conditions under which the nonholonomic system would be conditionally variational were originally stated in [28]. There the author shows that the necessary and sufficient conditions for the equivalence between the two formalisms, in the notation adopted here, are

$$\mu_a E_J(\phi^a) \delta q^J = 0, \quad J = 1, \dots, n, \quad a = 1, \dots, k. \quad (3.1)$$

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 2, 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 (1.7) to relate (3.1) to the curvature (1.4) of the Ehresmann connection A

$$E_J(\phi^a) dq^J = B_{\alpha\beta}^a \dot{r}^\beta dr^\alpha = d\omega^a(\dot{q}, \cdot). \quad (3.2)$$

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

$$\mu_a E_J(\phi^a) \delta q^J = \mu_a d\omega^a(\dot{q}, \delta q). \quad (3.3)$$

Comparing the right-hand side of the preceding with the vertical-vector-valued-two-form definition of curvature given in (1.4), we arrive at the geometric necessary and sufficient conditions for a nonholonomic system to be conditionally variational (see also [13]).

Proposition 2. *The nonholonomic system (1.6), (1.7) is conditionally variational with Lagrangian (2.1) if and only if*

$$\langle \mu, B(\dot{q}, \delta q) \rangle = 0. \quad (3.4)$$

Remark 3.1.1: Since the curvature is vertical-vector-valued, proposition 2 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 ($n - k = 2$) with k constraints, (3.4) reads $\mu_a B_{n-1,n}^a \dot{q}^2 = -\mu_a B_{n-1,n}^a \dot{q}^1 = 0$, and in this special case these two conditions are satisfied if and only if $\Lambda := \mu_a B_{n-1,n}^a = 0$.

We shall call the conditions (3.4) the *equivalence conditions*, as originally named in [28]. Proposition 2 then confirms that the extra term in (2.6) must vanish in order for the nonholonomic system (1.6) to be conditionally variational. Moreover, the computations above show how to obtain the conditions starting from the analytical mechanics viewpoint of [28], where the equivalence conditions first appeared.

Verifying (3.4) 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. Our contributions remove this need in special cases, in addition to showing how to verify equivalence simply by inspection of the nonholonomic system.

3.2. Chaplygin systems

We now specialize to the so-called *Chaplygin* nonholonomic systems, partly motivated by the fact that most of the known physical examples of nonholonomic systems can be cast [3] as Chaplygin systems (for example, the vertical rolling disc, the Chaplygin sleigh and the bowling ball), but also because these types of systems possess many interesting properties which will simplify our investigations.

Geometrically, a *Chaplygin* nonholonomic system [3, 10] is a triple (L, \mathcal{D}, G) where $L : TQ \rightarrow \mathbb{R}$ is a G -invariant regular Lagrangian with respect to the lifted action on TQ , and

the nonholonomic constraints are determined by the horizontal distribution \mathcal{D} of a principal connection \mathcal{A} on the principal G -bundle $\pi : Q \rightarrow Q/G$ associated with a free and proper action Ψ of G on Q . Such a definition is referred to by Koiller [19] as a *non-Abelian Chaplygin system*. The *Abelian Chaplygin* case we shall take to be when the configuration space Q is a product of copies of S^1 and \mathbb{R} , with the connection coefficients in (1.7) depending only on the r variables, and the Lagrangian L of section 1 being cyclic in the s variables. As we shall see below these additional simplifications will allow us to re-interpret the conditions (3.4) in a much more natural way.

3.3. Abelian Chaplygin systems

Suppose now that we are considering an Abelian Chaplygin nonholonomic system. The curvature (1.4) then simplifies to just the difference of the first two terms, and the last term on the right-hand side of (1.6) also vanishes. Thus, the equations of motion (1.6) reduce to the simpler form

$$E_\alpha(L_c) = F_{\gamma\beta\alpha}\dot{r}^\gamma\dot{r}^\beta, \tag{3.5}$$

where the $F_{\gamma\beta\alpha}$ are the components of $\Lambda_{\alpha\beta}$ from the end of section 1, i.e. $\Lambda_{\alpha\beta} = (g_{b\gamma} - g_{ba}A_\gamma^a)B_{\alpha\beta}^b\dot{r}^\gamma =: F_{\gamma\alpha\beta}\dot{r}^\gamma$. In this modified form, it becomes easier to state the first main result.

Proposition 3. *Suppose the nonholonomic system (1.6), (1.7) is (Abelian) Chaplygin. Then we have the following:*

(1) *The variational nonholonomic multipliers are given by:*

$$\mu_a = \left(\frac{\partial L}{\partial \dot{s}^a} \right)_c + C_a, \quad a = 1, \dots, k, \tag{3.6}$$

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 (1.6) 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 (3.4) reduces to the following conditions:*

$$F_{\alpha\alpha\beta} = 0 \quad \forall \alpha \neq \beta, \tag{3.7}$$

$$F_{\alpha\beta\gamma} + F_{\beta\alpha\gamma} = 0 \quad \text{for each } \gamma, \forall \alpha < \beta, \quad \alpha, \beta \neq \gamma. \tag{3.8}$$

Consider now a general nonholonomic system (not necessarily Chaplygin) as in section 1. 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 $\partial L / \partial \dot{s}^a = 0$ by the Abelian Chaplygin assumption. Then by (2.4) 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 (3.5) 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 (3.4) after taking into account (3.6). Conversely, suppose that the system is conditionally variational. Then by (3.4) the constrained equations (3.5) 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 (3.4). 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 (3.5) 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 (3.4) reads $\mu B_{n-1,n} = 0$ (taking into account remark 3.1.1). By assumption, the system is nonholonomic, meaning that it is not variational, so that μ is nonzero. Thus, the condition reduces to $B_{n-1,n} = 0$, which means that the system must actually be holonomic, in contradiction. \square

Remarks:

- (1) Part (5) applies in the Chaplygin case as well. As such, it prevents some common nonholonomic systems from being conditionally variational simply by inspection: the knife edge, the Chaplygin sleigh, the Heisenberg free particle, and the Euler–Poincare–Suslov system on $SO(3)$ (for details see [3]) 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 (3.6) into (3.4). We see at once that one needs $\mu_a(0) = (\partial L / \partial \dot{s}^a)_c(0) \forall a$ such that $B_{\alpha\beta}^a \neq 0$, and $\mu_a(0)$ may be chosen arbitrarily for each a such that $B_{\alpha\beta}^a = 0 \forall \alpha, \beta$.
- (3) The previous remark, along with conditions (3.4), now allows us to characterize conditionally variational systems in terms of the $\Lambda_{\alpha\beta}$ from (1.8). 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 3.1.1, which is simpler to handle. Moreover, for a three degree of freedom system it leaves nine conditions in (4.3) to be checked.
- (6) Although not part of proposition 3, 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 $B_{\alpha\beta}^b = 0 \forall \alpha, \beta$, then the new system will still be conditionally variational. We shall use this observation in analyzing the two-wheeled carriage in section 5.

Proposition 3 is the first main result which, for Abelian Chaplygin systems, completely characterizes when such systems are conditionally variational. We shall make use of the proposition in section 5 to identify some conditionally equivalent systems by example, but also use the proposition as the backbone in section 4 for the analysis of what relationship exists between the nonholonomic system’s invariant measure (if it has one) and its being conditionally variational.

3.4. Non-Abelian Chaplygin systems

In the non-Abelian Chaplygin case, the structure constants of the Lie algebra of G are no longer all zero. They enter into the curvature, now regarded as a Lie algebra valued two-form, as described in [3, 10]. The Lagrangian L induces the reduced lagrangian $l : TQ/G \rightarrow \mathbb{R}$,

which after substitution of the constraints (1.7) gives us the constrained reduced Lagrangian l_c . By choosing a local trivialization $Q = Q/G \times G$ with coordinates (r^α, g^a) the equations of motion and constraints are [3]

$$E_\alpha(l_c) = \left\langle \frac{\partial l}{\partial \xi}, \mathcal{B} \right\rangle_\alpha, \tag{3.9}$$

$$\xi = \mathcal{A}(r)\dot{q}, \tag{3.10}$$

where $\xi = g^{-1}\dot{g}$, with $g \in G$, and \mathcal{A} is the unique principal connection whose horizontal space is the distribution \mathcal{D} with curvature \mathcal{B} .

Now, to derive the analog of proposition 4 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 [8, 24]

$$\frac{d}{dt} \left(\frac{\partial l}{\partial \xi} - \mu \right) = \text{ad}_\xi^* \left(\frac{\partial l}{\partial \xi} - \mu \right), \quad E_\alpha(l_c) = \left\langle \mu - \frac{\partial l}{\partial \xi}, \mathcal{B} \right\rangle_\alpha, \tag{3.11}$$

along with (3.10).

We can now compare these equations with (2.6) and write the results of proposition 3 in this context. However, due to the non-Abelian character of the system, some aspects of proposition 3 no longer hold.

Corollary 4. *Suppose that we have a non-Abelian Chaplygin system given by (3.9)–(3.10) for which the right-hand side of (3.9) vanishes. Then a solution to (3.9), (3.10) with initial condition (q_0, ξ_0) is also a solution to the variational nonholonomic system (3.11) with*

$$\mu_b = \frac{\partial l}{\partial \xi} + C_b, \tag{3.12}$$

and initial condition (q_0, ξ_0, μ_0) , where μ_0 and C_b are subject to remark 2 of section 3.3. Moreover, under this condition part (3) of proposition 3 holds when the added potential is independent of ξ .

Proof. Suppose that the right-hand side of (3.9) vanishes, and consider a solution to the non-Abelian system (3.9), (3.10) with initial condition (q_0, ξ_0) . Then the condition (3.4) with μ_b chosen as above, taking into account remark 2 of section 3.3, is satisfied. Thus, the solution to (3.9) is also a solution to (3.11) with μ as in (3.12) and μ_0 and C_b subject to remark 2 of section 3.3. Moreover, the last part of the corollary follows by again observing that (3.4) is independent of the potential so long as the potential is independent of ξ . \square

Clearly non-Abelian Chaplygin systems are more complicated than their Abelian counterparts. For example, part (1) of proposition 3 no longer applies, and here the variational multipliers are not given by the fiber derivative *a priori*. However, corollary 4 provides one with an alternative to solving the variational nonholonomic non-Abelian problem to check for equivalence. The corollary shows that if one can find a non-Abelian 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 (3.11) with initial conditions (q_0, ξ_0, μ_0) , taking into account remark 1 of section 3.3 and (3.12).

3.5. 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. 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 can 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 3.

Suppose then that we again have an Abelian Chaplygin nonholonomic system with regular mechanical Lagrangian, as in section 1, and consider the variational Lagrangian (2.1)

$$L_V = L - \frac{\partial L}{\partial \dot{s}^a} \phi^a. \tag{3.13}$$

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

$$\begin{aligned} 0 = E_\alpha(L_V) &= E_\alpha(L) + \frac{\partial L}{\partial \dot{s}^a} B_{\beta\alpha}^a \dot{r}^\beta - A_\alpha^a \left(\frac{\partial L}{\partial \dot{s}^a} \right)' + \left[\frac{\partial^2 L}{\partial r^\alpha \partial \dot{s}^a} \phi^a - (g_{\alpha a} \phi^a)' \right], \\ \implies E_\alpha(L_V) &= (g_{\alpha a} \phi^a)' - \frac{\partial^2 L}{\partial r^\alpha \partial \dot{s}^a} \phi^a, \end{aligned} \tag{3.14}$$

$$0 = E_a(L_V) = (g_{ab} \phi^b)', \quad \implies (g_{ab} \phi^b)(t) = g_{ab}(0) \phi^b(0), \tag{3.15}$$

where we remind the reader that all lower case Roman indices range from 1 through k , all Greek indices range from $k + 1$ through n , and all capital Roman indices range from 1 through n , and where g^{IJ} is the inverse of the assumed Riemannian metric g on Q from section 1.

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 right-hand side of (3.14) vanishes, then the constrained Euler–Lagrange equations are Lagrangian, and by proposition 3, part (2) we would then know that the Abelian nonholonomic Chaplygin system under consideration is actually conditionally variational with Lagrangian given by (3.13). Moreover, we can easily compute the Hessian of L_V to be the matrix

$$Hess(L_V) = g_{IJ}^V = \begin{pmatrix} g_{\alpha\beta} - g_{\alpha a} A_\beta^a - g_{\beta a} A_\alpha^a & -A_\alpha^a g_{ab} \\ -g_{ac} A_\alpha^c & g_{ab} \end{pmatrix}, \tag{3.16}$$

which is not automatically singular, as was the case for the $L(q, \dot{q}, \mu; t)$ Lagrangian (2.1). We can thus summarize the preceding to give:

Proposition 5. *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 (3.13) and with the initial data satisfying the constraints (1.7). Moreover, (3.13) will be regular if and only if (3.16) is invertible.*

Before discussing proposition 5, we should mention that its origins are rooted in momentum conservation. Assuming the Hessian (3.16) is invertible we can effect the Legendre transform and define $H : T^*Q \rightarrow \mathbb{R}$ as usual

$$H_V(p_I, q^I) = p_I \dot{q}^I - L_V. \tag{3.17}$$

We can then examine the momenta conjugate to the s variables

$$p_a := (\mathbb{F}L_V)_a = \frac{\partial L_V}{\partial \dot{s}^a} = g_{ab} \phi^b, \tag{3.18}$$

and from here it is clear that (3.15) is nothing but a statement of the conservation of momenta. However, proposition 5 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.

The main accomplishment of this proposition is to enable the description of conditionally variational nonholonomic systems without the use of the variational multipliers, hence possibly regaining regularity in the instances in which (3.16) is invertible. In fact, one such instance where (3.16) is nonsingular is the vertical rolling disc (see section 6).

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, (3.18) 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 (3.17) 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 disc (see, section 6.1), and is also true for some other systems, though by a different method (see [14]).

4. Conditionally variational systems, Hamiltonization and invariant measures

Two main differences between nonholonomic and variational nonholonomic systems are their derivability from stationary action principles, and the question of existence of invariant measures. Nonholonomic systems are in general not derivable from a stationary action principle, and also do not, in general, possess invariant measures. Variational nonholonomic dynamics, on the other hand, can be derived from Hamilton’s principle using the augmented Lagrangian (2.1) as we have done. Moreover, because the resulting system is Hamiltonian, it naturally preserves any nonzero constant multiple of the associated standard measure. However, intuitively nonholonomic systems which *do* possess invariant measures are in some sense closer to Hamiltonian systems (for which we have Liouville’s theorem), 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 (in the sense defined below) the nonholonomic system.

4.1. Abelian Chaplygin systems with invariant measures

Consider the nonholonomic system (L, \mathcal{D}) from section 1 and assume it is of Abelian Chaplygin type. We will need the corresponding constrained Hamilton equations in what follows so we will review them here [3].

By assumption L is regular, so we may pass to the cotangent bundle T^*Q via the Legendre transform. One can then write down the constrained Hamilton equations on the constrained phase space $\mathcal{M} := \mathbb{F}L(\mathcal{D}) \subset T^*Q$

$$\dot{r}^\alpha = \frac{\partial H_c}{\partial \tilde{p}_\alpha}, \quad \alpha = k + 1, \dots, n, \tag{4.1}$$

$$\dot{\tilde{p}}_\alpha = -\frac{\partial H_c}{\partial r^\alpha} - \Lambda_{\alpha\beta} \frac{\partial H_c}{\partial \tilde{p}_\beta}, \quad \alpha, \beta = k + 1, \dots, n, \tag{4.2}$$

$$\dot{s}^a = -A_\beta^a \frac{\partial H_c}{\partial \tilde{p}_\beta}, \quad a = 1, \dots, k, \tag{4.3}$$

where $H_c = \tilde{p}_\alpha \dot{r}^\alpha - L_c$ is the constrained Hamiltonian, with $\tilde{p}_\alpha = \partial L_c / \partial \dot{r}^\alpha$ the constrained momenta. Here $\Lambda_{\alpha\beta}$ from (1.8) is now written in terms of \tilde{p} as

$$\Lambda_{\alpha\beta} = (g_{b\gamma} - g_{ba} A_\gamma^a) B_{\alpha\beta}^b \partial H_c / \partial \tilde{p}_\gamma = F_{\gamma\alpha\beta} \partial H_c / \partial \tilde{p}_\gamma =: F_{\gamma\alpha\beta} K_\gamma^\delta \tilde{p}_\delta =: C_{\alpha\beta}^\delta \tilde{p}_\delta, \quad (4.4)$$

where we have used the definition of $F_{\gamma\alpha\beta}$ from (3.5) in the second equality. These are a set of $2n - m$ equations on the submanifold \mathcal{M} with induced coordinates $(r^\alpha, s^a, \tilde{p}_\alpha)$, where (4.1)–(4.2) are the Hamiltonian analog of (3.5), and (4.3) represents the constraints (1.7).

We now consider the case when the nonholonomic system possesses an invariant measure. The conditions for the existence of this invariant measure are well studied [10, 22, 30], but we shall just assume here that the system already has an invariant measure. If we denote by X the vector field which solves (4.1)–(4.2) and further assume that the system has an invariant measure $N(r) dr^\alpha \wedge d\tilde{p}_\alpha = N(r)\omega^\alpha$, where $\alpha = k + 1, \dots, n$, and we denote the standard measure on \mathcal{M} by ω^α , then by definition $\mathcal{L}_X(N\omega^k) = 0$, where \mathcal{L} denotes the Lie derivative. From this, we have

$$0 = \text{div}_{N\omega^k}(X) = \text{div}_{\omega^k}(X) + \frac{1}{N} X(N) = \frac{\partial \dot{r}^\alpha}{\partial r^\alpha} + \frac{\partial \dot{\tilde{p}}_\alpha}{\partial \tilde{p}_\alpha} + \frac{\dot{r}^\alpha}{N} \frac{\partial N}{\partial r^\alpha}. \quad (4.5)$$

Using (4.1) and (4.2), we have

$$\left(\frac{1}{N} \frac{\partial N}{\partial r^\beta} - \frac{\partial \Lambda_{\alpha\beta}}{\partial \tilde{p}_\alpha} \right) \dot{r}^\beta = 0, \quad (4.6)$$

and from (4.4) we see explicitly that Λ 's depend linearly on the momenta, so that the quantities in parentheses in (4.6) depend only on the coordinates. Thus, the only way (4.6) vanishes is if the parenthetical terms vanish identically

$$\frac{1}{N} \frac{\partial N}{\partial r^\beta} - \frac{\partial \Lambda_{\alpha\beta}}{\partial \tilde{p}_\alpha} = 0, \quad (4.7)$$

$$\implies \frac{1}{N} \frac{\partial N}{\partial r^\beta} - C_{\alpha\beta}^\alpha = 0, \quad \alpha, \beta = k + 1, \dots, n. \quad (4.8)$$

Equations (4.8) and (4.4) describe explicitly how the measure density, metric and curvature arising from the nonholonomic constraints interact. More importantly though, it gives an explicit relationship between the invariant measure density N and $\Lambda_{\alpha\beta}$, which are directly related to the property of being conditionally variational (see remark 3 of section 3.3).

4.2. Conditionally variational systems and invariant measures

Recall that remark 3 from section 3.3 tells us when a system is conditionally variational, and it is clear from that remark and (4.7) that only nonholonomic systems with constant measure densities can hope to be conditionally variational. Although this restricts the set of possible systems, the vertical rolling disc [3, 22, 30] possesses a constant density invariant measure, and we shall make use of this in section 5.1 below.

We now wish to rephrase remark 3 from section 3.3 by instead asking under what conditions the nonholonomic Abelian Chaplygin system is conditionally variational if it is known to possess an invariant measure N with coordinate-only dependence? The answer is simple for two-degree of freedom systems, but becomes increasingly complicated with the number of degrees of freedom, and is given in the following proposition:

Proposition 6. *Suppose now that the constrained nonholonomic Abelian Chaplygin system (4.1)–(4.2) has an invariant measure with constant density N . Then the system satisfies*

$$C_{\alpha\beta}^\alpha = 0. \quad (4.9)$$

In addition, if the systems has two constraints ($k = 2$ in (4.3)), and

- (a) Two degrees of freedom, it is also conditionally variational.
- (b) Three degrees of freedom, if $K_4^5, K_3^5, K_3^4 \neq 0$ and the system satisfies the conditions (3.7), which in this case become

$$F_{334} = F_{335} = F_{443} = F_{445} = F_{553} = F_{554} = 0, \tag{4.10}$$

then it is also conditionally variational. For each of the aforementioned K_γ^α which are zero we must add to (4.10) the condition $F_{\gamma\alpha\beta} + F_{\alpha\gamma\beta} = 0$, for $\gamma \neq \alpha \neq \beta$. Also, for each K_α^α which is zero, we may omit the conditions $F_{\gamma\gamma\beta}, \gamma \neq \beta$ from (4.10).

Proof. First, we have from (4.8) that having an invariant measure of constant density gives

$$C_{\alpha\beta}^\alpha = 0. \tag{4.11}$$

For the two degree of freedom case of part (a), (4.8) becomes (writing $\Lambda := \Lambda_{34}$)

$$\frac{1}{N} \frac{\partial N}{\partial r^3} + \frac{\partial \Lambda}{\partial \tilde{p}_4} = 0, \quad \frac{1}{N} \frac{\partial N}{\partial r^4} - \frac{\partial \Lambda}{\partial \tilde{p}_3} = 0. \tag{4.12}$$

However, a straightforward computation using (4.4) shows 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

$$\Lambda = \frac{\partial \Lambda}{\partial \tilde{p}_3} \tilde{p}_3 + \frac{\partial \Lambda}{\partial \tilde{p}_4} \tilde{p}_4, \quad \implies \Lambda = \frac{1}{N} \left(\tilde{p}_3 \frac{\partial N}{\partial r^4} - \tilde{p}_4 \frac{\partial N}{\partial r^3} \right), \tag{4.13}$$

where the last line follows by (4.12). By assumption of constant N (4.13) then shows that Λ vanishes, which by remark 1 of section 3.3 implies that the system is conditionally variational.

For part (b), note that there are $n - k$ equations in (4.11), and $\frac{(n-k)^2}{2}(n - k - 3) + (n - k)$ equations in (3.8). Thus, for $n - k = 3$ (the three degree of freedom case) these two give the same number of equations, and we can write out (4.11) by using the symmetry of \tilde{g}

$$K_4^5 (F_{453} + F_{543}) = - [K_3^4 F_{343} + K_3^5 F_{353} + K_4^4 F_{443} + K_5^5 F_{553}], \tag{4.14}$$

$$K_3^5 (F_{354} + F_{534}) = - [K_3^3 F_{334} + K_4^3 F_{434} + K_4^5 F_{454} + K_5^5 F_{554}], \tag{4.15}$$

$$K_3^4 (F_{343} + F_{435}) = - [K_3^3 F_{335} + K_4^4 F_{445} + K_5^3 F_{535} + K_5^4 F_{545}]. \tag{4.16}$$

The parenthetical quantities are precisely (3.8) for the three degree of freedom case, and vanish precisely under the assumptions (4.10), by using $F_{\gamma\alpha\beta} = -F_{\gamma\beta\alpha}$. Moreover, it is clear that for each K_α^γ in (4.14)–(4.16) which is zero, we must add the extra conditions as stated in part (b). □

Remark 4.2.1: A similar result to part (b) holds for the greater than three degree of freedom case, and also for the $k > 2$ case, by using the same argument starting from (4.11), but is not very computationally useful.

4.3. Hamiltonization and conditionally variational systems

Another method which has been used with great success to find an almost-Hamiltonian structure for nonholonomic systems is called *Hamiltonization*, and dates back to Chaplygin [9]. This procedure historically grew out of Chaplygin’s *reducing multiplier theorem* [6], which defines a new time parameterization τ by $d\tau = N(r) dt$, where $N(r)$ is the density of the assumed existent invariant measure of the Abelian Chaplygin nonholonomic system, and states that for two degree of freedom systems, the constrained mechanics (1.6) become Lagrangian in τ -time (see [6] and [12] for more details).

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 Hamiltonization.

Proposition 7. *Suppose that the constrained Abelian Chaplygin nonholonomic system (1.6), (1.7) has two degrees of freedom. Then if the system is conditionally variational, it is also Hamiltonizable. Also, if the system is 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 [15]) and part (a) of proposition 6. \square

Remark 4.3.1: proposition 7 gives a 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 [6] 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 7.

5. Examples: Conditionally variational nonholonomic systems

5.1. The vertical rolling disk

Consider the nonholonomic vertical rolling disk (see [3]) pictured in figure 1 below 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 disc, θ is the angle that a point fixed on the disc makes with respect to the vertical, and φ is measured from the positive x -axis. This system has Lagrangian and constraints given by

$$\begin{aligned} L &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2, \\ \phi^1 &= \dot{x} - R \cos \varphi \dot{\theta} = 0, \\ \phi^2 &= \dot{y} - R \sin \varphi \dot{\theta} = 0, \end{aligned} \tag{5.1}$$

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

This is an Abelian Chaplygin system with $(s^1, s^2, r_1^\alpha, r_2^\alpha) = (x, y, \theta, \varphi)$ and as such proposition 3 gives

$$\begin{aligned} \Lambda_{34} &= (-mR\dot{x} \sin \varphi + mR\dot{y} \cos \varphi) \\ &= mR(-(R \cos \varphi \dot{\theta}) \sin \varphi + (R \sin \varphi \dot{\theta}) \cos \varphi) = 0 \end{aligned} \tag{5.2}$$

and by remark 3.1.1 this shows that the system is conditionally variational. Moreover, the variational nonholonomic Lagrangian (3.13) is given by

$$\begin{aligned} L_V &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2 - m\dot{x}(x - R \cos \varphi \dot{\theta}) - m\dot{y}(y - R \sin \varphi \dot{\theta}) \\ &= -\frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2 + mR\dot{\theta}(x \cos \varphi + y \sin \varphi). \end{aligned} \tag{5.3}$$

The results of proposition 5 apply here, and thus one could apply the Euler–Lagrange equations to L_V , and by imposing the constraints (5.1) only initially, recover the full nonholonomic dynamics. It is also worth noting that a straightforward computation of (3.16)

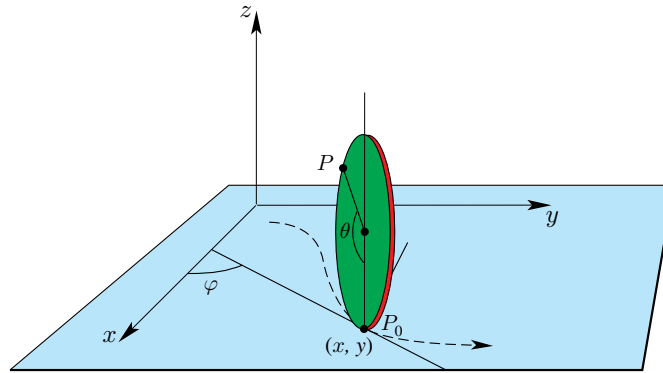


Figure 1. The vertical penny.
(This figure is in colour only in the electronic version)

shows that L_V is regular, unlike in the variational approach wherein we have the extra μ variables, and there 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 (3.18)

$$\begin{aligned} p_x &= -m\dot{x} + mR \cos \varphi \dot{\theta}, & p_y &= -m\dot{y} + mR \sin \varphi \dot{\theta}, \\ p_\theta &= I\dot{\theta} + mR (\dot{x} \cos \varphi + \dot{y} \sin \varphi), & p_\varphi &= J\dot{\varphi}. \end{aligned} \tag{5.4}$$

With these momenta, the constraints (5.1) are simply $p_x = p_y = 0$, and the Hamiltonian becomes

$$H_V = \frac{p_\varphi^2}{2J} + \frac{1}{2m\beta} [m^2 p_\theta^2 - (a^2 \sin^2 \varphi + Im) p_x^2 - (a^2 \cos^2 \varphi + Im) p_y^2 + (a^2 \sin 2\varphi) p_x p_y], \tag{5.5}$$

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.

5.2. The mobile robot with fixed orientation and the two-wheeled carriage

These two examples are variations of the vertical disc. First, consider a mobile robot whose body maintains a fixed orientation with respect to its environment [10]. It has three wheels which roll without slipping, with the position of the center of mass at (x, y) , the steering angle θ , and the rotation angle ψ of the wheels. The system's Lagrangian and constraints are given by

$$\begin{aligned} L &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}J\dot{\theta}^2 + \frac{3}{2}J_w\dot{\psi}^2, \\ 0 &= \dot{x} \sin \theta - \dot{y} \cos \theta, \\ R\dot{\psi} &= \dot{x} \cos \theta + \dot{y} \sin \theta. \end{aligned} \tag{5.6}$$

where again R denotes the common radius of the three wheels. However, after rewriting the constraints as

$$\dot{x} - R \cos \theta \dot{\psi} = 0, \quad \dot{y} - R \sin \theta \dot{\psi} = 0, \tag{5.7}$$

we see that this system is Abelian Chaplygin, and equivalent to the vertical rolling disc under the identification $J \rightarrow 3J_w$. The calculation showing that condition (3.4) is satisfied is in this case then identical to (5.2).

Another similar example found in the literature [13, 19] is the two-wheeled carriage. This system consists of a body with two wheels which roll without slipping. The position of the center of mass of the ensemble is again denoted (x, y) , and the rotation angles of the two wheels are denoted by ϕ_1 and ϕ_2 , with distance between the two wheels denoted by $2w$. Moreover, denote by m the total mass of the system (where we assume both wheels have equal mass), I the axial moment of inertia of both wheels, J the moment of inertia about a diameter of both wheels, and R the common radius of both wheels. The system's Lagrangian and constraints can be written as [13, 19]

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}J\dot{\theta}^2 + \frac{1}{2}I(\dot{\psi}_1^2 + \dot{\psi}_2^2), \tag{5.8}$$

$$0 = \dot{x} \sin \theta - \dot{y} \cos \theta, \quad R\dot{\psi}_1 = \dot{x} \cos \theta + \dot{y} \sin \theta, \quad R\dot{\psi}_2 = w\dot{\theta}, \tag{5.9}$$

where we have adopted the definitions $\psi_1 = (1/2)(\phi_1 + \phi_2)$ and $\psi_2 = (1/2)(\phi_1 - \phi_2)$ from [13].

Comparing the system with (5.6), we can view (5.8) as an extension of (5.6) to one more degree of freedom (with the associated constraint (5.9)). Moreover since (5.9) is actually a holonomic constraint, remark 6 from section 3.3 applies and this system is therefore conditionally variational as well.

5.3. Veselova's system

Veselova's system describes [6] the motion of a rigid body with a fixed point subject to the nonholonomic constraint $(\omega, \gamma) = 0$, where ω and γ are the body's angular velocity vector and unit vector of the space-fixed axis in the frame of reference fixed to the body, respectively. By using Euler angles, we can write these vectors as

$$\omega = (\dot{\psi} \sin \theta \sin \varphi + \dot{\theta} \cos \varphi, \dot{\psi} \sin \theta \cos \varphi - \dot{\theta} \sin \varphi, \dot{\psi} \cos \theta + \dot{\varphi}), \tag{5.10}$$

$$\gamma = (\sin \theta \sin \varphi, \sin \theta \cos \varphi, \cos \theta), \tag{5.11}$$

and the equation of constraint is

$$\phi^1 = \dot{\psi} + \cos \theta \dot{\varphi} = 0. \tag{5.12}$$

It is shown in [6] that this system possesses an invariant measure with density $N = (\gamma, \mathbf{I}\gamma)^{-1/2}$, where \mathbf{I} is the moment of inertia matrix. We may now use remark 4.3.1 to restrict to a case when the system might be conditionally variational. To that end, restrict to the case where the moments of inertia are all equal, so that the measure density reduces to a constant ($= I^{-1/2}$). The Lagrangian then becomes $L = \frac{1}{2}(\omega_1^2 + \omega_2^2 + \omega_3^2)$, and the system is Abelian Chaplygin. Hence the variational nonholonomic multiplier is given by proposition 3

$$\begin{aligned} \mu &= -\frac{\partial L}{\partial \dot{\psi}} \\ &= -I(\omega_1 \sin \theta \sin \varphi + \omega_2 \sin \theta \cos \varphi + \omega_3 \cos \theta) \\ &= -I(\dot{\psi} + \cos \theta \dot{\varphi}), \end{aligned} \tag{5.13}$$

and clearly by using the constraint (5.12) we see that $\mu = 0$, which is a special case of the condition (3.4).

Note that in this case the system was conditionally variational with vanishing variational multiplier, which in turn occurs because the nonholonomic constraint (5.12) in the equal moment of inertia case is actually an integral of the motion.

6. Examples: partially conditionally variational systems

Having introduced the idea of conditionally variational nonholonomic systems, we have now seen that such a situation is rare, owing to the restriction (3.4) places on such systems. The more common case which is still of interest is when only some (and not all) nonholonomic trajectories can be seen as variational ones. We present several of these examples here.

6.1. The nonholonomic free particle

The nonholonomic free particle is a prime example of a nonholonomic Chaplygin system which, although it has an invariant measure, is not conditionally variational (which is most immediately seen from part (5) of proposition 3). However, we may still view some of the nonholonomic solutions as variational ones.

The system 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 [3]

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \quad \dot{z} = -x\dot{y}. \tag{6.1}$$

As it stands the system is Abelian Chaplygin with only one constraint, and we may apply the results of proposition 3 to see that the variational multiplier is given by $\mu = m\dot{z}$, choosing $C = 0$ (see remark 2 of section 3.3). Since this is not identically zero, the system is not conditionally variational. However, in this case we have, by remark 3.1.1, that $\Lambda = -m\dot{z}$, and hence the constrained two degree of freedom system (1.6) has the invariant measure density $N = \frac{1}{\sqrt{1+x^2}}$, as can be verified directly through (4.12). This again verifies that the system is not conditionally variational, since N is non-constant.

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 above).

6.2. The Chaplygin sphere

The Chaplygin sphere is a sphere rolling without slipping on a horizontal plane (see [3]) 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

$$L = \frac{I_1}{2}(\dot{\theta} \cos \psi + \dot{\varphi} \sin \psi \sin \theta)^2 + \frac{I_2}{2}(-\dot{\theta} \sin \psi + \dot{\varphi} \cos \psi \sin \theta)^2 + \frac{I_3}{2}(\dot{\psi} + \dot{\varphi} \cos \theta)^2 + \frac{m}{2}(\dot{x}^2 + \dot{y}^2), \tag{6.2}$$

$$\phi^1 = \dot{x} - \dot{\theta} \sin \varphi + \dot{\psi} \cos \varphi \sin \theta = 0,$$

$$\phi^2 = \dot{y} + \dot{\theta} \cos \varphi + \dot{\psi} \sin \varphi \sin \theta = 0.$$

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 have an Abelian Chaplygin system. The system also has an invariant measure whose density is in general non-constant (see [6]). 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 (4.10). However, a quick computation shows that $F_{\psi\theta\psi} = \sin \theta \cos \theta$, which fails to satisfy (4.10).

There are still nonholonomic solutions that can be seen as variational ones though. In this case, the only nonzero components of the Ehresmann connection are

$$\begin{aligned}
 b_\theta^x &= -\sin \varphi, & b_\psi^x &= \cos \varphi \sin \theta, \\
 \frac{\partial b_\theta^x}{\partial \varphi} &= -\cos \varphi, & \frac{\partial b_\psi^x}{\partial \theta} &= \cos \varphi \cos \theta, & \frac{\partial b_\psi^x}{\partial \varphi} &= -\sin \varphi \sin \theta, \\
 b_\theta^y &= \cos \varphi, & b_\psi^y &= \sin \varphi \sin \theta, \\
 \frac{\partial b_\theta^y}{\partial \varphi} &= -\sin \varphi, & \frac{\partial b_\psi^y}{\partial \theta} &= \sin \varphi \cos \theta, & \frac{\partial b_\psi^y}{\partial \varphi} &= \cos \varphi \sin \theta.
 \end{aligned}
 \tag{6.3}$$

Then condition (3.4) becomes the matrix

$$\begin{pmatrix} m\dot{\psi} \sin \theta (\dot{\varphi} + \dot{\psi} \cos \theta) \\ -m\dot{\theta} \sin \theta (\dot{\varphi} + \dot{\psi} \cos \theta) \\ -m\dot{\psi} \dot{\theta} \sin \theta + m\dot{\psi} \dot{\theta} \sin \theta \end{pmatrix} = 0.
 \tag{6.4}$$

Clearly the last entry of (6.4) vanishes, and a straightforward computation shows that the variational nonholonomic Euler–Lagrange equation for φ reduces to the momentum conservation law

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\varphi}} = 0.
 \tag{6.5}$$

Now, for $I_1 = I_2 = I_3$ (6.5) 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 (6.4). 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.

Summarizing, we have thus seen that nonholonomic trajectories can be viewed as variational trajectories in many very different ways. For the trivial cases, the constraints may in fact be dynamic nonholonomic constraints [3] (i.e. constants of motion), as in the case of Veselova’s system restricted to the case of constant measure density. For kinematic nonholonomic constraints one may only be able to realize some of the nonholonomic trajectories as variational ones, as in the case of the homogeneous Chaplygin sphere, and the nonholonomic free particle. Lastly, for kinematic constraints satisfying (3.4) one can actually view all of the nonholonomic trajectories as variational ones, and the system is then conditionally variational, as is the case for the vertical rolling disc.

7. Conclusion

As we have seen, the dynamics of the variational nonholonomic formalism in most cases differs from the mechanics of nonholonomic systems. Variational nonholonomic systems are by definition Hamiltonian, and carry along with them natural properties (such as measure preservation) which nonholonomic systems generally fail to satisfy. This is not to say that there are no instances when the two dynamics coincide (which has largely been the subject of this paper). However, attempting to directly verify the equivalence conditions of [28] leads to time consuming calculations of little geometric meaning. Thus, our study into instances when the two coincide has stressed geometric ideas, as well as simple computational tools to verify conditions (3.4). Although most of the results are contained in proposition 3, which only holds for Abelian Chaplygin systems, corollary 4 equally addresses the non-Abelian

Chaplygin case, and in both cases we have stressed the notion that to a certain extent it is the structure of the constrained equations which qualifies a system to be conditionally variational.

We originally approached the idea of a system being conditionally variational by examining the associated variational nonholonomic system on \tilde{Q} in section 2. However, we showed in section 3.5 that in some cases one can eliminate the multipliers altogether and still regain the nonholonomic mechanics by restricting to certain initial data. Indeed, we showed this to be the case for the vertical rolling disc in section 5.1, and a similar situation is satisfied by other nonholonomic systems which are actually not conditionally variational (see [14] for more details).

Our final main development was the quantification of how nonholonomic systems possessing an invariant measure are closer to being Hamiltonian. We showed in proposition 6 under what circumstances the existence of a constant density invariant measure for a given nonholonomic system gives rise to the system being conditionally variational. Moreover, in the two degree of freedom case we related this to the procedure of Hamiltonization.

Along the way we have also highlighted the vast array of possibilities for which nonholonomic and variational nonholonomic systems can have identical trajectories. From the conditionally variational vertical disc to the zero set of the momentum map for the Chaplygin sphere, we can see that identical trajectories vary considerably depending on the geometric properties of the system. In doing so we have hopefully quantified the difference (and similarity) of nonholonomic and variational nonholonomic dynamics which has sometimes been a point of confusion [16].

Acknowledgments

This research was supported in part by the Rackham Graduate School of the University of Michigan, through the Rackham Science award, and through NSF grants DMS-0604307 and CMS-0408542. We would also like to thank Professor Alejandro Uribe for many helpful discussions.

References

- [1] Arnold V I, Kozlov V V and Neishtadt A I 1997 *Mathematical Aspects of Classical and Celestial Mechanics* (Berlin: Springer)
- [2] Bates L 1998 Examples of singular nonholonomic reduction *Rep. Math. Phys.* **42** 231–47
- [3] Bloch A M 2003 *Nonholonomic Mechanics and Control* (New York: Springer)
- [4] Bloch A M and Crouch P E 1993 Nonholonomic and vakonomic control systems on Riemannian manifolds *Fields Institute Commun.* **1** 25–52
 Bloch A M and Crouch P E 1996 Optimal control and geodesic flows *Syst. Control Lett.* **28** 65–72
 Bloch A M and Crouch P E 1998 Newton’s law and integrability of nonholonomic systems *SIAM J. Control Optim.* **36** 2020–39
- [5] Bloch A M, Krishnaprasad P S, Marsden J E and Ratiu T S 1996 The Euler-Poincare equations and double bracket dissipation *Comm. Math. Phys.* **175** 1–42
 Bloch A M, Krishnaprasad P S, Marsden J E and Ratiu T S 1996 Nonholonomic mechanical systems with symmetry *Arch. Rat. Mech. An.* **136** 21–99
- [6] Borisov A V and Mamaev I S 2005 Hamiltonization of nonholonomic systems *Preprint nlin/0509036*
- [7] Cardin F and Favretti M 1996 On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints *J. Geom. Phys.* **18** 295–325
- [8] Cendra H, Marsden J E and Ratiu T S Geometric mechanics, Lagrangian reduction, and nonholonomic systems *Mathematics Unlimited-2001 and Beyond* pp 221–73
- [9] Chaplygin S A 1903 On a rolling sphere on a horizontal plane *Mat. Sbornik* **24** 139–68 (in Russian)
- [10] Cortes J 2002 *Geometric, Control and Numerical Aspects of Nonholonomic Systems* (Berlin: Springer)

- [11] Cortes J, de Leon M, de Diego D M and Martinez S 2003 Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions *SIAM J. Control Optim.* **41** 1389–412
- [12] Ehlers K, Koiller J, Montgomery R and Rice P M 2004 Nonholonomic systems via moving frames: Cartan's equivalence and Chaplygin Hamiltonization *The Breadth of Symplectic and Poisson Geometry, Festschrift in Honor of Alan Weinstein (Progress in Mathematics vol 232)* (Boston: Birkhauser) pp 75–120
- [13] Favretti M 1998 Equivalence of dynamics for nonholonomic systems with transverse constraints *J. Dyn. Differ. Eqns.* **10** 511–36
- [14] Fernandez O E, Bloch A M and Mestdag T The Pontryagin maximum principle applied to nonholonomic mechanics *47th IEEE Proc. on Decision and Control* submitted
- [15] Federov Y N and Jovanovic B 2006 Quasi-Chaplygin systems and nonholonomic rigid body dynamics *Lett. Math. Phys.* **76** 215–30
- Federov Y N and Jovanovic B 2005 Integrable nonholonomic geodesic flows on compact lie groups *Topol. Methods Theory Integrable Syst.* 115–52
- Federov Y N and Jovanovic B 2004 Nonholonomic LR systems as generalized Chaplygin systems with an invariant measure and geodesic flows on homogeneous spaces *J. Non. Sci.* **14** 341–81
- [16] Flannery M R 2005 The enigma of nonholonomic constraints *Am. J. Phys.* **73** 265–72
- [17] Guo Y X, Wang Y, Chee G Y and Mei F X 2005 Nonholonomic versus vakonomic dynamics on a Riemann-cartan manifold *J. Math. Phys.* **46** 062902
- [18] Guo Y X, Luo S K, Shang M and Mei F X 2001 Birkhoffian formulations of nonholonomic constrained systems *Rep. Math. Phys.* **47** 313–32
- [19] Koiller J 1992 Reduction of some classical nonholonomic systems with symmetry *Arch. Ration. Mech. Anal.* **118** 113–48
- [20] Koon W S and Marsden J E 1997 The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems *Rep. Math Phys.* **40** 21–62
- [21] Koon W S and Marsden J E 1998 The Poisson reduction of nonholonomic mechanical systems *Rep. Math. Phys.* **42** 101–34
- [22] Kozlov V V 2003 *Dynamical Systems X: General Theory of Vortices* (New York: Springer)
- Kozlov V V Invariant measures of the Euler-Poincare equations on Lie algebras *Funct. Anal. Appl.* **22** 69–70
- [23] Lewis A and Murray R 1995 Variational principles for constrained systems: theory and experiment *Int. J. Nonlinear Mechanics* **30** 793–815
- [24] Marsen J and Scheurle J 1993 The reduced Euler-Lagrange equations *Fields Inst. Com.* **1** 139–64
- [25] Martinez S, Cortes J and de Leon M 2000 The geometrical theory of constraints applied to the dynamics of vakonomic mechanical systems: the vakonomic bracket *J. Math. Phys.* **43** 2090–120
- [26] Montgomery R 2001 *A Tour of Subriemannian Geometries, Their Geodesics, and Applications* (Providence, RI: American Mathematical Society)
- [27] Ramos A 2004 Poisson structures for reduced non-holonomic systems *J. Phys. A: Math. Gen.* **37** 4821–42
- [28] Rumiantsev V V 1978 On Hamilton's principle for nonholonomic systems *P. M. M. USSR* **42** 387–99
- [29] van der Schaft A J and Maschke B M 1994 On the Hamiltonian formulation of nonholonomic mechanical systems *Rep. Math. Phys.* **34** 225–33
- [30] Zenkov D V and Bloch A M 2003 Invariant measures of nonholonomic flows with internal degrees of freedom *Nonlinearity* **16** 1793–807