Nonholonomic Dynamics

Anthony M. Bloch, Jerrold E. Marsden, Dmitry V. Zenkov

Introduction

Nonholonomic systems are, roughly speaking, mechanical systems with constraints on their velocity that are not derivable from position constraints. They arise, for instance, in mechanical systems that have rolling contact (for example, the rolling of wheels without slipping) or certain kinds of sliding contact (such as the sliding of skates). They are a remarkable generalization of classical Lagrangian and Hamiltonian systems in which one allows position constraints only.

There are some fascinating differences between nonholonomic systems and classical Hamiltonian or Lagrangian systems. Among other things: nonholonomic systems are nonvariational—they arise from the Lagrange-d'Alembert principle and not from Hamilton's principle; while energy is preserved for nonholonomic systems, momentum is not always preserved for systems with symmetry (i.e., there is nontrivial dynamics associated with the nonholonomic generalization of Noether's

theorem); nonholonomic systems are almost Poisson but not Poisson (i.e., there is a bracket that together with the energy on the phase space defines the motion, but the bracket generally does not satisfy the Jacobi identity); and finally, unlike the Hamiltonian setting, volume may not be preserved in the phase space, leading to interesting asymptotic stability in some cases, despite energy conservation. The purpose of this article is to engage the reader's interest by highlighting some of these differences along with some current research in the area. There has been some confusion in the literature for quite some time over issues such as the variational character of nonholonomic systems, so it is appropriate that we begin with a brief review of the history of the subject.

Some History

The term "nonholonomic system" was coined by Hertz (1894). The oldest publication that addresses the dynamics of a rolling rigid body known to the authors is Euler (1734), in which small oscillations of a rigid body moving without slipping on a horizontal plane were studied. Later, the dynamics of a rigid body rolling on a surface was studied in Routh (1860), Slesser (1861), Vierkandt (1892), and Walker (1896).

The derivation of the equations of motion of a nonholonomic system in the form of the Euler-Lagrange equations corrected by some additional terms to take into account the constraints (but without Lagrange multipliers), was outlined by Ferrers (1872). The formal derivation of this form of equations was performed in Voronetz (1901). In the case in which some of the configuration variables

Anthony M. Bloch is professor of mathematics at the University of Michigan, Ann Arbor. His email address is `abloch@umich.edu`. *Research partially supported by NSF grants DMS 0103895 and 0305837.*

Jerrold E. Marsden is professor in Control and Dynamical Systems at the California Institute of Technology. His email address is `marsden@cds.caltech.edu`. *Research partially supported by NSF grant DMS-0204474.*

Dmitry V. Zenkov is assistant professor of mathematics at North Carolina State University. His email address is `dvzenkov@unity.ncsu.edu`. *Research partially supported by NSF grant DMS-0306017.*

are cyclic, such equations (now called Chaplygin equations) were obtained by Chaplygin in 1895 (and published two years later). This result of Chaplygin eventually gave rise to the modern technique of nonholonomic reduction. Chaplygin also was first to realize the importance of an invariant measure in nonholonomic dynamics.

One of the more interesting historical events was the paper of Korteweg (1899). Up to that point (and even persisting until recently) there was some confusion in the literature between nonholonomic mechanical systems and variational nonholonomic systems (also called "vakonomic" systems). The latter are appropriate for optimal control problems. One of the purposes of Korteweg's paper was to straighten out this confusion, and in doing so, he pointed out a number of errors in papers up to that point. We refer the reader to Cendra, Marsden, and Ratiu [2001] for an elaboration on some of these points and a more comprehensive historical review.

Classic books in mechanics, as well as their modern counterparts, have discussed in detail the geometry of Hamiltonian and Lagrangian systems; on the other hand, there has not been much work until recently on the geometry of nonholonomic systems. The geometry and reduction of such systems is discussed in the recent book by Bloch [2003], in which a fairly comprehensive survey is given together with a discussion of the natural connections to control theory. A comprehensive set of references to the literature may be found in this reference together with many other topics not touched on here. In the last section of the present paper, we do, however, give a brief discussion of interesting topics that still await further investigation, such as integrability.

Toys and Warnings

Figure 1 shows the famous physicists Wolfgang Pauli and Niels Bohr examining the "tippe top" toy undergoing its interesting inversion. It is simply a half-sphere, with a cylindrical stem mounted on the flat part of the half-sphere used to spin the toy. If one spins it fast enough, then it undergoes a 180-degree flip of its axis of rotation. There are similar toys, such as the rattleback which we discuss below, that also undergo rather nonintuitive motions.

However, one has to be quite careful about how one models such systems. For example, while it might seem quite appealing to model the initial motion of the tippe top as a sphere rolling on a flat surface, in this and some similar situations (such as the "rising egg") it turns out that sliding friction (which would mean using a *holonomic* mechanical model) plays a very important role, and so modeling it as a nonholonomic system is too simplistic a view. For further discussion and simulations, see Bou-Rabee, Marsden, and Romero [2004].

Figure 1. Wolfgang Pauli and Niels Bohr— examining "tippe top" inversion (at the Institute of Physics at Lund, Sweden in 1955).

The Lagrange-d'Alembert Principle

We now describe the equations of motion for a nonholonomic system. We confine our attention to nonholonomic constraints that are homogeneous in the velocity. Accordingly, we consider a mechanical system with a configuration manifold Q, whose local coordinates are denoted q^i, and an $(n - p)$-dimensional nonintegrable constraint distribution $\mathcal{D} \subset TQ$. The distribution $\mathcal{D}$ can be described locally by equations of the form

$$(1) \qquad \dot{s}^a + A^a_\alpha(r, s)\dot{r}^\alpha = 0, \quad a = 1, \ldots, p,$$

where $q = (r, s) \in \mathbb{R}^{n-p} \times \mathbb{R}^p$ are appropriately chosen local coordinates in Q, which we write as $q^i = (r^\alpha, s^a)$, $1 \le \alpha \le n - p$ and $1 \le a \le p$. Note that we only consider here constraints that are linear in the velocities. These linear constraints cover essentially all physical systems of interest. Nonlinear constraints are of interest, however—a discussion and history may be found, for example, in Marle [1998].

Consider, in addition to the constraint distribution, a given Lagrangian $L: TQ \to \mathbb{R}$. As in holonomic mechanics, the Lagrangian for many systems is the kinetic energy minus the potential energy. The equations of motion are then given by the following Lagrange-d'Alembert principle.

Definition 1. *The Lagrange-d'Alembert equations of motion for the system with the Lagrangian L and constraint distribution $\mathcal{D}$ are those determined by*

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

where we choose variations $\delta q(t)$ of the curve $q(t)$ that satisfy the constraints for each $t \in [a, b]$ and vanish at the endpoints, i.e., $\delta q(a) = \delta q(b) = 0$. This principle is supplemented by the condition that the curve itself satisfies the constraints; that is, we require that $\dot{q} \in \mathcal{D}$.

It is also of interest to consider the role of Dirac structures in nonholonomic mechanics. Interestingly, this point of view enables one to formulate the variation of the Lagrangian and constraint as one condition (see Yoshimura and Marsden [2004] and references therein).

Note carefully that in the above definition, we take the variation *before* imposing the constraints; that is, we do not impose the constraints on the family of curves defining the variation. These operations *do not commute*, and this fact is a central reason that nonholonomic mechanics is *nonvariational* in the usual sense of the word. This distinction, already remarked on in our historical introduction, is well known to be important for obtaining the correct mechanical equations (see Bloch, Krishnaprasad, Marsden, and Murray [1996] and Bloch [2003] for a discussion and references).

The usual arguments in the calculus of variations show that the Lagrange-d'Alembert principle is equivalent to the equations

$$(2) \qquad -\delta L = \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right) \delta q^i = 0$$

for all variations $\delta q^i = (\delta r^\alpha, \delta s^a)$ satisfying the constraints at each point of the underlying curve $q(t)$, i.e., such that $\delta s^a + A^a_\alpha \delta r^\alpha = 0$. Substituting variations of this type, with δr^α arbitrary, into (2) gives

$$(3) \qquad \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{r}^\alpha} - \frac{\partial L}{\partial r^\alpha} \right) = A^a_\alpha \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a} \right)$$

for all $\alpha = 1, \ldots, n - p$. One can equivalently write these equations in terms of Lagrange multipliers. Equations (3), combined with the constraint equations (1), give the complete equations of motion of the system.

A useful way of reformulating equations (3) is to define a *constrained Lagrangian* by substituting the constraints (1) into the Lagrangian:

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

The equations of motion can be written in terms of the constrained Lagrangian in the following way, as a direct coordinate calculation shows:

$$(4) \qquad \frac{d}{dt} \frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot{s}^b} B^b_{\alpha\beta} \dot{r}^\beta,$$

where $B^b_{\alpha\beta}$ is defined by

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

There is a beautiful geometric interpretation of these equations: the constraints define an Ehresmann connection on the tangent bundle TQ and B is the curvature of the connection which vanishes precisely when the constraints are integrable; that is, are holonomic.
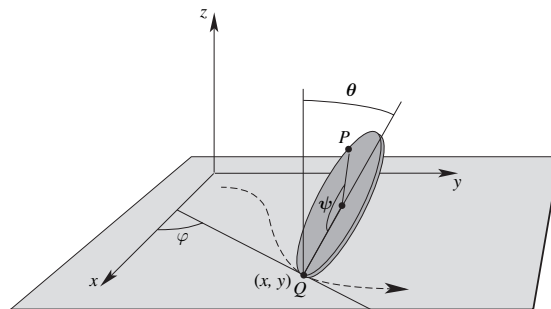

Figure 2. The geometry for the rolling disk.

The Falling Rolling Disk

The falling rolling disk is a simple but instructive example to consider. We consider a disk (such as a coin) that rolls without slipping on a horizontal plane and that can "tilt" as it rolls.

As Figure 2 indicates, we denote the coordinates of contact of the disk with the xy-plane by (x, y) and let θ, φ, and ψ denote the angle between the plane of the disk and the vertical axis, the "heading angle" of the disk, and the "self-rotation" angle of the disk, respectively.[1]

While the equations of motion are straightforward to develop, they are somewhat complicated. One can show that this example is, in an appropriate sense, an integrable system and that it conserves volume in the phase space and, in addition, it exhibits stability but not asymptotic stability. See Zenkov, Bloch, and Marsden [1998] and Bloch [2003] for more details.

This system demonstrates unusual conservation laws, but ones that are typical for nonholonomic systems. One can check that while φ and ψ are cyclic variables (that is, they do not appear explicitly in the constrained Lagrangian), their associated momenta

---

[1]*A classical reference for the rolling disk is Vierkandt [1892], who showed something very interesting: On an appropriate symmetry-reduced space, namely, the constrained velocity phase space modulo the action of the group of Euclidean motions of the plane, all orbits of the system are periodic.*

$$p_1 = \frac{\partial L_c}{\partial \dot{\varphi}} \quad \text{and} \quad p_2 = \frac{\partial L_c}{\partial \dot{\psi}}$$

are *not conserved*.

However, there exist two independent vector fields $\eta_1(\theta)$ and $\eta_2(\theta)$ such that the momentum components along these fields are preserved by the dynamics. We emphasize that the vector fields $\eta_1(\theta)$ and $\eta_2(\theta)$ *do not equal* the fields $\partial/\partial\varphi$ and $\partial/\partial\psi$. See Zenkov [2003] and the references therein for details.

Momentum Equation

Assume there is a Lie group G (with Lie algebra denoted $\mathfrak{g}$) that acts freely and properly on the configuration space Q. A Lagrangian system is called *G-invariant* if its Lagrangian L is invariant under the induced action of G on TQ. Recall the definition of the *momentum map* for an unconstrained Lagrangian system with symmetry: *The **momentum map** $\mathbf{J}: TQ \to \mathfrak{g}_Q^*$ is the bundle map taking TQ to the bundle $\mathfrak{g}_Q^*$ whose fiber over the point q is the dual Lie algebra $\mathfrak{g}^*$ that is defined by*

(5) $\qquad \langle \mathbf{J}(v_q), \xi \rangle = \langle \mathbb{F}L(v_q), \xi_Q \rangle := \frac{\partial L}{\partial \dot{q}^i}(\xi_Q)^i,$

where $\xi \in \mathfrak{g}$, $v_q \in TQ$, and where $\xi_Q \in TQ$ is the generator associated with the Lie algebra element ξ.

A nonholonomic system is called *G-invariant* if both the Lagrangian L and the constraint distribution $\mathcal{D}$ are invariant under the induced action of G on TQ. Let $\mathcal{D}_q$ denote the fiber of the constraint distribution $\mathcal{D}$ at $q \in Q$.

Definition 2. *The **nonholonomic momentum map** $\mathbf{J}^{nhc}$ is defined as the collection of the components of the ordinary momentum map $\mathbf{J}$ that are **consistent with the constraints**, i.e., the Lie algebra elements ξ in equation (5) are chosen from the subspace $\mathfrak{g}^q$ of Lie algebra elements in $\mathfrak{g}$ whose infinitesimal generators evaluated at q lie in the intersection $\mathcal{D}_q \cap T_q(\mathrm{Orb}(q))$.*

Unlike Hamiltonian systems, G-invariant nonholonomic systems often do not have associated momentum conservation laws. Besides the rolling penny, the rattleback and the snakeboard are well-known examples (see Bloch, Krishnaprasad, Marsden, and Murray [1996] and Zenkov, Bloch, and Marsden [1998]). The rattleback is discussed further below.

It is easy to see why the momentum quantities are generally not conserved from the Lagrange–d'Alembert equations of motion. The simplest situation would be the case in which the Lagrangian and the constraint have a cyclic variable (more general definitions of cyclic symmetry that apply to problems like the falling disk are possible). Recall

that the equations of motion have the form (4). If these equations had a cyclic variable, say r^1, then all the quantities L, L_c, and $B^b_{\alpha\beta}$ would be independent of r^1. This is equivalent to saying that there is a translational symmetry in the r^1 direction. Let us also suppose, as is often the case, that the s variables are also cyclic. Then the equation for the momentum $p_1 = \partial L_c / \partial \dot{r}^1$ becomes

$$\dot{p}_1 = -\frac{\partial L}{\partial \dot{s}^b} B^b_{1\beta} \dot{r}^\beta.$$

This fails to be a conservation law in general since the right-hand side need not vanish. Note that the right-hand side is linear in $\dot{r}$, and the equation does not depend on r^1 itself. This is a very special case of what is called the *momentum equation*. For systems with a noncommutative symmetry group, such as the Chaplygin sleigh discussed below, the above analysis for cyclic variables, while giving the right idea, fails to capture the full story.

Thus the nonholonomic momentum is a dynamically evolving quantity. The momentum dynamics is specified in Theorem 3 (see Bloch, Krishnaprasad, Marsden, and Murray [1996]). Let $\mathfrak{g}^D$ be the bundle over Q whose fiber at the point q is given by $\mathfrak{g}^q$.

Theorem 3. *Assume that the Lagrangian is invariant under the group action and that ξ^q is a section of the bundle $\mathfrak{g}^D$. Then a solution $q(t)$ of the Lagrange-d'Alembert equations for a nonholonomic system must satisfy the **momentum equation***

(6) $\qquad \dfrac{d}{dt} \langle \mathbf{J}^{nhc}, (\xi^{q(t)}) \rangle = \dfrac{\partial L}{\partial \dot{q}^i} \left[\dfrac{d}{dt}(\xi^{q(t)}) \right]^i_Q.$

We thus have the following *nonholonomic Noether theorem*:

Corollary 4. *If ξ is a **horizontal symmetry**, i.e., if $\xi_Q(q) \in \mathcal{D}_q$ for all $q \in Q$, then the following conservation law holds:*

(7) $\qquad \dfrac{d}{dt} \langle \mathbf{J}^{nhc}, (\xi) \rangle = 0.$

A somewhat restricted version of the momentum equation was given by Kozlov and Kolesnikov [1978], and the corollary was given by Arnold, Kozlov, and Neishtadt [1988], page 82.

The Poisson Geometry of Nonholonomic Systems

So far we have adopted the philosophy of Lagrangian mechanics; now in this section, we consider the Hamiltonian description of nonholonomic systems. Because of the necessary replacement of conservation laws with the momentum equation, it is natural to let the value of the momentum be a variable, and for this reason it is natural to take a Poisson viewpoint. Some of this theory was

initiated in van der Schaft and Maschke [1994]. What follows builds on their work, further develops the theory of nonholonomic Poisson reduction, and ties this theory to other work in the area. See also Koon and Marsden [1997].

The following two complications make this effort especially interesting. First of all, as we have mentioned, symmetry need not lead to conservation laws but rather to a momentum equation. Second, the natural Poisson bracket fails to satisfy the Jacobi identity. In fact, the so-called *Jacobiator* (the cyclic sum that vanishes when the Jacobi identity holds) or, equivalently, the Schouten bracket, is an interesting expression involving the curvature of the underlying distribution describing the nonholonomic constraints. Thus in the nonholonomic setting we have an *almost Poisson* structure.

Poisson Formulation

The approach of van der Schaft and Maschke [1994] starts on the Lagrangian side with a configuration space Q and a Lagrangian L (possibly of the form kinetic energy minus potential energy, i.e.,

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

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

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

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

where $a = 1, \ldots, k$. Recall that the cotangent bundle T^*Q is equipped with a canonical Poisson bracket which is expressed in the canonical coordinates (q, p) as

$$\{F, G\}(q, p) = \frac{\partial F}{\partial q^i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q^i}$$

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

Here J is the canonical Poisson tensor

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

As in the Lagrangian setting it is desirable to model the Hamiltonian equations without the Lagrange multipliers by a vector field on a submanifold of T^*Q. In van der Schaft and Maschke [1994] it is done through a clever change of coordinates. In Bloch [2003] we recall how they do this. Here we just present the results.

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

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

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

(8) $$\tilde{p}_\alpha = X_\alpha^i p_i, \qquad \tilde{p}_a = \omega_a^i p_i.$$

It is shown in van der Schaft and Maschke [1994] that in the new (generally not canonical) coordinates $(q, \tilde{p}_\alpha, \tilde{p}_a)$, the Poisson tensor becomes

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

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

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

van der Schaft and Maschke [1994] use $(q, \tilde{p}_\alpha)$ as induced local coordinates for $\mathcal{M}$. It is easy to show that

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

$$\frac{\partial \tilde{H}}{\partial \tilde{p}_\beta}(q, \tilde{p}_\alpha, \tilde{p}_a) = \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta}(q, \tilde{p}_\alpha),$$

where $H_{\mathcal{M}}$ is the constrained Hamiltonian on $\mathcal{M}$ expressed in the induced coordinates. We can also truncate the Poisson tensor $\tilde{J}$ in (9) by leaving out its last k columns and last k rows and then describe the constrained dynamics on $\mathcal{M}$ expressed in the induced coordinates $(q^i, \tilde{p}_\alpha)$ as follows:

(10)
$$\begin{pmatrix} \dot{q}^i \\ \dot{\tilde{p}}_\alpha \end{pmatrix} = J_{\mathcal{M}}(q, \tilde{p}_\alpha) \begin{pmatrix} \frac{\partial H_{\mathcal{M}}}{\partial q^j}(q, \tilde{p}_\alpha) \\ \frac{\partial H_{\mathcal{M}}}{\partial \tilde{p}_\beta}(q, \tilde{p}_\alpha) \end{pmatrix}, \qquad \begin{pmatrix} q^i \\ \tilde{p}_\alpha \end{pmatrix} \in \mathcal{M}.$$

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

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

$$\{F_{\mathcal{M}}, G_{\mathcal{M}}\}_{\mathcal{M}}(q, \tilde{p}_\alpha) :=$$

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

for any two smooth functions $F_{\mathcal{M}}, G_{\mathcal{M}}$ on the constraint submanifold $\mathcal{M}$. Clearly, this bracket satisfies the first two defining properties of a Poisson bracket, namely, skew symmetry and the Leibniz rule, and one can show that it satisfies the Jacobi identity if and only if the constraints are holonomic. Furthermore, the constrained Hamiltonian $H_{\mathcal{M}}$ is an integral of motion for the constrained dynamics on $\mathcal{M}$ due to the skew symmetry of the bracket.

A Formula for the Constrained Hamilton Equations

In holonomic mechanics, it is well known that the Poisson and the Lagrangian formulations are equivalent via a Legendre transform. And it is natural to ask whether the same relation holds for the nonholonomic mechanics as developed in van der Schaft and Maschke [1994] and in Bloch, Krishnaprasad, Marsden, and Murray [1996].

We can use the general procedures of van der Schaft and Maschke [1994] to write down a compact formula for the nonholonomic equations of motion.

Theorem 5. *Let $q^i = (r^\alpha, s^a)$ be the local coordinates in which ω^a has the form*

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

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

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

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

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

where $B_{\alpha\beta}^b$ are the coefficients of the curvature of the Ehresmann connection. Here p_b should be understood as p_b restricted to $\mathcal{M}$ and, more precisely, should be denoted by $(p_b)_{\mathcal{M}}$.

One can show that the equations in this theorem are equivalent to those in the Lagrange-d'Alembert formulation (see Bloch [2003]).

We remark that the theory of reduction for nonholonomic systems is elegant and interesting—one can formulate the equations in intrinsic fashion on the *constrained reduced velocity phase space*

$\mathcal{D}/G$ under appropriate conditions. The Lagrangian induces a well-defined function, the *constrained reduced Lagrangian*

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

on this phase space. We do not discuss this here for reason of space but refer the reader to Bloch, Krishnaprasad, Marsden, and Murray [1996], Bloch [2003], and Cendra, Marsden, and Ratiu [2001] for both the Lagrangian and Hamiltonian analysis of the equations of motion and the reduced equations of motion. The last paper cited, in particular, gives an intrinsic, coordinate-free formulation that also gives a very neat interpretation to the momentum equation in terms of parallel transport on the appropriate bundle. In particular, the coordinate form of the reduced equations is quite complicated, whereas the intrinsic formulation reveals their structure more clearly.

Measure-Preserving Systems on Lie Groups and Asymptotic Dynamics

In this section we demonstrate that nonholonomic dynamics is not necessarily measure-preserving. This is in contrast to the volume-preserving nature of Hamiltonian systems and follows from the fact that nonholonomic systems are only almost Poisson. Energy, however, is preserved. This illustrates the very special nature of Hamiltonian systems in which both energy and volume are preserved.

The existence of an invariant measure as a necessary condition for integrability of a nonholonomic system was pointed out by Kozlov. The procedure of integration of a measure-preserving dynamical system goes back to Jacobi (1866).

Euler-Poincaré-Suslov Equations

An important special case of the (reduced) nonholonomic equations is the dynamics of a constrained generalized rigid body.

The configuration space for a generalized rigid body is a Lie group G. The Lagrangian $L : TG \to \mathbb{R}$ is a left-invariant metric on G, i.e., $L(g, \dot{g}) = l(g^{-1}\dot{g})$, where $l : \mathfrak{g} \to \mathbb{R}$ is the *reduced Lagrangian* defined by the formula $l(\Omega) = \frac{1}{2} I_{ab} \Omega^a \Omega^b$, $\Omega = (\Omega^1, \dots, \Omega^n)$ lie in a Lie algebra $\mathfrak{g}$ and I_{ab} are the components of the positive-definite *inertia tensor* $\mathbb{I} : \mathfrak{g} \mapsto \mathfrak{g}^*$. The reduced dynamics of the generalized rigid body are governed by the *Euler-Poincaré equations*

(12) $$\dot{p}_b = C_{ab}^c I^{ad} p_c p_d = C_{ab}^c p_c \Omega^a,$$

where $p_b = I_{ab} \Omega^b$ are the components of the momentum and C_{ab}^c are the structure constants of the Lie algebra $\mathfrak{g}$. The system (12) is Hamiltonian. Nonetheless, it can fail the phase volume preservation property, as the following theorem states.

Theorem 6 (*Kozlov [1988]*). *The Euler-Poincaré equations (12) have an invariant measure if and only if the group G is unimodular.*[2]

The constrained generalized rigid body is the dynamical system (12) subject to the left-invariant nonholonomic constraint

$$(13) \qquad \langle a, \Omega \rangle = a_i \Omega^i = 0,$$

where a is a fixed element of the dual Lie algebra $\mathfrak{g}^*$ and $\langle \cdot , \cdot \rangle$ denotes the natural pairing between the Lie algebra and its dual (multiple constraints may be imposed as well). The two classical examples of such systems are the *Chaplygin sleigh* (Chaplygin, 1911) and the *Suslov problem* (Suslov, 1902) discussed below.

The reduced dynamics of the constrained generalized rigid body is governed by the *Euler-Poincaré-Suslov* equations

$$(14) \qquad \dot{p}_b = C^c_{ab} I^{ad} p_c p_d + \lambda a_b = C^c_{ab} p_c \Omega^a + \lambda a_b$$

together with the constraint (13). If the Lagrange multiplier λ is eliminated, (14) becomes the momentum equation.

Next, we formulate a condition for the existence of an invariant measure of the Euler-Poincaré-Suslov equations:

Theorem 7. *Equations (14) have an invariant measure if and only if*

$$(15) \qquad KC^k_{ij} I^{ig} a_g a_k + C^k_{jk} = \mu a_j, \quad where$$

$$K = 1/\langle a, \mathbb{I}^{-1} a \rangle \quad and \quad \mu \in \mathbb{R}.$$

This result was proved by Kozlov [1988] for compact algebras and by Jovanović [1998] for arbitrary algebras. To prove the theorem, one eliminates the multiplier λ and obtains a system of differential equations with quadratic right-hand sides. According to Kozlov [1988], a system of differential equations with homogeneous polynomial right-hand sides is measure preserving if and only if it is divergence free. The condition (15) is then obtained by setting the divergence of the right-hand side of (14) equal to zero.

According to the definition, for a unimodular group, C^k_{jk} vanishes. In particular, if the group is compact or semisimple, it is unimodular, and we can identify $\mathfrak{g}^*$ with $\mathfrak{g}$ and rewrite condition (15) as

$$(16) \qquad [\mathbb{I}^{-1} a, a] = \mu a, \qquad \mu \in \mathbb{R}.$$

Pairing a with itself (via the Killing form or a multiple of the trace) we have

---

[2] *Recall that a Lie group is called **unimodular** if the structure constants satisfy the equations $C^c_{ac} = 0$. A standard fact is that a unimodular group has a bilaterally invariant measure.*

$$(17) \qquad \langle [\mathbb{I}^{-1} a, a], a \rangle = \langle \mu a, a \rangle$$

and, since the left-hand side is zero, μ must be zero. Thus in this case only constraint vectors a that commute with $\mathbb{I}^{-1} a$ allow the measure to be preserved. This means that a and $\mathbb{I}^{-1} a$ must lie in the same maximal commuting subalgebra. In particular, if a is an eigenstate of the inertia tensor, measure is preserved. When the maximal commuting subalgebra is one-dimensional, this is a necessary condition. This is the case for groups such as $SO(3)$ (see below).

Theorem 7 can be restated as the following symmetry requirement imposed on the constraints:

Theorem 8. *A compact Euler-Poincaré-Suslov system is measure preserving if the constraint vectors a are eigenvectors of the inertia tensor, or if the constrained system is $\mathbb{Z}_2$ symmetric about all principal axes. If the maximal commuting subalgebra is one-dimensional, this condition is necessary.*

The Euler-Poincaré-Suslov Problem on SO(3).

As an illustration, consider the classical Suslov problem, which can be formulated as the standard Euler top dynamics subject to the constraint

$$(18) \qquad \langle a, \Omega \rangle = a_1 \Omega^1 + a_2 \Omega^2 + a_3 \Omega^3 = 0,$$

where $\Omega = (\Omega^1, \Omega^2, \Omega^3) \in so(3)$ is the angular velocity of the top.

Constraint (18) forces the projection of the angular velocity along the direction $a = (a_1, a_2, a_3)$ relative to the body frame to vanish. The reduced nonholonomic equations of motion are then given by (14) with

$$C^3_{12} = C^1_{23} = C^2_{31} = -C^3_{21} = -C^1_{32} = -C^2_{13} = 1$$

$$\text{and} \quad C^k_{ij} = 0 \quad \text{otherwise.}$$

As (17) implies, the momentum dynamics is measure (phase volume) preserving if and only if the constraint direction a is an eigenvector of the inertia tensor $\mathbb{I}$.

An alternative way to obtain this conclusion is to compute the eigenvalues of the linearized momentum flow at the equilibria. If $a_2 = a_3 = 0$ (a constraint that is an eigenstate of the moment of inertia operator) one gets zero eigenvalues while in general one gets a real non-zero eigenvalue and two zero eigenvalues, which is incompatible with measure preservation.

The Chaplygin Sleigh

One of the simplest mechanical systems that illustrates the possible "dissipative nature" of nonholonomic systems, even though they are energy preserving, is the Chaplygin sleigh. This system consists of a rigid body sliding on a plane. The body 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 orthogonal to this edge.

To analyze the system, one can use a coordinate system Oxy fixed in the plane and a coordinate system $A\xi\eta$ fixed in the body with its origin at the point of support of the knife edge and the axis $A\xi$ through the center of mass C of the rigid body. The configuration of the body is described by the coordinates x, y of the contact point and the angle θ between the moving and fixed sets of axes, i.e., the configuration space is the group $SE(2)$. Let m be the mass and I the moment of inertia of the body about the center of mass. Let a be the distance from A to C (see Figure 3). The nonholonomic momentum has two components: p_1, the angular momentum of the system relative to the contact point, and p_2, the projection of the linear momentum of the system on the ξ axis.

The momentum equations written relative to the body frame become

$$(19) \qquad \dot{p}_1 = -\frac{a\,p_1 p_2}{I + ma^2}, \qquad \dot{p}_2 = \frac{ma\,p_1^2}{(I + ma^2)^2}.$$

This dynamics has a family of equilibria (i.e., points at which the right-hand sides vanish) given by $\{(p_1, p_2) \mid p_1 = 0, p_2 = \text{const}\}$.

Assuming $a > 0$ and linearizing about any of these equilibria, one finds a zero eigenvalue, and a negative eigenvalue if $p_2 > 0$ or a positive eigenvalue if $p_2 < 0$.

Thus the volume in the momentum plane is preserved if and only if $a = 0$, which is equivalent to (15).[3] In fact, the solution curves are ellipses in the $p_1 p_2$-plane with the positive p_2-axis attracting all solutions (see Figure 4).

If $a = 0$, the dynamics is integrable, and in particular, the *body* momentum relative to the group $SE(2)$ is preserved. Recall that a free rigid body on the plane conserves the *spatial* momentum. This illustrates how different momentum conservation laws are in the case of nonhorizontal symmetry.

The Rattleback

We end with a brief discussion of one of the most fascinating nonholonomic systems—the rattleback top or Celtic stone. A rattleback is a convex asymmetric rigid body rolling without sliding on a horizontal plane (see Figure 5). It is known for its ability to spin in one direction and to resist spinning in the opposite direction for some parameter values, and for other values to exhibit multiple reversals in clear violation of conservation of angular momentum or of damped angular momentum. In fact, this phenomenon may be viewed as a re-
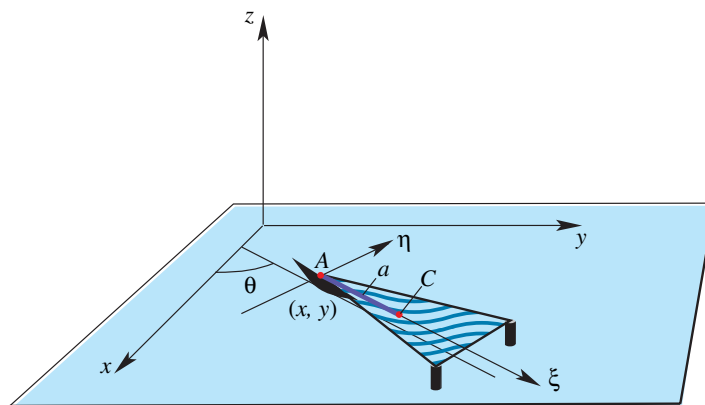

Figure 3. The Chaplygin sleigh is a rigid body moving on two sliding posts and one knife edge.
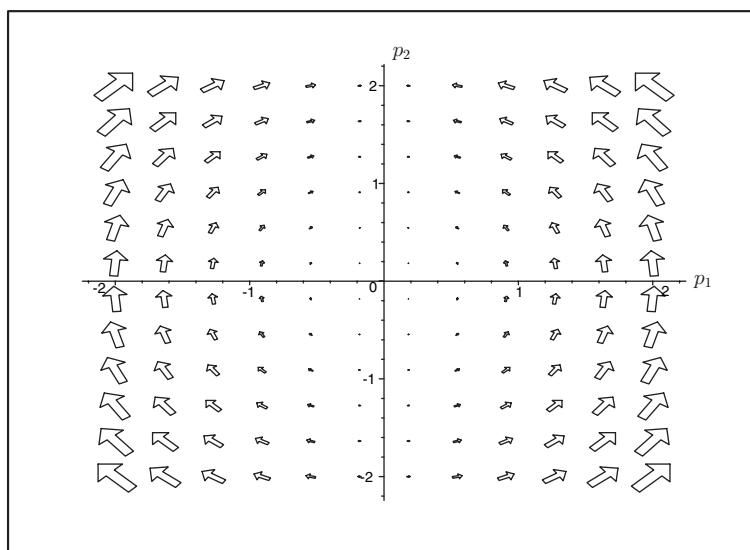

Figure 4. Chaplygin sleigh phase portrait.

markable demonstration of the nontriviality of the momentum equation. Moreover, the stable spin direction is in fact asymptotically stable.

We adopt the ideal model (with no energy dissipation and no sliding) and within that context no approximations are made. In particular, the rattleback's shape need not be ellipsoidal. Walker did some initial stability and instability investigations by computing the spectrum, while Bondi extended this analysis and also used what we now recognize as the momentum equation. See Bloch [2003] and Zenkov, Bloch, and Marsden [1998] for the explicit form of the momentum for the rattleback. A discussion of the momentum equation for the rattleback may also be found in Burdick, Goodwine, and Ostrowski [1994]. Karapetyan carried out a stability analysis of the relative equilibria, while Markeev's and Pascal's main contributions were to the study of spin reversals using small-parameter and averaging techniques. Energy methods were used
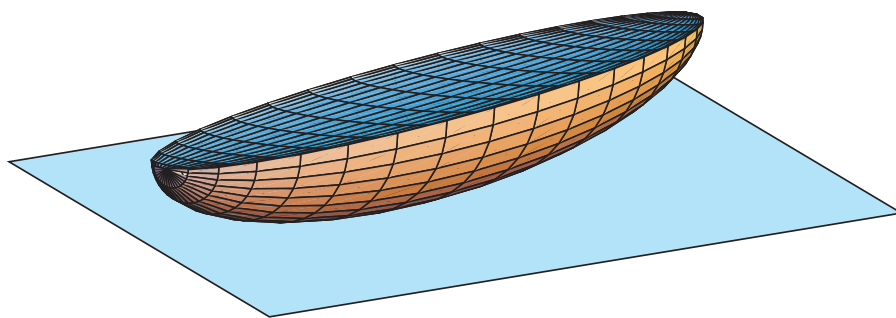
---

[3] *This calculation is performed for the nonholonomic momentum, whereas in the Suslov problem example the full momentum is used.*

Figure 5. The rattleback.

to analyze the problem in Zenkov, Bloch, and Marsden [1998].

There are many other remarkable nonholonomic systems and for these we refer the reader to Bloch [2003], the references therein, and many other papers.

Further Topics

This review has touched on just a few of the fascinating aspects of nonholonomic mechanics. There are many other topics of interest, and we conclude by mentioning some of these.

One question of interest is when a nonholonomic system is integrable. There is no known analogue of the Liouville-Arnold theorem well-known from holonomic mechanics. One can show that nonholonomic systems are integrable if the dimension of the phase space of the system is n and if there exist $(n - 2)$ integrals of motion and an invariant measure (see, e.g., Arnold, Kozlov, and Neishtadt [1998]). Examples of integrable nonholonomic systems include the rolling disk discussed above, Routh's problem of a homogeneous sphere rolling on a surface of revolution (see, e.g., Zenkov [1995]), and Chaplygin's sphere—a balanced inhomogeneous sphere rolling on a plane (see, e.g., Arnold, Kozlov, and Neishtadt [1988]). One of the interesting aspects of such systems is that one can obtain invariant tori as in the Hamiltonian case, but the dynamics on these tori may be nonuniform. It is possible in some such cases to make the system Hamiltonian by a trajectory-dependent time reparameterization. This is discussed in the work of Kozlov and more recently in Ehlers, Koiller, Montgomery, and Rios [2004], who denote the process "Hamiltonization". (This process of Hamiltonization need not necessarily be applied to the integrable case.) These systems conserve a measure, but other systems such as the Chaplygin sleigh do not and are still solvable.

Analysis of the stability of nonholonomic motion is also of interest, and there is a natural generalization of the energy-momentum method of Arnold and of Marsden and collaborators. This is discussed in Zenkov, Bloch, and Marsden [1998]. This method makes use of integrals similar to those discussed for the rolling penny earlier in this paper.

An important topic is the control of nonholonomic systems. This is discussed in detail in Bloch [2003], where many references are given. There is a natural link between nonlinear control systems and nonholonomic distributions: the control vector fields in a control system provide controllability precisely when the distribution they span is nonintegrable, thus giving rise to new directions of motion. A key example is the "nonholonomic integrator", a system with two controls defined on the Heisenberg group introduced and studied in Brockett [1981]. The role of sub-Riemannian geometry in the optimal control of the nonholonomic integrator is discussed in Bloch [2003]. For more on sub-Riemannian geometry see Montgomery [2002]. In sub-Riemannian geometry one has an evolution of a variational or Hamiltonian system subject to a nonholonomic constraint—this should not be confused with nonholonomic mechanical systems. The differences are very interesting and are exposed in detail in Bloch [2003].

Another topic of interest is numerical integration of nonholonomic systems. The idea is to preserve key mechanical quantities of interest such as momentum conservation laws. For a survey of these ideas in the Hamiltonian case see Marsden and West [2001].

Remark. A complete set of references for this paper may be found at `http://www.cds.caltech.edu/mechanics_and_control/`.

References

V. I. ARNOLD, V. V. KOZLOV, and A. I. NEISHTADT, Dynamical systems III, *Encyclopedia of Mathematics*, vol. 3, Springer, Berlin, 1988.

L. BATES and J. ŚNIATYCKI, Nonholonomic reduction, *Reports on Math. Phys.* **32** (1993), 99–115.

A. M. BLOCH, with J. Baillieul, P. E. Crouch and J. E. Marsden, *Nonholonomic Mechanics and Control*, Springer, Berlin, 2003.

A. M. BLOCH, P. S. KRISHNAPRASAD, J. E. MARSDEN, and R. MURRAY, Nonholonomic mechanical systems with symmetry, *Arch. Rat. Mech. An.* **136** (1996), 21–99.

N. M. BOU-RABEE, J. E. MARSDEN, and L. N. ROMERO, Tippe top inversion as a dissipation induced instability, *SIAM J. on Appl. Dyn. Systems* **3** (2004), 352–377.

R. W. BROCKETT, Control theory and singular Riemannian geometry, *New Directions in Applied Mathematics* (P. J. Hilton and G. S. Young, eds.), Springer-Verlag, 1981, 11–27.

J. BURDICK, B. GOODWINE, and J. P. OSTROWSKI, *The Rattleback Revisited*, preprint, 1994.

H. CENDRA, J. E. MARSDEN, and T. S. RATIU, Geometric mechanics, Lagrangian reduction and nonholonomic systems, *Mathematics Unlimited: 2001 and Beyond* (B. Enguist

and W. Schmid, eds.), Springer-Verlag, New York, 2001, 221–273.

K. EHLERS, J. KOILLER, R. MONTGOMERY, and P. M. RIOS, *Nonholonomic Mechanics via Moving Frames: Cartan's Equivalence and Hamiltonizable Chaplygin Systems*, in "The Breadth of Symplectic and Poisson Geometry: Festschrift for Alan Weinstein", preprint, 2004.

B. JOVANOVIĆ, Nonholonomic geodesic flows on Lie groups and the integrable Suslov problem on $SO(4)$, *J. Phys. A: Math. Gen.* **31** (1998), 1415–1422.

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

V. V. KOZLOV, Invariant measures of the Euler–Poincaré equations on Lie algebras, *Functional An. Appl.* **22** (1998), 69–70.

V. V. KOZLOV and N. N. KOLESNIKOV, On theorems of dynamics, *J. Appl. Math. Mech.* **42** (1978), 28–33.

A. LEWIS, J. P. OSTROWSKI, R. M. MURRAY, and J. BURDICK, Nonholonomic mechanics and locomotion: The snakeboard example, *IEEE Intern. Conf. on Robotics and Automation* (1994).

C.-M. MARLE, Various approaches to conservative and nonconservative nonholonomic systems, *Reports on Mathematical Physics* **42** (1998), 211–229.

J. E. MARSDEN and T. S. RATIU, *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics, vol. 17; first edition 1994, second edition, 1999.

J. E. MARSDEN and M. WEST, Discrete mechanics and variational integrators, *Acta Numerica* **10** (2001), 357–514.

R. MONTGOMERY, *A Tour of Sub-Riemannian Geometries, their Geodesics and Applications*, Mathematical Surveys and Monographs, vol. 91, Amer. Math. Soc. 2002.

A. J. VAN DER SCHAFT and B. M. MASCHKE, On the Hamiltonian formulation of nonholonomic mechanical systems, *Rep. on Math. Phys.* **34** (1994), 225–233.

A. VIERKANDT, Über gleitende und rollende Bewegung, *Monatshefte der Math. und Phys.* **III**, (1982) 31–54.

H. YOSHIMURA and J. E. MARSDEN, *Variational principles, Dirac structures, and implicit Lagrangian systems*, Preprint, 2004.

D. V. ZENKOV, The geometry of the Routh problem, *J. Nonlinear Sci.* **5** (1995), 503–519.

_____ , Linear conservation laws of nonholonomic systems with symmetry, *Discrete and Continuous Dynamical Systems* (supplementary volume), 2003, 963–972.

D. V. ZENKOV and A. M. BLOCH, Dynamics of the *n*-dimensional Suslov problem, *Journal of Geometry and Physics* **34** (2000), 121–136.

D. V. ZENKOV, A. M. BLOCH, and J. E. MARSDEN, The energy-momentum method for the stability of nonholonomic systems, *Dyn. Stab. of Systems* **13** (1998), 123–166.