

# Hamel's Formalism for Infinite-Dimensional Mechanical Systems

Donghua Shi\*

School of Mathematics and Statistics  
Beijing Institute of Technology  
Beijing 100081, China  
dshi@bit.edu.cn

Dmitry V. Zenkov  
Department of Mathematics  
North Carolina State University  
Raleigh, NC 27695, USA  
dvzenkov@ncsu.edu

Yakov Berchenko-Kogan  
Department of Mathematics  
Massachusetts Institute of Technology  
Cambridge, MA 02139, USA  
yashabk@math.mit.edu

Anthony M. Bloch  
Department of Mathematics  
University of Michigan  
Ann Arbor, MI 48109, USA  
abloch@umich.edu

Keywords: Hamel equations, nonholonomic constraints, momentum, symmetry  
AMS Subject Classification: 70F25, 37J60, 70H33

## Abstract

In this paper we introduce Hamel's formalism for infinite-dimensional mechanical systems and in particular consider its applications to the dynamics of nonholonomically constrained systems. This development is a nontrivial extension of its finite-dimensional counterpart. The analysis is applied to several continuum-mechanical systems of interest, including coupled systems and systems with infinitely-many constraints.

## 1 Introduction

Hamel's formalism is a natural extension of Euler's ideas of using nonmaterial velocity in mechanics. Nonmaterial velocity carries information about system's velocity, but is not the rate of change of system's configuration with respect to time. For a system with finite number of degrees of freedom, nonmaterial velocity is a collection of velocity components relative to a set of vector fields that span the fibers of the tangent bundle of the configuration space. In the finite-dimensional setting, this development was carried out by Hamel himself in [27]. Here we introduce Hamel's formalism for infinite-dimensional mechanical systems.

One of the reasons for using nonmaterial velocity is that the Euler–Lagrange equations are not always effective for analyzing the dynamics of a mechanical system of interest. For example, it is difficult to study the motion of the Euler top if the Euler–Lagrange equations are used to represent the dynamics. On the other hand, the use of the angular velocity components relative to a body frame as pioneered by Euler [18] results in a much simpler representation of dynamics. In a similar fashion, Euler [19, 20, 21] uses *convective velocity* to represent the dynamics of ideal incompressible fluid. Euler's approach was further developed by Lagrange [36] for reasonably general Lagrangians on the rotation group and by Poincaré [43] for arbitrary Lie groups (see [39] for details and history).

The nonmaterial velocity used in Lagrange [36] and Poincaré [43] is associated with a group action. Hamel [27] obtained the equations of motion in terms of nonmaterial velocity that is unrelated to a group action on the configuration space. Hamel's equations include both the Euler–Lagrange and Euler–Poincaré equations (for the rigid body for example) as special cases.

---

\*Corresponding author

As clearly seen from his paper, Hamel was particularly motivated by nonholonomic mechanics. His formalism features the simplicity of an analytic representation of constraints and the intrinsic absence of Lagrange multipliers in equations of motion. It is exceptionally effective for studying (finite-dimensional) constrained systems and understanding their dynamics, both analytically and numerically; see e.g. Bloch *et al.* [9], Ball and Zenkov [4], and Zenkov *et al.* [53] and references therein.

As mentioned above, in finite dimensions nonmaterial velocity refers to velocity components relative to a set of vector fields that span the fibers of the tangent bundle of the configuration space. Alternatively, one interprets nonmaterial velocity as a result of configuration-dependent velocity substitution. In the infinite-dimensional setting, the former interpretation fails to work as generically bases fail to exist.

General properties of infinite-dimensional constrained systems are studied from the field-theoretic viewpoint in e.g. Binz *et al.* [5] and Vankerschaver [47, 48, 49]. Infinite-dimensional nonholonomic mechanics, in relation to electromechanical systems, diffeomorphism groups, and optimal transport, has been utilized in Neimark and Fufaev [40] and Khesin and Lee [33], respectively. At the moment, the general theory of infinite-dimensional nonholonomic systems in the form of ordinary differential equations on an infinite-dimensional configuration space does not exist. As demonstrated by Ebin and Marsden [16] (see also Ebin [15]), the use of the infinite-dimensional formalism in combination with nonmaterial (spatial) velocity is crucial for proving existence of solutions in fluid mechanics. Motivated by this and as a part of the development of Hamel’s formalism, we introduce the general theory of infinite-dimensional nonholonomic systems.

Our paper develops the general formalism of Hamel’s equations in infinite-dimensions with careful attention to the analytic setting in which the equations are well defined. We also discuss mechanical examples, including a string, a constrained string, and strings attached to rigid bodies. The differential equations describing these systems are derived and analyzed.

Having in mind possible future applications to the dynamics of systems with non-Banach configuration manifolds (such as infinite-dimensional Lie groups), we develop the formalism on *convenient manifolds*. These are (infinite-dimensional) manifolds modeled on *convenient spaces*. A convenient space is a locally-convex space with ‘tweaked’ topology. By definition, the  $c^\infty$ -topology on a locally convex space  $E$  is the final topology with respect to all smooth curves  $\gamma : \mathbb{R} \rightarrow E$ , i.e., it is the finest topology on  $E$  with respect to which the aforementioned curves are continuous. Recall that smoothness is a bornological concept and is independent of the choice of topology of a locally convex space. In general, the  $c^\infty$ -topology is finer than any locally convex topology with the same collection of bounded sets. A locally convex vector space is said to be  $c^\infty$ -complete or *convenient* if it is  $c^\infty$ -closed in any locally convex space. For a *Fréchet space*, the  $c^\infty$ -topology coincides with the given locally convex topology. In general, a convenient space is not a topological vector space.

A mapping between two convenient spaces is called *smooth* if it maps smooth curves to smooth curves. The important property of the smooth mappings,

$$C^\infty(E \times F, G) \cong C^\infty(E, C^\infty(F, G)),$$

is known as the *Cartesian closedness*. We refer the reader to Kriegl and Michor [34] for details and proofs.

The use of such general spaces is natural in variational calculus and helps to clarify various technical assumptions one needs to impose in the infinite-dimensional setting with velocity constraints. Meanwhile, the complexity of proofs is unaffected by the use of convenient analysis. Moreover, the results remain correct after a rollback to ‘simpler’ (e.g., Fréchet or Banach) manifolds.

Based on the earlier observations of the authors in the finite-dimensional setting and recent publications on geometric continuum mechanics, we expect that the proposed formalism will be useful in:

- Systematic derivation of simple representations of equations of motion in continuum mechanics, which includes elimination of unnecessary Lagrange multipliers from the equations of motion and singling out control directions when they are inconsistent with configuration coordinates and/or do not commute.
- Multibody dynamics of systems with continuum components.
- Dynamics of contacting elastic rods (Gay-Balmaz and Putkaradze [24, 25]), molecular strands (Ellis *et al.* [17]), and certain non-Newtonian fluids (Gay-Balmaz and Yoshimura [26]).

- Stability and qualitative analysis, including existence of conservation laws in constrained systems and Lyapunov function construction strategies for partial stability analysis, with applications to stability of rigid bodies with fins moving in a fluid, fluid-filled bodies, and rolling elastic bodies.
- Development of structure-preserving integrators for constrained infinite-dimensional systems.

The paper is organized as follows: In Section 2 we review the finite-dimensional Euler–Lagrange, Euler–Poincaré, and Hamel equations. In Section 3, the infinite-dimensional Hamel formalism is introduced. In Sections 4 and 5, systems with constraints and symmetry are treated. In particular, systems with infinitely-many constraints are studied, which, in the presence of symmetry, requires a somewhat different approach than the standard formalism for systems with symmetry. Numerous illustrative examples are given.

## 2 Preliminaries

Lagrangian mechanics provides a systematic approach to deriving the equations of motion as well as establishes the equivalence of force balance and variational principles.

### 2.1 The Euler–Lagrange Equations

A *Lagrangian mechanical system* is specified by a smooth manifold  $Q$  called the *configuration space* and a function  $L : TQ \rightarrow \mathbb{R}$  called the *Lagrangian*. In many cases, the Lagrangian is the kinetic minus potential energy of the system, with the kinetic energy defined by a Riemannian metric and the potential energy being a smooth function on the configuration space  $Q$ . If necessary, non-conservative forces can be introduced (e.g., gyroscopic forces that are represented by terms in  $L$  that are linear in the velocity), but this is not discussed in detail in this paper.

In local coordinates  $q = (q^1, \dots, q^n)$  on the configuration space  $Q$  we write  $L = L(q, \dot{q})$ . The dynamics is given by the *Euler–Lagrange equations*

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial q^i}, \quad i = 1, \dots, n.$$

These equations were originally derived by Lagrange [36] in 1788 by requiring that simple force balance be *covariant*, i.e., expressible in arbitrary generalized coordinates. A variational derivation of the Euler–Lagrange equations, namely *Hamilton’s principle* (see Theorem 2.1 below), came later in the work of Hamilton [28, 29] in 1834/35.

Let  $q(t)$ ,  $a \leq t \leq b$ , be a smooth curve in  $Q$ . A *variation* of the curve  $q(t)$  is a smooth map  $\beta : [a, b] \times [-\varepsilon, \varepsilon] \rightarrow Q$  that satisfies the condition  $\beta(t, 0) = q(t)$ . This variation defines the vector field

$$\delta q(t) = \left. \frac{\partial \beta(t, \tau)}{\partial \tau} \right|_{\tau=0}$$

along the curve  $q(t)$ .

**Theorem 2.1.** *The following statements are equivalent:*

- (i) *The curve  $q(t)$ , where  $a \leq t \leq b$ , is a critical point of the **action functional***

$$\int_a^b L(q, \dot{q}) dt$$

*on the space of curves in  $Q$  connecting  $q_a$  to  $q_b$  on the interval  $[a, b]$ , where we choose variations of the curve  $q(t)$  that satisfy  $\delta q(a) = \delta q(b) = 0$ .*

- (ii) *The curve  $q(t)$  satisfies the Euler–Lagrange equations.*

We point out here that this principle assumes that a variation of the curve  $q(t)$  induces the variation  $\delta\dot{q}(t)$  of its velocity vector according to the formula

$$\delta\dot{q}(t) := \frac{d}{dt}\delta q(t).$$

For more details and a proof, see e.g. [6] and [39].

## 2.2 The Euler–Poincaré Equations

The classical *Euler equations* for freely rotating rigid body read

$$J\dot{\Omega} = J\Omega \times \Omega,$$

where  $\Omega$  is the body angular velocity and  $J$  is the inertia tensor. First derived by Euler [18], these equations, as well as the Euler equation for an incompressible fluid flow,

$$\frac{\partial v}{\partial t} + \nabla_v v = -\nabla p, \quad \operatorname{div} v = 0,$$

were generalized by Poincaré [43, 44] to *any* Lie algebra. These *Euler–Poincaré equations* for a Lagrangian  $l(\xi)$  defined on a Lie algebra  $\mathfrak{g}$  are

$$\frac{d}{dt} \frac{\partial l}{\partial \xi} = \pm \operatorname{ad}_\xi^* \frac{\partial l}{\partial \xi}. \quad (2.1)$$

These equations are variational, with variations satisfying certain constraints, as the following theorem clarifies.

**Theorem 2.2.** *Let  $\mathfrak{g}$  be a Lie algebra and  $l : \mathfrak{g} \rightarrow \mathbb{R}$  be a Lagrangian. The following statements are equivalent:*

(i) *The variational principle*

$$\delta \int_a^b l(\xi(t)) dt = 0$$

*holds on  $\mathfrak{g}$ , using variations of the form*

$$\delta\xi = \dot{\eta} \pm \operatorname{ad}_\xi \eta,$$

*where  $\eta$  vanishes at the endpoints.*

(ii) *The Euler–Poincaré equations (2.1) hold.*

See Bloch *et al.* [8], Cendra *et al.* [10], Marsden [37], Holm *et al.* [31], and Marsden and Ratiu [39] for details, history, and proofs.

## 2.3 Lagrangian Mechanics in Non-Coordinate Frames

In many cases the Lagrangian and the equations of motion have a simpler structure when the velocity components are measured against a frame that is unrelated to the system’s local configuration coordinates. An example of such a system is the rigid body.

Let  $q = (q^1, \dots, q^n)$  be local coordinates on the configuration space  $Q$  and  $u_i \in TQ$ ,  $i = 1, \dots, n$ , be smooth independent *local* vector fields on  $Q$  defined in the same coordinate neighborhood. In certain cases, some or all of the fields  $u_i$  can be chosen to be *global* vector fields on  $Q$ .

Let  $\xi = (\xi^1, \dots, \xi^n) \in \mathbb{R}^n$  be the components of the velocity vector  $\dot{q} \in TQ$  relative to the frame  $u(q) = (u_1(q), \dots, u_n(q))$ , i.e.,

$$\dot{q} = \xi^i u_i(q).$$

The Lagrangian of the system written in the local coordinates  $(q, \xi)$  on the velocity phase space  $TQ$  reads

$$l(q, \xi) := L(q, \xi^i u_i(q)).$$

The coordinates  $(q, \xi)$  are Lagrangian analogue of non-canonical variables in Hamiltonian dynamics.

Define the **structure functions**  $c_{ij}^a(q)$  by the equations

$$[u_i(q), u_j(q)] = c_{ij}^k(q) u_k(q),$$

$i, j, k = 1, \dots, n$ . These quantities vanish if and only if the vector fields  $u_i(q)$ ,  $i = 1, \dots, n$ , commute.

Viewing  $u_i$  as vector fields on  $TQ$  whose fiber components equal 0 (that is, taking the vertical lift of the frame vector fields), one defines the directional derivatives  $u_i[l]$  for a function  $l : TQ \rightarrow \mathbb{R}$  in a usual way.

The evolution of the variables  $(q, \xi)$  is governed by the **Hamel equations**

$$\frac{d}{dt} \frac{\partial l}{\partial \xi^j} = c_{ij}^k \frac{\partial l}{\partial \xi^k} \xi^i + u_j[l], \quad (2.2)$$

$i, j, k = 1, \dots, n$ , coupled with equation  $\dot{q} = \xi^i u_i(q)$ . These equations were introduced in [27] (see also [40] and [9] for details, history, and contemporary geometric exposition). If  $u_i = \partial/\partial q^i$ , the Hamel equations become the Euler–Lagrange equations.

## 2.4 Ideal Constraints

Assume now that there are *velocity constraints* imposed on the system. We confine our attention to constraints that are linear and homogeneous in the velocity. Accordingly, we consider a configuration space  $Q$  and a distribution  $\mathcal{D}$  on  $Q$  that describes these constraints. Recall that a distribution  $\mathcal{D}$  is a collection of linear subspaces of the tangent spaces of  $Q$ ; we denote these subspaces by  $\mathcal{D}_q \subset T_q Q$ , one for each  $q \in Q$ .

A curve  $q(t) \in Q$  is said to **satisfy the constraints** if  $\dot{q}(t) \in \mathcal{D}_{q(t)}$  for all  $t$ . This distribution will, in general, be nonintegrable, i.e., the constraints will be, in general, **nonholonomic**.<sup>1</sup>

As discussed in e.g. Suslov [45] and Chetaev [13], it is assumed in classical mechanics that the constraints imposed on the system can be replaced with the **reaction force**. This means that after the force is imposed on the *unconstrained* system, the constraint distribution  $\mathcal{D} \subset TQ$  becomes a *conditional invariant manifold* of the *forced unconstrained* Lagrangian system whose dynamics on this invariant manifold is identical to that of the constrained system.

**Definition 2.3.** Constraints (either holonomic or nonholonomic) are called **ideal** if their reaction force at each  $q \in Q$  belong to the null space  $\mathcal{D}_q^\circ \subset T_q^* Q$  of  $\mathcal{D}_q \subset T_q Q$ .

As shown in Suslov [45] and Chetaev [13], the reaction force of ideal constraints is defined uniquely at each state  $(q, \dot{q}) \in TQ$ .

In summary, for a system subject to ideal constraints, the forced dynamics is equivalent to the Lagrange–d’Alembert principle. We refer the reader to books [45] and [13] for a more detailed exposition and history of the concept of ideal constraints.

Utilizing Hamel’s formalism and assuming the ideal velocity constraints read  $\xi^{m+1} = \dots = \xi^n = 0$ , the dynamics of the constrained system is given by (2.2) for  $j = 1, \dots, m$ . The remaining  $n - m$  equations serve for computing the reaction force and do not affect the dynamics of the system.

For the early development of these equations see [43] and [27]. We refer the readers to [39] and [9] for the history and development of variational principles for the Euler–Lagrange, Euler–Poincaré, and Hamel equations, and to [3] for the Hamilton–Pontryagin principle for the Hamel equations.

<sup>1</sup>Constraints are nonholonomic if and only if they cannot be rewritten as *position* constraints.

## 2.5 The Chaplygin Sleigh

Here we describe the *Chaplygin sleigh* which is one of the simplest but nonetheless instructive example of a nonholonomic mechanical system. The sleigh is essentially a vertical blade moving on a horizontal plane, with no motion perpendicular to the blade allowed. There is a single contact point of the blade and the plane, and the center of mass of the blade coincides with this contact point. The sleigh is often thought of as a balanced platform on the top of the blade. See Figure 2.1 where the platform, the blade, and the contact point are depicted as an oval, a bold segment, and a bold dot, respectively.

Let  $\theta$  be the angular orientation of the sleigh and  $(x, y)$  be the coordinates of the contact point as shown in Figure 2.1. The configuration space for the sleigh is the Euclidean group  $\text{SE}(2) = \text{SO}(2) \ltimes \mathbb{R}^2$ . We

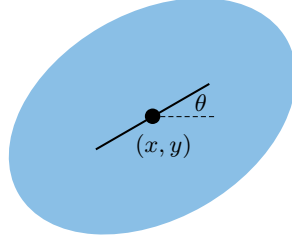


Figure 2.1: The Chaplygin sleigh.

parametrize the elements of  $\text{SE}(2)$  as  $(\theta, x, y)$ . The **body frame** is

$$\frac{\partial}{\partial \theta}, \quad \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y}, \quad -\sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y}.$$

Using this frame,

$$\dot{\theta} = \omega, \quad \dot{x} = v^1 \cos \theta - v^2 \sin \theta, \quad \dot{y} = v^1 \sin \theta + v^2 \cos \theta, \quad (2.3)$$

where  $\omega$  is the angular velocity of the sleigh relative to the vertical line through the contact point and where  $(v^1, v^2)$  are the components of the linear velocity of the contact point in the directions along and orthogonal to the blade, respectively. Thus, the constraint reads  $v^2 = 0$ .

Having in mind infinite-dimensional generalizations of the Chaplygin sleigh, we start using the complex configuration variable  $z = x + iy$  on the plane. Similarly, the linear velocity relative to the body frame is written as  $v = v^1 + iv^2$ . Formulae (2.3) become

$$\dot{\theta} = \omega, \quad \dot{z} = e^{i\theta} v, \quad (2.4)$$

whereas the constraint in this complex representation reads

$$v = \bar{v}.$$

Denote the mass and the moment of inertia of the sleigh by  $m$  and  $J$ . The Lagrangian is just the kinetic energy of the sleigh, which is the sum of the kinetic energies of the linear and rotational modes of the sleigh. Therefore the reduced Lagrangian of the Chaplygin sleigh is

$$l = \frac{1}{2}(J\omega^2 + mv\bar{v}).$$

The elements of the algebra  $\mathfrak{se}(2) = \mathfrak{so}(2) \ltimes \mathbb{C}$  in the complex representation used here are written as  $(i\omega, v)$ , where  $i\omega \in \mathfrak{so}(2)$ ,  $\omega \in \mathbb{R}$ , and  $v \in \mathbb{C}$ . The bracket operation on the algebra  $\mathfrak{se}(2)$  reads

$$[(i\omega_1, v_1), (i\omega_2, v_2)] = (0, i\omega_1 v_2 - i\omega_2 v_1). \quad (2.5)$$

Using (2.5), the constrained Hamel equations for the Chaplygin sleigh are computed to be

$$\dot{\omega} = 0, \quad \dot{v} = 0. \quad (2.6)$$

Recall that in (2.6) the quantity  $v$  is real-valued. The reduced dynamics of the Chaplygin sleigh is given by the constrained Hamel equations (2.6) along with the kinematic equations (2.4).

Solving these equations gives

$$\omega = \text{const}, \quad v = \text{const}, \quad \theta = \theta_0 + \omega t, \quad z = \begin{cases} z_0 + e^{i\theta_0} vt & \text{if } \omega = 0 \\ z_0 - \frac{iv}{\omega} (e^{i(\theta_0 + \omega t)} - e^{i\theta_0}) & \text{if } \omega \neq 0 \end{cases}$$

and so generically the sleigh moves along a circle at a uniform rate.

### 3 Infinite-Dimensional Systems

Starting from this section, we no longer assume that systems have a finite number of degrees of freedom. As the use of frames and bases in the infinite-dimensional setting is unnatural, cumbersome, and not even always possible, we introduce a coordinate-free approach to Hamel's formalism. Thus, instead of frames, we use linear velocity substitutions. These substitutions, however, are not induced by a (local) configuration coordinate change.

#### 3.1 Lagrangian Mechanics

Let  $M$  be an infinite-dimensional smooth manifold modeled on a convenient vector space  $W$  and let  $TM$  be its *kinematic tangent bundle* with the projection  $\pi_M : TM \rightarrow M$ . Consider the *initial inclusion map*  $i : Q \rightarrow M$  and the pullback vector bundle  $P = i^*TM$ .<sup>2</sup> The importance of the manifold  $Q$  will be demonstrated shortly.

A *Lagrangian* is a smooth function  $L : P \rightarrow \mathbb{R}$ . The dynamics for this Lagrangian is defined in a usual way by *Hamilton's principle*: The curve  $\gamma : [a, b] \rightarrow Q$  is a trajectory if

$$\delta \int_a^b L dt = 0$$

along  $\gamma$ .

To demonstrate the necessity of the initial inclusion map in infinite-dimensional mechanics, consider the wave equation

$$\phi_{tt} = \nabla^2 \phi$$

on  $\mathbb{R}^n$ . This is the Euler–Lagrange equation for the Lagrangian

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

where  $\langle \cdot, \cdot \rangle$  is the standard Riemann metric on  $L^2(\mathbb{R}^n)$ . This Lagrangian is defined on the space  $P = H^1 \times L^2 \subset L^2 \times L^2 = TL^2$  and not on the entire  $TL^2$ . Using our notations,  $Q = H^1 \subset L^2 = M$ . See Marsden and Hughes [38] for details.

---

<sup>2</sup>The map  $i$  satisfies the following property: A map  $f : N \rightarrow Q$  is smooth if and only if  $i \circ f : N \rightarrow M$  is smooth. Note that  $Q$  is usually not a submanifold of  $M$ . See Kriegl and Michor [34] for details.

### 3.2 Hamel's Formalism and Hamilton's Principle

Turning to Hamel's formalism, let  $U$  be an open subset of  $M$  containing  $q \in Q$  and let

$$U \times W \ni (q, \xi) \mapsto (q, \Psi_q \xi) \in \pi_M^{-1}(U) \subset TM \quad (3.1)$$

be a fiber-preserving diffeomorphism that is linear in the second input. Hence, for each  $q \in U$ , both  $\Psi_q : W \rightarrow T_q M$  and  $\Psi_q^{-1} : T_q M \rightarrow W$  are invertible bounded linear operators smoothly dependent on  $q$  in an open subset  $i^{-1}(U) \subset Q$ . The latter is verified by selecting the map (3.1) as a bundle chart on  $TM$  and using the Cartesian closedness  $C^\infty(U \times W, W) \cong C^\infty(U, C^\infty(W, W))$  along with the fact that the space of all bounded linear operators from a convenient vector space  $E$  to a convenient vector space  $F$  is a closed linear subspace of  $C^\infty(E, F)$ .<sup>3</sup>

**Remark.** As pointed out in Marsden and Ratiu [39], the space  $W$  and its dual are chosen to be suitable, in the functional-analytic sense, to the problem in question. As the Lagrangian fails to be defined, in general, on  $TM$ , it is necessary to consider various forms of equations of motion, such as weak and strong forms. In particular, the weak form is important in understanding the existence of solutions as well as in numerical methods. Further, the objects involved may not even be defined on the entire  $TM$  but only on a dense subset already in the Banach case. See Marsden and Ratiu [39] for details and references.

For each  $q \in U$  and  $\xi \in W$ , the operator  $\Psi_q : W \rightarrow T_q M$  introduced in (3.1) outputs the vector  $\Psi_q \xi \in T_q M$ .<sup>4</sup> Thus, each  $\xi \in W$  defines the vector field

$$\Psi \xi := \cup_{q \in U} (q, \Psi_q \xi)$$

on  $U$ , which usually will be written as

$$\Psi \xi(q) := \Psi_q \xi.$$

Given two vectors  $\xi, \eta \in W$ , define an antisymmetric bilinear operation  $[\cdot, \cdot]_q : W \times W \rightarrow W$  by

$$\Psi_q[\xi, \eta]_q = [\Psi \xi, \Psi \eta](q), \quad (3.2)$$

where  $[\cdot, \cdot]$  is the Jacobi–Lie bracket of two vector fields on the manifold  $M$ . Next, for arbitrary  $\xi, \eta, \zeta \in W$ , we have

$$\begin{aligned} & \Psi_q([\xi, \eta]_q, \zeta]_q + [[\eta, \zeta]_q, \xi]_q + [[\zeta, \xi]_q, \eta]_q) \\ &= [[\Psi \xi, \Psi \eta], \Psi \zeta](q) + [[\Psi \eta, \Psi \zeta], \Psi \xi](q) + [[\Psi \zeta, \Psi \xi], \Psi \eta](q) = 0, \end{aligned}$$

implying, in view of invertibility of  $\Psi_q$ , the Jacobi identity for the bracket  $[\cdot, \cdot]_q$ . Therefore, for each  $q \in U$ , the space  $W$  with the operation  $[\cdot, \cdot]_q$  is a Lie algebra, denoted hereafter  $W_q$ .

The **dual** of  $[\cdot, \cdot]_q$  is, by definition, the operation  $[\cdot, \cdot]_q^* : W_q \times W_q^* \rightarrow W_q^*$  given by

$$\langle [\xi, \alpha]_q^*, \eta \rangle_W := \langle \alpha, [\xi, \eta]_q \rangle_W, \quad \xi, \eta \in W, \quad \alpha \in W^*.$$

As in the finite-dimensional setting, let  $\dot{q}$  and  $\delta q$  denote the velocity and the virtual displacement at  $q \in Q$ . From now on, the inverse images of  $\dot{q}$  and  $\delta q$  are written as  $\xi, \eta \in W$ , that is,  $\dot{q} = \Psi_q \xi$  and  $\delta q = \Psi_q \eta$ .

Interpreting  $\xi$  as an independent variable that replaces  $\dot{q}$  (locally) defines the Lagrangian as a smooth function of  $(q, \xi)$  on  $V \times W$ :

$$l(q, \xi) := L(q, \Psi_q \xi),$$

where we used the smoothness of the mapping  $V \times W \ni (q, \xi) \mapsto \Psi_q \xi \in \pi_M^{-1}(U)$ . The equations of motion written when  $(q, \xi)$  are selected as (local) coordinates on the velocity phase space are called the **Hamel equations**.

<sup>3</sup>Here,  $U$  can be thought of as an open subset of a convenient space.

<sup>4</sup>Each vector from  $T_q M$  can be represented this way.



Recall that, given a smooth curve  $q(t) \in Q$ ,  $t \in [a, b]$ , its **variation** is a smooth one-parameter family of curves

$$[a, b] \times [-\varepsilon, \varepsilon] \ni (t, \tau) \mapsto \beta(t, \tau) \in Q, \quad \text{such that } \beta(t, 0) = q(t).$$

An **infinitesimal variation**, also known as **variation field**  $\delta q$ , is defined by

$$\delta q(t, \tau) := \frac{\partial}{\partial \tau} \beta(t, \tau).$$

When this field is evaluated along the curve  $q(t)$ , we write  $\delta q(t)$ , i.e.,

$$\delta q(t) := \delta q(t, 0) = \left. \frac{\partial}{\partial \tau} \right|_{\tau=0} \beta(t, \tau). \quad (3.3)$$

Thus, a variation of a smooth curve  $q(t) \in Q$  defines a curve  $\eta(t) \in W$

$$\delta q(t) = \Psi_{q(t)} \eta(t).$$

**Theorem 3.1 (Hamilton's Principle for Hamel's Equations).** *Let  $L : P \rightarrow \mathbb{R}$  be a Lagrangian and  $l$  be its representation in local coordinates  $(q, \xi)$ . Then, the following statements are equivalent:*

- (i) *The curve  $q(t)$ , where  $a \leq t \leq b$ , is a critical point of the action functional*

$$\int_a^b L(q, \dot{q}) dt$$

*on the space of curves in  $Q$  connecting  $q_a$  to  $q_b$  on the interval  $[a, b]$ , where we choose variations of the curve  $q(t)$  that satisfy  $\delta q(a) = \delta q(b) = 0$ .*

- (ii) *The curve  $q(t)$  satisfies the **weak form of the Euler–Lagrange equations***

$$\int_a^b \left\langle \frac{\delta L}{\delta q} - \frac{d}{dt} \frac{\delta L}{\delta \dot{q}}, \delta q \right\rangle dt = 0. \quad (3.4)$$

*Furthermore, if  $i_* T_q Q$  is dense in  $T_q M$  for every  $q \in Q$ , the curve  $q(t)$  satisfies the **strong form of the Euler–Lagrange equations***

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} = 0. \quad (3.5)$$

- (iii) *The curve  $(q(t), \xi(t))$  is a critical point of the functional*

$$\int_a^b l(q, \xi) dt \quad (3.6)$$

*with respect to variations  $\delta \xi$ , induced by the variations*

$$\delta q = \Psi_q \eta,$$

*and given by*

$$\delta \xi = \dot{\eta} + [\xi, \eta]_q. \quad (3.7)$$

- (iv) *The curve  $(q(t), \xi(t))$  satisfies the **weak form of the Hamel equations***

$$\int_a^b \left\langle \Psi_q^* \frac{\delta l}{\delta q} + \left[ \xi, \frac{\delta l}{\delta \xi} \right]_q^* - \frac{d}{dt} \frac{\delta l}{\delta \xi}, \eta \right\rangle dt = 0, \quad \eta \in \Psi_q^{-1}(T_q Q) \quad (3.8)$$

*coupled with the equations  $\dot{q} = \Psi_q \xi$ . If  $i_* T_q Q$  is dense in  $T_q M$  for every  $q \in Q$ , the curve  $(q(t), \xi(t))$  satisfies the **strong form of the Hamel equations***

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \left[ \xi, \frac{\delta l}{\delta \xi} \right]_q^* + \Psi_q^* \frac{\delta l}{\delta q} \quad (3.9)$$

*coupled with the equation  $\dot{q} = \Psi_q \xi$ .*

For the early development of these equations in the finite-dimensional setting see [43] and [27].

*Proof.* The equivalence of (i) and the weak form of the Euler–Lagrange equations (3.4) is proved by integration by parts:

$$\delta \int_a^b L(q, \dot{q}) dt = \int_a^b \left( \left\langle \frac{\delta L}{\delta q}, \delta q \right\rangle + \left\langle \frac{\delta L}{\delta \dot{q}}, \delta \dot{q} \right\rangle \right) dt = \int_a^b \left\langle \frac{\delta L}{\delta q} - \frac{d}{dt} \frac{\delta L}{\delta \dot{q}}, \delta q \right\rangle dt.$$

The strong form of the Euler–Lagrange equations (3.5) follows easily from a standard contradiction argument,<sup>5</sup> which finishes the proof of the equivalence of (i) and (ii).

To prove the equivalence of (i) and (iii), we first compute the quantities  $\delta \dot{q}$  and  $d(\delta q)/dt$ . Recall that

$$\delta q(t) = \frac{\partial}{\partial \tau} \Big|_{\tau=0} \beta(t, \tau) = \Psi_{q(t)} \eta(t), \quad \text{where } \eta(t) \in W.$$

Using the definition (3.3) of the field  $\delta q$ , one concludes that

$$\delta \Psi_{q(t)} = \frac{\partial}{\partial \tau} \Big|_{\tau=0} \Psi_{\beta(t, \tau)} = \delta q(t) [\Psi_{q(t)}] = (\Psi_{q(t)} \eta(t)) [\Psi_{q(t)}]. \quad (3.10)$$

Hereafter,  $v[f]$  denotes the derivative of the function  $f$  along the vector field  $v$ ; in particular, in (3.10) an operator-valued function is differentiated.

Similarly,

$$\frac{d}{dt} \Psi_{q(t)} = \dot{q}(t) [\Psi_{q(t)}] = (\Psi_{q(t)} \xi(t)) [\Psi_{q(t)}],$$

and therefore

$$\delta \dot{q} = (\Psi_q \eta) [\Psi_q] \xi + \Psi_q \delta \xi, \quad \frac{d}{dt} \delta q = (\Psi_q \xi) [\Psi_q] \eta + \Psi_q \dot{\eta}.$$

From  $\delta \dot{q} = \frac{d}{dt} \delta q$ , we obtain

$$\Psi_q (\delta \xi - \dot{\eta}) = (\Psi \xi) [\Psi \eta](q) - (\Psi \eta) [\Psi \xi](q) = [\Psi \xi, \Psi \eta](q) = \Psi_q [\xi, \eta]_q,$$

which implies formula (3.7).

To prove the equivalence of (iii) and the weak form of Hamel’s equations (3.8), we use the above formula and compute the variation of the action (3.6):

$$\begin{aligned} \delta \int_a^b l(q, \xi) dt &= \int_a^b \left( \left\langle \frac{\delta l}{\delta q}, \delta q \right\rangle + \left\langle \frac{\delta l}{\delta \xi}, \delta \xi \right\rangle \right) dt \\ &= \int_a^b \left( \left\langle \frac{\delta l}{\delta q}, \Psi_q \eta \right\rangle + \left\langle \frac{\delta l}{\delta \xi}, \dot{\eta} + [\xi, \eta]_q \right\rangle \right) dt \\ &= \int_a^b \left\langle \Psi_q^* \frac{\delta l}{\delta q} + \left[ \xi, \frac{\delta l}{\delta \xi} \right]_q^* - \frac{d}{dt} \frac{\delta l}{\delta \xi}, \eta \right\rangle dt. \end{aligned}$$

If  $i_* T_q Q$  is dense in  $T_q M$  for every  $q \in Q$ , then for each  $t$  the subspace  $\Psi_{q(t)}^{-1}(i_* T_{q(t)} Q)$  is dense in  $W$ , and the variational derivative vanishes if and only if the strong form of the Hamel equations (3.9) is satisfied.  $\square$

**Example 3.2.** For an incompressible fluid flow in a compact domain  $D \subset \mathbb{R}^3$  with a smooth boundary the configuration space is the group  $\text{Diff}(D)$  of (volume-preserving) diffeomorphism of  $D$ , which is a regular Lie group in convenient setting. Let  $q(t)$  be a curve in this group, one may think of  $q(t)$  as a particular fluid flow. Following Euler [19, 20, 21], one typically uses the *spatial velocity*  $\xi := \dot{q} \circ q^{-1} \in T_e \text{Diff}(D)$ . Selecting

<sup>5</sup> If  $\frac{\delta L}{\delta q} - \frac{d}{dt} \frac{\delta L}{\delta \dot{q}} \neq 0$  at some  $t_0$ , and  $i_* T_{q(t_0)} Q$  is dense in  $T_{q(t_0)} M$ , there exists  $X \in T_{q(t_0)} Q$  such that  $\langle \frac{\delta L}{\delta q} - \frac{d}{dt} \frac{\delta L}{\delta \dot{q}}, X \rangle > 0$ . Using continuity, it is straightforward to construct a variation of the curve  $q(t)$  for which  $\int_a^b \langle \frac{\delta L}{\delta q} - \frac{d}{dt} \frac{\delta L}{\delta \dot{q}}, \delta q(t) \rangle dt > 0$ , which is a contradiction.

$W = T_e \text{Diff}(D) = \mathcal{X}(D)$ , the space of smooth vector fields on  $D$  tangent to the boundary, and  $\Psi_q = TR_q$  gives  $\dot{q} = \Psi_q \xi$ . Therefore, the use of spatial velocity in fluid dynamics is an instance of infinite-dimensional Hamel's formalism. The variation formula (3.7) becomes

$$\delta\xi = \dot{\eta} - \text{ad}_\xi \eta \equiv \dot{\eta} + [\xi, \eta],$$

where  $[\cdot, \cdot]$  is the Jacobi–Lie bracket on  $D$ . The dynamics, in the form of Hamel's equations, reads

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} + \text{ad}_\xi^* \frac{\delta l}{\delta \xi} = 0.$$

The latter are the Euler–Poincaré equations, as established by Poincaré [43] and Arnold [1].

**Remark.** The *convective representation* of fluid dynamics is straightforward to obtain by setting  $\Psi_q = TL_q$ . See Gay-Balmaz *et al.* [23] for details on the convective representation in continuum mechanics.

**Example 3.3.** For an inextensible string moving in the plane the configuration manifold is the space of smooth embeddings  $\text{Emb}([0, 1], \mathbb{R}^2)$ . We will view  $\mathbb{R}^2$  as a complex plane.

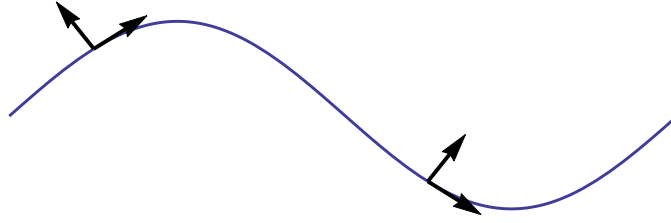


Figure 3.1: An inextensible planar string.

Given  $z \in \text{Emb}([0, 1], \mathbb{C})$ , the *inextensibility constraint* reads  $\|z_s\| = 1$ ,  $0 \leq s \leq 1$ . For simplicity, we assume no resistance to bending. Therefore, the Lagrangian reads

$$L(z) = \int_0^1 \frac{1}{2} (\|\dot{z}\|^2 - \lambda(\|z_s\|^2 - 1)) ds,$$

where  $\lambda : [0, 1] \rightarrow \mathbb{R}$  is the Lagrange multiplier (tension). The boundary conditions for the Lagrange multiplier are a part of the requirement  $\delta L = 0$ . For a free motion of a string, these conditions read

$$\lambda|_{s=0} = \lambda|_{s=1} = 0. \quad (3.11)$$

Let

$$\dot{z} = \Psi_z \xi = z_s \xi, \quad (3.12)$$

so the velocity components to be used to construct Hamel's equations are represented by a complex-valued function  $\xi = \xi(s, t)$ . The real and imaginary parts of  $\xi$  are the tangent and normal velocity components of the points of the string.

The Lagrangian becomes

$$l = \int_0^1 \frac{1}{2} (\bar{z}_s z_s \bar{\xi} \xi - \lambda(\bar{z}_s z_s - 1)) ds,$$

in which the density should be understood as a function of  $(z_s, \bar{z}_s, \xi, \bar{\xi})$  and the Lagrange multiplier  $\lambda$ .

Next, the bracket formula (3.2) for the string becomes

$$[\Psi \xi, \Psi \eta](z) = \frac{d}{d\tau} \Big|_{\tau=0} ((z + \tau z_s \xi)_s \eta - (z + \tau z_s \eta)_s \xi) = z_s (\xi_s \eta - \eta_s \xi) = \Psi_z [\xi, \eta]_z.$$

That is,

$$[\xi, \eta]_z = \xi_s \eta - \xi \eta_s. \quad (3.13)$$

Instead of establishing the formulae for the dual bracket and dual operator  $\Psi^*$ , it is more efficient in this example to directly work with the variational principle. We have:

$$\frac{\delta l}{\delta z} \delta z + \frac{\delta l}{\delta \xi} \delta \xi + \frac{\delta l}{\delta \bar{z}} \delta \bar{z} + \frac{\delta l}{\delta \bar{\xi}} \delta \bar{\xi}, \quad (3.14)$$

and since  $l$  is real-valued, the two last terms are obtained from the first two by conjugation. Thus, it is sufficient to evaluate the last two terms in (3.14):

$$\begin{aligned} \frac{\delta l}{\delta \bar{z}} \delta \bar{z} + \frac{\delta l}{\delta \bar{\xi}} \delta \bar{\xi} &= \frac{\delta l}{\delta \bar{z}} \delta \bar{z} + \frac{\delta l}{\delta \bar{\xi}} (\bar{\xi}_s \bar{\eta} - \bar{\xi} \bar{\eta}_s) - \frac{d}{dt} \frac{\delta l}{\delta \bar{\xi}} \bar{\eta} \\ &= \int_0^1 \frac{1}{2} \left( (\lambda z_s - z_s \bar{\xi} \xi)_s \delta \bar{z} + \bar{z}_s z_s \xi (\bar{\xi}_s \bar{\eta} - \bar{\xi} \bar{\eta}_s) - \frac{d}{dt} (\bar{z}_s z_s \xi) \bar{\eta} \right) ds - \frac{1}{2} \bar{z}_s z_s (\lambda - \bar{\xi} \xi) \bar{\eta} \Big|_{s=0}^{s=1} \\ &= \int_0^1 \frac{1}{2} \left( z_s \bar{z}_{ss} \xi \bar{\xi} + z_s \bar{z}_s \xi \bar{\xi}_s + \lambda_s \bar{z}_s z_s + \lambda z_{ss} \bar{z}_s - \frac{d}{dt} (\bar{z}_s z_s \xi) \right) \bar{\eta} ds - \frac{1}{2} \bar{z}_s z_s \lambda \bar{\eta} \Big|_{s=0}^{s=1}, \end{aligned}$$

which, after imposing the constraint

$$\bar{z}_s z_s = 1, \quad (3.15)$$

implies *Hamel's string equation*

$$\dot{\xi} = \xi \bar{\xi}_s + \lambda_s + i\kappa(\lambda - \bar{\xi} \xi) \quad (3.16)$$

as well as the tension conditions (3.11). Here  $\kappa = iz_s \bar{z}_{ss}$  is the signed curvature of the curve  $[0, 1] \ni s \mapsto z(s) \in \mathbb{C}$ . One of the implications of (3.16) is

$$\lambda_s = \text{Re}(\dot{\xi}_t - \xi \bar{\xi}_s), \quad (3.17)$$

it will be used to identify the Lagrange multiplier  $\lambda$  in terms of the state of the string.

Using the identity

$$\xi_s + \bar{\xi}_s + \bar{z}_s z_{ss} (\xi - \bar{\xi}) = 0 \quad (3.18)$$

that follows from (3.12) and (3.15), one obtains an alternative representation of equation (3.16),

$$\dot{\xi} = -\xi \xi_s + \lambda_s + i\kappa(\lambda - \xi \xi). \quad (3.19)$$

Either of the equations (3.16) and (3.19) is equivalent to the Euler–Lagrange equations for the string.

The Lagrange multiplier as a function of string's state is obtained by solving the differential equation

$$\lambda_{ss} - \kappa^2 \lambda + |\xi_s + i\kappa \xi|^2 = 0 \quad (3.20)$$

subject to the boundary conditions (3.11). To derive equation (3.20), one differentiates (3.17) with respect to  $s$  to obtain

$$\lambda_{ss} = -|\xi_s|^2 + \text{Re}(\xi_{st} - \xi \bar{\xi}_{ss}). \quad (3.21)$$

Formula (3.21) becomes (3.20) after a transformation using the identity

$$\text{Re}(\xi_{st} - \xi \bar{\xi}_{ss} + i\kappa(\xi_t - \bar{\xi} \xi_s)) = 0$$

that follows from (3.12) and (3.15).

**Remark.** Alternatively, one defines the operator  $\Psi$  by

$$\Psi_z \xi := \frac{z_s}{|z_s|} \xi.$$

Because of the constraint  $|z_s| = 1$ , the resulting Hamel equation is (3.16). However, the bracket  $[\xi, \eta]_z$  is now given not by (3.13) but by a slightly different formula. This latter bracket is in fact induced by the (standard) bracket of the Lie algebra of the infinite-dimensional group

$$G = \{[0, 1] \ni s \mapsto g(s) \in \text{SE}(2)\}.$$

To see that, the string dynamics should be interpreted as a motion on  $G$  specified by the degenerate Lagrangian

$$l = \int_0^1 \frac{1}{2} (\bar{\xi} \xi - \lambda (\bar{z}_s z_s - 1)) ds$$

subject to the constraint

$$\frac{z_s}{|z_s|} = e^{i\theta}, \quad (3.22)$$

where  $(\theta, z)$  are the (standard) coordinates on  $\text{SE}(2) = \text{SO}(2) \otimes \mathbb{C}$  and  $\xi$  is the  $\mathbb{C}$ -component of the body velocity  $g^{-1}\dot{g}$ . One then composes the  $\mathbb{C}$ -component of the Euler–Poincaré equations on  $G$  and imposes constraint (3.22). This results in the string equations (3.16). It is worth noticing that the derivation of these equations using formula (3.12) is simpler and more efficient.

### 3.3 The Hamilton–Pontryagin Principle

Here we utilize the Hamilton–Pontryagin principle of Yoshimura and Marsden [50, 51, 52] for the derivation of Hamel’s equations. This approach provides an alternative interpretation of Hamel’s formalism in general, and of the bracket term in particular. The Hamilton–Pontryagin principle for finite-dimensional systems of Hamel type was introduced in Ball, Zenkov, and Bloch [3].

Recall that the Hamilton–Pontryagin principle identifies the trajectories for the Lagrangian  $L : TQ \rightarrow \mathbb{R}$  with the critical points of the functional

$$\int_a^b (L(q, v) + \langle p, \dot{q} - v \rangle) dt,$$

and so the curves the functional is calculated along belong to the Whitney sum  $TQ \oplus T^*Q$ . See [50, 51, 52] for details, motivation, and history.

**Theorem 3.4 (The Hamilton–Pontryagin Principle for Hamel’s Equations).** *Let  $L : P \rightarrow \mathbb{R}$  be a Lagrangian and  $l$  be its representation in local coordinates  $(q, \xi)$  on  $P$ . Then, the following statements are equivalent:*

- (i) *The curve  $(q(t), \xi(t), \mu(t)) \in V \times (W \oplus W^*)$  is a critical point of the action functional*

$$\int_a^b (l(q, \xi) + \langle \mu, \Psi_q^{-1} \dot{q} - \xi \rangle) dt$$

*with respect to independent variations  $\delta q = \Psi_q \eta$ ,  $\delta \xi$ , and  $\delta \mu$ , with  $\eta(a) = \eta(b) = 0$ .*

- (ii) *The curve  $(q(t), \xi(t), \mu(t))$  satisfies the **weak form of the implicit Hamel equations***

$$\int_a^b \left\langle \Psi_q^* \frac{\delta l}{\delta q} - \dot{\mu} + [\xi, \mu]_q^*, \eta \right\rangle = 0, \quad \frac{\delta l}{\delta \xi} = \mu, \quad \text{where } \eta \in \Psi_q^{-1}(T_q Q),$$

*together with the constraint  $\xi = \Psi_q^{-1} \dot{q}$ .*

If  $i_*T_qQ$  is dense in  $T_qM$  for every  $q \in Q$ , the curve  $(q(t), \xi(t), \mu(t))$  satisfies the **strong form of the implicit Hamel equations**

$$\Psi_q^* \frac{\delta l}{\delta q} - \dot{\mu} + [\xi, \mu]_q^* = 0, \quad \frac{\delta l}{\delta \xi} = \mu,$$

together with the constraint  $\xi = \Psi_q^{-1} \dot{q}$ .

*Proof.* From  $\Psi_q^{-1} \Psi_q = \text{id}$  we have

$$\delta(\Psi_q^{-1} \Psi_q) = \delta \Psi_q^{-1} \Psi_q + \Psi_q^{-1} \delta \Psi_q = 0,$$

which implies

$$\delta \Psi_q^{-1} = -\Psi_q^{-1} \delta \Psi_q \Psi_q^{-1}.$$

Therefore, using earlier calculation,

$$\begin{aligned} \delta(\Psi_q^{-1} \dot{q}) &= (\delta \Psi_q^{-1}) \dot{q} + \Psi_q^{-1} \delta \dot{q} \\ &= -\Psi_q^{-1} \delta \Psi_q \Psi_q^{-1} \dot{q} - \left(\frac{d}{dt} \Psi_q^{-1}\right) \delta q + \frac{d}{dt} (\Psi_q^{-1} \delta q) \\ &= -\Psi_q^{-1} \delta \Psi_q \Psi_q^{-1} \dot{q} + \Psi_q^{-1} \frac{d}{dt} \Psi_q \Psi_q^{-1} \delta q + \dot{\eta} \\ &= [\Psi_q^{-1} \dot{q}, \eta]_q + \dot{\eta}, \end{aligned}$$

where  $\eta = \Psi_q^{-1} \delta q \in \Psi_q^{-1}(T_qQ)$ .

Finally, taking the variation and using the above formula along with integration by parts gives

$$\begin{aligned} &\delta \int_a^b (l(q, \xi) + \langle \mu, \Psi_q^{-1} \dot{q} - \xi \rangle) dt \\ &= \int_a^b \left( \left\langle \frac{\delta l}{\delta q}, \delta q \right\rangle + \left\langle \frac{\delta l}{\delta \xi} - \mu, \delta \xi \right\rangle - \langle \delta \mu, \xi - \Psi_q^{-1} \dot{q} \rangle + \langle \mu, \delta(\Psi_q^{-1} \dot{q}) \rangle \right) dt \\ &= \int_a^b \left( \left\langle \frac{\delta l}{\delta q}, \Psi_q \eta \right\rangle - \langle \dot{\mu}, \eta \rangle + \langle \mu, [\Psi_q^{-1} \dot{q}, \eta]_q \rangle + \left\langle \frac{\delta l}{\delta \xi} - \mu, \delta \xi \right\rangle - \langle \delta \mu, \xi - \Psi_q^{-1} \dot{q} \rangle \right) dt \\ &= \int_a^b \left( \left\langle \Psi_q^* \frac{\delta l}{\delta q} - \dot{\mu} + [\Psi_q^{-1} \dot{q}, \mu]_q^*, \eta \right\rangle + \left\langle \frac{\delta l}{\delta \xi} - \mu, \delta \xi \right\rangle - \langle \delta \mu, \xi - \Psi_q^{-1} \dot{q} \rangle \right) dt. \end{aligned}$$

Setting the latter equal to zero and noting that  $\delta \xi$  is an arbitrary element of  $W$  yields the desired results.  $\square$

## 4 Systems with Constraints

In this section we consider systems with ideal velocity constraints, both holonomic and nonholonomic. The dynamics of such systems is equivalent to the Lagrange–d'Alembert principle. To simplify the exposition, in the rest of the paper we assume that  $i_*T_qQ$  is dense in  $T_qM$  for every  $q \in Q$ , and thus all results will be stated for strong equations of motion. Similar statements for weak equations are straightforward to obtain.

### 4.1 The Lagrange–d'Alembert Principle

Recall that in this paper the constraints imposed on the system are assumed linear and homogeneous in the velocity, and thus are specified by a distribution  $\mathcal{D}$  on the manifold  $M$ . This distribution will, in general, be nonintegrable. Verification of integrability of a distribution in the infinite-dimensional setting may be nontrivial as the Frobenius theorem has not been established in the general infinite-dimensional setting. See [34, 30, 46, 22] for some of the infinite-dimensional versions of Frobenius theorem.

The condition for a curve to satisfy the constraints is of course insufficient for the development of constrained mechanics. One needs a mechanism for constructing the vector field that captures the dynamics of the constrained system. For the ideal constraints in the finite-dimensional setting, this is accomplished by a projection. Such constraints define a submanifold of the velocity phase space and a projection onto this submanifold.

For a projection to be meaningful in the infinite-dimensional case, a submanifold has to be splitting. Thus, in the infinite-dimensional setting, we require that  $\mathcal{D}$  is a **locally splitting subbundle** of  $P$ . That is, for each  $q \in Q$  there exists a chart  $(U, h)$  of  $M$  with  $i(q) \in U$  such that

$$Th(\pi_M^{-1}(U) \cap \mathcal{D}) = h(U) \times W^{\mathcal{D}},$$

where the closed subspace  $W^{\mathcal{D}}$  of the model space  $W$  is **splitting, or complemented**, i.e., there is a closed subspace  $W^{\mathcal{U}}$  of  $W$  such that  $W^{\mathcal{D}} \oplus W^{\mathcal{U}} = W$  and the projection  $\pi^{\mathcal{D}}$  uniquely determined by setting  $(\text{Ker } \pi^{\mathcal{D}}, \text{Im } \pi^{\mathcal{D}}) = (W^{\mathcal{U}}, W^{\mathcal{D}})$  is continuous. Note that the continuous projection  $\pi^{\mathcal{D}}$  in the convenient space  $W$  is automatically smooth as a bounded linear mapping.<sup>6</sup>

The following Lagrange–d’Alembert principle is known to be equivalent to the dynamics of systems with ideal holonomic and nonholonomic constraints:

**Definition 4.1.** *The **Lagrange–d’Alembert equations of motion** for the system are those determined by*

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

where we choose variations  $\delta q(t)$  of the curve  $q(t)$  that satisfy  $\delta q(a) = \delta q(b) = 0$  and  $\delta q(t) \in \mathcal{D}_{q(t)}$  for each  $t$  where  $a \leq t \leq b$ .

This principle is supplemented by the condition that the curve  $q(t)$  itself satisfies the constraints. Note that we take the variation *before* imposing the constraints; that is, we do not impose the constraints on the family of curves defining the variation. This is well known to be important to obtain the correct mechanical equations (see [2] and [6] for a discussion of other types of constraints and references).

The Lagrange–d’Alembert principle is equivalent to the equations

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} \in \mathcal{D}_q^\circ, \quad \dot{q} \in \mathcal{D}_q. \quad (4.1)$$

Here,

$$\mathcal{D}_q^\circ = \{a \in T_q^*M \mid \langle a, v \rangle = 0, v \in \mathcal{D}_q\}.$$

One way to give a more explicit representation of dynamics (4.1), under certain technical conditions, is to make use of the **Euler–Lagrange equations with multipliers**. Let the distribution  $\mathcal{D}$  be locally written as

$$\mathcal{D} = \{\dot{q} \in TM \mid A(q)\dot{q} = 0\},$$

where  $A(q)$  for each  $q \in Q$  is a continuous linear operator on  $T_qM$  with values in the vector space  $W^{\mathcal{U}}$ .

For instance, if there is a single constraint imposed on the system,  $A(q)$  is a linear functional. We assume smooth dependence  $q \mapsto A(q)$ .

The **constrained variations**  $\delta q(t) \in T_{q(t)}Q$  satisfy the condition

$$A(q)\delta q = 0. \quad (4.2)$$

Using Definition 4.1 and formula (4.2), one writes the equations of motion as

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} \in \overline{\text{Im } A^*(q)}, \quad A(q)\dot{q} = 0,$$

---

<sup>6</sup>Nonsplitting closed subspaces already exist in Banach spaces; for more information on splitting subspaces, see Domański and Mastyło [14] and references therein. The continuity of  $\pi^{\mathcal{D}}$  in a Banach space is a consequence of the closed graph theorem. For more spaces with this property see Jarchow [32].

where  $A^*(q) : W^{\mathcal{U}^*} \rightarrow T_q^*M$  is the adjoint of  $A(q)$  and where the closure should be understood as weak\* closure. Thus, if  $A^*(q)$  has a weak\* closed range,<sup>7</sup> there exist **Lagrange multipliers**  $\lambda \in W^{\mathcal{U}^*}$  such that

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} = A^*(q) \lambda, \quad A(q) \dot{q} = 0.$$

## 4.2 The Constrained Hamel Equations

Given a nonholonomic system, that is, a Lagrangian  $L : P \rightarrow \mathbb{R}$  and constraint distribution  $\mathcal{D}$ , select the operators  $\Psi_q : W \rightarrow T_qM$ ,  $q \in U \subset Q$ , such that there exist closed subspaces  $W^{\mathcal{D}}, W^{\mathcal{U}} \subset W$  with the properties  $W = W^{\mathcal{D}} \oplus W^{\mathcal{U}}$  and  $\Psi_q = \Psi_q^{\mathcal{D}} \oplus \Psi_q^{\mathcal{U}}$ , where  $\Psi_q^{\mathcal{D}} : W^{\mathcal{D}} \rightarrow \mathcal{D}_q$  and  $\Psi_q^{\mathcal{U}} : W^{\mathcal{U}} \rightarrow \mathcal{U}_q$  and their inverses are bounded linear operators smoothly dependent on  $q \in U$ . One way to choose the operators  $\Psi_q$  is to use the above subbundle chart. In general,  $U \neq Q$ , as numerous finite-dimensional examples demonstrate.

Each  $\dot{q} \in TM$  is then uniquely written as

$$\dot{q} = \Psi_q \xi^{\mathcal{D}} + \Psi_q \xi^{\mathcal{U}}, \quad \text{where } \Psi_q \xi^{\mathcal{D}} \in \mathcal{D}_q, \quad (4.3)$$

i.e.,  $\Psi_q \xi^{\mathcal{D}}$  is the component of  $\dot{q}$  along  $\mathcal{D}_q$ . Similarly, each  $\alpha \in W^*$  can be uniquely decomposed as

$$\alpha = \alpha_{\mathcal{D}} + \alpha_{\mathcal{U}},$$

where  $\alpha_{\mathcal{D}}$  and  $\alpha_{\mathcal{U}}$  denote the component of  $\alpha$  along the duals of  $W^{\mathcal{D}}$  and  $W^{\mathcal{U}}$ , respectively. Actually, we have

$$\alpha_{\mathcal{D}} = (\pi^{\mathcal{D}})^* \circ \alpha|_{W^{\mathcal{D}}} \quad \text{and} \quad \alpha_{\mathcal{U}} = \left( \text{id} - (\pi^{\mathcal{D}})^* \right) \circ \alpha|_{W^{\mathcal{U}}},$$

where  $(\pi^{\mathcal{D}})^*$  is the adjoint of  $\pi^{\mathcal{D}}$ . Using (4.3), the constraints read

$$\xi = \xi^{\mathcal{D}} \quad \text{or} \quad \xi^{\mathcal{U}} = 0.$$

This implies

$$\delta \xi = \delta \xi^{\mathcal{D}} \quad \text{or} \quad \delta \xi^{\mathcal{U}} = 0.$$

The Lagrange–d’Alembert principle then implies the following theorem:

**Theorem 4.2.** *The dynamics of a nonholonomic system is represented by the **strong form of constrained Hamel equations***

$$\left( \frac{d}{dt} \frac{\delta l}{\delta \xi} - \left[ \xi^{\mathcal{D}}, \frac{\delta l}{\delta \xi} \right]_q^* - \Psi_q^* \frac{\delta l}{\delta q} \right)_{\mathcal{D}} = 0, \quad \xi^{\mathcal{U}} = 0, \quad \dot{q} = \Psi_q \xi^{\mathcal{D}}. \quad (4.4)$$

**Example 4.3.** Consider an inextensible string moving in the plane subject to the vanishing normal velocity constraint. See Figure 3.1. One may think of a motion of a ‘sharp’ string on the horizontal ice. Using the notations introduced in Example 3.3, the constraint reads

$$\xi = \bar{\xi}, \quad (4.5)$$

i.e.,  $\xi \in \mathbb{R}$ . Equations (4.4) for the constrained string thus become

$$\dot{\xi} = \xi \xi_s + \lambda_s, \quad (4.6)$$

$$\dot{z} = z_s \xi, \quad (4.7)$$

supplemented by the inextensibility condition.

Identity (3.18) in the presence of constraint (4.5) implies

<sup>7</sup>This holds if  $T_qM$  is a Fréchet space and  $A(q)$  has closed range, see Jarchow [32] for details.



$$\xi_s = 0.$$

That is, all points of the string have the same speed, and (4.6) becomes

$$\dot{\xi} = \lambda_s. \tag{4.8}$$

Using the boundary conditions (3.11), we conclude that

$$\dot{\xi} = 0.$$

Summarizing,  $\xi = \text{const}$  throughout the motion. This is in agreement with the motion of the Chaplygin sleigh for which the velocity of the contact point relative to the body frame is constant. Unlike the sleigh, the constrained string motion is not completely determined by the initial state: Any solution of (4.7) is of the form

$$z = \phi(s + \xi t),$$

where  $\phi$  is an arbitrary twice-differentiable complex-valued function. The initial conditions define  $\phi$  on the segment  $[0, 1]$ . Outside this segment, the function  $\phi$  is unknown, unless, for example, the motion of the front end of the string is prescribed. The motion of the constrained string is therefore *purely kinematic*: The string follows its front end at a constant speed.

This behavior is similar to that of the degenerate Chaplygin sleigh specified by the Lagrangian  $l = \frac{1}{2}\xi\bar{\xi}$  and the constraint  $\xi = \bar{\xi}$ , where  $\xi = e^{-i\theta}\dot{z}$ . The Lagrangian is degenerate as the term quadratic in  $\theta$  is absent, i.e., the moment of inertia of the sleigh equals zero.

For the degenerate sleigh, the dynamics reads

$$\dot{\xi} = 0, \quad \dot{z} = e^{i\theta}\xi,$$

where  $\theta(t)$  is an arbitrary function. Thus, the motions are not identified by the initial conditions.<sup>8</sup>

Another remark about the string on ice is that, similar to SE(2)-snakes of Krishnaprasad and Tsakiris [35], it is *almost holonomic*: Adding a single (holonomic) constraint  $z(0, t) = z(1, t)$  renders the system holonomic. The string becomes closed and will slide, at a constant speed, along its initial shape.

## 5 Systems with Symmetry

Here we present infinite-dimensional analogues of some of the results of Bloch *et al.* [7, 9] on systems with symmetry. The mechanical and nonholonomic connections, the momentum equation, and mechanics of systems with infinitely-many velocity constraints are discussed. Recall that  $i_*T_qQ$  is dense in  $T_qM$  for every  $q \in Q$ , and thus all results are stated for strong equations of motion.

### 5.1 The Lagrange–Poincaré Equations

Recall that in finite-dimensional mechanics with symmetry one starts with an action of a Lie group  $G$  on the configuration space  $Q$  and Lagrangian and constraints (if any) that are invariant with respect to the lifted action on the velocity phase space. The quotient space  $Q/G$ , whose points are the group orbits, is called the *shape space*. It is known that if the group action is free and proper, the shape space is a smooth manifold and the projection  $\pi : Q \rightarrow Q/G$  is a smooth surjective map with a surjective derivative at each point. The configuration space thus has the structure of a principal fiber bundle, with the group acting on the fibers by (left) multiplication.

We thus assume that in the general infinite-dimensional setting in the presence of symmetry the manifold  $M$  is a principal fiber bundle (see Kriegl and Michor [34] for details in the infinite-dimensional case). The base of this bundle is a smooth manifold, and the group  $G$  involved in the construction of this bundle

---

<sup>8</sup>Unlike the regular sleigh, for which  $\dot{\theta} = \text{const}$ .

may be finite- or infinite-dimensional. In the latter case, the group is assumed regular (see Kriegl and Michor [34] and Omori [41]).<sup>9</sup>

We denote the bundle coordinates  $(r, g)$  where  $r$  is a local coordinate in the base, or shape space  $M/G$ , and  $g$  is a group coordinate. Such a local trivialization is characterized by the fact that in such coordinates the group does not act on the factor  $r$  but acts on the group coordinate by group multiplication. Thus, locally in the base, the space  $M$  is isomorphic to the product  $M/G \times G$ , and in this local trivialization the map  $\pi$  becomes the projection onto the first factor. The model space of the manifold  $M/G$  is denoted  $W_B$ .

Recall that in general in the infinite-dimensional setting Lagrangians are defined on the bundle  $P = i^*TM$ , where  $i : Q \rightarrow M$  is the initial inclusion map and  $Q$  is a manifold. We will assume in the rest of the paper that  $Q$  is a principal fiber bundle with the same group  $G$  and that in the aforementioned local trivialization the group component of the initial inclusion map is the identity map.

**Definition 5.1.** *We say that the Lagrangian is **G**-invariant if  $L$  is invariant under the induced action of  $G$  on  $P$ .*

Both left and right actions may be of interest. Below we develop the formalism assuming that the action is left; the case of the right action is similar.

We start with the construction of the operator  $\Psi$ . Let  $\mathfrak{g}$  be the Lie algebra, and thus the model space, of the group  $G$ . When the configuration space is a principal fiber bundle, the group component of the operator  $\Psi$  can be defined globally, as shown below. For the fibers of  $T(M/G)$  we write the corresponding operator as  $\psi_r : W_B \rightarrow T_r(M/G)$ , i.e.,  $\dot{r} = \psi_r \xi$ , whereas for the fibers of  $TG$  we set  $\dot{g} := L_{g*} \circ \varphi_r \zeta$ , where  $q = (r, g)$  in a local trivialization and  $\varphi_r : \mathfrak{g}_r \rightarrow \mathfrak{g}$  is a linear operator. The operators  $\psi_r$  and  $\varphi_r$  and their inverses are, for each  $r \in Q/G$ , bounded, continuous, and depend smoothly on  $r$ . This implies that the algebras  $\mathfrak{g}$  and  $\mathfrak{g}_r$  are isomorphic. The bracket on the space  $W_B$  is

$$[\xi_1, \xi_2]_r := \psi_r^{-1}[\psi_r \xi_1, \psi_r \xi_2],$$

whereas the bracket on  $\mathfrak{g}_r$  is defined by

$$[\zeta_1, \zeta_2]_{\mathfrak{g}_r} := \varphi_r^{-1}[\varphi_r \zeta_1, \varphi_r \zeta_2]_{\mathfrak{g}}.$$

The use of the nonmaterial shape velocity is of importance already in the finite-dimensional setting in some of the problem involving the rolling rigid body, such as the rattleback. In a number of interesting finite-dimensional instances it suffices to use the shape velocity  $\dot{r}$ . The version of the infinite-dimensional formalism developed here is motivated by the utility of nonmaterial velocity in accounting for fluid- and elastic-type shapes.

The operator  $\varphi_r$  is important for accounting for various subbundle structures associated with subspaces of the Lie algebra  $\mathfrak{g}$ . Such structures originate in the presence of control inputs and/or velocity constraints, particularly in the motion generation problems (see Ostrowski *et al.* [42] and Bloch *et al.* [7] for the details in the finite-dimensional setting). In general, the position of these subspaces in the Lie algebra  $\mathfrak{g}$  is shape-dependent. While this operator may not appear to be important in the current part of the paper, it is more straightforward to introduce it here in order to make the exposition of the rest of the paper more straightforward.

Let  $\mathcal{A}_s$  be a principal connection on the bundle  $Q \rightarrow Q/G$  (see Kriegl and Michor [34] for details on principal connections in the infinite-dimensional setting). Below we will discuss how the structure of the Lagrangian facilitates the selection of a connection. Tangent vectors in a local trivialization  $Q = Q/G \times G$  at the point  $(r, g)$  are denoted  $(x, y)$ . We write the action of  $\mathcal{A}_s$  on this vector as  $\mathcal{A}_s(x, y)$ .<sup>10</sup> Using this notation, we write the connection form in the local trivialization as

$$\mathcal{A}_s(x, y) = \text{Ad}_g(y_c + \mathcal{A}_c x),$$

<sup>9</sup>According to Kriegl and Michor [34], all finite-dimensional Lie groups and all known infinite-dimensional Lie groups are regular.

<sup>10</sup>Here and below, the subscripts ‘c’ and ‘s’ stand for ‘convective’ and ‘spatial’, respectively.

where  $y_c$  is the left translation of  $y$  to the identity and, since we are working locally in shape space, we regard  $\mathcal{A}_c$  as a  $\mathfrak{g}$ -valued one-form on  $Q/G$ .

The *curvature* of the principal connection  $\mathcal{A}_s$  is the  $\mathfrak{g}$ -valued two-form

$$\mathcal{B}_s(X_1, X_2) := d\mathcal{A}_s(\text{hor } X_1, \text{hor } X_2),$$

where  $X_1$  and  $X_2$  are two vector fields and  $\text{hor } X_1$  and  $\text{hor } X_2$  are their horizontal components. In a local trivialization one writes

$$\mathcal{B}_s((x_1, y_1), (x_2, y_2)) = \text{Ad}_g \mathcal{B}_c(x_1, x_2). \quad (5.1)$$

where  $x_1, x_2 \in T(Q/G)$ ,  $y_1, y_2 \in \mathfrak{g}$ , and where

$$\mathcal{B}_c(x_1, x_2) = d\mathcal{A}_c(x_1, x_2) - [\mathcal{A}_c x_1, \mathcal{A}_c x_2]_{\mathfrak{g}}.$$

The quantity  $y_c + \mathcal{A}_c x$  regarded as an element of  $\mathfrak{g}_r$  will be denoted  $\zeta + \mathcal{A}x$ , i.e.,

$$y_c + \mathcal{A}_c x = \varphi_r(\zeta + \mathcal{A}x), \quad x \in T_r(Q/G), \quad \zeta \in \mathfrak{g}_r.$$

Thus, the velocity vector  $\dot{q}$  in such a representation is characterized by the components

$$\dot{r} \quad \text{and} \quad \Omega = \zeta + \mathcal{A}\dot{r},$$

where  $\varphi_r \zeta = g^{-1} \dot{g} \in \mathfrak{g}$ . The curvature of the connection  $\mathcal{A}$  on  $Q/G$  in this representation is hereafter denoted  $\mathcal{B}$ .

With nonmaterial shape velocity being used, the velocity components become

$$\xi \quad \text{and} \quad \Omega = \zeta + \mathcal{A}\psi_r \xi. \quad (5.2)$$

Summarizing,

$$\dot{q} = \Psi_q(\xi, \Omega) = (\psi_r \xi, L_{g^*} \varphi_r(\Omega - \mathcal{A}\psi_r \xi)). \quad (5.3)$$

**Theorem 5.2.** *The strong Lagrange–Poincaré equations<sup>11</sup> for a system with a  $G$ -invariant Lagrangian  $L : P \rightarrow \mathbb{R}$  are*

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} - \psi_r^* \frac{\delta l}{\delta r} = \left[ \xi, \frac{\delta l}{\delta \xi} \right]_r^* - \left\langle \frac{\delta l}{\delta \Omega}, \mathbf{i}_\xi \psi_r^* \mathcal{B} + \langle \mathbf{i}_\xi \psi_r^* \gamma, \psi_r^* \mathcal{A} \rangle - \langle \psi_r^* \gamma, \mathbf{i}_\xi \psi_r^* \mathcal{A} \rangle + \langle \psi_r^* \mathcal{E}, \Omega \rangle \right\rangle, \quad (5.4)$$

$$\frac{d}{dt} \frac{\delta l}{\delta \Omega} = \left[ \Omega, \frac{\delta l}{\delta \Omega} \right]_{\mathfrak{g}_r}^* + \left\langle \frac{\delta l}{\delta \Omega}, \mathbf{i}_\xi \psi_r^* \mathcal{E} \right\rangle, \quad (5.5)$$

where  $\mathcal{B}$  is the curvature of  $\mathcal{A}$ ,  $\gamma := \varphi_r^{-1} \circ d\varphi_r$ , and  $\mathcal{E} := \gamma - \text{ad}_{\mathcal{A}}$ , so that  $\gamma$  and  $\mathcal{E}$  are  $\mathfrak{g}_r \otimes \mathfrak{g}_r^*$ -valued one-forms on  $Q/G$ . The **shape equation** (5.4) and the **momentum equation** (5.5) govern the reduced dynamics for the Lagrangian  $L$ . The full dynamics is given by equations (5.4) and (5.5) along with the **kinematic shape equation**

$$\dot{r} = \psi_r \xi$$

and **reconstruction equation**

$$\dot{g} = g\varphi_r(\Omega - \mathbf{i}_\xi \psi_r^* \mathcal{A}). \quad (5.6)$$

*Proof.* The principal step is to uncover the structure of the bracket terms in the Hamel equations. Recall that these terms are induced by the map  $\Psi$  defined in (5.3). Evaluating the Jacobi–Lie brackets of two vector fields  $X_1$  and  $X_2$  whose components in local trivialization are  $(\xi_i, \zeta_i)$ , that is,

$$X_i = (\psi_r \xi_i, L_{g^*} \varphi_r(\zeta_i - \mathcal{A}\psi_r \xi_i)), \quad \xi_i = \text{const}, \quad \zeta_i = \text{const}, \quad i = 1, 2,$$

<sup>11</sup>Often,  $\partial l / \partial \Omega$  is viewed as an independent variable in these equations. See Cendra *et al.* [11, 12], and Bloch *et al.* [9] for the history, motivation, and finite-dimensional version of these equations.

gives:

$$[X_1, X_2] = [(\psi_r \xi_1, 0), (\psi_r \xi_2, 0)] + [(\psi_r \xi_1, 0), (0, L_{g^*} \varphi_r(\zeta_2 - \mathcal{A}\psi_r \xi_2))] \\ + [(0, L_{g^*} \varphi_r(\zeta_1 - \mathcal{A}\psi_r \xi_1)), (\psi_r \xi_2, 0)] + [(0, L_{g^*} \varphi_r(\zeta_1 - \mathcal{A}\psi_r \xi_1)), (0, L_{g^*} \varphi_r(\zeta_2 - \mathcal{A}\psi_r \xi_2))].$$

Next,

$$[(\psi_r \xi_1, 0), (\psi_r \xi_2, 0)] = (\psi_r[\xi_1, \xi_2]_r, 0), \\ [(\psi_r \xi_1, 0), (0, L_{g^*} \varphi_r(\zeta_2 - \mathcal{A}\psi_r \xi_2))] = (0, L_{g^*} d_{\psi_r \xi_1}(\varphi_r(\zeta_2 - \mathcal{A}\psi_r \xi_2))), \\ [(0, L_{g^*} \varphi_r(\zeta_1 - \mathcal{A}\psi_r \xi_1)), (0, L_{g^*} \varphi_r(\zeta_2 - \mathcal{A}\psi_r \xi_2))] = (0, L_{g^*} \varphi_r[\zeta_1 - \mathcal{A}\psi_r \xi_1, \zeta_2 - \mathcal{A}\psi_r \xi_2]_{\mathfrak{g}_r}).$$

Summarizing the components of  $[X_1, X_2]$  are

$$\psi_r[\xi_1, \xi_2]_r$$

and

$$L_{g^*} \varphi_r(\langle \mathbf{i}_{\psi_r \xi_1} \mathcal{E}, \zeta_2 \rangle - \langle \mathbf{i}_{\psi_r \xi_2} \mathcal{E}, \zeta_1 \rangle + \langle \mathbf{i}_{\psi_r \xi_2} \gamma, \mathcal{A}\psi_r \xi_1 \rangle - \langle \mathbf{i}_{\psi_r \xi_1} \gamma, \mathcal{A}\psi_r \xi_2 \rangle) \\ - (\psi_r \xi_1)[\mathcal{A}\psi_r \xi_2] + (\psi_r \xi_2)[\mathcal{A}\psi_r \xi_1] + [\mathcal{A}\psi_r \xi_1, \mathcal{A}\psi_r \xi_2]_{\mathfrak{g}_r} + [\zeta_1, \zeta_2]_{\mathfrak{g}_r}.$$

Applying the inverse of  $\Psi_q$  outputs the components

$$[\xi_1, \xi_2]_r \in W_B$$

and

$$\langle \mathbf{i}_{\psi_r \xi_1} \mathcal{E}, \zeta_2 \rangle - \langle \mathbf{i}_{\psi_r \xi_2} \mathcal{E}, \zeta_1 \rangle + \langle \mathbf{i}_{\psi_r \xi_2} \gamma, \mathcal{A}\psi_r \xi_1 \rangle - \langle \mathbf{i}_{\psi_r \xi_1} \gamma, \mathcal{A}\psi_r \xi_2 \rangle - (\psi_r \xi_1)[\mathcal{A}\psi_r \xi_2] \\ + (\psi_r \xi_2)[\mathcal{A}\psi_r \xi_1] + \mathcal{A}([\psi_r \xi_1, \psi_r \xi_2]) + [\mathcal{A}\psi_r \xi_1, \mathcal{A}\psi_r \xi_2]_{\mathfrak{g}_r} + [\zeta_1, \zeta_2]_{\mathfrak{g}_r} \in \mathfrak{g}_r.$$

Utilizing the formula

$$d\omega(X_1, X_2) = X_1[\omega(X_2)] - X_2[\omega(X_1)] - \omega([X_1, X_2])$$

for the exterior derivative of one-forms, the definition of the curvature, and the Cartan structure equation, the bracket  $[(\xi_1, \zeta_1), (\xi_2, \zeta_2)]_r$  on  $W_B \oplus \mathfrak{g}_r$  reads

$$([\xi_1, \xi_2]_r, \langle \mathbf{i}_{\xi_1} \psi_r^* \mathcal{E}, \zeta_2 \rangle - \langle \mathbf{i}_{\xi_2} \psi_r^* \mathcal{E}, \zeta_1 \rangle + \langle \mathbf{i}_{\xi_2} \psi_r^* \gamma, \mathbf{i}_{\xi_1} \psi_r^* \mathcal{A} \rangle - \langle \mathbf{i}_{\xi_1} \psi_r^* \gamma, \mathbf{i}_{\xi_2} \psi_r^* \mathcal{A} \rangle - \psi_r^* \mathcal{B}(\xi_1, \xi_2) + [\zeta_1, \zeta_2]_{\mathfrak{g}_r}). \quad (5.7)$$

From (5.7), the components of the dual bracket  $[(\xi, \zeta), (\alpha, \beta)]_q^*$  are given by the formulae

$$[\xi, \alpha]_r^* - \langle \beta, \mathbf{i}_{\xi} \psi_r^* \mathcal{B} + \langle \mathbf{i}_{\xi} \psi_r^* \gamma, \psi_r^* \mathcal{A} \rangle - \langle \psi_r^* \gamma, \mathbf{i}_{\xi} \psi_r^* \mathcal{A} \rangle + \langle \psi_r^* \mathcal{E}, \zeta \rangle \rangle \quad \text{and} \quad [\zeta, \beta]_{\mathfrak{g}_r}^* + \langle \beta, \mathbf{i}_{\xi} \psi_r^* \mathcal{E} \rangle. \quad (5.8)$$

Equations (5.4) and (5.5) then follow from Theorem 3.1 and formula (5.8).  $\square$

In certain situations, it may be useful to know the formulae for variations. For systems considered in this part of the paper, the general formula (3.7) transforms into

$$\delta\xi = \dot{\eta} + [\xi, \eta]_r, \\ \delta\Omega = \dot{\Sigma} + [\Omega, \Sigma]_{\mathfrak{g}_r} + \langle \mathbf{i}_{\xi} \psi_r^* \mathcal{E}, \Sigma \rangle - \langle \mathbf{i}_{\eta} \psi_r^* \mathcal{E}, \Omega \rangle + \langle \mathbf{i}_{\eta} \psi_r^* \gamma, \mathbf{i}_{\xi} \psi_r^* \mathcal{A} \rangle - \langle \mathbf{i}_{\xi} \psi_r^* \gamma, \mathbf{i}_{\eta} \psi_r^* \mathcal{A} \rangle - \psi_r^* \mathcal{B}(\xi, \eta),$$

where the curves  $\eta(t) \in W_B$  and  $\varphi_{r(t)} \Sigma(t) \in \mathfrak{g}$  satisfy the boundary conditions  $\eta(a) = \eta(b) = 0$  and  $\Sigma(a) = \Sigma(b) = 0$ , respectively. The reconstruction equation (5.6) is just the group component of the equation  $\dot{q} = \Psi_q \xi$ .

In the case that the material shape velocity is utilized, equations (5.4) and (5.5) read

$$\begin{aligned} \frac{d}{dt} \frac{\delta l}{\delta \dot{r}} - \frac{\delta l}{\delta r} &= - \left\langle \frac{\delta l}{\delta \Omega}, \mathbf{i}_r \mathcal{B} + \langle \mathbf{i}_r \gamma, \mathcal{A} \rangle - \langle \gamma, \mathbf{i}_r \mathcal{A} \rangle + \langle \mathcal{E}, \Omega \rangle \right\rangle, \\ \frac{d}{dt} \frac{\delta l}{\delta \Omega} &= \text{ad}_\Omega^* \frac{\delta l}{\delta \Omega} + \left\langle \frac{\delta l}{\delta \Omega}, \mathbf{i}_r \mathcal{E} \right\rangle. \end{aligned}$$

whereas the reconstruction equation becomes

$$\dot{g} = g\varphi_r(\Omega - \mathbf{i}_r \mathcal{A}).$$

Now assume that the  $G$ -invariant Lagrangian equals the kinetic minus potential energy of the system and that the kinetic energy is given by a (weak) Riemannian metric  $\langle\langle \cdot, \cdot \rangle\rangle$  on the configuration space  $Q$ . In such a setting the isomorphism  $\varphi_r$  is usually selected to be the identity map.

**Definition 5.3.** The *mechanical connection*  $\mathcal{A}^{\text{mech}}$  is, by definition, the connection on  $Q$  regarded as a bundle over shape space  $Q/G$  that is defined by declaring its horizontal space at a point  $q \in Q$  to be the subspace that is the orthogonal complement to the tangent space to the group orbit through  $q \in Q$  using the kinetic energy metric. The *locked inertia tensor*  $\mathcal{I}(q) : \mathfrak{g} \rightarrow \mathfrak{g}^*$  is defined by  $\langle \mathcal{I}(q) \xi, \eta \rangle = \langle\langle \xi_Q(q), \eta_Q(q) \rangle\rangle$ , where  $\xi_Q$  is the infinitesimal generator of  $\xi \in \mathfrak{g}$  and where  $\langle\langle \cdot, \cdot \rangle\rangle$  is the kinetic energy inner product.

Of course, in the infinite-dimensional setting the mechanical connection may fail to exist when the group metric is not strong.<sup>12</sup> It is certain to exist in the important case of a finite-dimensional symmetry group.

Given a system with symmetry, one may use the mechanical connection to set up equations (5.4) and (5.5). This choice of a connection changes the Lie algebra variables from  $\zeta$  to the local version of the *locked group velocity*  $\Omega$ . With this connection choice the kinetic energy metric becomes *block-diagonal*, that is,

$$\langle\langle \dot{q}, \dot{q} \rangle\rangle = \langle\langle \xi, \xi \rangle\rangle + \langle\langle \Omega, \Omega \rangle\rangle.$$

For a rotating rigid body, the locked group velocity has the physical interpretation of the body angular velocity.

**Example 5.4.** Here we revisit the planar string motion, but now derive the dynamics with an emphasis on the  $\text{SE}(2)$  symmetry. We also assume that one of the ends of the string is attached to a platform (a flat rigid body). Both the platform and string move in the same horizontal plane.

The elements of the Euclidean group  $\text{SE}(2)$  are written using complex notation,  $g = (e^{i\theta}, w)$ , i.e., the points of the  $\mathbb{R}^2$  are identified with complex numbers, as before. The group  $\text{SE}(2)$  is the configuration space of the platform and simultaneously the symmetry group of the entire system.

The angular and linear velocities of the platform (measured relative to the platform) are denoted  $\omega \in \mathbb{R}$  and  $v \in \mathbb{C}$ , so that

$$(i\omega, v) := (i\dot{\theta}, e^{-i\theta}\dot{w}) = g^{-1}\dot{g}$$

and the operator  $\varphi$  in (5.3) is the identity operator. The position of the string relative to the platform is given by the embedding map  $z \in \text{Emb}([0, 1], \mathbb{C})$  subject to the conditions  $z(0) = 0$  and  $z_s(0) = 1$ .

The absolute position of the points of the string is

$$w + e^{i\theta}z(s), \tag{5.9}$$

and thus the velocity of the points of the string is given by

$$\dot{w} + e^{i\theta}(iz\dot{\theta} + \dot{z}). \tag{5.10}$$

Written relative to the platform's frame, the absolute velocity of the point  $z$  of the string reads

$$v + i\omega z + \dot{z} \equiv v + i\omega z + z_s \xi.$$

---

<sup>12</sup>The orthogonal complement may not exist in non-Hilbert spaces.

That is, the operator  $\psi$  is defined by  $\dot{z} = \psi_z \xi = z_s \xi$ .

The Lagrangian reads

$$\int_0^1 \frac{1}{2} \left( J\omega^2 + m\bar{v}v + (\bar{v} - i\omega\bar{z} + \bar{z}_s\bar{\xi})(v + i\omega z + z_s\xi) - \lambda(\bar{z}_s z_s - 1) \right) ds,$$

where  $m$  and  $J$  are the mass and inertia of the platform. Recall that the quantities  $\omega$  and  $v$  are  $s$ -independent. The components of the mechanical connection denoted here  $\mathcal{A}_\theta$ ,  $\mathcal{A}_w$ , and  $\mathcal{A}_{\bar{w}}$ , so that

$$\omega = \Omega - \mathcal{A}_\theta \dot{z} - \bar{\mathcal{A}}_\theta \dot{\bar{z}}, \quad v = V - \mathcal{A}_w \dot{z} - \bar{\mathcal{A}}_{\bar{w}} \dot{\bar{z}},$$

are computed to be

$$\mathcal{A}_\theta = -\frac{im\bar{z}}{2(J(m+1) + m\bar{z}z)}, \quad \mathcal{A}_w = \frac{1 - iz\mathcal{A}_\theta}{m+1}, \quad \mathcal{A}_{\bar{w}} = \frac{i\bar{z}\mathcal{A}_\theta}{m+1}.$$

The reduced dynamics is given by equations (5.4) and (5.5), where  $[\cdot, \cdot]_r^*$  is the dual of the bracket (3.13) and  $[\cdot, \cdot]_{\mathfrak{g}_r}^*$  is the operator  $\text{ad}^*$  on the algebra  $\mathfrak{se}(2)$ .

## 5.2 Nonholonomic Systems with Symmetry

Recall that the manifolds  $M$  and  $Q$  are principal fiber bundles. The Lagrangian  $L$  and constraint distribution  $\mathcal{D}$  are now invariant with respect to the induced action of  $G$  on  $P$ . Recall also that  $\mathcal{D}$  is a (locally) splitting subbundle of  $P$ .

Although it is not needed for everything that we will be doing, the examples and the theory are somewhat simplified if we make the following assumption:

**Assumption 5.5.** *The constraints and the orbit directions span the entire tangent space to the configuration space:*

$$\mathcal{D}_q + T_q \text{Orb}(q) = T_q Q,^{13}$$

for each  $q \in Q$ . If this condition is satisfied, we say that the **principal case** holds.

Let  $\mathcal{S}$  be the subbundle of  $\mathcal{D}$  whose fiber at  $q$  is  $\mathcal{S}_q = \mathcal{D}_q \cap T_q \text{Orb}(q)$ . We assume here that  $\mathcal{S}_q \neq \{0\}$ .<sup>14</sup> The subbundle  $\mathcal{S}$  is invariant with respect to the action of  $G$  on  $P$  induced by the left action of  $G$  on  $M$ . We assume that  $\mathcal{S}$  is a (locally) splitting subbundle.<sup>15</sup> Therefore, for each  $q \in Q$  there exist a subspace  $\mathcal{U}_q \subset \text{Orb}(q)$  such that  $\text{Orb}(q) = \mathcal{S}_q \oplus \mathcal{U}_q$  and the subbundle  $\mathcal{U}$  is (locally) splitting and  $G$ -invariant. Since the distributions  $\mathcal{S}$  and  $\mathcal{U}$  are left-invariant, for each  $r \in Q/G$  there exist subspaces  $\mathfrak{b}_r^{\mathcal{S}}$  and  $\mathfrak{b}_r^{\mathcal{U}}$  of the Lie algebra  $\mathfrak{g}_r = \mathfrak{b}_r^{\mathcal{S}} \oplus \mathfrak{b}_r^{\mathcal{U}}$  and an operator  $\varphi_r : \mathfrak{g}_r \rightarrow \mathfrak{g}$  such that in a local trivialization

$$\mathcal{S}_q = L_{g^*} \varphi_r \mathfrak{b}_r^{\mathcal{S}} \quad \text{and} \quad \mathcal{U}_q = L_{g^*} \varphi_r \mathfrak{b}_r^{\mathcal{U}}.$$

The operators  $\varphi_r$  are constructed in such a way that  $\mathfrak{b}_r^{\mathcal{S}}$  and  $\mathfrak{b}_r^{\mathcal{U}}$  are fixed subspaces of  $\mathfrak{g}_r$ . Such an arrangement is necessary for accounting for constraint subspaces that change their location in the Lie algebra  $\mathfrak{g}$  of the symmetry group depending on system's shape configuration.

Let  $\mathfrak{b}^{\mathcal{S}}$  and  $\mathfrak{b}^{\mathcal{U}}$  be the bundles over  $Q/G$  whose fibers are the subspaces  $\mathfrak{b}_r^{\mathcal{S}}$  and  $\mathfrak{b}_r^{\mathcal{U}}$  of  $\mathfrak{g}_r$ . Given  $\zeta \in \mathfrak{g}_r$ , we write its components along these subspaces as  $\zeta^{\mathcal{S}}$  and  $\zeta^{\mathcal{U}}$  so that for each  $r \in Q/G$  we have

$$g^{-1} \dot{g} = \varphi_r \zeta \equiv \varphi_r \zeta^{\mathcal{S}} + \varphi_r \zeta^{\mathcal{U}}, \quad \text{where} \quad \zeta^{\mathcal{S}} \in \mathfrak{b}_r^{\mathcal{S}} \quad \text{and} \quad \zeta^{\mathcal{U}} \in \mathfrak{b}_r^{\mathcal{U}}.$$

<sup>13</sup>This is called the *dimension assumption* in the finite-dimensional setting, see Bloch *et al.* [7] for details.

<sup>14</sup>If  $\mathcal{S}_q = \{0\}$ , a set of nonholonomic constraints is said to be *purely kinematic*.

<sup>15</sup>Note that the intersection of two splitting subspaces may fail to be splitting already in a Banach space. For the intersection of two subspaces to be splitting, additional assumptions are necessary. According to Bill Johnson, asking that two subspaces are norm one complemented and the space itself is uniformly convex is sufficient. See <http://mathoverflow.net/questions/85492/intersection-of-complemented-subspaces-of-a-banach-space> for details.

Let  $\mathcal{A}_s$  be a connection defined in a local trivialization by the formula

$$\mathcal{A}_s(\dot{r}, \dot{g}) = \text{Ad}_g(\varphi_r(\zeta + \mathcal{A}\dot{r})),$$

where  $\zeta \in \mathfrak{g}_r$  and where  $\mathcal{A}$  is a  $\mathfrak{g}_r$ -valued one-form on  $Q/G$ . This form is such that the constraints in a local trivialization read

$$\Omega^{\mathcal{U}} \equiv \zeta^{\mathcal{U}} + \mathcal{A}^{\mathcal{U}}\dot{r} = 0. \quad (5.11)$$

That is, the  $\mathcal{U}$ -component of the connection  $\mathcal{A}$  is defined by the constraints. The  $\mathcal{S}$ -component of  $\mathcal{A}$  is arbitrary at the moment; later on we will see how the structure of the Lagrangian affects the choice of this component.

As in §5.1, introduce the operator  $\psi_r : W_B \rightarrow T_r(M/G)$  (locally), so that  $\dot{r} = \psi_r \xi$  and  $\xi$  is now the shape velocity component used instead of  $\dot{r}$ . Utilizing the constrained Hamel equations (4.4) and the  $G$ -invariance of the constraint distribution  $\mathcal{D}$ , one obtains the reduced nonholonomic equations of motion from (5.4) and (5.5) by projecting equation (5.5) onto the fibers of the bundle  $\mathfrak{b}_S^*$  and imposing constraints, i.e., setting  $\Omega = \Omega^S$ . Summarizing, we have:

**Theorem 5.6.** *The strong form of the reduced nonholonomic dynamics is given by the equations*

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} - \psi_r^* \frac{\delta l_c}{\delta r} = \left[ \xi, \frac{\delta l}{\delta \xi} \right]_r^* - \left\langle \frac{\delta l}{\delta \Omega}, \mathbf{i}_\xi \psi_r^* \mathcal{B} + \langle \mathbf{i}_\xi \psi_r^* \gamma, \psi_r^* \mathcal{A} \rangle - \langle \psi_r^* \gamma, \mathbf{i}_\xi \psi_r^* \mathcal{A} \rangle + \langle \psi_r^* \mathcal{E}, \Omega^S \rangle \right\rangle, \quad (5.12)$$

$$\left[ \frac{d}{dt} \frac{\delta l}{\delta \Omega} \right]_S = \left[ \left[ \Omega^S, \frac{\delta l}{\delta \Omega} \right]_{\mathfrak{g}_r}^* + \left\langle \frac{\delta l}{\delta \Omega}, \mathbf{i}_\xi \psi_r^* \mathcal{E} \right\rangle \right]_S, \quad (5.13)$$

coupled with the kinematic shape equation

$$\dot{r} = \psi_r \xi. \quad (5.14)$$

In the above, the Lagrangian  $l$  is written as a function of  $(r, \xi, \Omega)$ ,  $\mathcal{B}$  is the curvature of  $\mathcal{A}$  on  $Q/G$  (see (5.1)), and the quantities  $\gamma$  and  $\mathcal{E}$  are defined as in Theorem 5.2. Note that the partial derivatives of  $l$  in (5.12) and (5.13) are evaluated *before* setting  $\Omega = \Omega^S$ , i.e., before imposing the constraints.

The shape and momentum equations (5.12) and (5.13) can be rewritten as

$$\begin{aligned} \frac{d}{dt} \frac{\delta l_c}{\delta \xi} - \psi_r^* \frac{\delta l_c}{\delta r} &= \left[ \xi, \frac{\delta l}{\delta \xi} \right]_r^* - \left\langle \frac{\delta l}{\delta \Omega}, \mathbf{i}_\xi \psi_r^* \mathcal{B} + \langle \mathbf{i}_\xi \psi_r^* \gamma, \psi_r^* \mathcal{A} \rangle - \langle \psi_r^* \gamma, \mathbf{i}_\xi \psi_r^* \mathcal{A} \rangle + \langle \psi_r^* \mathcal{E}, \Omega^S \rangle \right\rangle, \\ \frac{d}{dt} \frac{\delta l_c}{\delta \Omega^S} &= \left[ \left[ \Omega^S, \frac{\delta l}{\delta \Omega} \right]_{\mathfrak{g}_r}^* + \left\langle \frac{\delta l}{\delta \Omega}, \mathbf{i}_\xi \psi_r^* \mathcal{E} \right\rangle \right]_S, \end{aligned}$$

where  $l_c(r, \xi, \Omega^S) := l(r, \xi, \Omega^S)$  is the **constrained reduced Lagrangian**. These equations follow directly from (5.12) and (5.13) as

$$\frac{\delta l_c}{\delta \xi} = \frac{\delta l}{\delta \xi} \Big|_{\Omega=\Omega^S}, \quad \frac{d}{dt} \frac{\delta l_c}{\delta \xi} = \frac{d}{dt} \frac{\delta l}{\delta \xi} \Big|_{\Omega=\Omega^S}, \quad \frac{\delta l_c}{\delta r} = \frac{\delta l}{\delta r} \Big|_{\Omega=\Omega^S}, \quad \frac{\delta l_c}{\delta \Omega^S} = \frac{\delta l}{\delta \Omega^S} \Big|_{\Omega=\Omega^S}.$$

Note that in general

$$\frac{\delta l_c}{\delta \Omega^S} \neq \frac{\delta l}{\delta \Omega} \Big|_{\Omega=\Omega^S}.$$

### 5.3 The Nonholonomic Connection

In the rest of the section we assume that the Lagrangian equals the kinetic minus potential energy, that the kinetic energy is given by a (weak) Riemannian metric on the manifold  $M$ , and that the principal case holds (see Assumption 5.5).

**Definition 5.7.** The *nonholonomic connection*  $\mathcal{A}^{\text{nhc}}$  is, by definition, the connection on the principal bundle  $Q \rightarrow Q/G$  whose horizontal space at  $q \in Q$  is given by the orthogonal complement to the space  $\mathcal{S}_q = \mathcal{D}_q \cap T_q \text{Orb}(q)$  within the space  $\mathcal{D}_q$ .

Similar to the remark on the existence of the mechanical connection, in the infinite-dimensional setting, the nonholonomic connection may fail to exist when the group metric is not strong, but is certain to exist when the symmetry group is finite-dimensional.

Under the assumption that the distribution  $\mathcal{D}$  is invariant and from the fact that the group action preserves orthogonality (since it is assumed to preserve the Lagrangian and hence the kinetic energy metric), it follows that the distribution and the horizontal spaces transform to themselves under the group action. Therefore, the nonholonomic connection in a local trivialization is defined by the formula

$$\mathcal{A}^{\text{nhc}}(\dot{r}, \dot{g}) = \text{Ad}_g(y_c + \mathcal{A}_c \dot{r}), \quad (5.15)$$

where  $y_c \in \mathfrak{g}$  and where  $\mathcal{A}_c$  is a  $\mathfrak{g}$ -valued one-form on  $Q/G$ . Given  $\dot{q} = (\dot{r}, \dot{g}) \in T_q Q$ , the horizontal and vertical components of  $\dot{q}$  in a local trivialization are

$$(\dot{r}, -\mathcal{A}_c \dot{r}) \quad \text{and} \quad (0, y_c + \mathcal{A}_c \dot{r}).$$

In order to reflect the bundle structure associated with velocity constraints, we use Definition 5.7 and the map  $\varphi_r$  and rewrite  $y_c + \mathcal{A}_c \dot{r}$  in (5.15) as  $\varphi_r(\zeta + \mathcal{A}\dot{r})$ , where

$$\zeta = \zeta^{\mathcal{S}} + \zeta^{\mathcal{U}}, \quad \mathcal{A} = \mathcal{A}^{\mathcal{S}} + \mathcal{A}^{\mathcal{U}}, \quad \zeta^{\mathcal{S}} + \mathcal{A}^{\mathcal{S}} \dot{r} \in \mathfrak{b}^{\mathcal{S}}, \quad \zeta^{\mathcal{U}} + \mathcal{A}^{\mathcal{U}} \dot{r} \in \mathfrak{b}^{\mathcal{U}}.$$

Using the nonholonomic connection, define the *group velocity*  $\Omega \in \mathfrak{g}$  by the formula

$$\Omega = \zeta + \mathcal{A}\dot{r}.$$

The constraints then are represented by formula (5.11).

Let the kinetic energy metric in a local trivialization be written as

$$\langle\langle \dot{q}, \dot{q} \rangle\rangle = \langle G(r) \dot{r}, \dot{r} \rangle + 2\langle \mathcal{K}(r) \dot{r}, \zeta \rangle + \langle \mathcal{I}(r) \zeta, \zeta \rangle. \quad (5.16)$$

The *constrained locked inertia tensor*  $\mathcal{I}_{\mathcal{S}} : \mathfrak{b}^{\mathcal{S}} \rightarrow (\mathfrak{b}^{\mathcal{S}})^*$  is given in a local trivialization by

$$\langle \mathcal{I}_{\mathcal{S}}(r) \zeta_1, \zeta_2 \rangle = \langle\langle L_{g^*} \varphi_r \zeta_1, L_{g^*} \varphi_r \zeta_2 \rangle\rangle, \quad \zeta_1, \zeta_2 \in \mathfrak{b}^{\mathcal{S}}.$$

Similarly, define  $\mathcal{I}_{\mathcal{U}}(r) : \mathfrak{b}^{\mathcal{U}} \rightarrow (\mathfrak{b}^{\mathcal{U}})^*$ ,  $\mathcal{I}_{\mathcal{SU}}(r) : \mathfrak{b}^{\mathcal{U}} \rightarrow (\mathfrak{b}^{\mathcal{S}})^*$ , and  $\mathcal{I}_{\mathcal{US}}(r) : \mathfrak{b}^{\mathcal{S}} \rightarrow (\mathfrak{b}^{\mathcal{U}})^*$  by

$$\begin{aligned} \langle \mathcal{I}_{\mathcal{U}}(r) \zeta_1, \zeta_2 \rangle &= \langle\langle L_{g^*} \varphi_r \zeta_1, L_{g^*} \varphi_r \zeta_2 \rangle\rangle, \quad \zeta_1, \zeta_2 \in \mathfrak{b}^{\mathcal{U}}, \\ \langle \mathcal{I}_{\mathcal{SU}}(r) \zeta_1, \zeta_2 \rangle &= \langle\langle L_{g^*} \varphi_r \zeta_1, L_{g^*} \varphi_r \zeta_2 \rangle\rangle, \quad \zeta_1 \in \mathfrak{b}^{\mathcal{U}}, \quad \zeta_2 \in \mathfrak{b}^{\mathcal{S}}, \\ \langle \mathcal{I}_{\mathcal{US}}(r) \zeta_1, \zeta_2 \rangle &= \langle\langle L_{g^*} \varphi_r \zeta_1, L_{g^*} \varphi_r \zeta_2 \rangle\rangle, \quad \zeta_1 \in \mathfrak{b}^{\mathcal{S}}, \quad \zeta_2 \in \mathfrak{b}^{\mathcal{U}}, \end{aligned}$$

respectively.

Definition 5.7 implies that the constrained kinetic energy metric written as a function of  $(\dot{r}, \Omega^{\mathcal{S}})$  is *block-diagonal*, that is, it reads

$$\langle G(r) \dot{r}, \dot{r} \rangle + \langle \mathcal{I}_{\mathcal{S}}(r) \Omega^{\mathcal{S}}, \Omega^{\mathcal{S}} \rangle.$$

Substituting  $\zeta = \Omega^{\mathcal{S}} - \mathcal{A}\dot{r}$  in (5.16), one concludes that the  $\mathcal{S}$ -component of the nonholonomic connection satisfies the equation

$$\langle (\mathcal{I}_{\mathcal{S}}(r) \mathcal{A}^{\mathcal{S}} - \mathcal{K}_{\mathcal{S}}(r) + \mathcal{I}_{\mathcal{SU}}(r) \mathcal{A}^{\mathcal{U}}) \dot{r}, \Omega^{\mathcal{S}} \rangle = 0.$$

This equation may or may not have a solution in the infinite dimensional setting. If the restriction of the metric on the subspace  $\mathfrak{b}^{\mathcal{S}}$  is strong, the  $\mathcal{S}$ -component of the nonholonomic connection is well-defined and reads

$$\mathcal{A}^{\mathcal{S}} = \mathcal{I}_{\mathcal{S}}^{-1}(r) \mathcal{K}_{\mathcal{S}}(r) - \mathcal{I}_{\mathcal{S}}^{-1}(r) \mathcal{I}_{\mathcal{SU}}(r) \mathcal{A}^{\mathcal{U}}.$$

The  $\mathcal{U}$ -component of the nonholonomic connection is defined by the constraints. As in §5.2, the reduced dynamics is given by equations (5.12)–(5.14).



**Example 5.8.** Consider the planar motion of an inextensible string coupled with the two identical Chaplygin sleighs. Assume that the ends of the string are attached at the contact points of the sleighs and the plane. See Figure 5.1 for the illustration of the basic configuration.

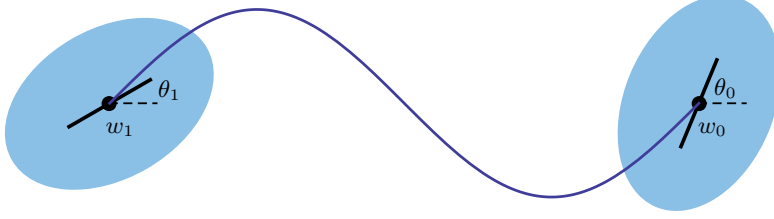


Figure 5.1: The string with Chaplygin sleighs attached at the ends.

The manifold  $M(= Q)$  for this system is  $\text{SE}(2) \times \text{SE}(2) \times \text{Emb}([0, 1], \mathbb{C})$ , and the system is invariant with respect to the diagonal action of the group  $\text{SE}(2)$  on  $M$ . We identify the symmetry group with configurations of the sleigh at  $s = 0$ . In such a representation, the position of the sleigh at the front end of the string (i.e., at  $s = 0$ ) is characterized by the element  $(e^{i\theta_0}, w_0) \in \text{SE}(2)$ , while the relative position of the sleigh at the opposite end of the string is  $(e^{i\phi_1}, z_1) \in \text{SE}(2)$ , so that the absolute position of that sleigh is

$$(e^{i\theta_1}, w_1) = (e^{i\theta_0}, w_0) \cdot (e^{i\phi_1}, z_1) = (e^{i(\theta_0+\phi_1)}, w_0 + e^{i\theta_0} z_1).$$

The angular and linear velocity (relative to the body frame) of the sleigh at the front end of the string are

$$\omega_0 = \dot{\theta}_0 \quad \text{and} \quad v_0 = e^{-i\theta_0} \dot{w}_0,$$

while the angular and linear velocity of the other sleigh are

$$\omega_0 + \dot{\phi}_1 \quad \text{and} \quad e^{-i\phi_1} (v_0 + iz_1 \omega_0 + \dot{z}_1).$$

Thus, the constraints are

$$\bar{v}_0 = v_0 \quad \text{and} \quad e^{-i\phi_1} (v_0 + iz_1 \omega_0 + \dot{z}_1) = e^{i\phi_1} (\bar{v}_0 - i\bar{z}_1 \omega_0 + \dot{\bar{z}}_1). \quad (5.17)$$

Formulae (5.17) define the operator  $\varphi$  and the constraint component of the nonholonomic connection. Both are nontrivial for this coupled system.

Recall that the nonholonomic connection modifies the elements of the Lie algebra of the symmetry group. The equations of motion that use such ‘shifted’ Lie algebra elements are useful in numerous situations, including stability analysis of relative equilibria. However, one may verify that this form of equations is rather complicated for the system considered here. Below we give a simpler representation of system’s dynamics using the operators introduced earlier for the sleigh and string.

Using the notations introduced in §2.5 and Example 3.3, parametrize the configuration spaces of the sleighs as  $(\theta_j, w_j)$ ,  $j = 0, 1$ , where  $w_j = x_j + iy_j$ , and denote the *absolute* position of the string by  $w \in \text{Emb}([0, 1], \mathbb{C})$ . The Lagrangian reads

$$l = \frac{1}{2}(J\omega_0^2 + m\bar{v}_0 v_0) + \frac{1}{2}(J\omega_1^2 + m\bar{v}_1 v_1) + \int_0^1 \frac{1}{2}(\bar{w}_s w_s \bar{\xi} \xi - \lambda(\bar{w}_s w_s - 1)) ds. \quad (5.18)$$

Coupling is accomplished by matching the position of the sleighs and the string’s ends:

$$w_0 = w|_{s=0} \quad \text{and} \quad w_1 = w|_{s=1}. \quad (5.19)$$

That is, the string is attached at and can rotate freely around the contact point of each sleigh.

The constrained Hamel equations for this system read

$$J\dot{\omega}_0 = 0, \tag{5.20}$$

$$m\dot{v}_0 = \frac{1}{2}\lambda_0(\bar{w}_s^0 e^{i\theta_0} + w_s^0 e^{-i\theta_0}), \tag{5.21}$$

$$J\dot{\omega}_1 = 0, \tag{5.22}$$

$$m\dot{v}_1 = -\frac{1}{2}\lambda_1(\bar{w}_s^1 e^{i\theta_1} + w_s^1 e^{-i\theta_1}), \tag{5.23}$$

$$\dot{\xi} = \xi\bar{\xi}_s + \lambda_s + i\mathcal{X}(\lambda - \bar{\xi}\xi), \tag{5.24}$$

where, by definition,  $\lambda_0 := \lambda|_{s=0}$ ,  $\lambda_1 := \lambda|_{s=1}$ ,  $w_s^0 := w_s|_{s=0}$ , and  $w_s^1 := w_s|_{s=1}$ . Equations (5.20), (5.22), (5.24) and the left-hand sides of (5.21), (5.23) have been obtained earlier. The right-hand sides of (5.21) and (5.23) account for the coupling constraints (5.19) and are obtained as projections of the external terms (i.e., those due to integration by parts) associated with the Lagrange multiplier  $\lambda$  in (5.18) along the sleigh directions.

## 5.4 Systems with Shape Constraints

Here we study systems with a ‘slim’ constraint distribution. That is, at each  $q \in Q$ , the complimentary subspace of  $\mathcal{D}_q \subset T_qM$  is ‘larger’ than the tangent space to the group orbit at  $q$ . While this may happen when the symmetry group is infinite-dimensional as well as when the system is finite-dimensional, our main objects of interest here are systems with finite-dimensional symmetry group and constraint distributions of infinite codimension. Assumption 5.5 does not hold for such systems, and the formalism of Bloch *et al.* [7] and its infinite-dimensional analogue presented in §5.2 should be modified.

**Example 5.9.** Consider the Chaplygin sleigh with an inextensible string attached at and allowed to rotate around the contact point of the sleigh and the plane, as shown in Figure 5.2. Assume that the string is constrained as in Example 4.3. This system is SE(2)-invariant. Using formulae (5.9) and (5.10) there, the

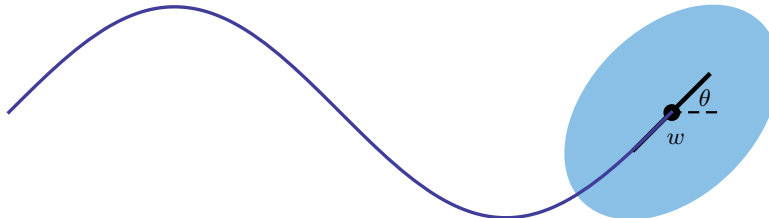


Figure 5.2: The Chaplygin sleigh coupled to a constrained string.

condition of vanishing normal velocity component is equivalent to

$$\bar{z}_s(v + i\omega z + \dot{z}) = z_s(\bar{v} - i\omega\bar{z} + \dot{\bar{z}}). \tag{5.25}$$

Here the string position  $z$  is measured relative to the sleigh, and  $\omega$  and  $v$  are the sleigh’s angular and linear velocity measured relative to its frame. Formula (5.25) imposes infinitely-many constraints on the system, one for each  $s \in [0, 1]$ .

For a generic string placement, constraints (5.25) exhaust the Lie algebra  $\mathfrak{se}(2)$ . Indeed, the vectors that span the constraint subspaces for  $s_0 = 0$  and  $0 < s_1 < s_2 < 1$  are verified to be  $\mathbb{R}$ -independent if and only if

$$(\bar{z}_s z + z_s \dot{\bar{z}})|_{s=s_1} (z_s - \bar{z}_s)|_{s=s_2} - (\bar{z}_s z + z_s \dot{\bar{z}})|_{s=s_2} (z_s - \bar{z}_s)|_{s=s_1} \neq 0,$$

which generically holds. The rest of the constraints are effectively imposed on the shape degrees of freedom of the system, and thus Assumption 5.5 does not hold.

We begin by developing a special version of the formalism for unconstrained  $G$ -invariant systems. Using notations introduced in §5.1, this system is characterized by the Lagrangian  $l(r, \xi, \zeta)$ . Recall that  $\dot{r} = \psi_r \xi$ . In order to simplify the exposition, we assume here that  $\varphi_r$  is the identity operator so that  $\dot{g} = L_{g*} \zeta$ . Instead of the velocity components (5.2), we will use

$$\eta = \xi + \mathcal{C}\zeta \in W_B \quad \text{and} \quad \zeta \in \mathfrak{g},$$

where  $\mathcal{C} : \mathfrak{g} \rightarrow W_B$  is a linear map. It depends parametrically and smoothly on  $r \in U$ , where  $U$  is an open subset of  $Q/G$ . This is motivated by the structure of the constraints (5.25). Thus, for the operator  $\Psi_q$  we have

$$\dot{q} = (\dot{r}, \dot{g}) = \Psi_q(\eta, \zeta) = (\psi_r(\eta - \mathcal{C}\zeta), L_{g*}\zeta).$$

The inverse of  $\Psi_q$  is

$$(\eta, \zeta) = \Psi_q^{-1}(\dot{r}, \dot{g}) = (\psi_r^{-1}\dot{r} + \mathcal{C}L_{g*}^{-1}\dot{g}, L_{g*}^{-1}\dot{g}).$$

One can think of  $\mathcal{C}$  as a connection on the local fiber bundle whose base and fiber are open subsets of  $G$  and  $Q/G$ , respectively.

**Theorem 5.10.** *The strong Lagrange–Poincaré equations are*

$$\begin{aligned} \frac{d}{dt} \frac{\delta l}{\delta \eta} - \psi_r^* \frac{\delta l}{\delta r} &= \left[ \eta - \mathcal{C}\zeta, \frac{\delta l}{\delta \eta} \right]_r^* + \left\langle \frac{\delta l}{\delta \eta}, \psi_r^* d\mathcal{C}\zeta \right\rangle, \\ \frac{d}{dt} \frac{\delta l}{\delta \zeta} &= \text{ad}_\zeta^* \left( \frac{\delta l}{\delta \zeta} + \mathcal{C}^* \frac{\delta l}{\delta \eta} \right) - \left\langle \left[ \eta - \mathcal{C}\zeta, \frac{\delta l}{\delta \eta} \right]_r^*, \mathcal{C} \right\rangle - \left\langle \frac{\delta l}{\delta \eta}, \mathbf{i}_{(\eta - \mathcal{C}\zeta)} \psi_r^* d\mathcal{C} + \mathbf{i}_{\mathcal{C}} \psi_r^* d\mathcal{C}\zeta \right\rangle. \end{aligned}$$

*Proof.* The key step is to establish the bracket operation associated with the operator  $\Psi_q$ . One begins with the computation of the Jacobi–Lie bracket

$$[\Psi(\eta_1, \zeta_1), \Psi(\eta_2, \zeta_2)] = [(\psi(\eta_1 - \mathcal{C}\zeta_1), L_*\zeta_1), (\psi(\eta_2 - \mathcal{C}\zeta_2), L_*\zeta_2)],$$

where  $\eta_i$  and  $\zeta_i$  are constant vectors. For the  $M/G$ -component of the Jacobi–Lie bracket, we have, using the chain rule:

$$[\psi(\eta_1 - \mathcal{C}\zeta_1), \psi(\eta_2 - \mathcal{C}\zeta_2)]_{M/G} = \psi[\eta_1 - \mathcal{C}\zeta_1, \eta_2 - \mathcal{C}\zeta_2]_r - \psi(\psi\eta_1 - \psi\mathcal{C}\zeta_1)[\mathcal{C}\zeta_2] + \psi(\psi\eta_2 - \psi\mathcal{C}\zeta_2)[\mathcal{C}\zeta_1].$$

For the  $G$ -component, we obtain

$$[L_*\zeta_1, L_*\zeta_2]_G = L_*[\zeta_1, \zeta_2]_{\mathfrak{g}} = L_* \text{ad}_{\zeta_1} \zeta_2.$$

Applying the inverse operator  $\Psi_q$  outputs the bracket operation on  $W_B \oplus \mathfrak{g}$ :

$$[(\eta_1, \zeta_1), (\eta_2, \zeta_2)] = ([\eta_1 - \mathcal{C}\zeta_1, \eta_2 - \mathcal{C}\zeta_2]_r - (\psi\eta_1 - \psi\mathcal{C}\zeta_1)[\mathcal{C}\zeta_2] + (\psi\eta_2 - \psi\mathcal{C}\zeta_2)[\mathcal{C}\zeta_1] + \mathcal{C} \text{ad}_{\zeta_1} \zeta_2, \text{ad}_{\zeta_1} \zeta_2).$$

Pairing with an element  $(\alpha, \beta) \in W_B^* \oplus \mathfrak{g}^*$ , we obtain:

$$\begin{aligned} &\langle \alpha, [\eta_1 - \mathcal{C}\zeta_1, \eta_2 - \mathcal{C}\zeta_2]_r - (\psi\eta_1 - \psi\mathcal{C}\zeta_1)[\mathcal{C}\zeta_2] + (\psi\eta_2 - \psi\mathcal{C}\zeta_2)[\mathcal{C}\zeta_1] + \mathcal{C} \text{ad}_{\zeta_1} \zeta_2 \rangle_{W_B} + \langle \beta, \text{ad}_{\zeta_1} \zeta_2 \rangle_{\mathfrak{g}} \\ &= \langle [\eta_1 - \mathcal{C}\zeta_1, \alpha]_r^* + \langle \alpha, \psi^* d\mathcal{C}\zeta_1 \rangle_{W_B}, \eta_2 \rangle_{W_B} \\ &\quad + \langle \text{ad}_{\zeta_1}^* (\beta + \mathcal{C}^* \alpha) - \langle [\eta_1 - \mathcal{C}\zeta_1, \alpha]_r^*, \mathcal{C} \rangle_{W_B} - \langle \alpha, \mathbf{i}_{(\eta_1 - \mathcal{C}\zeta_1)} \psi^* d\mathcal{C} \rangle_{W_B} - \langle \alpha, \mathbf{i}_{\mathcal{C}} \psi^* d\mathcal{C}\zeta_1 \rangle_{W_B}, \zeta_2 \rangle_{\mathfrak{g}}, \end{aligned}$$

which yields the components of the dual bracket,

$$\langle [\eta_1 - \mathcal{C}\zeta_1, \alpha]_r^* + \langle \alpha, \psi_r^* d\mathcal{C}\zeta_1 \rangle_{W_B}, \tag{5.26}$$

$$\langle \text{ad}_{\zeta_1}^* (\beta + \mathcal{C}^* \alpha) - \langle [\eta_1 - \mathcal{C}\zeta_1, \alpha]_r^*, \mathcal{C} \rangle_{W_B} - \langle \alpha, \mathbf{i}_{(\eta_1 - \mathcal{C}\zeta_1)} \psi_r^* d\mathcal{C} \rangle_{W_B} - \langle \alpha, \mathbf{i}_{\mathcal{C}} \psi_r^* d\mathcal{C}\zeta_1 \rangle_{W_B}. \tag{5.27}$$

Utilizing the dual bracket (5.26) and (5.27) and composing equations (3.9) completes the proof.  $\square$

Next, assume that there exists a splitting distribution  $\mathcal{D}$  along with its complimentary distribution  $\mathcal{U}$  such that the constraint  $\eta \in \mathcal{D}_q$  reads

$$\xi^{\mathcal{U}} + \mathcal{C}^{\mathcal{U}}\zeta = 0. \quad (5.28)$$

Thus, the constraint is defined by the operator  $\mathcal{C}^{\mathcal{U}} : \mathfrak{g} \rightarrow \mathcal{U}_q$ . This representation is, in general, *local*, i.e., (5.28) represents the constraint on an open subset  $U$  of the shape space  $Q/G$ . The operator  $\mathcal{C}^{\mathcal{U}}$  is then extended to  $\mathcal{C} : \mathfrak{g} \rightarrow W_B$  at each  $r \in U$ .

There are, of course, infinitely many such extensions, and one needs to utilize additional requirements to single out the desired one. Utilizing the general Hamel's equations established in §4.2, the reduced dynamics becomes

$$\left( \frac{d}{dt} \frac{\delta l}{\delta \eta} - \psi_r^* \frac{\delta l}{\delta r} - \left[ \eta - \mathcal{C}\zeta, \frac{\delta l}{\delta \eta} \right]_r^* - \left\langle \frac{\delta l}{\delta \eta}, \psi_r^* d\mathcal{C}\zeta \right\rangle \right)_{\mathcal{D}} = 0, \quad (5.29)$$

$$\left( \frac{d}{dt} \frac{\delta l}{\delta \zeta} - \text{ad}_{\zeta}^* \left( \frac{\delta l}{\delta \zeta} - \mathcal{C}^* \frac{\delta l}{\delta \eta} \right) + \left\langle \left[ \eta - \mathcal{C}\zeta, \frac{\delta l}{\delta \eta} \right]_r^*, \mathcal{C} \right\rangle + \left\langle \frac{\delta l}{\delta \eta}, \mathbf{i}_{(\eta - \mathcal{C}\zeta)} \psi_r^* d\mathcal{C} + \mathbf{i}_{\mathcal{C}} \psi_r^* d\mathcal{C}\zeta \right\rangle \right)_{\mathcal{D}} = 0. \quad (5.30)$$

If desired, one can give the representation of these equations using the constrained Lagrangian and constrained versions of the velocity components  $\eta$  and  $\zeta$ , as in the approach in §5.2.

**Example 5.9 (continued).** Consider again the motion of the Chaplygin sleigh with the flexible string attached. The normal velocity component of the string at each point is constrained to be zero, as in (5.25). The system is SE(2)-invariant and the symmetry group is also the configuration space of the sleigh. The Lagrangian of the system is the sum of the kinetic energy of the sleigh and the string Lagrangian and reads

$$\frac{1}{2}(J\omega^2 + m\bar{v}v) + \int_0^1 \frac{1}{2} \left( (\bar{v} - i\omega\bar{z} + \bar{z}_s\bar{\xi})(v + i\omega z + z_s\xi) - \lambda(\bar{z}_s z_s - 1) \right) ds,$$

where  $\xi = \bar{z}_s \dot{z}$  is string's shape velocity. The coupling is accomplished by setting

$$z|_{s=0} = 0, \quad z_s|_{s=0} = 1, \quad \xi|_{s=0} = 0,$$

and the constraint is given by formula (5.25). We assume from now on that  $v > 0$ , i.e., the string trails the sleigh.

Define the quantity  $\eta$  and the operator  $\mathcal{C}$  by the formula

$$\eta = \xi + \bar{z}_s(v + i\omega z),$$

so that  $\eta$  is string's *absolute* velocity. Doing so results in the Lagrangian

$$l = \frac{1}{2}(J\omega^2 + m\bar{v}v) + \int_0^1 \frac{1}{2} (\bar{\eta}\eta - \lambda(\bar{z}_s z_s - 1)) ds.$$

The constraint in this representation reads  $\eta = \bar{\eta}$ , which is formula (5.28) written for the string. Equations (5.29) and (5.30) for the sleigh-string system become

$$J\dot{\omega} = 0, \quad (5.31)$$

$$m\dot{v} = \lambda|_{s=0}, \quad (5.32)$$

$$\dot{\eta} = \eta\eta_s + \lambda_s, \quad (5.33)$$

where  $v$  and  $\eta$  are real-valued. These equations should be amended with the coupling conditions

$$\eta|_{s=0} = v. \quad (5.34)$$

Note also that the position of the front end of the string coincides with the position of the contact point of the sleigh and the plane.

Arguing as in Example 4.3, one concludes that  $\eta$  is independent of  $s$ . Thus, equation (5.33) becomes

$$\dot{\eta} = \lambda_s. \tag{5.35}$$

Equation (5.31) implies  $\omega = \text{const}$ .

The tension  $\lambda$  is obtained by integrating (5.35) with respect to  $s$  and, since  $\lambda|_{s=1} = 0$ , we conclude that

$$\lambda = (s - 1)\dot{\eta}. \tag{5.36}$$

Therefore,  $\lambda|_{s=0} = -\dot{\eta}$ , which in combination with (5.32) and (5.34) yields  $v = \text{const}$ . The velocity coupling condition then implies that the blade moves at the constant speed  $\eta = v$ . Using (5.36), we conclude that  $\lambda = 0$ .

Summarizing, the Chaplygin sleigh with the constrained string attached generically undergoes uniform circular motion. Nongeneric trajectories are straight lines. The string (possibly after some period of time) follows the trajectory of the contact point of the sleigh.

It is interesting to point out that in this example the shape dynamics (string’s motion) is modulated by the group dynamics (skate’s motion). This is the opposite of typical reconstruction in finite-dimensional constrained systems discussed in [7].

We note also that the qualitative dynamics of this system—uniform circular or straight line motion—is consistent with the behavior of integrable Hamiltonian systems. One may raise the question of whether it is integrable in a more precise sense—with infinitely-many conserved quantities. We intend to return to this issue in a forthcoming publication.

## Conclusions

This paper introduced Hamel’s formalism for infinite-dimensional mechanical systems, proved some general results, and illustrated them with examples. A number of topics remain outside the scope of the paper. For instance, the nature of nonholonomic constraints in the infinite-dimensional setting should be studied in more detail and will be addressed in a forthcoming publication. We also intend to develop Hamel’s formalism for field theories and apply it for constructing structure-preserving integrators for constrained continuum mechanical systems.

Examples in this paper were intentionally kept relatively simple for pedagogical reasons, and solutions there are classical, i.e., not generalized. More realistic examples, including systems with generalized solutions, will be treated in the forthcoming paper on the field-theoretic Hamel formalism.

## Acknowledgments

We would like to thank Professors Yongxin Guo, Francois Gay-Balmaz, Vakhtang Putkaradze, and Tudor Ratiu for valuable discussions, and the reviewers for helpful remarks.

The research of AMB was partially supported by NSF grants DMS-1207893, DMS-1613819, INSPIRE-1363720, and the Simons Foundations. The research of DS was partially supported by the China Scholarship Council. DS wishes to thank support and hospitality of North Carolina State University during his visit. The research of YBK was partially supported by NDSEG Fellowship. The research of DVZ was partially supported by NSF grants DMS-0908995 and DMS-1211454. DVZ would like to acknowledge support and hospitality of the Beijing Institute of Technology, where a part of this work was carried out.

## References

- [1] Arnold, V. I. [1966], Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l’hydrodynamique des fluides parfaits, *Ann. Inst. Fourier, Grenoble* **16**, 319–361.

- [2] Arnold, V. I., V. V. Kozlov, and A. I. Neishtadt [2006], *Mathematical Aspects of Classical and Celestial Mechanics*, 3rd ed., Springer-Verlag, Berlin.
- [3] Ball, K., D. V. Zenkov, and A. M. Bloch [2012], Variational Structures for Hamel’s Equations and Stabilization, *Proc. IFAC*, 178–183.
- [4] Ball, K. R. and D. V. Zenkov [2015], Hamel’s Formalism and Variational Integrators, in D. E. Chang, D. D. Holm, G. Patrick, and T. Ratiu, eds., *Geometry, Mechanics, and Dynamics: The Legacy of Jerry Marsden*, *Fields Institute Communications* **73**, Springer, 477–506.
- [5] Binz, E., M. de León, D. Martín de Diego, and D. Socolescu [2002], Nonholonomic Constraints in Classical Field Theories, *Reports on Mathematical Physics* **49**, 151–166.
- [6] Bloch, A. M. [2015], *Nonholonomic Mechanics and Control*, *Interdisciplinary Appl. Math.* **24**, 2nd ed., Springer-Verlag, New York.
- [7] Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and R. Murray [1996], Nonholonomic Mechanical Systems with Symmetry, *Arch. Rational Mech. Anal.* **136**, 21–99.
- [8] Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and T. S. Ratiu [1996], The Euler–Poincaré Equations and Double Bracket Dissipation, *Communications in Mathematical Physics* **175**, 1–42.
- [9] Bloch, A. M., J. E. Marsden, and D. V. Zenkov [2009], Quasivelocities and Symmetries in Nonholonomic Systems, *Dynamical Systems: An International Journal* **24**, 187–222.
- [10] Cendra, H., D. D. Holm, J. E. Marsden, and T. S. Ratiu [1988], Lagrangian Reduction, the Euler–Poincaré Equations, and Semidirect Products, *Amer. Math. Soc. Transl.* **186**, 1–25.
- [11] Cendra, H., J. E. Marsden, and T. S. Ratiu [2001], Geometric Mechanics, Lagrangian Reduction, and Nonholonomic Systems, in B. Enguist and W. Schmid, eds., *Mathematics Unlimited–2001 and Beyond*, Springer-Verlag, New York, 221–273.
- [12] Cendra, H., J. E. Marsden, and T. S. Ratiu [2001], *Lagrangian Reduction by Stages*, *Memoirs of the Amer. Math. Soc.* **152**, AMS, Providence, Rhode Island.
- [13] Chetaev, N. G. [1989], *Theoretical Mechanics*, Springer-Verlag, New York.
- [14] Domański, P. and M. Mastyło [2007], Characterization of Splitting for Fréchet–Hilbert Spaces via Interpolation, *Mathematische Annalen* **338**, 317–340.
- [15] Ebin, D. G. [2015], Groups of Diffeomorphisms and Fluid Motion: Reprise, in D. E. Chang, D. D. Holm, G. Patrick, and T. Ratiu, eds., *Geometry, Mechanics, and Dynamics: The Legacy of Jerry Marsden*, *Fields Institute Communications* **73**, Springer, 477–506.
- [16] Ebin, D. G. and J. E. Marsden [1970], Groups of Diffeomorphisms and the Motion of an Incompressible Fluid, *Ann. of Math.* **92**, 102–163.
- [17] Ellis, D. C. P., F. Gay-Balmaz, D. D. Holm, V. Putkaradze, and T. S. Ratiu [2010], Symmetry Reduced Dynamics of Charged Molecular Strands, *Archive for Rational Mechanics and Analysis* **197**, 811–902.
- [18] Euler, L. [1752], Decouverte d’un nouveau principe de Mecanique, *Mémoires de l’académie des sciences de Berlin* **6**, 185–217.
- [19] Euler, L. [1757], Principes généraux de l’état d’équilibre des fluides, *Mémoires de l’académie des sciences de Berlin* **11**, 217–273.
- [20] Euler, L. [1757], Principes généraux du mouvement des fluides, *Mémoires de l’académie des sciences de Berlin* **11**, 274–315.

- [21] Euler, L. [1761], Principia motus fluidorum, *Novi Commentarii Academiae Scientiarum Petropolitanae* **6**, 271–311.
- [22] Filipović, D. and J. Teichmann [2003], Existence of Invariant Manifolds for Stochastic Equations in Infinite Dimension, *Journal of Functional Analysis* **197**, 398–432.
- [23] Gay-Balmaz, F., J. E. Marsden, and T. S. Ratiu [2012], Reduced Variational Formulations in Free Boundary Continuum Mechanics, *J. Nonlinear Sci.* **22**, 463–497.
- [24] Gay-Balmaz, F. and V. Putkaradze [2012], Dynamics of Elastic Rods in Perfect Friction Contact, *Phys. Rev. Lett.* **109**, 244303.
- [25] Gay-Balmaz, F. and V. Putkaradze [2015], Dynamics of Elastic Strands with Rolling Contact, *Physica D* **294**, 6–23.
- [26] Gay-Balmaz, F. and H. Yoshimura [2015], Dirac Reduction for Nonholonomic Mechanical Systems and Semidirect Products, *Advances in Applied Mechanics* **63**, 131–213.
- [27] Hamel, G. [1904], Die Lagrange–Eulersche Gleichungen der Mechanik, *Z. Math. Phys.* **50**, 1–57.
- [28] Hamilton, W. R. [1834], On a General Method in Dynamics, Part I, *Phil. Trans. Roy. Soc. Lond.* **124**, 247–308.
- [29] Hamilton, W. R. [1835], On a General Method in Dynamics, Part II, *Phil. Trans. Roy. Soc. Lond.* **125**, 95–144.
- [30] Hiltunen, S. [2000], A Frobenius Theorem for Locally Convex Global Analysis, *Monatshefte für Mathematik* **129**, 109–117.
- [31] Holm, D. D., J. E. Marsden, and T. S. Ratiu [1998], The Euler–Poincaré Equations and Semidirect Products with Applications to Continuum Theories, *Adv. Math.* **137**, 1–81.
- [32] Jarchow, H. [1981], *Locally Convex Spaces*, Teubner, Stuttgart.
- [33] Khesin, B. and P. Lee [2009], A Nonholonomic Moser Theorem and Optimal Transport, *Journal of Symplectic Geometry* **7**, 381–414.
- [34] Kriegl, A. and P. W. Michor [1997], *The Convenient Setting of Global Analysis*, *Mathematical Surveys and Monographs* **53**, AMS, Providence, Rhode Island.
- [35] Krishnaprasad, P. S. and D. P. Tsakiris [1994], *G-Snakes: Nonholonomic Kinematic Chains on Lie Groups*, *ISR Technical Report*, 26 pp.
- [36] Lagrange, J. L. [1788], *Mécanique Analytique*, Chez la Veuve Desaint.
- [37] Marsden, J. E. [1992], *Lectures on Mechanics*, *London Mathematical Society Lecture Note Series* **174**, Cambridge University Press.
- [38] Marsden, J. E. and T. J. R. Hughes [1983], *Mathematical Foundations of Elasticity*, Prentice Hall.
- [39] Marsden, J. E. and T. S. Ratiu [1999], *Introduction to Mechanics and Symmetry*, *Texts in Appl. Math.* **17**, 2nd ed., Springer-Verlag, New York.
- [40] Neimark, Ju. I. and N. A. Fufaev [1972], *Dynamics of Nonholonomic Systems*, *Translations of Mathematical Monographs* **33**, AMS, Providence, Rhode Island.
- [41] Omori, H. [1997], *Infinite-Dimensional Lie Groups*, *Translations of Mathematical Monographs* **158**, AMS, Providence, Rhode Island.

- [42] Ostrowski, J., A. Lewis, R. Murray, and J. Burdick [1994], Nonholonomic Mechanics and Locomotion: The Snakeboard Example, *Proc. IEEE Intern. Conf. on Robotics and Automation*, 2391–2397.
- [43] Poincaré, H. [1901], Sur une forme nouvelle des équations de la mécanique, *CR Acad. Sci.* **132**, 369–371.
- [44] Poincaré, H. [1910], Sur la precession des corps deformables, *Bull. Astron.* **27**, 321–356.
- [45] Suslov, G. K. [1946], *Theoretical Mechanics*, 3rd ed., GITTL, Moscow–Leningrad.
- [46] Teichmann, J. [2001], A Frobenius Theorem on Convenient Manifolds, *Monatshefte für Mathematik* **134**, 159–167.
- [47] Vankerschaver, J. [2005], The Momentum Map for Nonholonomic Field Theories with Symmetry, *Int. J. Geom. Meth. Mod. Phys.* **2**, 1029–1041.
- [48] Vankerschaver, J. [2007], A Class of Nonholonomic Kinematic Constraints in Elasticity, *J. Phys. A: Math. Theor.* **40**, 3889–3913.
- [49] Vankerschaver, J. [2007], *Continuous and Discrete Aspects of Lagrangian Field Theories with Nonholonomic Constraints*, Ph.D. thesis, Ghent University.
- [50] Yoshimura, H. and J. E. Marsden [2006], Dirac Structures in Lagrangian Mechanics. Part I: Implicit Lagrangian Systems, *Journal of Geometry and Physics* **57**, 133–156.
- [51] Yoshimura, H. and J. E. Marsden [2006], Dirac Structures in Lagrangian Mechanics. Part II: Variational Structures, *Journal of Geometry and Physics* **57**, 209–250.
- [52] Yoshimura, H. and J. E. Marsden [2007], Reduction of Dirac Structures and the Hamilton–Pontryagin Principle, *Reports on Math. Phys.* **60**, 381–426.
- [53] Zenkov, D. V., M. Leok, and A. M. Bloch [2012], Hamel’s Formalism and Variational Integrators on a Sphere, *Proc. CDC* **51**, 7504–7510.