

Hamel's Formalism and Variational Integrators

Kenneth R. Ball and Dmitry V. Zenkov

Abstract Hamel's formalism is a representation of Lagrangian mechanics obtained by measuring the velocity components relative to a frame that generically is not induced by configuration coordinates. The use of this formalism often leads to a simpler representation of dynamics. Utilizing the variational discretization approach, this paper develops a discrete Hamel's formalism with applications to nonholonomic integrators.

1 Introduction

This paper introduces the discrete Hamel formalism along with some of its applications. Besides being of a pure theoretical interest, this development is motivated by restoring the concept of ideal constraints in the discrete setting and by an attempt to better understand structural stability of variational and nonholonomic integrators. A loss of structural stability has been recently observed in [25, 26, 34].

Hamel's formalism is a version of Lagrangian mechanics in which the velocity components are measured relative to a set of independent vector fields on the configuration space. These vector fields are not associated with configuration coordinates and therefore do not commute, leading to the so-called 'bracket terms' in the equations of motion.

One of the reasons for using Hamel's formalism is that the Euler–Lagrange equations written in generalized coordinates, while universal, are not always the best tool for analyzing the dynamics of mechanical systems. For example, it is difficult to study the motion of the Euler top if the Euler–Lagrange equations (either intrinsically or in generalized coordinates) are used to represent the dynamics. On the other hand, the use of the angular velocity components relative to a body frame pioneered by Euler [13] results in a much simpler representation of dynamics. Euler's approach led to the development of the Euler–Poincaré equations by Lagrange [24] for reasonably general Lagrangians on the rotation group and by Poincaré [35] for arbitrary Lie groups (see [27] for details and history). An extension

K.R. Ball • D.V. Zenkov (✉)

Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA

e-mail: kball@ncsu.edu; dvzenkov@ncsu.edu

of this formalism from Lie groups to arbitrary configuration manifolds was carried out by Hamel [16]. Hamel's formalism is especially useful in nonholonomic mechanics. See e.g. [5, 30, 33] for the history and contemporary exposition of Hamel's formalism.

Discrete Lagrangian mechanics is obtained by discretizing Hamilton's variational principle. This approach leads to symplectic- and, for systems with symmetry, momentum-preserving integrators. By discretizing the Lagrange–d'Alembert principle, nonconservative forces (see Kane et al. [20] and Marsden and West [28]) and nonholonomic constraints (see Cortés and Martínez [12]) can be incorporated as well. Recall that, in the continuous-time setting, the dynamics of a Lagrangian system with nonholonomic constraints may be reformulated as the dynamics of an unconstrained system by adding the constraint reaction force. See Suslov [37] and Chetaev [11] for details and precise statements. However, as pointed out in Cortés and Martínez [12], the discretizations of these two representations, as a rule, are not the same, which makes the versions of the discrete Lagrange–d'Alembert principle of [20, 28] and [12] incompatible. In other words, the notion of an ideal constraint of continuous-time mechanics is not retained by the discretization of Cortés and Martínez.

Following the variational discretization approach, we develop discrete Hamel's formalism by discretizing Hamilton's principle for Hamel's equations. The principal difficulty in extending this program to Hamel's setting is caused by the bracket terms, as a discrete analogue of the Jacobi–Lie bracket is known only for left- or right-invariant vector fields on Lie groups (Moser and Veselov [32], Marsden, Pekarsky, and Shkoller [29], Bobenko and Suris [6, 7]). In this paper we resolve the bracket term discretization issue for systems on vector spaces.

When a continuous-time system is discretized, we first select the vector fields that are used to measure the velocity components, and then set up the discrete variational principle. In general, the outcome is a somewhat different discrete dynamical system than the outcome of the usual variational discretization procedure. Remarkably, a modification of our formalism for systems with nonholonomic constraints resolves, at least for Chaplygin systems, the ideal constraint issue of Cortés and Martínez. That is, the discrete Lagrange–d'Alembert principle for Hamel's equations in the presence of nonholonomic constraints is identical to the discrete Lagrange–d'Alembert principle of Kane et al. [20] and Marsden and West [28] written after replacing the constraints with their reactions.

Our formalism also contributes to the study of structural stability of nonholonomic integrators. Recently, Lynch and Zenkov [25, 26] discovered that the nonholonomic integrator of Cortés and Martínez, in general, is not structure-preserving, as it is capable of changing the dimension and stability of manifolds of relative equilibria of continuous-time systems. A similar effect was observed in the holonomic setting in [34]. This lack of structural stability is a serious issue as it alters the α - and ω -limit sets, thus making the asymptotic dynamics of the integrator different from the asymptotic dynamics of the underlying continuous-time system. Such an integrator, in principle, is not suitable for long-term numerical simulations of continuous-time nonholonomic systems. Discrete Hamel's equations are certain

to preserve the manifolds of relative equilibria and their stability, and thus are a better candidate for good quality long-term integrators.

The paper is organized as follows: Continuous-time Lagrangian mechanics and Hamel’s formalism, Hamilton’s variational principle, and discrete mechanics are reviewed in Sections 2–4. Discrete Hamel’s formalism is introduced in Section 5. Applications of discrete Hamel’s formalism to nonholonomic mechanics and to global energy-momentum numerical integration of the spherical pendulum are exposed in Sections 6 and 7.

2 Lagrangian Mechanics

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*

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

These equations were originally derived by Lagrange [24] 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 1 below), came later in the work of Hamilton [17, 18] 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 gives rise to the vector field

$$\delta q(t) = \left. \frac{\partial \beta(t, s)}{\partial s} \right|_{s=0} \tag{2}$$

along the curve $q(t)$.

Theorem 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 $a \leq t \leq b$, where we choose variations of the curve $q(t)$ that satisfy the condition $\delta q(a) = \delta q(b) = 0$.

- (ii) *The curve $q(t)$ satisfies the Euler–Lagrange equations (1).*

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 according to the formula

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

For more details and a proof, see e.g. [2, 27], and Theorem 2 below.

3 Lagrangian Mechanics in Non-coordinate Frames

In this section we discuss the continuous-time Hamel formalism and a relevant variational principle, following the exposition of [5].

3.1 The Hamel Equations

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 not necessarily induced by 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 hereafter denoted U . In certain cases, some or all of u_i can be chosen to be *global* vector fields on Q . The components of u_i relative to the coordinate-induced basis $\partial/\partial q^j$ are written as ψ_i^j ; that is,

$$u_i(q) = \psi_i^j(q) \frac{\partial}{\partial q^j},$$

where $i, j = 1, \dots, n$.

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.,

$$\dot{q} = u(q) \cdot \xi, \tag{3}$$

where, by definition,

$$u(q) \cdot \xi := \xi^i u_i(q). \tag{4}$$

When convenient, we reverse the order of factors in (4), i.e., we assume that

$$u(q) \cdot \xi = \xi \cdot u(q).$$

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

$$l(q, \xi) := L(q, u(q) \cdot \xi). \tag{5}$$

The coordinates (q, ξ) are a Lagrangian analogue of non-canonical variables in Hamiltonian dynamics.

Given two elements $\xi, \zeta \in \mathbb{R}^n$, define the antisymmetric bracket operation $[\cdot, \cdot]_q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$u(q) \cdot [\xi, \zeta]_q = [u(q) \cdot \xi, u(q) \cdot \zeta],$$

where $[\cdot, \cdot]$ is the Jacobi–Lie bracket of vector fields on Q . That is, $[\xi, \zeta]_q$ consists of the components of $[u_i \xi^i, u_j \zeta^j](q)$ relative to the frame u_1, \dots, u_n .

Therefore, each tangent space $T_q U$ is isomorphic to the Lie algebra $W_q := (\mathbb{R}^n, [\cdot, \cdot]_q)$, and the tangent bundle TU is diffeomorphic to a Lie algebra bundle over U .

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

$$\langle [\xi, \alpha]_q^*, \zeta \rangle := \langle \alpha, [\xi, \zeta]_q \rangle.$$

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

$$[u_i(q), u_j(q)] = c_{ij}^a(q) u_a(q),$$

$i, j, a = 1, \dots, n$. 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, one defines the directional derivatives $u_i[l]$ for a function $l : TQ \rightarrow \mathbb{R}$ in a usual way. It is straightforward to show that

$$u_i[l] = \psi_i^j \frac{\partial l}{\partial q^j}.$$

For a frame $u = (u_1, \dots, u_n)$, define $u[l]$ by the formula

$$u[l] = (u_1[l], \dots, u_n[l]).$$

The evolution of the variables (q, ξ) is governed by the *Hamel equations*

$$\frac{d}{dt} \frac{\partial l}{\partial \xi^j} = c_{ij}^a \xi^i \frac{\partial l}{\partial \xi^a} + u_j[l], \tag{6}$$

coupled with equation (3). If $u_i = \partial/\partial q^i$, equations (6) become the Euler–Lagrange equations (1). Equations (6) were introduced in [16] (see also [33] and [5] for details and some history).

3.2 Hamilton’s Principle for Hamel’s Equations

The variational derivation of Hamel’s equations in this section mostly follows [5]. We refer the readers to [27] for the related history of the development of variational principles for the Euler–Lagrange, Euler–Poincaré, and Hamel equations, and to [1] for the Hamilton–Pontryagin principle for the Hamel equations.

Theorem 2 (Zenkov, Bloch, and Marsden [5]). *Let $L : TQ \rightarrow \mathbb{R}$ be a Lagrangian and l be its representation in local coordinates (q, ξ) . 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*

$$\int_a^b L(q, \dot{q}) dt \tag{7}$$

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*

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}.$$

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

$$\int_a^b l(q, \xi) dt \tag{8}$$

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

$$\delta q = u(q) \cdot \zeta \equiv u_i(q)\zeta^i, \tag{9}$$

and given by

$$\delta\xi = \dot{\zeta} + [\xi, \zeta]_q. \tag{10}$$

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

$$\frac{d}{dt} \frac{\partial l}{\partial \xi} = \left[\xi, \frac{\partial l}{\partial \xi} \right]_q^* + u[l]$$

coupled with the equations $\dot{q} = u(q) \cdot \xi \equiv \xi^i u_i(q)$.

For the early development of these equations see [35] and [16].

Proof. The equivalence of (i) and (ii) is proved by computing the variation of the action functional (7):

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

Recall that we denote the components of $\delta q(t)$ relative to the frame $u(q(t)) = (u_1(q(t)), \dots, u_n(q(t)))$ by $\zeta(t) = (\zeta^1(t), \dots, \zeta^n(t))$; that is,

$$\delta q(t) = u(q(t)) \cdot \zeta(t) \equiv u_i(q(t))\zeta^i(t).$$

To prove the equivalence of (i) and (iii), we first compute the quantities $\delta\dot{q}$ and $d(\delta q)/dt$. Using the definition (2) of the field δq , one concludes that

$$\delta u_a(q(t)) = \left. \frac{\partial u_a(\beta(t, s))}{\partial s} \right|_{s=0} = \zeta \cdot u[u_a] = \delta q[u_a] \equiv \zeta^b u_b[u_a]. \tag{11}$$

Similarly,

$$\frac{d}{dt} u_b(q(t)) = \dot{q}[u_b] = \xi \cdot u[u_b] \equiv \xi^a u_a[u_b].$$

¹If \mathcal{Q} is a Lie group, this formula is derived in Bloch, Krishnaprasad, Marsden, and Ratiu [4].

Next,

$$\begin{aligned}\delta\dot{q} &= \delta u(q(t)) \cdot \xi(t) + u(q(t)) \cdot \delta\xi(t), \\ \frac{d(\delta q)}{dt} &= \frac{du(q(t))}{dt} \cdot \zeta(t) + u(q(t)) \cdot \dot{\zeta}(t).\end{aligned}$$

Equivalently, in coordinates,

$$\begin{aligned}\delta\dot{q} &= \delta(\xi^i(t)u_i(q(t))) = \delta\xi^i(t)u_i(q(t)) + \xi^i(t)\frac{\partial u_i}{\partial q^j}\delta q^j, \\ \frac{d(\delta q)}{dt} &= \frac{d}{dt}(\zeta^i(t)u_i(q(t))) = \dot{\zeta}^i(t)u_i(q(t)) + \zeta^i(t)\frac{\partial u_i}{\partial q^j}\dot{q}^j.\end{aligned}$$

Since $\delta\dot{q} = d(\delta q)/dt$, we obtain

$$\begin{aligned}u(q(t)) \cdot (\delta\xi(t) - \dot{\zeta}(t)) &= \xi^i(t)\zeta^j(t)(u_i(q(t))[u_j(q(t))] - u_j(q(t))[u_i(q(t))]) \\ &= \xi^i(t)\zeta^j(t)[u_i(q(t)), u_j(q(t))] \equiv [u(q(t)) \cdot \xi(t), u(q(t)) \cdot \zeta(t)],\end{aligned}$$

which implies formula (10).

To prove the equivalence of (iii) and (iv), we use the above formula and compute the variation the functional (8):

$$\begin{aligned}\delta \int_a^b l(q, \xi) dt &= \int_a^b \left(\frac{\partial l}{\partial q} \delta q + \frac{\partial l}{\partial \xi} \delta \xi \right) dt \\ &= \int_a^b \left(\zeta \cdot u[l] + \frac{\partial l}{\partial \xi} \left(\dot{\zeta} + [\xi, \zeta]_{q(t)} \right) \right) dt \\ &= \int_a^b \left(u[l] + \left[\xi, \frac{\partial l}{\partial \xi} \right]_{q(t)}^* - \frac{d}{dt} \frac{\partial l}{\partial \xi} \right) \zeta dt.\end{aligned}$$

The latter vanishes if and only if the Hamel equations are satisfied. \square

3.3 Remarks on the Frame Selection

As discussed in [2, 3], and [5], constraints and symmetry naturally define subbundles of the velocity phase space TQ . For underactuated mechanical systems, the controlled directions define a subbundle of the momentum phase space T^*Q . It may be beneficial to select a frame in such a way that suitable subframes of the

frame and its dual span the mentioned subbundles. Such frames lead to a simpler representation of dynamics and clarify the structure of the mechanical system under consideration (subsystems, interconnections, etc.).

4 Discrete Mechanics

A discrete analogue of Lagrangian mechanics can be obtained by discretizing Hamilton’s principle; this approach underlies the construction of variational integrators. See Marsden and West [28], and references therein, for a more detailed discussion of discrete mechanics.

A key notion is that of the *discrete Lagrangian*, which is a map $L^d : Q \times Q \rightarrow \mathbb{R}$ that approximates the action integral along an exact solution of the Euler–Lagrange equations joining the configurations $q_k, q_{k+1} \in Q$,

$$L^d(q_k, q_{k+1}) \approx \operatorname{ext}_{q \in \mathcal{C}([0,h],Q)} \int_0^h L(q, \dot{q}) dt,$$

where $\mathcal{C}([0, h], Q)$ is the space of curves $q : [0, h] \rightarrow Q$ with $q(0) = q_k, q(h) = q_{k+1}$, and ext denotes extremum.

In the discrete setting, the action integral of Lagrangian mechanics is replaced by an *action sum*

$$S^d(q_0, q_1, \dots, q_N) = \sum_{k=0}^{N-1} L^d(q_k, q_{k+1}),$$

where $q_k \in Q, k = 0, 1, \dots, N$, is a finite sequence in the configuration space. The equations are obtained by the *discrete Hamilton principle*, which extremizes the discrete action given fixed endpoints q_0 and q_N . Taking the extremum over q_1, \dots, q_{N-1} gives the *discrete Euler–Lagrange equations*

$$D_1 L^d(q_k, q_{k+1}) + D_2 L^d(q_{k-1}, q_k) = 0 \tag{12}$$

for $k = 1, \dots, N - 1$. Here and below, $D_i F$ denotes the partial derivative of the function F with respect to its i th input. Equations (12) implicitly define the *update map* $\Phi : Q \times Q \rightarrow Q \times Q$, where $\Phi(q_{k-1}, q_k) = (q_k, q_{k+1})$ and $Q \times Q$ replaces the velocity phase space TQ of continuous-time Lagrangian mechanics.

In the case that Q is a vector space, it may be convenient to use $(q_{k+1/2}, v_{k,k+1})$, where $q_{k+1/2} = \frac{1}{2}(q_k + q_{k+1})$ and $v_{k,k+1} = \frac{1}{h}(q_{k+1} - q_k)$, as a state of a discrete

mechanical system. In such a representation, the discrete Lagrangian becomes a function of $(q_{k+1/2}, v_{k,k+1})$, and the discrete Euler–Lagrange equations read

$$\begin{aligned} & \frac{1}{2}(D_1 L^d(q_{k-1/2}, v_{k-1,k}) + D_1 L^d(q_{k+1/2}, v_{k,k+1})) \\ & + \frac{1}{h}(D_2 L^d(q_{k-1/2}, v_{k-1,k}) - D_2 L^d(q_{k+1/2}, v_{k,k+1})) = 0. \end{aligned}$$

These equations are equivalent to the variational principle

$$\delta S^d = \sum_{k=0}^{N-1} (D_1 L^d(q_{k+1/2}, v_{k,k+1}) \delta q_{k+1/2} + D_2 L^d(q_{k+1/2}, v_{k,k+1}) \delta v_{k,k+1}) = 0, \quad (13)$$

where the variations $\delta q_{k+1/2}$ and $\delta v_{k,k+1}$ are induced by the variations δq_k and are given by the formulae

$$\delta q_{k+1/2} = \frac{1}{2}(\delta q_{k+1} + \delta q_k), \quad \delta v_{k,k+1} = \frac{1}{h}(\delta q_{k+1} - \delta q_k).$$

The discrete Hamel formalism introduced below may be interpreted as a generalization of the representation (13) of discrete mechanics.

5 Discrete Hamel’s Equations

In the rest of the paper we assume that Q is a vector space. Start with a sequence of configurations $\{q_k\}_{k=0}^N$. Given a parameter $\tau \in [0, 1]$, define the points $q_{k+\tau} := (1 - \tau)q_k + \tau q_{k+1}$ for each $0 \leq k \leq N - 1$. The velocity components relative to the frame $u(q)$ at $q_{k+\tau}$ are denoted $\xi_{k,k+1} = (\xi_{k,k+1}^1, \dots, \xi_{k,k+1}^n)$. Similar to [8, 22], the phase space for the suggested discretization of Hamel’s equation is the tangent bundle TQ . In local coordinates (q, ξ) on TQ , the discrete Lagrangian $l^d : TQ \rightarrow \mathbb{R}$ reads $l^d = l^d(q_{k+\tau}, \xi_{k,k+1})$. To discretize a continuous-time system, we suggest the following procedure:

- (i) Select a frame $u(q)$ and identify the continuous-time Lagrangian $l(q, \xi)$, as in (5).
- (ii) Construct the discrete Lagrangian using the formula

$$l^d(q_{k+\tau}, \xi_{k,k+1}) = hl(q_{k+\tau}, \xi_{k,k+1}).$$

The action sum then is

$$s^d = \sum_{k=0}^{N-1} l^d(q_{k+\tau}, \xi_{k,k+1}), \quad (14)$$

which is an approximation of the action integral (8) of the continuous-time system.

Given $\tau \in [0, 1]$, define $\zeta_{k+\tau}$ by the formula

$$\zeta_{k+\tau} = (1 - \tau)\zeta_k + \tau\zeta_{k+1}. \tag{15}$$

The quantities ζ_k , ζ_{k+1} , and $\zeta_{k+\tau}$ will be used below to establish the discrete analogues of the variation formulae (9) and (10).

Define the *discrete conjugate momentum* by

$$\mu_{k,k+1} := D_2 l^d(q_{k+\tau}, \xi_{k,k+1}). \tag{16}$$

Below, we use the notations

$$u_{k+\tau} := u(q_{k+\tau}), \quad l^d_{k+\tau} := l^d(q_{k+\tau}, \xi_{k,k+1}), \quad u[l^d]_{k+\tau} := u[l^d](q_{k+\tau}, \xi_{k,k+1}),$$

etc.

Theorem 3. *The sequence $(q_{k+\tau}, \xi_{k,k+1}) \in TQ$ satisfies the discrete Hamel equations*

$$\begin{aligned} & \frac{1}{h}(\mu_{k-1,k} - \mu_{k,k+1}) + \tau u[l^d]_{k-1+\tau} + (1 - \tau)u[l^d]_{k+\tau} \\ & + \tau [\xi_{k-1,k}, \mu_{k-1,k}]^*_{q_{k-1+\tau}} + (1 - \tau)[\xi_{k,k+1}, \mu_{k,k+1}]^*_{q_{k+\tau}} = 0 \end{aligned} \tag{17}$$

if and only if

$$\delta s^d = \delta \sum_{k=0}^{N-1} l^d(q_{k+\tau}, \xi_{k,k+1}) = 0,$$

where

$$\delta q_{k+\tau} = u(q_{k+\tau}) \cdot \zeta_{k+\tau}, \tag{18}$$

$$\delta \xi_{k,k+1} = \frac{1}{h}(\zeta_{k+1} - \zeta_k) + [\xi_{k,k+1}, \zeta_{k+\tau}]_{q_{k+\tau}}. \tag{19}$$

Here $\zeta_0 = \zeta_N = 0$, and $\zeta_{k+\tau}$ is defined in (15), $k = 0, \dots, N - 1$.

In order to obtain a complete system of equations, one supplements (17) with a discrete analogue of the kinematic equation $\dot{q} = u(q) \cdot \xi$. There is a certain freedom in doing that. For now, we assume this discrete analogue to be

$$\frac{\Delta q_k}{h} = u_{k+\tau} \cdot \xi_{k,k+1}.$$

We will use a different discretization of the kinematic equation to construct an integrator for the spherical pendulum in Section 7.

In the coordinate form, the discrete Hamel equations and the formulae for variations read

$$\frac{1}{h}(\mu_{k-1,k;j} - \mu_{k,k+1;j}) + \tau u_j [l^d]_{k-1+\tau} + (1 - \tau) u_j [l^d]_{k+\tau} + \tau c_{ij}^a(q_{k-1+\tau}) \xi_{k-1,k}^i \mu_{k-1,k;a} + (1 - \tau) c_{ij}^a(q_{k+\tau}) \xi_{k,k+1}^i \mu_{k,k+1;a} = 0,$$

and

$$\delta q_{k+\tau}^i = \psi_b^i(q_{k+\tau}) \zeta_{k+\tau}^b,$$

$$\delta \xi_{k,k+1}^b = \frac{1}{h}(\zeta_{k+1}^b - \zeta_k^b) + c_{ij}^b(q_{k+\tau}) \xi_{k,k+1}^i \zeta_{k+\tau}^j,$$

respectively.

Remark. Unlike the continuous-time case, the formulae for variations (18) and (19) cannot be derived in a manner presented in the proof of Theorem 2. The situation here is somewhat similar to the issue encountered and resolved by Chetaev in his work [10] on the equivalence of the Lagrange–d’Alembert and Gauss principles for systems with nonlinear nonholonomic constraints. Recall that Chetaev’s approach was to *define* variations in such a way that the two principles become equivalent.

Proof. Using formulae (18) and (19) and computing the variation of the action sum (14), one obtains

$$\begin{aligned} \delta S^d &= \sum_{k=0}^{N-1} D_1 l^d(q_{k+\tau}, \xi_{k,k+1}) \delta q_{k+\tau} + D_2 l^d(q_{k+\tau}, \xi_{k,k+1}) \delta \xi_{k,k+1} \\ &= \sum_{k=0}^{N-1} \left\langle D_1 l_{k+\tau}^d, u_{k+\tau} \cdot \zeta_{k+\tau} \right\rangle \\ &\quad + \left\langle D_2 l_{k+\tau}^d, (\zeta_{k+1} - \zeta_k)/h + [\xi_{k,k+1}, \zeta_{k+\tau}]_{q_{k+\tau}} \right\rangle \\ &= \sum_{k=1}^{N-1} \left\langle \frac{1}{h}(\mu_{k-1,k} - \mu_{k,k+1}), \zeta_k \right\rangle \\ &\quad + \left\langle u[l^d]_{k+\tau} + [\mu_{k,k+1}, \xi_{k,k+1}]_{q_{k+\tau}}^*, (1 - \tau)\zeta_k + \tau\zeta_{k+1} \right\rangle \\ &= \sum_{k=1}^{N-1} \left\langle \frac{1}{h}(\mu_{k-1,k} - \mu_{k,k+1}), \zeta_k \right\rangle + \left\langle \tau u[l^d]_{k-1+\tau} + (1 - \tau)u[l^d]_{k+\tau}, \zeta_k \right\rangle \\ &\quad + \left\langle \tau [\mu_{k-1,k}, \xi_{k-1,k}]_{q_{k-1+\tau}}^* + (1 - \tau)[\mu_{k,k+1}, \xi_{k,k+1}]_{q_{k+\tau}}^*, \zeta_k \right\rangle. \end{aligned}$$

Thus, vanishing of δS^d for arbitrary ζ_k , $k = 1, \dots, N - 1$, is equivalent to discrete Hamel's equations (17). \square

The formulae for variations (18) and (19) in the discrete setting are motivated by the following observations. First, recall that in the continuous-time setting the formula (10) for $\delta\xi$ follows from the formula

$$\delta(u \cdot \xi) - \frac{d}{dt}(u \cdot \zeta) = 0. \tag{20}$$

A discrete analogue of $\delta(u \cdot \xi)$ is relatively straightforward to obtain. Indeed, using the formula

$$\delta q_{k+\tau} = u_{k+\tau} \cdot \zeta_{k+\tau} \equiv u_{k+\tau} \cdot ((1 - \tau)\zeta_k + \tau\zeta_{k+1})$$

and the interpretation of the operator δ as a directional derivative, just like in formula (11), one obtains

$$\delta u_{k+\tau} = (\zeta_{k+\tau} \cdot u[u])_{k+\tau},$$

and therefore

$$\begin{aligned} \delta(u_{k+\tau} \cdot \xi_{k+1}) &= \delta u_{k+\tau} \cdot \xi_{k,k+1} + u_{k+\tau} \cdot \delta \xi_{k,k+1} \\ &= u_{k+\tau} \cdot \delta \xi_{k,k+1} + (\zeta_{k+\tau} \cdot u[\xi_{k,k+1} \cdot u])_{k+\tau}. \end{aligned}$$

However, a discrete analogue of the formula $\frac{d}{dt}(u \cdot \zeta)$ is not immediately available, as the operation of time differentiation is not intrinsically present in the discrete setting. A workaround that we suggest is to view the transition from q_k to q_{k+1} as a motion along a straight line segment at a uniform rate:

$$q_{k+\tau} = (1 - \tau)q_k + \tau q_{k+1}, \quad 0 \leq \tau \leq 1, \tag{21}$$

so that $q_{k+\tau} = q_k$ when $\tau = 0$ and $q_{k+\tau} = q_{k+1}$ when $\tau = 1$. Since the time step is h , the analogue of continuous-time velocity is $\Delta q_k / h$. From (21),

$$\frac{\Delta q_k}{h} = \frac{1}{h} \frac{dq_{k+\tau}}{d\tau},$$

leading to an interpretation of the operator

$$\frac{1}{h} \frac{d}{d\tau}$$

as a discrete analogue of time differentiation of continuous-time mechanics.

The discrete analogue of the term $\frac{d}{dt}(u \cdot \zeta)$ thus is

$$\begin{aligned} \frac{1}{h} \frac{d}{d\tau}(u_{k+\tau} \cdot \zeta_{k+\tau}) &= \frac{1}{h} \frac{du_{k+\tau}}{d\tau} \cdot \zeta_{k+\tau} + u_{k+\tau} \cdot \frac{1}{h} \frac{d\zeta_{k+\tau}}{d\tau} \\ &= u_{k+\tau} \cdot \frac{1}{h} \frac{d\zeta_{k+\tau}}{d\tau} + (\xi_{k,k+1} \cdot u[\zeta_{k+\tau} \cdot u])_{k+\tau} \\ &= u_{k+\tau} \cdot \frac{\zeta_{k+1} - \zeta_k}{h} + (\xi_{k,k+1} \cdot u[\zeta_{k+\tau} \cdot u])_{k+\tau}. \end{aligned}$$

Summarizing, the discrete analogue of (20) reads

$$u_{k+\tau} \cdot \delta \xi_{k,k+1} = u_{k+\tau} \cdot \frac{\zeta_{k+1} - \zeta_k}{h} + [u \cdot \xi_{k,k+1}, u \cdot \zeta_{k+\tau}]_{q_{k+\tau}},$$

which implies formula (19) for variation $\delta \xi$.

6 Hamel's Formalism and Nonholonomic Integrators

In this section we study some of the structure-preserving properties of discrete Hamel's formalism in the presence of velocity constraints.

6.1 The Lagrange–d'Alembert Principle

Assume now that there are *velocity constraints* imposed on the system. We confine our attention to constraints that are 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 spaces 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 is, in general, *nonintegrable*; i.e., the constraints are, in general, *nonholonomic*.²

Consider a Lagrangian $L : TQ \rightarrow \mathbb{R}$. The equations of motion are given by the following **Lagrange–d'Alembert principle**.

²Constraints are nonholonomic if and only if they cannot be rewritten as *position* constraints.

Definition 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 $t \in [a, b]$.

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 [23] and [3] for discussions and references).

6.2 Ideal Constraints

As discussed in e.g. Suslov [37] and Chetaev [11], it is assumed in classical mechanics that the constraints imposed on the system can be replaced with the *reaction forces*. This means that after the forces are imposed on the *unconstrained* system, the constraint distribution 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. Constraints (either holonomic or nonholonomic) are called *ideal* if their reaction forces at each $q \in Q$ belong to the null space $\mathcal{D}_q^\circ \subset T_q^*Q$ of \mathcal{D}_q .

As shown in Suslov [37] and Chetaev [11], the reaction forces of ideal constraints are 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 [37] and [11] for a more detailed exposition and history of the concept of ideal constraints.

6.3 The Constrained Hamel Equations

Given a system with velocity constraints, that is, a Lagrangian $L : TQ \rightarrow \mathbb{R}$ and constraint distribution \mathcal{D} , select the independent local vector fields

$$u_i : Q \rightarrow TQ, \quad i = 1, \dots, n,$$

such that $\mathcal{D}_q = \text{span}\{u_1(q), \dots, u_m(q)\}$, $m < n$. Each $\dot{q} \in TQ$ can be uniquely written as

$$\dot{q} = u(q) \cdot \xi^{\mathcal{D}} + u(q) \cdot \xi^{\mathcal{U}}, \tag{22}$$

where $u(q) \cdot \xi^{\mathcal{D}}$ is the component of \dot{q} along \mathcal{D}_q and $u(q) \cdot \xi^{\mathcal{U}}$ is the complementary component. Similarly, each $a \in T^*Q$ can be uniquely decomposed as

$$a = a_{\mathcal{D}} \cdot u^*(q) + a_{\mathcal{U}} \cdot u^*(q),$$

where $a_{\mathcal{D}} \cdot u^*(q)$ is the component of a along the dual of \mathcal{D}_q , where $a_{\mathcal{U}} \cdot u^*(q)$ is the complementary component, and where $u^*(q) \in T^*Q \times \dots \times T^*Q$ denotes the dual frame of $u(q)$. Using the decomposition (22), the constraints read

$$\xi = \xi^{\mathcal{D}} \quad \text{or} \quad \xi^{\mathcal{U}} = 0. \tag{23}$$

Similar to (22), we write

$$\delta q = u(q) \cdot \zeta = u(q) \cdot \zeta^{\mathcal{D}} + u(q) \cdot \zeta^{\mathcal{U}}.$$

Recall that $\delta q(t) \in \mathcal{D}_{q(t)}$, which is equivalent to

$$\zeta = \zeta^{\mathcal{D}} \quad \text{or} \quad \zeta^{\mathcal{U}} = 0. \tag{24}$$

The Lagrange–d’Alembert principle in combination with (24) proves the following theorem:

Theorem 4. *The dynamics of a system with velocity constraints is represented by the **constrained Hamel equations***

$$\left(\frac{d}{dt} \frac{\partial l}{\partial \xi} - \left[\xi^{\mathcal{D}}, \frac{\partial l}{\partial \xi} \right]_q^* - u[l] \right)_{\mathcal{D}} = 0, \quad \xi^{\mathcal{U}} = 0,$$

coupled with the kinematic equation

$$\dot{q} = u(q) \cdot \xi^{\mathcal{D}}.$$

The **constrained Lagrangian** is the restriction of the Lagrangian to the constraint distribution. Thus, using Hamel’s formalism, the constrained Lagrangian reads

$$l_c(q, \xi^{\mathcal{D}}) = l(q, \xi^{\mathcal{D}}, 0) \equiv l(q, \xi)|_{\xi^{\mathcal{U}}=0}.$$

It is straightforward to check that an alternative form of the constrained Hamel equations is

$$\frac{d}{dt} \frac{\partial l_c}{\partial \xi^{\mathcal{D}}} - \left(\left[\xi^{\mathcal{D}}, \frac{\partial l}{\partial \xi} \right]_q^* \right)_{\mathcal{D}} - u_{\mathcal{D}}[l_c] = 0, \quad \xi^{\mathcal{U}} = 0. \tag{25}$$

6.4 Continuous-Time Chaplygin Systems

As an important special case, consider *commutative Chaplygin systems*, which are nonholonomic systems with a commutative symmetry group H , $\dim H = n - m$, and subject to the condition that at each $q \in Q$ the tangent space $T_q Q$ is the direct sum of the fiber of the constraint distribution and the tangent space to the orbit $\text{Orb}_H(q)$ of H through q :

$$T_q Q = \mathcal{D}_q \oplus T_q \text{Orb}_H(q). \tag{26}$$

To avoid technical difficulties, assume that the group H acts freely and properly on the configuration space Q , so that $\pi : Q \rightarrow Q/H$ is a principal fiber bundle, where π is the projection. Elements of Q/H and H are denoted x and s , respectively.

Following [3], define an *Ehresmann connection* by requiring that \mathcal{D}_q and $T_q \text{Orb}_H(q)$ are the *horizontal and vertical spaces* at $q \in Q$, respectively. These spaces are denoted H_q and V_q .

In other words, the nonholonomic kinematic constraints provide an Ehresmann connection on the principal bundle $\pi : Q \rightarrow Q/H$. Under the assumptions made above, the equations of motion drop to the reduced space \mathcal{D}/H , which in this special case is the same as $T(Q/H)$.

Recall that an Ehresmann connection A on a bundle Q is a *vertical-valued one-form* that is a *projection*; i.e., $A_q : T_q Q \rightarrow V_q$ is a linear map for each $q \in Q$ and $A(v) = v$ for all $v \in V_q$. In the bundle coordinates (x, s) introduced above, the form A reads

$$A = \omega^a \frac{\partial}{\partial s^a}, \quad \text{where} \quad \omega^a(q) = A^a_\alpha(x) dx^\alpha + ds^a, \tag{27}$$

where $\alpha = 1, \dots, m$ and $a = m + 1, \dots, n$. Recall also that the horizontal space $H_q = \ker A_q$, so that $T_q Q = H_q \oplus V_q$, in full agreement with (26).

The *curvature* of A is the vertical-valued two-form defined by

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

where $\text{hor } X$ and $\text{hor } Y$ are the horizontal parts of the vectors $X, Y \in T_q Q$. In the bundle coordinates (x, s) ,

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

where

$$B_{\alpha\beta}^a = \frac{\partial A_\alpha^a}{\partial r^\beta} - \frac{\partial A_\beta^a}{\partial r^\alpha}.$$

Recall that the constrained Lagrangian is the restriction of the Lagrangian onto the constraint distribution: $L_c = L|_{\mathcal{D}}$. For Chaplygin systems, L and L_c naturally reduce to the functions on TQ/H and \mathcal{D}/H , respectively. In the bundle coordinates (x, s) , this simply means that L is independent of s ,³ i.e., $L = L(x, \dot{x}, \dot{s})$, and the constrained Lagrangian reads

$$L_c(x, \dot{x}) = L(x, \dot{x}, -A(x)\dot{x}).$$

The equations of motion for Chaplygin systems,

$$\frac{d}{dt} \frac{\partial L_c}{\partial \dot{x}} - \frac{\partial L_c}{\partial x} = \left\langle \frac{\partial L}{\partial \dot{s}}, \mathbf{i}_x B \right\rangle, \tag{28}$$

or, in coordinates,

$$\frac{d}{dt} \frac{\partial L_c}{\partial \dot{x}^\alpha} - \frac{\partial L_c}{\partial x^\alpha} = -\frac{\partial L}{\partial \dot{s}^a} B_{\alpha\beta}^a \dot{x}^\beta,$$

$\alpha, \beta = 1, \dots, m, a = m + 1, \dots, n$, were first derived, through a coordinate calculation, by Chaplygin in [9]. They are called the **Chaplygin equations**.

Following [30], we now obtain equations (28) using Hamel’s formalism. Recall that connection (27) is defined by the constraint distribution. Equivalently, the constraints read

$$\dot{s} + A(x)\dot{x} = 0.$$

Associated with the constraint distribution are the vector fields

$$u_\alpha = \text{hor } \partial_{x^\alpha} = \partial_{x^\alpha} - A_\alpha^a \partial_{s^a} \quad \text{and} \quad u_a = \partial_{s^a}. \tag{29}$$

³For a noncommutative symmetry group, L depends on (s, \dot{s}) through the combination $s^{-1}\dot{s}$.

Using this frame,

$$\dot{q} = \dot{x}^\alpha u_\alpha + (\dot{s}^a + A_\alpha^a \dot{x}^\alpha) u_a,$$

$\alpha = 1, \dots, m, a = m + 1, \dots, n$, or, equivalently,

$$\xi^{\mathcal{D}} = \dot{x}, \quad \xi^{\mathcal{U}} = \dot{s} + A(x) \dot{x}, \quad \dot{q} = u_{\mathcal{D}} \cdot \xi^{\mathcal{D}} + u_{\mathcal{U}} \cdot \xi^{\mathcal{U}},$$

and

$$l(x, \xi) = L(x, \xi^{\mathcal{D}}, \xi^{\mathcal{U}} - A(x)\xi^{\mathcal{D}}), \quad l_c(x, \xi^{\mathcal{D}}) = L(x, \xi^{\mathcal{D}}, -A(x)\xi^{\mathcal{D}}). \tag{30}$$

Evaluating the Jacobi–Lie brackets of the fields (29), one obtains

$$[u_\alpha, u_\beta] = \left(\frac{\partial A_\alpha^a}{\partial x^\beta} - \frac{\partial A_\beta^a}{\partial x^\alpha} \right) \frac{\partial}{\partial s^a} \equiv B_{\alpha\beta}^a \frac{\partial}{\partial s^a}, \quad [u_\alpha, u_a] = [u_a, u_b] = 0,$$

which implies

$$\left(\left[\xi^{\mathcal{D}}, \frac{\partial l}{\partial \xi} \right]_q^* \right)_{\mathcal{D}} = \left\langle \frac{\partial L}{\partial \dot{s}}, \mathbf{i}_{\dot{x}} B \right\rangle,$$

and thus (28) are just the constrained Hamel equations (25). Recall that B is the curvature of the form A .

An important remark is that, from Chaplygin’s prospective, equations (28) are the Euler–Lagrange equations on the configuration space Q/H subject to a nonconservative force

$$\left\langle \frac{\partial L}{\partial \dot{s}}, \mathbf{i}_{\dot{x}} B \right\rangle.$$

This force may be interpreted as a shape component of the constraint reaction.

Another important remark is that \dot{x}^α in the classical literature are viewed as the reduced configuration velocities, whereas from the point of view of Hamel’s formalism \dot{x}^α represent the velocity components along the non-commuting fields $u_\alpha, \alpha = 1, \dots, m$.

6.5 Discrete Nonholonomic Systems

Discrete nonholonomic systems (nonholonomic integrators) were introduced by Cortés and Martínez in [12].

Let Q be a configuration space. According to Cortés and Martínez, a **discrete nonholonomic mechanical system** on Q is characterized by:

- (i) A discrete Lagrangian $L^d : Q \times Q \rightarrow \mathbb{R}$;
- (ii) A constraint distribution \mathcal{D} on Q ;
- (iii) A **discrete constraint manifold** $\mathcal{D}^d \subset Q \times Q$ which has the same dimension as \mathcal{D} and satisfies the condition $(q, q) \in \mathcal{D}^d$ for all $q \in Q$.

The dynamics is given by the following **discrete Lagrange–d’Alembert principle** (see [12]):

$$\sum_{k=1}^{N-1} \left(D_1 L^d(q_k, q_{k+1}) + D_2 L^d(q_{k-1}, q_k) \right) \delta q_k = 0, \quad \delta q_k \in \mathcal{D}_{q_k}, \quad (q_k, q_{k+1}) \in \mathcal{D}^d.$$

As pointed out in [14, 15], the discrete constraint manifold should be carefully selected when a continuous-time nonholonomic system is discretized. For the details on the properties of discrete nonholonomic systems we refer the reader to papers [12, 14, 15, 31]. In a recent paper [22], a somewhat different approach to discretizing nonholonomic systems has been suggested.

Cortés and Martínez also study the dynamics of discrete Chaplygin systems. In particular, given a continuous-time Chaplygin system, they discretize the *Euler–Lagrange equations with constraint reactions*, and conclude that, in general, the resulting discrete system is inconsistent with the outcome of their discrete Lagrange–d’Alembert principle. In other words, *the concept of ideal constraints is not acknowledged by their discretization procedure*.

Lynch and Zenkov [25, 26] proved that the discrete dynamics defined by the Lagrange–d’Alembert principle of Cortés and Martínez may lack *structural stability*. For example, it is possible for the discretization of a continuous-time Chaplygin system to change the dimension and/or stability of manifolds of relative equilibria of the said continuous-time system.

Below, we shall show that a different definition of the discrete Lagrange–d’Alembert principle exists that is free of the aforementioned issues. In particular, the dimension and stability of manifolds of relative equilibria are kept intact if this new version of the Lagrange–d’Alembert principle is utilized.

6.6 Hamel’s Formalism for Discrete Nonholonomic Systems

Recall that the Lagrange–d’Alembert principle for continuous-time nonholonomic systems assumes that the variation of action is carried out before imposing the constraints. The outcome is the constrained Hamel equations, as discussed in Section 6.3. In a similar manner, we accept that the dynamics of a discrete nonholonomic system is determined by the **discrete Lagrange–d’Alembert principle**, obtained by *first* taking the variation of the discrete action (14) using variations (18)

and (19) subject to the discrete analogue of (24), and *then* imposing the discrete constraints. We emphasize that the definition of the discrete Lagrange–d’Alembert principle given here *is not* the same as the definition of Cortés and Martínez reproduced in Section 6.5.

In the continuous-time setting, the constraints are represented by formula (23). We thus suggest that, under the same assumptions on the frame selection as in Section 6.3, the discrete constraints are

$$\xi_{k,k+1} = \xi_{k,k+1}^{\mathcal{D}} \quad \text{or} \quad \xi_{k,k+1}^{\mathcal{U}} = 0.$$

The discrete analogue of (24) is

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

Arguing like in Section 6.3, one proves the discrete analogue of Theorem 4:

Theorem 5. *The dynamics of a discrete system with velocity constraints is given by the **constrained discrete Hamel equations***

$$\begin{aligned} & \frac{1}{h}(\mu_{k-1,k} - \mu_{k,k+1})_{\mathcal{D}} + (\tau u[l^d]_{k-1+\tau} + (1-\tau)u[l^d]_{k+\tau})_{\mathcal{D}} \\ & + (\tau[\xi_{k-1,k}^{\mathcal{D}}, \mu_{k-1,k}]_{q_{k-1+\tau}}^* + (1-\tau)[\xi_{k,k+1}^{\mathcal{D}}, \mu_{k,k+1}]_{q_{k+\tau}}^*)_{\mathcal{D}} = 0, \end{aligned} \quad (31)$$

where $\mu_{k,k+1}$ is given by formula (16).

Of a special interest is the value $\tau = 1/2$, in which case one verifies that the order of approximation of (31) is 2.

6.7 Discrete Chaplygin Systems

Given a continuous-time Chaplygin system, we construct its discretization by utilizing the discrete Hamel formalism. Using the frame (29) and the continuous-time Lagrangians (30) introduced in Section 6.4, the discrete Lagrangian and the discrete constrained Lagrangian read

$$\begin{aligned} l^d(x_{k+\tau}, \xi_{k,k+1}) &= hl(x_{k+\tau}, \xi_{k,k+1}), \\ l_c^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) &= l^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) \equiv hl_c(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}). \end{aligned}$$

The dynamics is then given by equation (31), with

$$(\mu_{k,k+1})_{\mathcal{D}} = D_2 l_c^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) \equiv D_2 l^d(x_{k+\tau}, \xi_{k,k+1})|_{\xi_{k,k+1}^{\mathcal{U}}=0}$$

and $\mu_{k,k+1}$ defined as in (16).

We now convert the dynamics into a discrete analogue of the Chaplygin equations (28). Following the general discretization procedure, we obtain the formulae

$$\xi_{k,k+1}^{\mathcal{D}} = \Delta x_k / h, \quad \xi_{k,k+1}^{\mathcal{U}} = \Delta s_k / h + A(x_{k+\tau}) \Delta x_k / h.$$

Then, invoking (30), it is straightforward to see that

$$\begin{aligned} l^d(x_{k+\tau}, \xi_{k,k+1}) &= hl(x_{k+\tau}, \xi_{k,k+1}) \\ &= hL(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}, \xi_{k,k+1}^{\mathcal{U}} - A(x_{k+\tau})\xi_{k,k+1}^{\mathcal{D}}) \end{aligned} \quad (32)$$

and

$$\begin{aligned} l_c^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) &= l^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) = hl_c(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) \\ &= hL_c(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) = hL_c(x_{k+\tau}, \Delta x_k / h) \\ &= hL(x_{k+\tau}, \Delta x_k / h, -A(x_{k+\tau})\Delta x_k / h), \end{aligned} \quad (33)$$

where $L(x, \dot{x}, \dot{s})$ is the Lagrangian of the continuous-time Chaplygin system. From formulae (32), (33), and (29), one obtains

$$\begin{aligned} \mu_{k,k+1} &= D_2 l^d(x_{k+\tau}, \xi_{k,k+1}), \\ (\mu_{k,k+1})_{\mathcal{D}} &= D_2 l_c^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) \\ &= hD_2 L_c(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) = hD_2 L_c(x_{k+\tau}, \Delta x_k / h), \\ (\mu_{k,k+1})_{\mathcal{U}} &= D_3 l^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}, \xi_{k,k+1}^{\mathcal{U}}) \\ &= hD_3 L(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}, \xi_{k,k+1}^{\mathcal{U}} - A(x_{k+\tau})\xi_{k,k+1}^{\mathcal{D}}) \\ &= hD_3 L(x_{k+\tau}, \Delta x_k / h, \Delta s_k / h). \end{aligned}$$

Next, since we utilize the frame (29) just like in the continuous-time setting, the formula

$$\begin{aligned} \left([\xi_{k,k+1}^{\mathcal{D}}, \mu_{k,k+1}]_{q_{k+\tau}}^* \right)_{\mathcal{D}} &= \langle \mu_{k,k+1}, \mathbf{i}_{\xi_{k,k+1}^{\mathcal{D}}} B_{q_{k+\tau}} \rangle \\ &= \langle (\mu_{k,k+1})_{\mathcal{U}}, \mathbf{i}_{\xi_{k,k+1}^{\mathcal{D}}} B_{q_{k+\tau}} \rangle = \langle (\mu_{k,k+1})_{\mathcal{U}}, \mathbf{i}_{\Delta x_k / h} B_{q_{k+\tau}} \rangle \\ &= \langle hD_3 L(x_{k+\tau}, \Delta x_k / h, -A(x_{k+\tau})\Delta x_k / h), \mathbf{i}_{\Delta x_k / h} B_{q_{k+\tau}} \rangle \end{aligned}$$

is established with an aid of the arguments of Section 6.4. To keep the formulae shorter, we write the latter expression as

$$\langle hD_3 L, \mathbf{i}_{\Delta x_k / h} B \rangle_{k+\tau}.$$

Finally,

$$\begin{aligned} (u[l^d](q_{k+\tau}, \xi_{k,k+1}))_{\mathcal{D}} &= D_1 l^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) \\ &= D_1 l_c^d(x_{k+\tau}, \xi_{k,k+1}^{\mathcal{D}}) = h D_1 L_c(x_{k+\tau}, \Delta x_k/h). \end{aligned}$$

Summarizing, the dynamics of the discrete Chaplygin system reads

$$\begin{aligned} \frac{1}{h}((D_2 L_c)_{k+\tau} - (D_2 L_c)_{k-1+\tau}) &= \tau(D_1 L_c)_{k-1+\tau} + (1 - \tau)(D_1 L_c)_{k+\tau} \\ &+ \tau(D_3 L, \mathbf{i}_{\Delta x_{k-1}/h} B)_{k-1+\tau} + (1 - \tau)(D_3 L, \mathbf{i}_{\Delta x_k/h} B)_{k+\tau}, \end{aligned} \tag{34}$$

where $(D_i L_c)_{k+\tau} := D_i L_c(x_{k+\tau}, \Delta x_k/h)$. Remarkably, the discrete Chaplygin equations (34) are identical to the discretization of continuous-time Chaplygin equations (28) viewed as forced Euler–Lagrange dynamics. For more details on this latter discretization of the Chaplygin equations see [12] and [26].

6.8 Stability

In this section we link up stability of relative equilibria of Chaplygin systems with structural stability of nonholonomic integrators.

Consider a commutative Chaplygin system characterized by the Lagrangian L and constraint distribution \mathcal{D} , as discussed in Section 6.4. Assume that the dynamics of the Chaplygin system (28) is invariant with respect to the action of a commutative group G on Q/H .⁴ Often such a situation is the result of the original system being invariant with respect to the semidirect product $G \ltimes H$ of the groups G and H . The elements of the group G are denoted g , and we assume that the action of G on Q/H is free and proper, so that Q/H has the structure of a principal fiber bundle with the structure group G . Thus, locally, there exist the bundle coordinates $x = (r, g)$ on Q/H .

Under certain assumptions (see e.g. [21] and [39]), the dynamics has a manifold (whose dimension equals $\dim G$) of relative equilibria. These relative equilibria are the solutions of (28) that in the bundle coordinates (r, g) read

$$r = r_e, \quad \dot{g} = \eta_e.$$

As established in Karapetyan [21], some of these relative equilibria may be *partially asymptotically stable*. Karapetyan justifies stability using the center manifold stability analysis, which, for nonholonomic systems under consideration, reduces

⁴The general noncommutative setting is not studied in this paper and will be the subject of a future publication.

to verifying that the nonzero spectrum of linearization of (28) at the relative equilibrium of interest belongs to the left half-plane.⁵

Partially asymptotically stable relative equilibria are a part of the ω -limit set of dynamics (28). Similarly, relative equilibria that become partially asymptotically stable after the time reversal are a part of the α -limit set of dynamics (28).

It is important for a long-term numerical integrator to preserve the manifold of relative equilibria and their stability types. Indeed, if the limit sets of an integrator are different from the limit sets of the continuous-time dynamics, this integrator will not adequately simulate the continuous-time dynamics over long time intervals.

As shown in [25, 26], the discrete Lagrange–d’Alembert principle of Cortés and Martínez may produce discretizations that fail to preserve the manifold of relative equilibria. For instance, it may change the dimension of this manifold, thus changing the structure of the limit sets. Informally, the origin of this effect can be explained as follows: The discrete Lagrange–d’Alembert principle of Cortés and Martínez is capable of introducing reactions that correspond to non-ideal constraints. A typical example would be a reaction force with a dissipative component, whose discrete counterpart causes the aforementioned changes of relative equilibria.

A relative equilibrium of a discrete Chaplygin system (34) with commutative symmetry is a solution

$$r_k = \text{const}, \quad \Delta g_k = \text{const}.$$

Assume now that $\tau = 1/2$ in equations (34). Let $h > 0$ be the time step.

Theorem 6 (Lynch and Zenkov [25, 26]). *Discretization (34)⁶ preserves the manifold of relative equilibria of the continuous-time Chaplygin system; that is, $r_k = r_e$, $\Delta g_k = h\eta_e$ is a relative equilibrium of the discretization (34) if and only if $r = r_e$, $\dot{g} = \eta_e$ is a relative equilibrium of the continuous-time system. The conditions for partial asymptotic stability of the equilibria of the continuous-time system and of its discretization are the same.*

Summarizing, the discrete Lagrange–d’Alembert principle proposed in this paper ensures the necessary conditions for structural stability of the associated nonholonomic integrator.

⁵The stability analysis of relative equilibria of nonholonomic systems has a long history, starting from the results of Walker [38] and Routh [36]; see [39] for some of this history and for the energy-momentum method for nonholonomic systems.

⁶Equations (34) were derived in [25, 26] without the use of the discrete Hamel formalism.

7 The Spherical Pendulum

Here we outline the results of Zenkov, Leok, and Bloch [40] on the applications of the discrete Hamel formalism to the energy-momentum-preserving integrator for the spherical pendulum.

7.1 The Spherical Pendulum as a Degenerate Rigid Body

Consider a spherical pendulum whose length is r and mass is m . We view the pendulum as a point mass moving on the sphere of radius r centered at the origin of \mathbb{R}^3 . The development here is based on the representation

$$\dot{\boldsymbol{\mu}} = \boldsymbol{\mu} \times \boldsymbol{\xi} + mg \boldsymbol{\gamma} \times \boldsymbol{a}, \quad (35)$$

$$\dot{\boldsymbol{\gamma}} = \boldsymbol{\gamma} \times \boldsymbol{\xi} \quad (36)$$

of pendulum's dynamics; that is, the pendulum is viewed as a rigid body rotating about a fixed point. This rigid body is of course *degenerate*, with the inertia tensor $\mathcal{I} = \text{diag}\{mr^2, mr^2, 0\}$. Here $\boldsymbol{\xi}$ is the angular velocity of the pendulum, $\boldsymbol{\mu}$ is its angular momentum, $\boldsymbol{\gamma}$ is the unit vertical vector (and thus the constraint $\|\boldsymbol{\gamma}\| = 1$ is imposed), and \boldsymbol{a} is the vector from the origin to the center of mass, which for the pendulum is its bob, all written relative to the body frame. Throughout the rest of the paper, the boldface characters represent three-dimensional vectors. The kinetic and potential energies of the pendulum are

$$K = \frac{1}{2} \langle \boldsymbol{\mu}, \boldsymbol{\xi} \rangle \equiv \frac{1}{2} \langle \mathcal{I} \boldsymbol{\xi}, \boldsymbol{\xi} \rangle, \quad U = mg \langle \boldsymbol{\gamma}, \boldsymbol{a} \rangle \equiv mgr \gamma^3,$$

and the Lagrangian reads

$$l(\boldsymbol{\xi}, \boldsymbol{\gamma}) = \frac{1}{2} \langle \mathcal{I} \boldsymbol{\xi}, \boldsymbol{\xi} \rangle - mg \langle \boldsymbol{\gamma}, \boldsymbol{a} \rangle. \quad (37)$$

This Lagrangian is invariant with respect to rotations about $\boldsymbol{\gamma}$, and therefore the vertical component of the *spatial angular momentum* is conserved.

There are two independent components in the vector equation (35). We emphasize that the representation (35) and (36) of the dynamics of the pendulum, though redundant, eliminates the use of local coordinates on the sphere, such as spherical coordinates. Spherical coordinates, while being a nice theoretical tool, introduce artificial singularities at the north and south poles. That is, the equations of motion written in spherical coordinates have denominators vanishing at the poles, but this has nothing to do with the physics of the problem and is solely caused by the geometry of the spherical coordinates. Thus, the use of spherical coordinates in calculations is not advisable.

Another important remark is that the length of the vector $\boldsymbol{\gamma}$ is a *conservation law* of equations (35) and (36), and thus adding the constraint $\|\boldsymbol{\gamma}\| = 1$ *does not* result in a system of differential-algebraic equations. The latter are known to be a nontrivial object for numerical integration.

Equations (35) and (36) may be interpreted in a number of ways. In the above, we viewed them as the dynamics of a degenerate rigid body. Since the moment of inertia relative to the direction of the vector \boldsymbol{a} is zero, the third component of the body angular momentum vanishes,

$$\mu_3 = \frac{\partial l}{\partial \xi^3} = 0,$$

and thus there are only two nontrivial equations in (35). Thus, one needs five equations to capture the pendulum dynamics. This reflects the fact that rotations about the direction of the pendulum have no influence on the pendulum's motion.

The dynamics then can be simplified by setting

$$\xi^3 = 0, \tag{38}$$

which leads to an interpretation of equations (35) and (36) as the dynamics of the *heavy Suslov top*⁷ with a rotationally-invariant inertia tensor and constraint (38).

Summarizing, the dynamics becomes

$$\dot{\boldsymbol{\mu}} = m\boldsymbol{g}\boldsymbol{\gamma} \times \boldsymbol{a}, \quad \dot{\boldsymbol{\gamma}} = \boldsymbol{\gamma} \times \boldsymbol{\xi}, \quad \langle \boldsymbol{\xi}, \boldsymbol{a} \rangle = 0. \tag{39}$$

These equations are in fact the constrained Hamel equations, the reconstruction equation, and the constraint, written in the redundant configuration coordinates $\boldsymbol{\gamma} = (\gamma^1, \gamma^2, \gamma^3)$; see [40] for details. Recall that the length of $\boldsymbol{\gamma}$ is the conservation law, so that the constraint $\|\boldsymbol{\gamma}\| = 1$ does not need to be imposed, but the appropriate level set of the conservation law has to be selected.

Our discretization is based on this point of view, i.e., the discrete dynamics will be written in the form of discrete Hamel's equations. The discrete dynamics will possess the discrete version of the conservation law $\|\boldsymbol{\gamma}\| = \text{const}$, so that the algorithm should be capable, in theory, of preserving the length of $\boldsymbol{\gamma}$ up to machine precision.

7.2 Variational Discretization for the Spherical Pendulum

The integrator for the spherical pendulum is constructed by discretizing equations (39).

⁷See [33, 37], and [2] for the Suslov top.

Let the positive real constant h be the time step. Applying the mid-point rule to (37), the discrete Lagrangian is computed to be

$$l^d(\boldsymbol{\xi}_{k,k+1}, \boldsymbol{\gamma}_{k+1/2}) = \frac{h}{2} \langle \mathcal{I} \boldsymbol{\xi}_{k,k+1}, \boldsymbol{\xi}_{k,k+1} \rangle - hU(\boldsymbol{\gamma}_{k+1/2}).$$

Here $\boldsymbol{\xi}_{k,k+1} = (\xi_{k,k+1}^1, \xi_{k,k+1}^2, 0)$ is the discrete analogue of the angular velocity $\boldsymbol{\xi} = (\xi^1, \xi^2, 0)$ and $\boldsymbol{\gamma}_{k+1/2} = \frac{1}{2}(\boldsymbol{\gamma}_{k+1} + \boldsymbol{\gamma}_k)$. The discrete dynamics then reads

$$\frac{1}{h} \mathcal{I} (\boldsymbol{\xi}_{k,k+1} - \boldsymbol{\xi}_{k-1,k}) = mg(\boldsymbol{\gamma}_{k+1/2} + \boldsymbol{\gamma}_{k-1/2}) \times \mathbf{a}, \tag{40}$$

$$\frac{1}{h} (\boldsymbol{\gamma}_{k+1/2} - \boldsymbol{\gamma}_{k-1/2}) = \frac{1}{2}(\boldsymbol{\gamma}_{k+1/2} + \boldsymbol{\gamma}_{k-1/2}) \times \frac{1}{2}(\boldsymbol{\xi}_{k,k+1} + \boldsymbol{\xi}_{k-1,k}). \tag{41}$$

We reiterate that there is a certain flexibility in setting up the discrete analogue (41) of the continuous-time reconstruction equation (36). Our choice may be justified in a number of ways, one of them being energy conservation by the discrete dynamics.

The structure-preserving properties of the proposed integrator for the spherical pendulum are summarized in the following theorem.

Theorem 7 (Zenkov, Leok, and Bloch [40]). *The discrete spherical pendulum dynamics (40) and (41) preserves the energy, the vertical component of the spatial angular momentum, and the length of $\boldsymbol{\gamma}$.*

We refer the readers to [40] for the proof and details.

7.3 Simulations

Here we present simulations of the dynamics of the spherical pendulum using the integrator constructed in Section 7.2. For simulations, we select the parameters of the system and the time step to be

$$m = 1 \text{ kg}, \quad r = 9.8 \text{ m}, \quad h = .2 \text{ s}.$$

The trajectory of the bob of the pendulum with the initial conditions

$$\xi_0^1 = .6 \text{ rad/s}, \quad \xi_0^2 = 0 \text{ rad/s}, \tag{42}$$

$$\gamma_0^1 = .3 \text{ m}, \quad \gamma_0^2 = .2 \text{ m}, \quad \gamma_0^3 = -\sqrt{1 - (\gamma_0^1)^2 - (\gamma_0^2)^2} \text{ m} \tag{43}$$

is shown in Figure 1a. As expected, it reveals the quasiperiodic nature of pendulum's dynamics.

Figure 1b shows pendulum's trajectory that crosses the equator. This simulation demonstrates the global nature of the algorithm, and also seems to do a good job of hinting at the geometric conservation properties of the method.

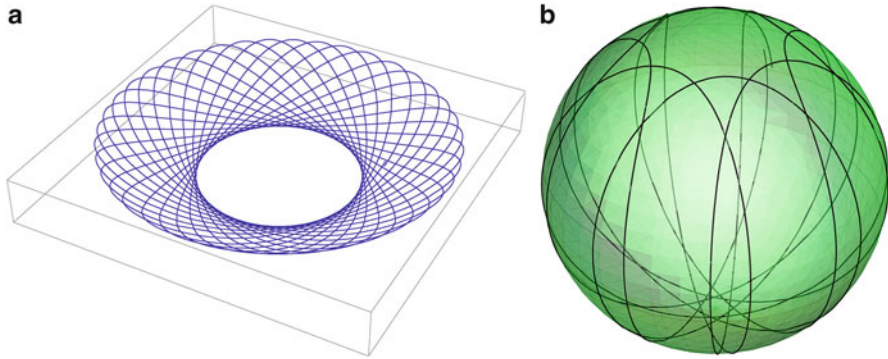


Fig. 1 Trajectories of the pendulum calculated with the Hamel integrator. **(a)** Pendulum's trajectory on S^2 for initial conditions (42) and (43). **(b)** Pendulum's trajectory on S^2 that crosses the equator

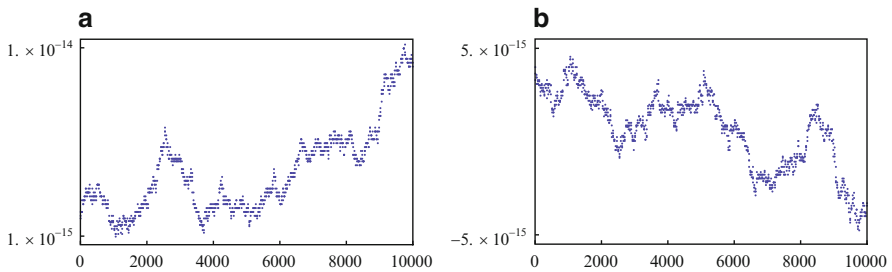


Fig. 2 Numerical properties of the Hamel integrator for the pendulum. **(a)** Preservation of the length of γ . **(b)** Conservation of energy

Theoretically, if one solves the nonlinear equations exactly, and in the absence of numerical roundoff error, the Hamel variational integrator should exactly preserve the length constraint and the energy. In practice, Figure 2a demonstrates that $\|\gamma\|$ stays to within unit length to about 10^{-14} after 10,000 iterations. Figure 2b demonstrates numerical energy conservation, and the energy error is to about $5 \cdot 10^{-15}$ after 10,000 iterations. Indeed, one notices that the energy error tracks the length error of the simulation, which is presumably due to the relationship between the length of the pendulum and the potential energy of the pendulum. The drift in both appear to be due to accumulation of numerical roundoff error, and could possibly be reduced through the use of compensated summation techniques.

For the comparison of the Hamel integrator with simulations using the generalized Störmer–Verlet method and the RATTLE method see [40]. We point out here that the energy error for the Hamel integrator is smaller than those of the Störmer–Verlet and RATTLE methods.

8 Conclusions

This paper introduced the discrete Hamel formalism and demonstrated its utility in nonholonomic mechanics. Future work will include further study of the properties of this formalism, as well as the development of discrete Hamel's formalism on manifolds in general, and on Lie groups and homogeneous spaces as important special cases. It would be also interesting to relate the discrete Hamel formalism to the results of Iglesias et al. [19].

Acknowledgements The authors would like to thank Jerry Marsden for his inspiration and helpful discussions in the beginning of this work, and the reviewers for useful remarks. The research of DVZ was partially supported by NSF grants DMS-0604108, DMS-0908995 and DMS-1211454. The research of KRB was partially supported by NSF grants DMS-0604108 and DMS-0908995. DVZ would like to acknowledge support and hospitality of Mathematisches Forschungsinstitut Oberwolfach, the Fields Institute, and the Beijing Institute of Technology, where a part of this work was carried out.

References

1. Ball, K., Zenkov, D.V., Bloch, A.M.: Variational structures for Hamel's equations and stabilization. In: Proceedings of the 4th IFAC, pp. 178–183 (2012)
2. Bloch, A.M.: Nonholonomic Mechanics and Control. Interdisciplinary Applied Mathematics, vol. 24. Springer, New York (2003)
3. Bloch, A.M., Krishnaprasad, P.S., Marsden, J.E., Murray, R.: Nonholonomic mechanical systems with symmetry. Arch. Ration. Mech. Anal. **136**, 21–99 (1996)
4. Bloch, A.M., Krishnaprasad, P.S., Marsden, J.E., Ratiu, T.S.: The Euler–Poincaré equations and double bracket dissipation. Commun. Math. Phys. **175**, 1–42 (1996)
5. Bloch, A.M., Marsden, J.E., Zenkov, D.V.: Quasivelocities and symmetries in nonholonomic systems. Dyn. Syst. Int. J. **24**(2), 187–222 (2009)
6. Bobenko, A.I., Suris, Yu.B.: Discrete Lagrangian reduction, discrete Euler–Poincaré equations, and semidirect products. Lett. Math. Phys. **49**, 79–93 (1999)
7. Bobenko, A.I., Suris, Yu.B.: Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top. Commun. Math. Phys. **204**(1), 147–188 (1999)
8. Bou-Rabee, N., Marsden, J.E.: Hamilton–Pontryagin integrators on Lie groups part I: Introduction and structure-preserving properties. Found. Comput. Math. **9**(2), 197–219 (2009)
9. Chaplygin, S.A.: On the motion of a heavy body of revolution on a horizontal plane. Phys. Sect. Imperial Soc. Friends Phys. Anthropol. Ethnograph. **9**, 10–16 (1897) (in Russian)
10. Chetaev, N.G.: On Gauss' principle. Izv. Fiz-Mat. Obsc. Kazan. Univ. Ser. 3 **6**, 68–71 (1932–1933) (in Russian)
11. Chetaev, N.G.: Theoretical Mechanics. Springer, New York (1989)
12. Cortés, J., Martínez, S.: Nonholonomic integrators. Nonlinearity **14**, 1365–1392 (2001)
13. Euler, L.: Decouverte d'un nouveau principe de Mécanique. Mémoires de l'académie des sciences de Berlin **6**, 185–217 (1752)
14. Fedorov, Yu.N., Zenkov, D.V.: Discrete nonholonomic LL systems on Lie groups. Nonlinearity **18**, 2211–2241 (2005)
15. Fedorov, Yu.N., Zenkov, D.V.: Dynamics of the discrete Chaplygin sleigh. Discrete Continuous Dyn. Syst. (extended volume), 258–267 (2005)

16. Hamel, G.: Die Lagrange–Eulersche Gleichungen der Mechanik. *Z. Math. Phys.* **50**, 1–57 (1904)
17. Hamilton, W.R.: On a general method in dynamics, part I. *Philos. Trans. R. Soc. Lond.* **124**, 247–308 (1834)
18. Hamilton, W.R.: On a general method in dynamics, part II. *Philos. Trans. R. Soc. Lond.* **125**, 95–144 (1835)
19. Iglesias, D., Marrero, J.C., Martín de Diego, D., Martínez, E.: Discrete nonholonomic Lagrangian systems on Lie groupoids. *J. Nonlinear Sci.* **18**(3), 221–276 (2008)
20. Kane, C., Marsden, J.E., Ortiz, M., West, M.: Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *Int. J. Numer. Math. Eng.* **49**(10), 1295–1325 (2000)
21. Karapetyan, A.V.: On the problem of steady motions of nonholonomic systems. *J. Appl. Math. Mech.* **44**, 418–426 (1980)
22. Kobilarov, M., Marsden, J.E., Sukhatme, G.S.: Geometric discretization of nonholonomic systems with symmetries. *Discrete Continuous Dyn. Syst. Ser. S* **3**(1), 61–84 (2010)
23. Kozlov, V.V.: The problem of realizing constraints in dynamics. *J. Appl. Math. Mech.* **56**(4), 594–600 (1992)
24. Lagrange, J.L.: *Mécanique Analytique*. Chez la Veuve Desaint (1788)
25. Lynch, C., Zenkov, D.: Stability of stationary motions of discrete-time nonholonomic systems. In: Kozlov, V.V., Vassilyev, S.N., Karapetyan, A.V., Krasovskiy, N.N., Tkhai, V.N., Chernousko, F.L. (eds.) *Problems of Analytical Mechanics and Stability Theory*. Collection of Papers Dedicated to the Memory of Academician Valentin V. Rumyantsev, pp. 259–271. Fizmatlit, Moscow (2009) (in Russian)
26. Lynch, C., Zenkov, D.V.: Stability of relative equilibria of discrete nonholonomic systems with semidirect symmetry. Preprint (2010)
27. Marsden, J.E., Ratiu, T.S.: *Introduction to Mechanics and Symmetry*. Texts in Applied Mathematics, vol. 17, 2nd edn. Springer, New York (1999)
28. Marsden, J.E., West, M.: Discrete mechanics and variational integrators. *Acta Numerica* **10**, 357–514 (2001)
29. Marsden, J.E., Pekarsky, S., Shkoller, S.: Discrete Euler–Poincaré and Lie–Poisson equations. *Nonlinearity* **12**, 1647–1662 (1999)
30. Maruskin, J.M., Bloch, A.M., Marsden, J.E., Zenkov, D.V.: A fiber bundle approach to the transpositional relations in nonholonomic mechanics. *J. Nonlinear Sci.* **22**(4), 431–461 (2012)
31. McLachlan, R., Perlmutter, M.: Integrators for nonholonomic mechanical systems. *J. Nonlinear Sci.* **16**, 283–328 (2006)
32. Moser, J., Veselov, A.: Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Commun. Math. Phys.* **139**(2), 217–243 (1991)
33. Neimark, Ju.I., Fufaev, N.A.: *Dynamics of Nonholonomic Systems*. Translations of Mathematical Monographs, vol. 33. AMS, Providence, RI (1972)
34. Peng, Y., Huynh, S., Zenkov, D.V., Bloch, A.M.: Controlled Lagrangians and stabilization of discrete spacecraft with rotor. *Proc. CDC* **51**, 1285–1290 (2012)
35. Poincaré, H.: Sur une forme nouvelle des équations de la mécanique. *CR Acad. Sci.* **132**, 369–371 (1901)
36. Routh, E.J.: *Treatise on the Dynamics of a System of Rigid Bodies*. MacMillan, London (1860)
37. Suslov, G.K.: *Theoretical Mechanics*, 3rd edn. GITTL, Moscow–Leningrad (1946) (in Russian)
38. Walker, G.T.: On a dynamical top. *Q. J. Pure Appl. Math.* **28**, 175–184 (1896)
39. Zenkov, D.V., Bloch, A.M., Marsden, J.E.: The energy-momentum method for the stability of nonholonomic systems. *Dyn. Stab. Syst.* **13**, 128–166 (1998)
40. Zenkov, D.V., Leok, M., Bloch, A.M.: Hamel’s formalism and variational integrators on a sphere. *Proc. CDC* **51**, 7504–7510 (2012)