ABSTRACT

BALL, KENNETH RAY. Structure Preserving Integrators and Hamel's Equations. (Under the direction of Dmitry Zenkov.)

Hamel's formalism is a representation of Lagrangian mechanics obtained by measuring the velocity components relative to a frame which is not related to a system's local configuration coordinates. The use of this formalism often leads to a simpler representation of dynamics. This dissertation extends Hamel's formalism to the discrete setting, utilizing the methods of variational integrators. The research presented is motivated by observations that, in some mechanical systems subject to non-integrable constraints and/or symmetries, an approach to discretization based on Hamel's formalism results in numerical integrators that preserve certain structures. These discrete Hamel's equations are demonstrated to be variational integrators, and their application in mechanical systems is examined.

Structure Preserving Integrators and Hamel's Equations

by
Kenneth Ray Ball

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Mathematics

Raleigh, North Carolina

2013

APPROVED BY:

Steve Campbell	David Brown
Irina Kogan	Dmitry Zenkov Chair of Advisory Committee

DEDICATION

To Larry Thomason and Mike Thompson

. . . two truly phenomenal teachers in very different fields.

BIOGRAPHY

Kenneth Ball was born in Tacoma, Washington, to James Ball and Carole Ball (née Barazotti) in 1985. He attended North Rowan High School in Spencer, NC, and had the opportunity to attend the North Carolina Governor's School in mathematics, where nontechnical and yet fascinating courses on manifold theory, non-Euclidean geometry, and matrix theory piqued his interest in mathematics.

Kenneth enrolled at North Carolina State University in the fall of 2003, and after enduring a short-lived stint as an engineering student, he switched majors and graduated in the spring of 2007 with degrees in physics and applied mathematics. As a junior Kenneth took an elective mathematics course on mechanics with his future advisor, further spurring his interest in mathematics and ultimately leading to his present course of research.

He re-enrolled as a doctoral candidate in the Department of Mathematics at NCSU in the fall of 2007. After six years of graduate study working under the direction of Dr. Dmitry Zenkov, Kenneth presented this dissertation for the defense of his Ph.D. in mathematics.

In his spare time, Kenneth likes to hunt waterfowl and upland game when they are in season, and shoot sporting clays when they are not. He also is an avid reader, computer programmer, outdoorsman, runner, painter, cook, and beer connoisseur. His past pursuits (and hopefully future now that his dissertation is finished!) have included music, diving, and hiking. He is an active parishioner at Sacred Heart Cathedral in Raleigh. While on a bit of hiatus, Kenneth is known for homebrewing, and he also regularly cooks great NC barbecue and throws awesome tailgates at NCSU football games.

ACKNOWLEDGEMENTS

First and foremost, I owe my advisor Dr. Dmitry Zenkov a debt of gratitude for his guidance, support, and encouragement.

Thanks to the NSF, which partially supported me through various grants, especially during my years as a graduate research assistant.

Also, many thanks to the NCSU Department of Mathematics, not only for engaging me and providing for my education over the last six years, but also for the REG funding (also partially from the NSF) which led to the research presented in this dissertation.

Many, many, thanks to my wonderful and supportive family, including my mom and dad Carole and James, my sister Virginia and her husband Dmitri, and all of my family in North Carolina, New York, Texas, and many other places.

Thanks to all of my friends in Raleigh, who after ten continuous years at N.C. State are as much a family to me as, well, my family.

Much of the work presented in this paper is inspired by the research of the late Jerry Marsden and his many collaborators around the world. Many thanks to those whose work I have cited in this dissertation. Especially many thanks to Drs. Marsden and West: you could make a good drinking game out of the number of times I cite your 2001 paper in here!

. . . and of course, thanks to you Astra for all of your love and support!

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

This dissertation provides an exposition and examination of structure preserving integrators that are motivated by Hamel's formalism of Lagrangian mechanical systems. The work presented is guided by two important observations in the study of mechanics. The first, dating back to Euler's work on the rigid body in the mid-18th century, is that the use of moving frames may greatly simplify the analysis of motion. The second is that integrators approximating the evolution of Lagrangian mechanical systems may be derived through a discretization of the variational principle that is equivalent to the differential equations of motion. Significantly, these integrators naturally conserve various geometric and mechanical structures.

1.1 Hamel's Equations: A Brief Introduction and History

As pointed out above, the first observation that motivates the work presented in this dissertation is that the equations of motion of a mechanical system and their analysis may be greatly simplified through the introduction of a non-coordinate frame in the velocity space. Euler [18] was probably the first to use body frames to measure the angular velocity components of a rotating rigid body; such frames tremendously simplify the study of rotational dynamics of a rigid body. Lagrange [29], and later Poincaré [43], generalized Euler's pioneering work to more general Lagrangians on the rotation group and to systems on arbitrary Lie groups, respectively. The general form of the differential equations of motion written with the aid of a frame has been studied in the work of Hamel [22] (and more recently in Bloch, Marsden, and Zenkov [10]) and is referred to as Hamel's equations.

The configuration of a mechanical system with n degrees of freedom can be represented (locally) by n generalized coordinates $(q^i)_{i=1}^n$ that are local coordinates in an n-dimensional configuration space Q (a differentiable manifold). The dynamics are dictated by a Lagrangian—a function on the tangent bundle of Q, which is the system's velocity phase space; for most physical systems the Lagrangian is the difference between kinetic and potential energy.

A coordinate-induced frame is a set of n vector fields that are tangent to the n coordinate lines at each point in a coordinate chart on Q, and as such they form a basis of the tangent space at each point in this neighborhood. We often use the notation $(q^i, \dot{q}^i)$ to denote the lift of the local coordinate chart to the tangent bundle; in doing so we assume that the components $\dot{q}$ of a velocity vector in T_qQ are measured against the coordinate-induced frame. When we use this frame as the basis to measure the velocity components the resulting equations of motion of our mechanical system are the well-known Euler–Lagrange equations.

There are circumstances in which an alternate frame can make the analysis of dynamics simpler. We have already mentioned the free rigid body: Euler's body frame approach uncouples kinematics and dynamics and makes analysis of the rigid body system manageable. This proves to be a vast simplification over the approach of calculating the Euler–Lagrange equations in local coordinates on the configuration space $SO(3)$.

Next, we follow Poincaré in his 1901 paper [43], and consider the Euler–Poincaré equations on a Lie algebra [10, 25, 37]. The free rigid body is then just a particular implementation of the Euler–Poincaré equations. The equations themselves, written in terms of generalized coordinates, include structure constants of the associated Lie algebra. These are actually consistent with the Euler–Poincaré equations as developed by Lagrange [29] for reasonably general Lagrangians on the rotation group (see Marsden and Ratiu [35] for details and history).

Our next historical leap is relatively brief: in the early 20th century, the German mathematician Georg Hamel was studying unlinked velocity and position measurements on an arbitrary configuration space, and in 1904 presented a particular generalization of the Euler–Poincaré equations to the case of a generalized differentiable manifold [22].

Remark 1.1. For the sake of historical accuracy, we point out that it is not clear whether or not Hamel was familiar with Poincaré's 1901 work [43] when he published his 1904 paper [22]. As a student of Hilbert who finished his thesis in 1901, it seems likely that Hamel would have been familiar with Poincaré's results, and by the end of his 1904 paper he does relate his findings to contemporary results in Lie group theory. However, Hamel does not mention Poincaré's paper directly, instead arriving at his conclusions through examination of non-commutativity of variations and time derivatives. Nonetheless, it is natural to consider Hamel's equations as a generalization of the aforementioned Euler–Poincaré equations.

An important observation that characterizes Hamel's result (as a generalization of Poincaré) is that the identification of TG with $G \times \mathfrak{g}$ described above can be thought of as the introduction of a *non-coordinate frame* onto each of the fibers comprising TG. From this perspective, the components of the vector $\dot{g}$ measured against the new frame are denoted by ξ.

An arbitrary (local) frame field over an n-dimensional differentiable manifold is a more general object, which is a set of n vector fields that, at each point q, define a basis of the vector space T_qQ. These frames need not be tied in any formal way to a coordinate chart. Furthermore, the structure constants of a Lie algebra become structure functions, as they will vary with the configuration of the system. We will see that Hamel's equations include a bilinear function, identified with the structure functions and varying with configuration. This term does not appear in the Euler–Lagrange equations and is a generalization of the bracket of the associated Lie algebra in the Euler–Poincaré formalism mentioned above.

Remark 1.2. Bloch, Marsden, and Zenkov [10] (working locally) identify each tangent fiber T_qQ with an individual Lie algebra V_q, and present a view of Hamel's formalism as the mechanics of a system on a Lie algebra bundle, formalizing the perspective that Hamel's equations are a generalization of the above formulation of the Euler–Poincaré equations. In this case, the Lagrangian function is viewed as a mapping from the Lie algebra bundle to the reals. From this perspective, we see that the Euler–Poincaré formalism is nothing but Hamel's formalism on a Lie group, where all fibers are the same Lie algebra.

This perspective is geometrically appealing, but not strictly necessary in the formulation of Hamel's equations. Indeed, we need not introduce any extra geometric structures in our description of Hamel's equations. In our development we will consider the components of velocity as either measured against a frame induced by coordinates, or the components of the same velocity vector measured against an alternate frame. More succinctly, each perspective is simply a different coordinate chart of TQ. Hamel's equations are the Euler–Lagrange equations, after a (linear) velocity substitution.

Hamel's formalism provides a natural framework for incorporating non-coordinate frames in the study of dynamics of complex mechanical systems. Study of Hamel's equations has been motivated by recent insight into their application in stabilization and control of systems with nonholonomic constraints (see, for example, recent publications by Bloch, Marsden, and Zenkov [9, 10]). See e.g. Neimark and Fufaev [40] and Bloch et al. [10] for the history and contemporary exposition of Hamel's formalism.

1.2 Variational Integrators: A Brief Introduction and History

Numerical integration, in one form or another, of the differential equations describing the evolution of mechanical systems has been an important topic of inquiry and a useful tool for mathematicians. An early, although most likely entirely incidental, example of the use of numerical integration occurs in the late 17th century. In the first book of his *Principia* [41], Newton provides a geometric proof of Kepler's second law (the well-known statement that the ray connecting two masses in the two-body problem sweeps out equal area in equal time), which from a modern-perspective is essentially geometric numerical integration of Newton's law for the two-body problem using the Störmer-Verlet scheme. In this early instance Newton was presenting a geometric justification of an analytic solution to a solveable differential equation to an audience largely ignorant of calculus: it should not be construed that Newton had geometric numerical integration in the modern sense on his mind. See Figure 1.1 for Newton's illustration, and the introductory section of Hairer, Lubich, and Wanner [21] for more details.

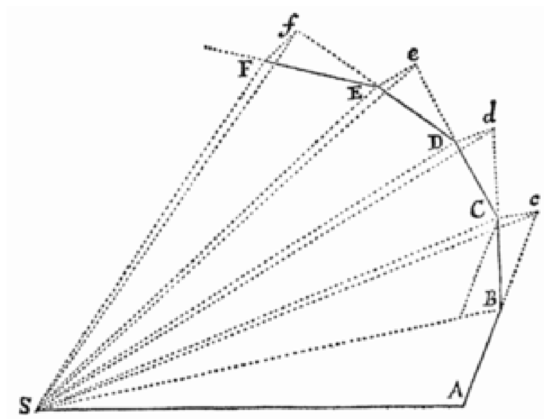


Figure 1.1: An illustration from Newton's *Principia* [41] describing his geometric analysis of Kepler's second law. Thanks to the Newton Project, `www.newtonproject.sussex.ac.uk`.

For any given system of ordinary differential equations, a variety of numerical integrators that approximate evolution of the system may be developed. While the integrators converge to exact solutions of the equations, for finite time steps the behavior of various integrators may be quite different. In this paper we examine a particular class of integrators of mechanical systems called *variational integrators*, so-named because rather than being derived via a direct discretization of the differential equations of motion, they are instead obtained by discretization of an underlying variational principle. This observation, that numerical integrators may be derived by discretization of an underlying variational principle, is the second important idea that gives rise to the results presented in this dissertation.

The general philosophy, illustrated in Figure 1.2, is that Hamilton's principle is discretized, and the resulting discrete variational principle will be a statement that discrete trajectories satisfying the variational integrator must be critical points of a discretized action. The resulting update maps are second-order difference equations, and, as we shall see, have structural similarities to the second-order ordinary differential equations they approximate.

A major advantage of variational integrators is that they are naturally structure-preserving,

$$
\begin{tikzcd}
\text{Hamilton's Principle} \arrow[r, "\text{Discretization}"] \arrow[dd, leftrightarrow] & \text{Discrete Hamilton's Principle} \arrow[d, leftrightarrow] \\
 & \boxed{\text{Variational Integrator}} \\
\text{Euler–Lagrange Equations} \arrow[r, "\text{Discretization}"] & \text{Other Numerical Integrators}
\end{tikzcd}
$$

Figure 1.2: Diagram of discretization approaches. It should be noted that particular discretizations of the Euler–Lagrange equations may inadvertently return an integrator that happens to be variational (an idea that we shall use later in this paper), but this is not guaranteed to be the case.

because they conserve a discrete symplectic two-form and display discrete momentum conservation in the presence of symmetry [36]. Conservation of a symplectic form is a feature of mechanics commonly presented in the context of the Hamiltonian formulation. In this case, a symplectic two-form over the phase space is conserved by solutions of Hamilton's equations. Marsden and West detail an analogous formulation in the context of Lagrangian mechanics, where Lagrangian flows conserve a symplectic two-form over the tangent bundle to the configuration space TQ [36]. In this paper we will favor the Lagrangian development, as Hamel's equations are viewed as a coordinate generalization of the Euler–Lagrange equations.

Thus, by focusing on variational integrators, we contribute to the process of developing structure-preserving integrators. In many circumstances, symplectic (and reversible) integrators display relatively good long-term behavior in simulation, often evidenced by energy evolution that is *bounded*, if not conserved [36].

The development of variational integrators that will be presented in the next chapter is based largely on the extensive review of the subject by Marsden and West [36] and in some aspects Hairer, Lubich, and Wanner [21]. Marsden and West themselves cite an extensive variety of sources in control literature dating from the 1960's as laying the groundwork for the understanding of variational integration presented in their paper.

1.3 Variational Integration of Hamel's Equations

Having presented the two observations in mechanics that are integral to our subsequent work, we now state the central questions addressed in this dissertation.

1. Can we develop, via a discrete variational principle, numerical integrators that take advantage of simplifications that arise in the continuous equations of motion when we select an appropriate, simplifying, non-coordinate frame?

2. Furthermore, are these discrete Hamel's equations variational integrators, and consequently are they structure-preserving integrators?

We will answer both of these questions in the affirmative by presenting geometrically-motivated discrete Hamel's equations that are variational integrators derived via a discretized action principle.

An outline of the remainder of this dissertation follows. Chapter 2 will present a more detailed background on Lagrangian mechanics, variational integrators, and Hamel's equations. In Chapter 3 we present the aforementioned geometrically-motivated derivation of the discrete Hamel's equations, documented in a preprint by Ball and Zenkov [5], and we will also describe an application to the derivation of integrators of nonholonomic systems. Chapter 4 documents results found in Ball, Zenkov, and Bloch [4] on a variational derivation of the Hamel's equations, and also an extension of the Hamilton–Pontryagin principle to the discretization of nonholonomic systems with symmetries. In Chapter 5 we demonstrate that the discrete Hamel's equations are, in fact, variational integrators, and we both examine and demonstrate the consequential geometric structure preservation. Finally, in Chapter 6 we summarize our results and suggest future avenues of work.

Chapter 2

Background

In this chapter we shall cover a variety of topics that serve as the foundation upon which we will study numerical integration of Hamel's equations. We will start with an overview of Lagrangian mechanics, variational calculus, and frames leading to a definition of Hamel's equations that illustrates their relationship to the Euler–Lagrange equations. We will present a more detailed account of how Euler's equations for the free rigid body are a special case of Hamel's equations before transitioning into a discussion of variational integrators following the work of Marsden and West [36].

2.1 Lagrangian Mechanics

A brief summary of Lagrangian mechanics follows, primarily intended to formally introduce our notation. This formalism originated in [29]. For a more thorough discussion of the topic we refer the reader to Arnold [3], Abraham and Marsden [1], and Marsden and Ratiu [35].

2.1.1 Lagrangian Mechanical Systems

A Lagrangian mechanical system is an abstract representation of a physical system the configuration of which may be described by a point in an n-dimensional differentiable manifold Q, called the ***configuration space***. The system is specified by a ***Lagrangian*** $L : TQ \to \mathbb{R}$, which for most physical systems is the difference of kinetic and potential energies.

The differentiable manifold Q is described by a covering atlas of compatible coordinate charts; each local coordinate chart mapping $q \mapsto (q^i)_{i=1}^n \in \mathbb{R}^n$ for all $q \in U$ where $U \subset Q$ is open. In the subsequent chapters we will work primarily in local coordinates; it will be understood that $(q^i)_{i=1}^n$ refers to coordinates of the point $q \in Q$ corresponding to an unspoken choice of coordinate chart.

A coordinate chart over an open set $U \subset Q$ induces a natural chart on the tangent bundle TU such that $(q, v) \mapsto (q^i, v^i)_{i=1}^n$. The induced charts on the fibers of the tangent bundle should be understood in terms of the coordinate-induced frame. For every $q \in U$, the chart induces n vector fields $\partial/\partial q^i$, $i = 1, \ldots, n$, such that at each point $q \in U$, $\text{span}\{\partial/\partial q^i\}_{i=1}^n = T_qQ$. In other words, the set $\{\partial/\partial q^i\}_{i=1}^n \subset T_qQ$ forms a basis of the vector space T_qQ, called the ***coordinate-induced basis***. A vector $\partial/\partial q^i \in T_qQ$ has a useful interpretation of a differential operator acting on smooth functions on Q and represented in local coordinates as a partial derivative.

For a vector $v \in T_qQ$, we denote the components of v measured against the basis $\partial/\partial q^i$ as $(v^i)_{i=1}^n$.

Likewise, a dual frame $(dq^i)_{i=1}^n$ is induced on the cotangent bundle T^*U; the n covectors at each point $q \in U$ span the fiber T_q^*Q of the cotangent bundle (note that $dq^i(\partial/\partial q^j) = \delta^i_j$). This in turn is associated with a chart on the cotangent bundle, $(q, p) \mapsto (q^i, p_i)_{i=1}^n$. In summary, for $(q, v) \in TQ$ and $(q, p) \in T^*Q$,

$$v = v^i \frac{\partial}{\partial q^i} \quad \text{and} \quad p = p_i dq^i, \tag{2.1}$$

where $i = 1, \ldots, n$ and where the pairs of repeated indices are understood to indicate summation as is typical with Einstein summation notation.

2.1.2 Path Space and the Action

Given an n-dimensional differentiable manifold Q and a time interval $[t_0, t_F] \subset \mathbb{R}$, we define the ***path space*** $\mathcal{C}(Q)$ to be the set

$$\mathcal{C}(Q) = \left\{ q : [t_0, t_F] \to Q \mid q \in C^2(Q) \right\}.$$

Furthermore suppose we select points $q_0, q_F \in Q$, then the set of paths in $\mathcal{C}(Q)$ between the points and parameterized on the interval $[t_0, t_F]$ will be referred to by the notation

$$\mathcal{C}(Q, q_0, q_F) = \{ q \in \mathcal{C}(Q) \mid q(t_0) = q_0, q(t_F) = q_F \}. \tag{2.2}$$

Given a Lagrangian mechanical system on Q specified with Lagrangian $L : TQ \to \mathbb{R}$, the ***action*** is defined as the functional $S : \mathcal{C}(Q) \to \mathbb{R}$, a path integral of the Lagrangian:

$$S(q) = \int_{t_0}^{t_F} L(q(t), \dot{q}(t)) dt. \tag{2.3}$$

2.1.3 Variations

Consider a path $q \in \mathcal{C}(Q, q_0, q_F)$. A ***variation*** of q is a one parameter family of paths $\beta : [t_0, t_F] \times [-\epsilon, \epsilon] \to Q$ in $\mathcal{C}(Q)$ such that $\beta(t, 0) = q(t)$. A ***virtual displacement*** at $q(t)$ is defined in terms of a variation to be the vector $\delta q(t) \in T_{q(t)}Q$ such that

$$\delta q(t) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \beta(t, \epsilon). \tag{2.4}$$

This definition of variations of curves on Q induces a definition of variations of curves on the tangent bundle TQ by differentiation of the parameterized path β. The t (time) derivative of q is denoted as $\dot{q} \in \mathcal{C}(TQ)$. We define the (induced) variation of $\dot{q}$ as a one parameter family of paths $\dot{\beta} : [t_0, t_F] \times [-\epsilon, \epsilon] \to TQ$ such that $\dot{\beta}(t, \epsilon) = \frac{\partial}{\partial t}\beta(t, \epsilon)$ and therefore $\dot{\beta}(t, 0) = (q(t), \dot{q}(t))$. A virtual displacement at $\dot{q}(t)$ is then a vector $(\delta q(t), \delta \dot{q}(t)) \in T_{(q(t),\dot{q}(t))}TQ$ such that $\delta q(t)$ is defined as in (2.4) and

$$\delta \dot{q}(t) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \frac{\partial}{\partial t}\beta(t, \epsilon).$$

By commutativity of mixed partial derivatives, we see that the we may either define virtual displacements $\delta \dot{q}(t)$ in terms of equivalence classes of $\dot{\beta}$'s, or equivalently take the time derivative of virtual displacements $\delta q(t) \in T_{q(t)}Q$ for $t \in [t_0, t_F]$. In other words, the time derivative and taking variations are commuting operators:

$$\begin{gathered} \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \frac{\partial}{\partial t}\beta(t, \epsilon) = \frac{\partial}{\partial t}\frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \beta(t, \epsilon) = \frac{d}{dt}\delta q(t), \\ \text{hence} \;\; \delta \dot{q}(t) = \frac{d}{dt}\delta q(t). \end{gathered} \tag{2.5}$$

2.1.4 Hamilton's Principle

Recall that in local coordinates on the configuration space Q, the dynamics of a mechanical system is given by the ***Euler–Lagrange equations***

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial q^i}, \quad i = 1, \ldots, n. \tag{2.6}$$

These equations were originally derived by Lagrange [29] in 1788 by requiring that simple force balance $F = ma$ be *covariant*, i.e. expressible in arbitrary generalized coordinates. A variational derivation of the Euler–Lagrange equations, namely Hamilton's principle, came later in the work of Hamilton [23, 24] in 1834/35.

Theorem 2.1 (**Hamilton's Principle**)**.** *The following statements are equivalent:*

1. *The path* $q(t)$*, where* $t \in [t_0, t_F]$*, is a critical point of the action functional* (2.3) *on* $\mathcal{C}(Q)$*,*

where we choose variations such that $\delta q(t_0) = \delta q(t_F) = 0$ (in other words, we restrict the one parameter families of paths β so that $\beta(t_0, \epsilon) = q_0$ and $\beta(t_F, \epsilon) = q_F$ for all ϵ).

2. *The path $q(t)$ satisfies the Euler–Lagrange equations* (2.6).

We refer the readers to Marsden and Ratiu [35] and Bloch [6] for details and proof.

2.1.5 Symplecticity of the Lagrangian Flow

In the subsequent section we follow the exposition of Marsden and West [36], and for the sake of consistency use some of their notations and definitions to illustrate the conservation of a symplectic two-form by the Lagrangian flow. Following Marsden and West [36], let $\ddot{Q}$ be the submanifold of TTQ defined as

$$\ddot{Q} = \{w \in TTQ \mid T\pi_Q(w) = \pi_{TQ}(w)\},$$

where $\pi_Q : TQ \to Q$ and $\pi_{TQ} : TTQ \to TQ$ are the canonical tangent bundle projections. In other words, elements of $\ddot{Q}$ are elements of the form $((q, \dot{q}), (\dot{q}, \ddot{q}))$. Define the *Euler–Lagrange map* $D_{EL}L : \ddot{Q} \to T^*Q$ by the formula

$$D_{EL}L = \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right) dq^i. \tag{2.7}$$

We define a second order vector field $X_L : TQ \to \ddot{Q}$ on TQ to be a *Lagrangian vector field* if it satisfies the condition that $X_L \circ D_{EL}L = 0$, and the flow of X_L is the *Lagrangian flow* $F_L : TQ \times \mathbb{R} \to TQ$. We will write the Lagrangian flow evaluated at the instantaneous time t as $F_L^t : TQ \to TQ$. Then if a path q satisfies the Euler–Lagrange equations (2.6), $(q, \dot{q})$ is an integral curve of X_L. Note that for an arbitrary Lagrangian, the field X_L may not be uniquely defined, under which circumstances the flow may not exist; this matter is resolved by requiring that L be *hyperregular*. The ***Legendre transform*** of the Lagrangian is defined as a mapping $\mathbb{F}L : TQ \to T^*Q$ and is expressed in coordinates as

$$\mathbb{F}L(q, \dot{q}) = \left(q, \frac{\partial L(q, \dot{q})}{\partial \dot{q}} \right) = (q, p)$$

A function L is ***regular*** if $\mathbb{F}L$ is a local isomorphism, and is ***hyperregular*** if $\mathbb{F}L$ is a global isomorphism. If L is regular then the Euler–Lagrange equations may be expressed in local coordinates such that $\ddot{q}$ is a well-defined function of $(q, \dot{q})$, and X_L will be uniquely defined locally and local flows exist for reasons discussed in Marsden and West [36].

We will define solution space $\mathcal{C}_L(Q) \subset \mathcal{C}(Q)$ to be the set of paths $q \in \mathcal{C}(Q)$ that are solutions of the Euler–Lagrange equations, or equivalently are integral curves of X_L. Such curves are uniquely determined by an initial condition (q_0, v_0), hence the solution space may be identified with the space of initial conditions TQ. We may therefore restrict the action S to integral curves of X_L, identified by TQ; define the restricted action $\hat{S} : TQ \to \mathbb{R}$ so that

$$\hat{S}(q_0, v_0) = S(q), \quad q \in \mathcal{C}_L(Q) \quad \text{and} \quad (q(t_0), \dot{q}(t_0)) = (q_0, v_0). \tag{2.8}$$

Given a C^k Lagrangian L (where $k \geq 2$), we will see that the Euler–Lagrange map and a unique C^{k-1} one-form $\Theta_L(q, \dot{q}) \in T^*_{(q,\dot{q})}(TQ)$ called the *Lagrangian one-form*, defined in coordinates as

$$\Theta_L(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}^i} dq^i,$$

are fundamentally related to the action (2.3). Suppose we arbitrarily select a variation of the curve $q \in \mathcal{C}(q)$, in other words we arbitrarily select smoothly varying virtual displacement $\delta q(t) \in T_q\mathcal{C}(Q)$. The ***variational derivative*** of the action $S(q)$ with respect to δq is defined as the change in S to the first order in δq. Because S is a real-valued function on an infinite dimensional space, we understand the derivative in a weak (Gâteaux) sense. Thus the variational derivative of S is expressible in coordinates as

$$\begin{aligned}
\delta S(q) &= \int_{t_0}^{t_F} \left(\frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \frac{d}{dt} \delta q^i \right) dt \\
&= \int_{t_0}^{t_F} \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right) \delta q^i dt + \left. \frac{\partial L}{\partial \dot{q}^i} \delta q^i \right|_{t_0}^{t_F},
\end{aligned}$$

by commutativity of mixed partial derivatives and and integration by parts. The term remaining in the integral is the Euler–Lagrange map (2.7), and we have integrated out the Lagrange one-form paired with the virtual displacement so that variation of the action may be written as

$$\delta S(q) = \int_{t_0}^{t_F} \langle D_{EL} L(\ddot{q}), \delta q \rangle \, dt + \left. \langle \Theta_L(\dot{q}), (\delta q, \delta \dot{q}) \rangle \right|_{t_0}^{t_F}.$$

Now, taking the variation of the restricted action $\hat{S}$ for arbitrary $(\delta q, \delta v) \in T_{(q,v)}(TQ)$ (of course satisfying $q \in \mathcal{C}_L(Q)$ and $(q(t_0), \dot{q}(t_0)) = (q, v)$) we see that because the path is required to be an integral curve of X_L, the Euler–Lagrange map term disappears and the variation of $\hat{S}$ is written in terms of Θ_L as

$$\begin{aligned}
d\hat{S}(q, v) \cdot (\delta q, \delta v) &= \Theta_L(q(t_F), \dot{q}(t_F)) \cdot (F_L^{t_F})_*(\delta q, \delta v) - \Theta_L(q, v) \cdot (\delta q, \delta v) \\
&= \Big((F_L^{t_F})^* \Theta_L(q, v) - \Theta_L(q, v) \Big) \cdot (\delta q, \delta v).
\end{aligned}$$

Then, because $d^2 \hat{S} = 0$ and by compatibility of the pullback with the exterior derivative, we see that

$$d\left((F_L^{t_F})^*\Theta_L - \Theta_L\right) = 0 \;\Rightarrow (F_L^{t_F})^* d\Theta_L - d\Theta_L = 0 \;\Rightarrow (F_L^{t_F})^*\Omega_L(q,\dot{q}) = \Omega_L(q,\dot{q}),$$

where $\Omega_L(q,\dot{q})$ is the *Lagrangian symplectic form*, given in coordinates as

$$\Omega_L(q,\dot{q}) = \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j} dq^i \wedge dq^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} d\dot{q}^i \wedge dq^j.$$

To summarize, the Lagrangian flow F_L preserves the Lagrangian symplectic form.

2.2 Hamel's Equations

A variety of problems in mechanics are best treated by the introduction of a *frame* for measuring the velocity components. In this section we shall endeavor to provide a concise and useful definition of a frame and the space that frames "live" in and introduce a generalized use of frames in mechanics that motivates our work. Finally we will illustrate the usefulness of Hamel's equations by examining its application in physical examples.

2.2.1 The Frame Bundle Associated with the Tangent Bundle

Suppose E is the total space of a vector bundle over base space M (a differentiable manifold), so that the projection $\pi : E \to M$ is a continuous surjection and the fiber $\pi^{-1}(p) = E_p$ over each $p \in M$ is an n-dimensional vector space. Then a ***frame*** at $p \in M$ is an isomorphism mapping $E_p \to \mathbb{R}^n$ [2]. Equivalently, a frame may be viewed as an ordered basis of the vector space E_p. The collection of all such frames related to all the fibers over M that constitute E defines the ***frame bundle*** FM, a principal fiber bundle associated to E. Formally,

$$FM = \{(p, f) \mid f \text{ is a frame at } p\}.$$

Because the tangent bundle TQ of an n-dimensional manifold Q is a vector bundle, Q has a special frame bundle associated with TQ that we denote as FQ.

A frame can be identified with a set of vectors $\{u_i(q)\}_{i=1}^n$ at point $q \in Q$ that forms a basis of the tangent fiber T_qQ. A smooth section of the frame bundle is therefore a frame field (one frame at each fiber) varying smoothly over the base space Q. Whether or not the frame bundle FQ is able to admit a global smooth section is closely related to the notion of whether or not the manifold is parallelizable; for the remainder of this section we shall assume that we are working on a local set in Q over which the tangent frame bundle is trivial.

A local coordinate chart induces a frame $\{\partial/\partial q^i\}_{i=1}^n$ that forms a basis of the tangent fiber

T_qQ, which we refer to as the ***coordinate induced frame***. Then, given a coordinate chart and a frame field $\{u_j\}_{j=1}^n$ (a section of FQ), there exists $\psi(q) \in GL(n)$ at each q such that, in coordinates,

$$u_j(q) = \psi_j^i(q) \frac{\partial}{\partial q^i}. \tag{2.9}$$

In this way we see that FQ is a principal $GL(n)$ bundle with the projection $\pi : FQ \to Q$; each fiber is isomorphic to $GL(n)$ so that an element of a fiber F_qQ may be identified with an invertible linear operator on the vector space T_qQ.

We shall hereon assume that a frame field $\{u_j\}_{j=1}^n$ is a C^2 section of the frame bundle FQ, i.e., the n^2 functions $\{\psi_j^i(q)\}_{i,j=1}^n$ are C^2 over the local coordinate neighborhood.

Just as we define FQ as the frame bundle associated to the vector bundle TQ, we likewise may define F^*Q to be the frame bundle associated with the cotangent bundle T^*Q, itself a vector bundle. We shall refer to frames in F^*Q as ***coframes***, paralleling the terminology of vectors and covectors. Thus, a coframe at q is a set of n covectors that form a basis of the cotangent space T_q^*Q. The ***coordinate induced coframe*** at q is the dual to the coordinate induced frame: it is denoted by the set of covectors $\{dq^i\}_{i=1}^n$ that are dual to the vectors comprising the coordinate induced frame.

The covectors $u^j(q)$ comprising a particular coframe may be defined in terms of the coordinate induced coframe as

$$u^j(q) = (\psi^{-1})_i^j(q) dq^i = \phi_i^j(q) dq^i. \tag{2.10}$$

For the remainder of this document we shall denote the inverse of ψ as ϕ. Notice that in this case the covectors $u^j(q)$ for $j = 1, \ldots, n$ are dual to the vectors $u_j(q)$.

2.2.2 Hamel's equations

Let $q \in \mathcal{C}(Q, q_0, q_F)$ be a path parameterized by time $t \in [t_0, t_F]$. Then given a coordinate chart on Q, $(q(t), \dot{q}(t)) \in TQ$ has natural coordinates expressed as

$$(q(t), \dot{q}(t)) \mapsto (q^i(t), \dot{q}^i(t)) \in \mathbb{R}^{2n},$$

where $\dot{q}^i(t)$ are the components of the velocity vector $\dot{q}(t)$ measured against the coordinate induced frame $\partial/\partial q^i$. Now suppose we have selected a C^2 section of the frame bundle $u_j(q)$ defined as in (2.9). Then the components of $\dot{q}(t)$ are likewise expressible in terms of the new frame according to the equation

$$\dot{q} = \dot{q}^i \frac{\partial}{\partial q^i} = \xi^j \psi_j^i(q) \frac{\partial}{\partial q^i} = \xi^j u_j(q)$$

so that

$$\dot{q}^i = \xi^j \psi^i_j(q).$$

We next rewrite the Lagrangian in terms of the new velocity components, so that in coordinates

$$L(q^i, \dot{q}^i) = L(q^i, \xi^j \psi^i_j(q)) = \ell(q^i, \xi^j). \tag{2.11}$$

Define the ***structure functions*** $c^m_{ij}(q)$ by the equations

$$c^m_{ij}(q) u_m(q) = [u_i(q), u_j(q)], \tag{2.12}$$

where $[\cdot, \cdot]$ is the Jacobi–Lie bracket of vector fields on Q. The structure functions at q are also expressible in terms of $\psi(q) \in GL(n)$ and its inverse $\phi(q)$ as

$$c^m_{ij}(q) = \psi^a_i(q)\psi^b_j(q)(\phi^m_{a,b}(q) - \phi^m_{b,a}(q)). \tag{2.13}$$

Given two vectors $v, w \in T_qQ$, we define the anti-symmetric bracket $[\cdot, \cdot]_q : T_qQ \times T_qQ \to T_qQ$ so that

$$[v, w]^m_q u_m(q) = [v^i u_i(q), w^j u_j(q)].$$

Then we see that the vector space T_qQ is isomorphic to an n-dimensional Lie algebra V_q equipped with the bracket operation $[\cdot, \cdot]_q$, and the tangent bundle is (locally) diffeomorphic to a Lie algebra bundle. Furthermore, the bracket operation induces a dual bracket $[\cdot, \cdot]_q : T_qQ \times T^*_qQ \to T^*_qQ$ by

$$\left\langle [v, p]^*_q, w \right\rangle = \langle p, [v, w]_q \rangle$$

for $v, w \in T_qQ$ and $p \in T^*_qQ$, and where $\langle \cdot, \cdot \rangle$ denotes the natural pairing between T_qQ and T^*_qQ [10].

We denote the directional derivatives of the function ℓ at (q, ξ) along the vectors $u(q)$ by the notation $u[\ell](q, \xi)$, so that in index notation

$$u_j[\ell](q, \xi) = \frac{\partial \ell(q, \xi)}{q^i} \psi^i_j(q). \tag{2.14}$$

The equations of motion expressed in this new coordinate chart on the velocity phase space TQ are known as ***Hamel's equations***, originally appearing in the work of Georg Hamel [22]. Hamel's equations in index notation read

$$\frac{d}{dt}\frac{\partial \ell}{\partial \xi^j} = \frac{\partial \ell}{\partial q^i}\psi^i_j(q) + \frac{\partial \ell}{\partial \xi^m}\xi^i c^m_{ij}(q). \tag{2.15}$$

Theorem 2.2 (Zenkov, Bloch, and Marsden [10])**.** *Let $L : TQ \to \mathbb{R}$ be a Lagrangian and ℓ be* *its representation in local coordinates* (q, ξ). *Then, the following statements are equivalent:*

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

$$\int_a^b L(q, \dot{q})\, dt \tag{2.16}$$

   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* $\delta q(a) = \delta q(b) = 0$.

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

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

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

$$\int_a^b \ell(q, \xi)\, dt \tag{2.17}$$

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

$$\delta q = u(q) \cdot \zeta \equiv u_i(q)\zeta^i, \tag{2.18}$$

   and given by

$$\delta\xi = \dot{\zeta} + [\xi, \zeta]_q.^1 \tag{2.19}$$

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

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

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

For the early development of these equations see [43] and [22]. More details, historical background, and proof can be found in Zenkov, Bloch, and Marsden [10]. We shall present a derivation of Hamel's equations from an alternate variational principle in Chapter 4.

2.2.3 The Free Rigid Body and Hamel's Equations

A fundamental result and a precursor of Hamel's formalism is found in Euler's equations describing the evolution of a freely rotating rigid body. This rich problem is found in a variety

[1]If Q is a Lie group, this formula is derived in Bloch, Krishnaprasad, Marsden, and Ratiu [11].

of texts on mathematics and mechanics, including Arnold [3], wherein the equations are often derived from and motivated by conservation laws and symmetries. We present a brief review of the problem not designed for its historicity, but rather to illustrate the usefulness of Hamel's equations in a well-understood setting.

A rigid body is a system of point masses, subject to holonomic constraints that state that the distance between each pair of points of the body is constant [3]. A free rigid body is neither subject to external forces, nor is it subject to external constraints. In the case of the free rigid body rotating about a fixed point, the configuration space is the Lie group $SO(3)$ as long as at least three points on the body are not in a straight line. Without loss of generality, we further consider the rigid body as rotating around its center of mass.

The angles $(q^i)_{i=1}^3 = (\theta, \phi, \psi)$, Euler angles (see Figure 2.1), locally parameterize the rotation group and thus are a local coordinate chart of the configuration manifold. The Lagrangian of our system is then defined as

$$L(q^i, \dot{q}^i) = \frac{1}{2} M_{ij}(q) \dot{q}^i \dot{q}^j \tag{2.20}$$

where $M_{ij}(q)$ are the components of the inertia tensor that are functions of the configuration coordinates, symmetric in its indices but not diagonal. Note the q dependence of M; indeed if we fix our coordinate system in space (as we have) and allow our body to rotate about its center of mass the distribution of mass of the body will change (relevant to our fixed coordinate system) and its moments of inertia must be updated to reflect the instantaneous distribution.

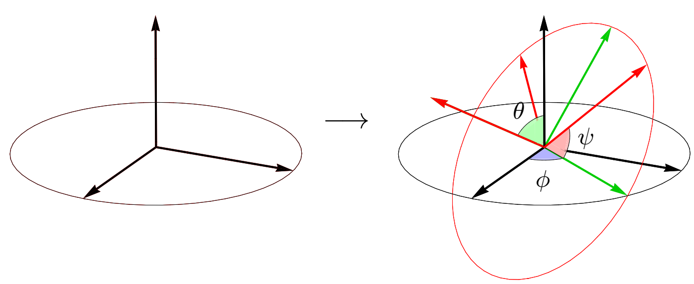


Figure 2.1: A depiction of the Euler angles, used as configuration coordinates for the free rigid body.

The Euler–Lagrange equations describing the evolution of the system are then

$$M_{ij,a}(q) \dot{q}^a \dot{q}^i + M_{ij}(q) \ddot{q}^i = \frac{1}{2} M_{ia,j}(q) \dot{q}^i \dot{q}^a,$$

which it first glance don't seem complicated, but recall we have refrained from explicitly writing out the functions comprising the inertia tensor $M_{ij}(q)$. In fact, expanded out in terms of the Euler angles, the Euler–Lagrange equations of the rigid body are

$$\begin{aligned}
&\big(J_1\cos^2\psi + J_2\sin^2\psi\big)\,\ddot{\theta} + (J_1 - J_2)\sin\theta\sin\psi\cos\psi\,\ddot{\phi} + (J_2 - J_1)\sin 2\psi\,\dot{\theta}\dot{\psi} \\
&-\big(J_1\sin^2\psi + J_2\cos^2\psi - J_3\big)\sin 2\theta\,\dot{\phi}^2 + (J_1\cos 2\psi - J_2\cos 2\psi + J_3)\sin\theta\,\dot{\phi}\dot{\psi} = 0, \\
&\big(J_1 - J_2\big)\sin\theta\sin\psi\cos\psi\,\ddot{\theta} + \big(J_1\sin^2\theta\sin^2\psi + J_2\sin^2\theta\cos^2\psi + J_3\cos^2\theta\big)\,\ddot{\phi} \\
&+J_3\cos\theta\,\ddot{\psi} + \big(J_1 - J_2\big)\cos\theta\sin 2\psi\,\dot{\theta}^2 + \big(J_1\sin^2\psi + J_2\cos^2\psi - J_3\big)\sin 2\theta\,\dot{\theta}\dot{\phi} \\
&\qquad +\big(J_1\cos 2\psi - J_2\cos 2\psi - J_3\big)\sin\theta\,\dot{\theta}\dot{\psi} + \big(J_1 - J_2\big)\sin^2\theta\sin 2\psi\,\dot{\phi}\dot{\psi} = 0, \\
&\qquad\qquad J_3\left(\cos\theta\,\ddot{\phi} + \ddot{\psi} - \sin\theta\,\dot{\theta}\dot{\phi}\right) = 0,
\end{aligned}$$

where J_1, J_2, J_3 are the moments of inertia about the principle axes.

This system of three second-order equations may be greatly simplified by the introduction of a moving frame; in this case the principle axes of inertia. Because $M(q)$ is a real symmetric matrix, it is always possible to find an orthogonal matrix $\phi(q)$ that diagonalizes M as follows

$$[M(q)] = [\phi(q)]^T \begin{bmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{bmatrix} [\phi(q)]$$

where J_i is the moment of inertia about the i-th principal axis, and we shall refer to the components of the above diagonal matrix as J_{ij}. Define (ξ^1, ξ^2, ξ^3) to be the components of $\dot{q} \in T_qSO(3)$ measured against the principal axes of inertia, our frame. Then we may rewrite our Lagrangian as

$$\begin{aligned}
L(q^i, \dot{q}^i) &= \frac{1}{2} J_{ij}\phi^i_a(q)\dot{q}^a\phi^j_b(q)\dot{q}^b \\
&= \frac{1}{2} J_{ij}\xi^i\xi^j = \ell(\xi^i).
\end{aligned}$$

Note that $\ell(q, \xi)$ is written as $\ell(\xi)$, reflecting the rotational invariance of the free rigid body.

Referring to (2.15), Hamel's equations written in the body frame describe the dynamics of the rigid body and are expressed by the system of equations

$$\begin{aligned}
\frac{d}{dt}\frac{\partial \ell}{\partial \xi^j} = J_{ij}\dot{\xi}^i = J_{am}\xi^a\xi^i c^m_{ij} &= \frac{\partial \ell}{\partial q^i}\psi^i_j(q) + \frac{\partial \ell}{\partial \xi^m}\xi^i c^m_{ij}(q) \Rightarrow \\
J_1\dot{\xi}^1 &= (J_2 - J_3)\xi^2\xi^3, \\
J_2\dot{\xi}^2 &= (J_3 - J_1)\xi^3\xi^1, \\
J_3\dot{\xi}^3 &= (J_1 - J_2)\xi^1\xi^2,
\end{aligned}$$

which are Euler's equations for a freely rotating rigid body. The full dynamics of the rigid body are described by these equations coupled with the kinematic reconstruction equations $\dot{q} = u(q) \cdot \xi$. The structure functions c_{ij}^m may be verified by direct computation.

Thus we see that Euler's equations of the free rigid body are a special case of Hamel's equations where the frame selected is the body frame. The advantage of the introduction of the body frame is apparent in the comparative simplicity of Hamel's equations to the Euler–Lagrange equations.

2.3 Variational Integrators

2.3.1 Discrete Euler–Lagrange Equations

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 [36], and references therein, for a more detailed discussion of discrete mechanics.

Our approach to discretization is motivated by the observation that the product of the configuration manifold with itself, $Q \times Q$, is *locally* isomorphic to TQ and hence local open subsets of $Q \times Q$ can be understood to contain the same information, locally, as TQ (again, see Marsden and West [36]).

The use of TQ instead of $Q \times Q$ for constructing variational integrators has been studied earlier in Bou-Rabee and Marsden [14] and in Kobilarov, Marsden, and Sukhatme [27]. Our approach to some extent develops these ideas even further, but we should emphasize that our version of the discrete Lagrange–d'Alembert principle is different from that of the aforementioned authors.

Assuming that we will use N integration steps in our approximation, we introduce a mesh on the time interval $[t_0, t_F]$ by defining $\{t_k = kh + t_0 \mid k = 0, \dots, N\}$, such that $t_N = t_F$. Then we may define the the discrete path space to be the space of discrete trajectories:

$$\mathcal{C}_d(Q) = \{q_d : \{t_k\}_{k=0}^N \to Q\}$$

and the discrete trajectories may be identified by $q_k = q_d(t_k)$ so that $q_d = \{q_k\}_{k=0}^N$.

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

$$L_d(q_k, q_{k+1}) \approx \underset{q \in \mathcal{C}(Q, q_k, q_{k+1})}{\text{ext}} \int_{t_k}^{t_{k+1}} L(q, \dot{q})\, dt, \tag{2.21}$$

where $\mathcal{C}(Q, q_k, q_{k+1})$ is defined as in (2.2), and ext denotes extremum.

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

$$S_d(q_d) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}),$$

where $q_d = \{q_k\}_{k=0}^{N} \subset Q$ is a finite sequence in the configuration space. Equations implicitly defining an update map are obtained by the ***discrete Hamilton's principle***, which extremizes the discrete action given fixed endpoints q_0 and q_N.

Variations in the discrete setting are defined similarly to the continuous case: a variation of a discrete trajectory q_d is a one-parameter family of trajectories $\beta_d : \{t_k\}_{k=0}^{N} \times [-\epsilon, \epsilon] \to Q$ such that $\beta_d(t_k, 0) = q_k$. Then virtual displacements in the discrete setting are defined as

$$\delta q_k = \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} \beta_d(t_k, \epsilon), \tag{2.22}$$

and we refer to a set of discrete displacements as $\delta q_d = \{(q_k, \delta q_k)\}_{k=0}^{N} \subset TQ$. In this manner the discrete Hamilton's principle may be written as

$$\delta S_d(q_d) = dS_d(q_d) \cdot \delta q_d = 0 \quad \text{for} \quad \delta q_0 = \delta q_N = 0.$$

Taking the extremum over $q_1, \ldots, q_{N-1}$ gives the ***discrete Euler–Lagrange equations***

$$D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) = 0, \tag{2.23}$$

for $k = 1, \ldots, N-1$. This implicitly defines the ***update map*** $F_{L_d} : Q \times Q \to Q \times Q$, where $F_{L_d}(q_{k-1}, q_k) = (q_k, q_{k+1})$ and $Q \times Q$ replaces the velocity phase space TQ of continuous-time Lagrangian mechanics.

The equations (2.23) motivate a definition of ***discrete fiber derivatives*** mappings $\mathbb{F}^+ L_d$, $\mathbb{F}^- L_d : Q \times Q \to T^*Q$ defined as

$$\begin{aligned} \mathbb{F}^+ L_d(q_k, q_{k+1}) \cdot v &= D_2 L_d(q_k, q_{k+1}) \cdot v, \\ \mathbb{F}^- L_d(q_k, q_{k+1}) \cdot v &= -D_1 L_d(q_k, q_{k+1}) \cdot v. \end{aligned} \tag{2.24}$$

The discrete fiber derivatives can also be understood as ***discrete Legendre transforms*** written as

$$\begin{aligned} \mathbb{F}^+ &: (q_k, q_{k+1}) \mapsto (q_{k+1}, p_{k+1}) = (q_{k+1}, D_2 L_d(q_k, q_{k+1})), \\ \mathbb{F}^- &: (q_k, q_{k+1}) \mapsto (q_k, p_k) = (q_k, -D_1 L_d(q_k, q_{k+1})). \end{aligned} \tag{2.25}$$

When both discrete fiber derivatives are local isomorphisms, the discrete Lagrangian is said to be ***regular***, and when these same fiber derivatives are *global* isomorphisms the discrete Lagrangian is said to be ***hyperregular*** (see Marsden and West [36]). In our later derivations we will assume that Q is a vector space and that L_d is *hyperregular* so that the update map is well-defined.

We have seen that $Q\times Q$ is a discrete analogue of the velocity phase space; in the same vein we replace submanifold $\ddot{Q}$ of TTQ with its discrete analogue, a submanifold $\ddot{Q}_d$ of $(Q\times Q)\times(Q\times Q)$, defined as

$$\ddot{Q}_d = \{((q_k, q_{k+1}), (\bar{q}_k, \bar{q}_{k+1})) \mid q_{k+1} = \bar{q}_k\} \subset (Q \times Q) \times (Q \times Q).$$

In other words, $\ddot{Q}_d$ is simply the subset of points in $(Q \times Q) \times (Q \times Q)$ that can be written as $((q_{k-1}, q_k), (q_k, q_{k+1}))$. Then, as verified in Marsden and West [36], given a discrete Lagrangian L_d that is C^k, there exists a unique C^{k-1} ***discrete Euler-Lagrange map*** $D_{DEL}L_d : \ddot{Q}_d \to T^*Q$ and two unique C^{k-1} one-forms $\Theta^-_{L_d}$ and $\Theta^+_{L_d}$ on the discrete velocity phase space $Q \times Q$ so that variation of the discrete action is expressible as

$$\begin{aligned} \delta S_d(q_d) = \sum_{k=1}^{N-1} & D_{DEL}L_d((q_{k-1}, q_k), (q_k, q_{k+1}))\delta q_k \\ & -\Theta^-_{L_d}(q_0, q_1)\delta q_0 + \Theta^+_{L_d}(q_{N-1}, q_N)\delta q_N, \end{aligned} \tag{2.26}$$

where the ***discrete Lagrangian one-forms*** in the coordinate representation are

$$\begin{aligned} \Theta^-_{L_d}(q_k, q_{k+1}) &= -D_1 L_d(q_k, q_{k+1}) dq_k, \\ \Theta^+_{L_d}(q_k, q_{k+1}) &= D_2 L_d(q_k, q_{k+1}) dq_{k+1}, \end{aligned}$$

and $D_{DEL}L_d((q_{k-1}, q_k), (q_k, q_{k+1})) = \Theta^+_{L_d}(q_{k-1}, q_k) - \Theta^-_{L_d}(q_k, q_{k+1})$. Thus, the discrete Euler–Lagrange equations (2.23) are equivalent to the condition $D_{DEL}L_d((q_{k-1}, q_k), (q_k, q_{k+1})) = 0$.

2.3.2 Update Maps are Symplectic

The update maps implicitly defined by (2.23) can be shown to conserve a discrete symplectic two-form over $Q \times Q$. In a manner similar to (2.8), we will define the restricted discrete action $\hat{S}_d$ as the discrete action restricted to the solution space of discrete Euler–Lagrange equations (2.23). A sequence $\{q_k\}_{k=0}^N \subset Q$ is formed by iteration of the update map F_{L_d}, so assuming the update map is well-defined, such a sequence may be uniquely identified by initial conditions $(q_0, q_1) \in Q \times Q$. The restricted action is then the map $\hat{S}_d : Q \times Q \to \mathbb{R}$ such that

$$\hat{S}_d(q_0, q_1) = S_d(\{q_k\}_{k=0}^N), \quad (q_k, q_{k+1}) = F_{L_d}(q_{k-1}, q_k), \quad \forall k = 1, \ldots, N-1.$$

Then variation of the restricted discrete action, that is restricting (2.26) to the solution space, results in

$$\begin{aligned} d\hat{S}_d(q_0, q_1) \cdot (\delta q_0, \delta q_1) = & - \Theta^-_{L_d}(q_0, q_1) \cdot (\delta q_0, \delta q_1) \\ & + \Theta^+_{L_d}(F^{N-1}_{L_d}(q_0, q_1)) \cdot (F^{N-1}_{L_d})_{*(q_0,q_1)}(\delta q_0, \delta q_1) \\ = & \left((F^{N-1}_{L_d})^* \Theta^+_{L_d} - \Theta^-_{L_d} \right) ((q_0, q_q), (\delta q_0, \delta q_1)) \end{aligned}$$

because $D_{DEL}L_d$ evaluated over the solution space is necessarily zero. In fact, evaluation of the discrete Lagrangian one-forms restricted to the solution space results in the condition $\Theta^+_{L_d}(q_{k-1}, q_k) = \Theta^-_{L_d}(q_k, q_{k+1})$, and therefore $d\Theta^+_{L_d}(q_{k-1}, q_k) = d\Theta^-_{L_d}(q_k, q_{k+1})$. Finally, because $d^2\hat{S} = 0$ and by compatibility of the pullback with the exterior derivative, we see that the above results imply conservation of a discrete two-form:

$$(F^{N-1}_{L_d})^*(\Omega_{L_d}) = \Omega_{L_d},$$

where $\Omega_{L_d} = d\Theta^+_{L_d} = d\Theta^-_{L_d}$ is defined to be the ***discrete Lagrangian symplectic form*** on $Q \times Q$. The above process also holds for any single step, and thus a single step conserves the discrete Lagrangian symplectic form: $(F_{L_d})^*\Omega_{L_d} = \Omega_{L_d}$. The update map $F_{L_d} : Q \times Q \to Q \times Q$ is therefore defined as a ***discrete symplectic map*** (see Marsden and West [36]). Ω_{L_d} is expressible in coordinates as

$$\Omega_{L_d}(q_k, q_{k+1}) = \frac{\partial^2 L_d(q_k, q_{k+1})}{\partial q^i_k \partial q^j_{k+1}} dq^i_k \wedge dq^j_{k+1}.$$

We have shown that the discrete Lagrangian flow F_{L_d} is a discrete symplectic map, conserving the discrete symplectic form Ω_{L_d}.

2.3.3 Discrete Noether's Theorem

In the presence of symmetry, variational integrators are symplectic-momentum integrators, meaning that, in addition to the discrete symplectic two-form, they conserve a discrete analogue of momentum associated with symmetry. We will describe the momentum conservation property by presenting a discrete formulation of Noether's theorem, following the exposition in Marsden and West [36].

Suppose a Lie group G with associated Lie algebra $\mathfrak{g}$ acts on Q via the left (or right) action $\Phi : G \times Q \to Q$. The lift of the action to the tangent bundle is a mapping $\Phi^{TQ} : G \times TQ \to TQ$

defined so that $\Phi^{TQ}(g,(q,v)) = \Phi_{*g}(q,v)$, or in index notation:

$$\Phi^{TQ}(g,(q,v)) = \left(\Phi^i(g,q), \frac{\partial \Phi^i}{\partial q^j}(g,q)v^j\right).$$

Furthermore, given a vector $\xi \in \mathfrak{g}$ we define the ***infinitesimal generator*** $\xi_Q : Q \to TQ$ in the following sense. Consider the usual exponential mapping $\exp : \mathfrak{g} \to G$. Then the vector ξ induces a one parameter subgroup of G as $t \mapsto \exp(t\xi) \in G$. The infinitesimal generator can be thought of in this case as

$$\xi_Q(q) = \frac{d}{dt}\Big|_{t=0}(\exp(t\xi)\cdot q),$$

in other words as the action of the subgroup of G on q. Marsden and West [36] define the infinitesimal generator in terms of the group action as

$$\xi_Q(q) = \frac{d}{dg}\left(\Phi(g,q)\right)\cdot \xi,$$

which can be expressed in index notation as:

$$\xi_Q(q) = \left(q^i, \frac{\partial \Phi^i}{\partial g^j}(e,q)\xi^j\right)$$

where $e \in G$ is the group identity. The span of the all such generators at q forms a vector subspace of T_qQ that we identify as a ***symmetry space*** of the mechanical system when L is ***invariant*** under the tangent lift of the group action: i.e. when $L \circ \Phi_g^{TQ} = L$.

The continuous-time Noether's theorem states that when a Lagrangian system with Lagrangian $L : TQ \to \mathbb{R}$ is *invariant* under the tangent lift of the action $\Phi : G \times Q \to Q$, the ***Lagrangian momentum map*** $J_L : TQ \to \mathfrak{g}^*$ defined by

$$J_L(q,\dot{q})\cdot\xi = \left\langle \frac{\partial L}{\partial \dot{q}}(q,\dot{q}), \xi_Q(q)\right\rangle,$$

is a conserved quantity of the Lagrangian flow F_L, i.e. for all $t \in [t_0, t_F]$, $J_L \circ F_L = J_L$. More details and proof can be found in Marsden and Ratiu [35] and Marsden and West [36].

The action Φ can likewise be extended to $Q \times Q$ by defining $\Phi^{Q\times Q} : G \times (Q \times Q) \to Q \times Q$ so that

$$\Phi^{Q\times Q}(g,(q_k,q_{k+1})) = \left(\Phi(g,q_k), \Phi(g,q_{k+1})\right). \tag{2.27}$$

We also extend our definition of an infinitesimal generator so that for $\xi \in \mathfrak{g}$, $\xi_{Q\times Q} : Q \times Q \to$

$T(Q \times Q)$ is a mapping given by

$$\xi_{Q\times Q}(q_k, q_{k+1}) = (\xi_Q(q_k), \xi_Q(q_{k+1})) = \left((q_k^i, q_{k+1}^i), \left(\frac{\partial \Phi^i}{\partial g^j}(e, q_k)\xi^j, \frac{\partial \Phi^i}{\partial g^j}(e, q_{k+1})\xi^j \right) \right). \quad (2.28)$$

We may then define *two* ***discrete Lagrangian momentum maps*** $J^{\pm}_{L_d} : Q \times Q \to \mathfrak{g}^*$ so that for some $\xi \in \mathfrak{g}$,

$$\begin{aligned} J^{-}_{L_d}(q_k, q_{k+1}) \cdot \xi &= \langle -D_1 L_d(q_k, q_{k+1}), \xi_Q(q_k) \rangle = \Theta^{-}_{L_d} \cdot \xi_{Q\times Q}(q_k, q_{k+1}), \\ J^{+}_{L_d}(q_k, q_{k+1}) \cdot \xi &= \langle D_2 L_d(q_k, q_{k+1}), \xi_Q(q_{k+1}) \rangle = \Theta^{+}_{L_d} \cdot \xi_{Q\times Q}(q_k, q_{k+1}). \end{aligned} \quad (2.29)$$

For reference, the discrete momentum maps may be written in index notation as

$$\begin{aligned} J^{-}_{L_d}(q_k, q_{k+1}) \cdot \xi &= (\Theta^{-}_{L_d})_i \frac{\partial \Phi^i}{\partial g^j}(e, q_k)\xi^j, \\ J^{+}_{L_d}(q_k, q_{k+1}) \cdot \xi &= (\Theta^{+}_{L_d})_i \frac{\partial \Phi^i}{\partial g^j}(e, q_{k+1})\xi^j. \end{aligned}$$

Now that we have formally defined our discrete Lagrangian momentum maps in terms of a Lie group action on $Q \times Q$, we are almost ready to formulate and prove the discrete version of Noether's theorem. However, first recall that we defined a discrete symplectic map, such as $f : Q\times Q \to Q\times Q$, by the condition that it conserved a discrete symplectic form: $f^*\Omega_{L_d} = \Omega_{L_d}$. We will define a ***special discrete symplectic map*** by a more restrictive condition: that such a map $f : Q \times Q \to Q \times Q$ satisfy the conditions $f^*\Theta^{+}_{L_d} = \Theta^{+}_{L_d}$ and $f^*\Theta^{-}_{L_d} = \Theta^{-}_{L_d}$.

Theorem 2.3 (Discrete Noether's Theorem)**.** *Suppose $L_d : Q \times Q \to \mathbb{R}$ is the discrete Lagrangian of a discrete Lagrangian mechanical system on $Q \times Q$, and furthermore suppose that L_d is invariant with respect to the action Φ of a Lie group G lifted to $Q\times Q$ as in* (2.27)*. Then:*

1. *$\Phi_g^{Q\times Q} : Q\times Q \to Q\times Q$ is a* special *discrete symplectic map. Also, the discrete Lagrangian momentum maps will be equivalent so that $J^{+}_{L_d} = J^{-}_{L_d} = J_{L_d}$ and can thus be referred to as one map $J_{L_d} : Q \times Q \to \mathfrak{g}^*$.*

2. *The discrete Lagrangian map F_{L_d} conserves the discrete Lagrangian momentum map J_{L_d}, i.e. $J_{L_d} \circ F_{L_d} = J_{L_d}$.*

Proof. A more detailed proof can be find in Marsden and West [36], who furthermore rely on results in Marsden and Ratiu [35]. Nonetheless, we provide a basic outline of a proof here, drawing from these sources but editing for clarity of exposition and relevance.

In the first statement, invariance of L_d with respect to Φ can be expressed by the equation

$L_d \circ \Phi_g^{Q\times Q} = L_d$. Differentiating with respect to q_k we see that

$$D_1 L_d\left(\Phi^{Q\times Q}(g,(q_k,q_{k+1}))\right)\cdot\frac{\partial\Phi(g,q_k)}{\partial q_k}dq^k = D_1L_d(q_k,q_{k+1})dq^k \quad\Rightarrow$$
$$\left(\Phi_g^{Q\times Q}\right)^*(-\Theta_{L_d}^-) = -\Theta_{L_d}^-.$$

Similarly we can show $\left(\Phi_g^{Q\times Q}\right)^*(\Theta_{L_d}^+) = -\Theta_{L_d}^+$. Thus $\Phi_g^{Q\times Q}$ is a special discrete symplectic map. Furthermore, invariance of L_d implies infinitesimal invariance, which is expressible by the condition $dL_d\cdot\xi_{Q\times Q}=0$ for any $\xi\in\mathfrak{g}$, so that

$$\begin{aligned} dL_d\cdot\xi_{Q\times Q}(q_k,q_{k+1}) &= D_1L_d(q_k,q_{k+1})\cdot\xi_Q(q_k) + D_2L_d(q_k,q_{k+1})\cdot\xi_Q(q_{k+1})\\ &= -\Theta_{L_d}^-(q_k,q_{k+1})\cdot\xi_Q(q_k) + \Theta_{L_d}^+(q_k,q_{k+1})\cdot\xi_Q(q_{k+1}) = 0.\end{aligned}$$

Next by applying the definition of the discrete Lagrangian momentum maps (2.29), the above implies that

$$J_{L_d}^-(q_k,q_{k+1})\cdot\xi = J_{L_d}^+(q_k,q_{k+1})\cdot\xi.$$

Hence, the discrete Lagrangian momentum maps are equivalent and we may refer to them as one map J_{L_d}.

To verify that F_{L_d} conserves J_{L_d}, we first note that Φ naturally induces an action of G on the space of trajectories $\mathcal{C}_d(Q)$ (the trajectories simply being a set of $N+1$ points in Q). Then, evaluating the discrete action with the infinitesimal generator, one obtains

$$dS_d(q_d)\cdot\xi_{\mathcal{C}_d(Q)}(q_d) = \sum_{k=0}^{N-1} dL_d(q_k,q_{k+1})\cdot\xi_{Q\times Q}(q_k,q_{k+1}) = 0,$$

because, as we have just seen, invariance of L_d implies infinitesimal invariance.

Next, if we restrict trajectories to solutions of the discrete Euler–Lagrange equations, the definition of the discrete restricted action in terms of initial conditions (q_0,q_1) implies that the above can also be written as

$$\begin{aligned} dS_d(q_d)\cdot\xi_{\mathcal{C}_d(Q)}(q_d) &= d\hat{S}_d(q_0,q_1)\cdot\xi_{Q\times Q}(q_0,q_1)\\ &= \left((F_{L_d}^N)^*\Theta_{L_d}^+ - \Theta_{L_d}^-\right)\cdot\xi_{Q\times Q}(q_0,q_1)\\ &= \left((F_{L_d}^N)^*J_{L_d}^+(q_0,q_1) - J_{L_d}^-(q_0,q_1)\right)\cdot\xi = 0.\end{aligned}$$

This holds for arbitrary $\xi\in\mathfrak{g}$ and, as before, the above argument is applicable to an arbitrary number of time-steps, including a single time-step. Thus, since $J_{L_d}^+ = J_{L_d}^- = J_{L_d}$, we see that $(F_{L_d})^*J_{L_d} = J_{L_d}\circ F_{L_d} = J_{L_d}$ and we have demonstrated that when the discrete Lagrangian is invariant with respect to the action Φ of group G on Q, solutions of the discrete Euler–Lagrange equations conserve discrete momentum maps.

$\square$

Chapter 3

Discrete Hamel's Equations

In the following chapter we propose and derive via a discrete variational principle a numerical integrator approximating the evolution of a Lagrangian mechanical system that we call the discrete Hamel's equations. The form of the integrators and their equivalent variational principle will first be motivated by a second examination at the of the discrete Euler–Lagrange equations within which we shall present the discrete Hamilton's principle in an alternate but equivalent form. This alternate viewpoint will motivate a discrete variational principle that is *not* equivalent to the discrete Hamilton's principle, but that gives rise to the discrete Hamel's equations.

The material presented in this chapter may be found in a preprint by Ball and Zenkov [5], submitted for publication in 2013.

As we have already seen, discrete Lagrangian mechanics is obtained by discretizing Hamilton's principle. This approach leads to symplectic- and, for systems with symmetry, momentum-preserving integrators. By discretizing the Lagrange–d'Alembert principle, nonconservative forces (see Marsden and West [36]) and nonholonomic constraints (see Cortés and Martínez [17]) can be incorporated as well. As pointed out in Cortés and Martínez [17], the versions of the discrete Lagrange–d'Alembert principle used in [36] and [17] are incompatible in the following sense: In the nonholonomic setting, discretizing constraints as opposed to discretizing their reactions generically results in different discrete models. In other words, the notion of an ***ideal constraint*** of continuous-time mechanics is not preserved by the discretization of Cortés and Martínez. Ideal constraints can be replaced by reaction forces, in other words work done by the constraint forces disappears under virtual displacements.

We develop discrete Hamel's formalism by discretizing Hamilton's principle for Hamel's equations. The principal difficulty in extending this program to the 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 [39], Marsden, Pekarsky, and

Shkoller [34], Bobenko and Suris [12,13]). In this chapter 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 the ideal constraint issue of Cortés and Martínez. That is, the discrete Lagrange–d'Alembert principle for Hamel equations is identical to the discrete Lagrange–d'Alembert principle of Marsden and West.

3.1 Discrete Lagrangian Mechanics Revistited

We have already seen how a discrete analogue of Lagrangian mechanics can be obtained by discretizing Hamilton's principle; this approach underlies the construction of variational integrators. In this section we present an alternate viewpoint on the same principle: this new viewpoint will subsequently guide our derivation of the discrete Hamel's equations.

Recall that the discrete Euler–Lagrange equations (2.23) are derived in the process of finding a discrete trajectory $q_d \in \mathcal{C}_d(Q)$ that extremizes the discrete action $S_d : \mathcal{C}_d(Q) \to \mathbb{R}$. The discrete action is the summation of a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$, and the discrete Lagrangian is itself understood as an approximation of the continuous action (a path integral of the Lagrangian function L) via an identification of the tangent bundle TQ with $Q \times Q$.

In the case that Q is a vector space, it may be convenient to use $(q_{k+1/2}, v_{k,k+1}) \in TQ$, 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 $(q_{k+1/2}, v_{k,k+1})$, which we will denote as $L^d : TQ \to \mathbb{R}$ so that

$$L^d(q_{k+1/2}, v_{k,k+1}) = L_d(q_k, q_{k+1}),$$

and the discrete Euler–Lagrange equations read

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

These equations are equivalent to the variational principle

$$\delta S^d = \sum_{k=0}^{N-1} \big(D_1 L^d(q_{k+1/2}, v_{k,k+1})\delta q_{k+1/2} + D_2 L^d(q_{k+1/2}, v_{k,k+1})\delta v_{k,k+1}\big) = 0, \tag{3.1}$$

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

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

It is straightforward to show that the principle (3.1) is equivalent to the discrete Hamilton's principle expressed as (2.26), and that the form of the discrete Euler–Lagrange equations presented here is likewise equivalent to (2.23). The discrete Hamel formalism introduced below may be interpreted as a generalization of the representation (3.1) of discrete mechanics.

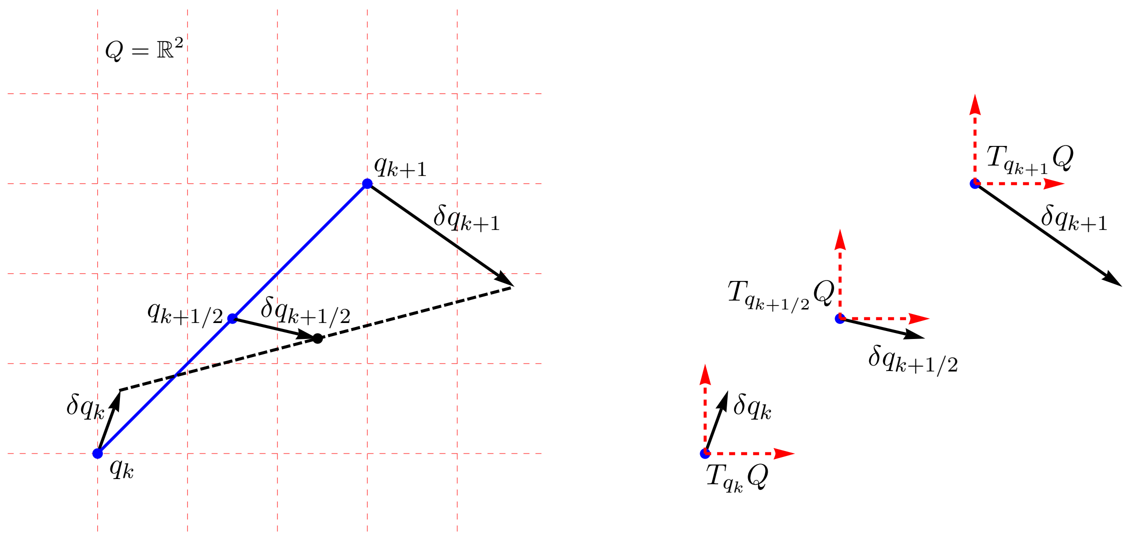


Figure 3.1: The virtual displacement vectors associated with an arbitrary variation at the step defined by $(q_k, q_{k+1}) \in Q \times Q$ in a flat two-dimensional vector space.

3.2 Discrete Hamel's Equations

In the rest of the chapter we assume that Q is a vector space. Start with a sequence of configurations specified by a trajectory $q_d = \{q_k\}_{k=0}^N$. Given a parameter $\tau \in [0,1]$, define the points $q_{k+\tau} := (1-\tau)q_k + \tau q_{k+1}$ for each $1 \le k \le N-1$ to be ***interpolation points***. We introduce a procedure for defining a "discrete velocity" $v_{k,k+1} \in T_{q_{k+\tau}}Q$ by taking the parameter τ-derivative of the interpolation point definition, so that

$$v_{k,k+1} = \frac{1}{h}\frac{d}{ds}\bigg|_{s=\tau}\big((1-s)q_k + sq_{k+1}\big) = \frac{1}{h}(q_{k+1} - q_k). \tag{3.2}$$

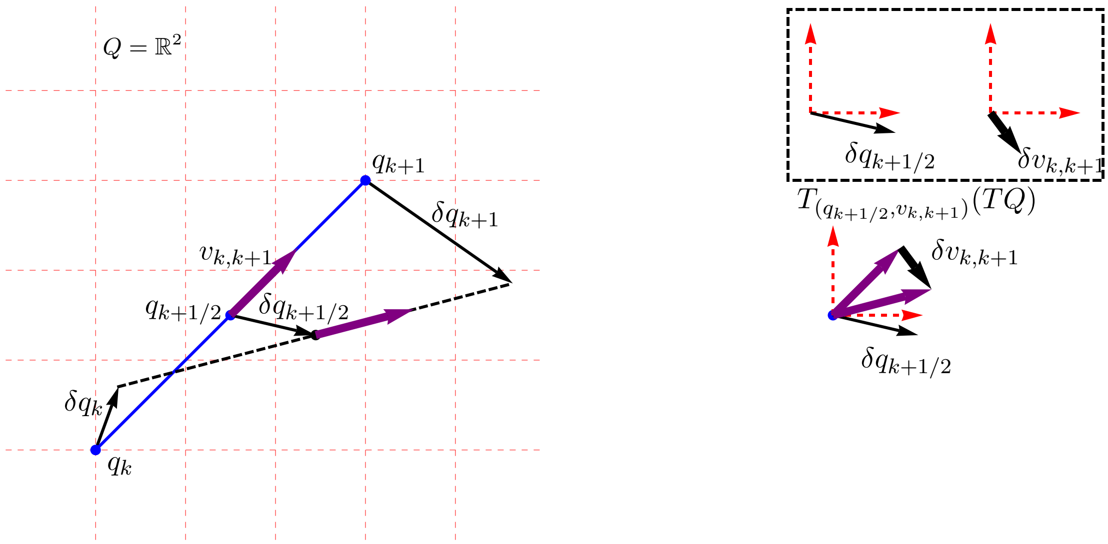


Figure 3.2: Virtual displacements $(\delta q_{k+1/2}, \delta v_{k,k+1}) \in T_{(q_{k+1/2}, v_{k,k+1})}(TQ)$, again where Q is a two-dimensional vector space.

Notice that in the case of $\tau = \frac{1}{2}$, the value of $(q_{k+1/2}, v_{k,k+1})$ under this definition is precisely the midpoint representation of the state of a discrete mechanical system presented in the previous section.

The velocity components relative to the frame $u(q)$ at $q_{k+\tau}$ will be denoted $\xi_{k,k+1} = (\xi^1_{k,k+1}, \ldots, \xi^n_{k,k+1})$. Similar to [14,27], the phase space for the suggested discretization of Hamel's equation is the tangent bundle TQ. That is, in local coordinates (q, ξ) on TQ the discrete Lagrangian $\ell^d : TQ \to \mathbb{R}$ reads $\ell^d = \ell^d(q_{k+\tau}, \xi_{k,k+1})$. To discretize a continuous-time system, we suggest the following procedure:

1. Select a frame $u(q)$ and identify the continuous-time Lagrangian $l(q, \xi)$, as in (2.11).

2. Construct the discrete Lagrangian using the formula

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

The action sum then is

$$s^d = \sum_{k=0}^{N-1} \ell^d(q_{k+\tau}, \xi_{k,k+1}), \tag{3.3}$$

which is an approximation of the action integral (2.3) of the continuous-time system.

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

$$\zeta_{k+\tau} = (1-\tau)\zeta_k + \tau\zeta_{k+1}. \tag{3.4}$$

The vectors ζ_k, ζ_{k+1}, and $\zeta_{k+\tau}$ in T_{q_k}, $T_{q_{k+1}}Q$, and $T_{q_{k+\tau}}Q$ will be used below to establish the discrete analogues of the variation formulae.

Define the ***discrete conjugate momentum*** by

$$\mu_{k,k+1} := D_2\ell^d(q_{k+\tau}, \xi_{k,k+1}). \tag{3.5}$$

Below, we use the notations

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

etc. Recall that the term $u[\ell^d](q,\xi)$ is defined to be to the directional derivatives of the (discrete) Lagrangian along the vectors $u(q)$ as in(2.14), so that

$$u_j[\ell^d](q,\xi) = \frac{\partial \ell^d(q,\xi)}{\partial q^i}\psi^i_j(q).$$

Theorem 3.1. *The sequence $\big(q_{k+\tau}, \xi_{k,k+1}\big) \in TQ$ for $k = 0,\dots,N-1$ satisfies the **discrete Hamel's equations***

$$\begin{aligned}\tfrac{1}{h}\big(\mu_{k-1,k} - \mu_{k,k+1}\big) + \tau u\big[\ell^d\big]_{k-1+\tau} + (1-\tau)u\big[\ell^d\big]_{k+\tau} \\ + \tau\big[\xi_{k-1,k}, \mu_{k-1,k}\big]^*_{q_{k-1+\tau}} + (1-\tau)\big[\xi_{k,k+1}, \mu_{k,k+1}\big]^*_{q_{k+\tau}} = 0\end{aligned} \tag{3.6}$$

if and only if

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

where

$$\delta q_{k+\tau} = u(q_{k+\tau})\cdot\zeta_{k+\tau} = \zeta^j_{k+\tau}u_j(q_{k+\tau}), \tag{3.7}$$

$$\delta\xi_{k,k+1} = \tfrac{1}{h}\big(\zeta_{k+1} - \zeta_k\big) + \big[\xi_{k,k+1}, \zeta_{k+\tau}\big]_{q_{k+\tau}}. \tag{3.8}$$

Here $\zeta_0 = \zeta_N = 0$, and $\zeta_{k+\tau}$ is defined in (3.4), $k = 0,\dots,N-1$.

Remark 3.1. In order to obtain a complete system of equations, one supplements (3.6) with a discrete analogue of the kinematic equation $\dot q = u(q)\cdot\xi$. There is a certain freedom in doing that.

Here we have already stated in (3.2) that the discrete velocity is related to the parameterization $q_{k+\tau} = (1-\tau)q_k + \tau q_{k+1}$. Thus, for now we assume this discrete analogue to be

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

so that in index notation

$$\xi^j_{k,k+1} = \left(\frac{q^i_{k+1} - q^i_k}{h} \right) \phi^j_i(q_{k+\tau}).$$

In coordinate form[1], the discrete Hamel's equations and the formulae for variations read

$$\begin{aligned} \tfrac{1}{h}\big(\mu_{k-1,k;j} - \mu_{k,k+1;j}\big) + \tau u_j \big[\ell^d\big]_{k-1+\tau} + (1-\tau) u_j \big[\ell^d\big]_{k+\tau} \\ + \tau c^a_{ij}(q_{k-1+\tau}) \xi^i_{k-1,k} \mu_{k-1,k;a} + (1-\tau) c^a_{ij}(q_{k+\tau}) \xi^i_{k,k+1} \mu_{k,k+1;a} = 0, \end{aligned}$$

and

$$\begin{aligned} \delta q^i_{k+\tau} &= \psi^i_b(q_{k+\tau}) \zeta^b_{k+\tau}, \\ \delta \xi^b_{k,k+1} &= \tfrac{1}{h}\big(\zeta^b_{k+1} - \zeta^b_k\big) + c^b_{ij}(q_{k+\tau})\, \xi^i_{k,k+1} \zeta^j_{k+\tau}, \end{aligned}$$

respectively.

Remark 3.2. Unlike the continuous-time case, the formulae for variations (3.7) and (3.8) *cannot* be derived in a manner presented in the proof of Theorem 2.2 (see Zenkov, Bloch, and Marsden [10] or more recently Ball and Zenkov [5] for the proof). The situation here is somewhat similar to the issue encountered and resolved by Chetaev in his work [16] 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 aforementioned principles become equivalent.

Proof. Using formulae (3.7) and (3.8) and computing the variation of the action sum (3.3), one

[1] We use a semi-colon to distinguish the index of the covector component from the index of the step $(k, k+1)$. This should not be confused with the covariant derivative. Hence, $\mu_{k,k+1;a}$ should be understood as the a-th component of the covector $\mu_{k,k+1} \in T^*_{q_{k+1/2}}Q$.

obtains

$$\begin{aligned}
\delta s^d &= \sum_{k=0}^{N-1} D_1\ell^d(q_{k+\tau},\xi_{k,k+1})\,\delta q_{k+\tau} + D_2\ell^d(q_{k+\tau},\xi_{k,k+1})\,\delta\xi_{k,k+1} \\
&= \sum_{k=0}^{N-1} \Big\langle D_1\ell^d_{k+\tau}, u_{k+\tau}\cdot\zeta_{k+\tau}\Big\rangle \\
&\qquad\qquad + \Big\langle D_2\ell^d_{k+\tau}, (\zeta_{k+1}-\zeta_k)/h + \big[\xi_{k,k+1},\zeta_{k+\tau}\big]_{q_{k+\tau}}\Big\rangle \\
&= \sum_{k=1}^{N-1} \Big\langle \tfrac{1}{h}(\mu_{k-1,k}-\mu_{k,k+1}),\zeta_k\Big\rangle \\
&\qquad\qquad + \Big\langle u\big[\ell^d\big]_{k+\tau} + \big[\mu_{k,k+1},\xi_{k,k+1}\big]^*_{q_{k+\tau}}, (1-\tau)\zeta_k + \tau\zeta_{k+1}\Big\rangle \\
&= \sum_{k=1}^{N-1} \Big\langle \tfrac{1}{h}(\mu_{k-1,k}-\mu_{k,k+1}),\zeta_k\Big\rangle + \Big\langle \tau u\big[\ell^d\big]_{k-1+\tau} + (1-\tau)u\big[\ell^d\big]_{k+\tau},\zeta_k\Big\rangle \\
&\qquad\qquad + \Big\langle \tau\big[\mu_{k-1,k},\xi_{k-1,k}\big]^*_{q_{k-1+\tau}} + (1-\tau)\big[\mu_{k,k+1},\xi_{k,k+1}\big]^*_{q_{k+\tau}},\zeta_k\Big\rangle.
\end{aligned}$$

Thus, vanishing of δs^d for arbitrary ζ_k, $k = 1,\ldots,N-1$, is equivalent to discrete Hamel's equations (3.6). $\square$

The formulae for variations (3.7) and (3.8) in the discrete setting are motivated by the following observations. First, recall that in the continuous-time setting the formula (2.19) for $\delta\xi$ follows from the formula

$$\delta(u\cdot\xi) - \frac{d}{dt}(u\cdot\zeta) = 0. \tag{3.9}$$

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

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

and the interpretation of the operator δ as a directional derivative, one obtains

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

and therefore

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

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, as we defined in the beginning of the current section:

$$q_{k+\tau} = (1-\tau)q_k + \tau q_{k+1}, \quad 0 \le \tau \le 1, \tag{3.10}$$

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 $v_{k,k+1} = (q_{k+1} - q_k)/h$. From (3.10),

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

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

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

To summarize, the discrete analogue of (3.9) reads

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

which implies formula (3.8) for variation $\delta\xi_{k,k+1}$.

Remark 3.3. The discrete Hamel's equations are, generally, distinct from the discrete Euler–Lagrange equations. That is, the discrete Hamel's equations and the discrete Euler–Lagrange equations are algebraically distinct second order difference equations. This will be discussed further in Chapter 5.

3.3 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.

3.3.1 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 $\mathcal{D}_q \subset T_qQ$, one for each $q \in Q$. A curve $q(t) \in Q$ will be 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 are, in general, ***nonholonomic***.[2]

Consider a Lagrangian $L : TQ \to \mathbb{R}$. Assume that the constraints are ***ideal***, that is, they can be replaced with ***reaction forces***[3] that at each $q \in Q$ belong to the null space $\mathcal{D}_q^\circ \subset T_q^*Q$ of $\mathcal{D}_q$. The equations of motion are given by the following ***Lagrange–d'Alembert principle***.

Definition 3.2. *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 $a \le t \le b$.

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 [28] and [8] for a discussion and references).

3.3.2 The Constrained Hamel's Equations

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

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

such that $\mathcal{D}_q = \operatorname{span}\{u_1(q), \dots, u_{n-p}(q)\}$. Each $\dot{q} \in TQ$ can be uniquely written as

$$\dot{q} = u(q) \cdot \xi^{\mathcal{D}} + u(q) \cdot \xi^{\mathcal{U}}, \quad \text{where} \quad u(q) \cdot \xi^{\mathcal{D}} \in \mathcal{D}_q, \tag{3.11}$$

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

[3] This means that the constraints are not present anymore, but forces are imposed on the unconstrained system, and the dynamics of the forced Lagrangian system is identical to that of the constrained system.

i.e., $u(q) \cdot \xi^{\mathcal{D}}$ 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 $u^*(q) \in T^*Q \times \cdots \times T^*Q$ denotes the dual frame of $u(q)$. Using (3.11), the constraints read

$$\xi = \xi^{\mathcal{D}} \quad \text{or} \quad \xi^{\mathcal{U}} = 0. \tag{3.12}$$

This implies

$$\delta\xi = \delta\xi^{\mathcal{D}} \quad \text{or} \quad \delta\xi^{\mathcal{U}} = 0. \tag{3.13}$$

The Lagrange–d'Alembert principle in combination with (3.13) proves the following theorem:

Theorem 3.3. *The dynamics of a nonholonomic system is represented by the* ***constrained Hamel's equations***

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

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

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

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

$$\frac{d}{dt} \frac{\partial \ell_c}{\partial \xi^{\mathcal{D}}} - \left(\left[\xi^{\mathcal{D}}, \frac{\partial \ell}{\partial \xi} \right]_q^* \right)_{\mathcal{D}} - u_{\mathcal{D}}[\ell_c] = 0, \quad \xi^{\mathcal{U}} = 0. \tag{3.14}$$

3.3.3 Continuous-Time Chaplygin Systems

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

$$T_qQ = \mathcal{D}_q \oplus T_q \,\mathrm{Orb}_H(q). \tag{3.15}$$

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

Following [8], define an ***Ehresmann connection*** by requiring that the group directions and the constraint distribution provide a vertical and horizontal spaces, respectively. These spaces are denoted V_q and H_q.

In other words, the nonholonomic kinematic constraints provide an Ehresmann connection on the principal bundle $\pi : Q \to Q/H$. 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 Q/H.

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

$$A = \omega^a \frac{\partial}{\partial s^a}, \quad \text{where} \quad \omega^a(q) = A^a_\alpha(x)\, dx^\alpha + ds^a, \tag{3.16}$$

where $\alpha = 1, \dots, (n-p)$, and where $a = (n-p+1), \dots, n$. Recall also that the *horizontal space* $H_q = \ker A$, so that $T_qQ = H_q \oplus V_q$, in full agreement with (3.15).

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

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

where $\mathrm{hor}\, X$ and $\mathrm{hor}\, Y$ are the horizontal parts of the vectors $X, Y \in T_qQ$, respectively. In the bundle coordinates (x^α, s^a),

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

where

$$B^a_{\alpha\beta} = \frac{\partial A^a_\alpha}{\partial r^\beta} - \frac{\partial A^a_\beta}{\partial r^\alpha}.$$

The ***constrained Lagrangian*** is the restriction of Lagrangian onto the constraint distribution, $L_c = L|_{\mathcal{D}}$. For Chaplygin systems, L and L_c naturally reduce to function on TQ/H and on $\mathcal{D}/H$. In the bundle coordinates (x, s) this simply means that L is independent of s,[4] i.e., $L = L(x, \dot{x}, \dot{s})$, and the constrained Lagrangian reads

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

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

The equations of motion for Chaplygin systems

$$\frac{d}{dt}\frac{\partial L_c}{\partial \dot{x}} - \frac{\partial L_c}{\partial x} = \left\langle \frac{\partial L}{\partial \dot{s}}, \mathbf{i}_{\dot{x}}\, B \right\rangle, \tag{3.17}$$

or, in coordinates,

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

$\alpha, \beta = 1, \ldots, (n-p)$, were first derived, by a coordinate calculation, by Chaplygin in [15]. They are called the ***Chaplygin equations***.

We now obtain equations (3.17) using Hamel's formalism. Recall that connection (3.16) 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

$$u_\alpha = \operatorname{hor} \partial_{x^\alpha} = \partial_{x^\alpha} - A^a_\alpha \partial_{s^a}, \quad u_a = \partial_{s^a}. \tag{3.18}$$

Using this frame,

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

$\alpha = 1, \ldots, (n-p)$, $a = (n-p+1), \ldots, n$, 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

$$\ell(x, \xi) = L(x, \xi^{\mathcal{D}}, \xi^{\mathcal{U}} - A(x)\xi^{\mathcal{D}}), \quad \ell_c(x, \xi^{\mathcal{D}}) = L(x, \xi^{\mathcal{D}}, -A(x)\xi^{\mathcal{D}}). \tag{3.19}$$

Evaluating the Jacobi–Lie brackets of the fields (3.18), one obtains

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

which implies

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

and thus (3.17) are just the constrained Hamel equations (3.14). Recall that B is the curvature of the form A.

An important remark is that, form Chaplygin's prospective, equations (3.17) 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 $\dot{x}^\alpha$ in the classical literature are viewed as the reduced configuration velocities, whereas from the point of view of Hamel's formalism $\dot{x}^\alpha$ represent the velocity components along the non-commuting fields u_α.

3.3.4 Discrete Nonholonomic Systems

Discrete nonholonomic systems (nonholonomic integrators) were introduced by Cortés and Martínez in [17].

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

- A discrete Lagrangian $L^d : Q \times Q \to \mathbb{R}$;
- An $(n-s)$-dimensional distribution $\mathcal{D}$ on TQ;
- 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 [17]):

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

Here $D_1 L^d$ and $D_2 L^d$ denote the partial derivatives of the discrete Lagrangian with respect to the first and the second inputs, respectively.

As pointed out in [19, 20], 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 [17,19,20]. In a recent paper [27], 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 [31, 32] 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 will 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.

3.3.5 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 3.3.2. In a similar manner, we accept that the dynamics of a discrete nonholonomic system is determined by the ***discrete Lagrange–d'Alembert equations***, obtained by *first* taking the variation of the discrete action, as in Section 3.2, and *then* imposing the discrete constraints. We point out 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 3.3.4.

In the continuous-time setting, the constraints are represented by formula (3.12). We thus suggest that, under the same assumptions on the frame selection as in Section 3.3.2, the discrete constraints are

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

The dynamics of a discrete nonholonomic system then is given by the ***constrained Hamel equations***

$$\begin{aligned} &\tfrac{1}{h}\big(\mu_{k-1,k} - \mu_{k,k+1}\big)_{\mathcal{D}} + \big(\tau u\big[\ell^d\big]_{k-1+\tau} + (1-\tau) u\big[\ell^d\big]_{k+\tau}\big)_{\mathcal{D}} \\ &\quad + \big(\tau\big[\xi^{\mathcal{D}}_{k-1,k}, \mu_{k-1,k}\big]^*_{q_{k-1+\tau}} + (1-\tau)\big[\xi^{\mathcal{D}}_{k,k+1}, \mu_{k,k+1}\big]^*_{q_{k+\tau}}\big)_{\mathcal{D}} = 0, \end{aligned} \tag{3.20}$$

where, as before, $\mu_{k,k+1}$ is given by formula (3.5). Of a special interest is the value $\tau = 1/2$, in which case one verifies that the order of approximation of (3.20) is 2.

3.3.6 Discrete Chaplygin Systems

Given a continuous-time Chaplygin system, we construct its discretization by utilizing the discrete Hamel formalism. Using the frame (3.18) and the continuous-time Lagrangians (3.19) introduced in Section 3.3.3, the discrete Lagrangian and the discrete constrained Lagrangian read

$$\ell^d(x_{k+\tau}, \xi_{k,k+1}) = h\ell(x_{k+\tau}, \xi_{k,k+1}),$$
$$\ell_c^d(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}) = \ell^d(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}) \equiv h\ell_c(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}).$$

The dynamics is then given by equation (3.20), with

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

and $\mu_{k,k+1}$ defined as in (3.5).

We now convert this dynamics into a discrete analogue of the Chaplygin equations (3.17). Following the general discretization procedure, we obtain the formulae

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

where $\Delta x_k = (x_{k+1} - x_k)$ and $\Delta s_k = (s_{k+1} - s_k)$. Then, invoking (3.19), it is straightforward to see that

$$\begin{aligned} \ell^d(x_{k+\tau}, \xi_{k,k+1}) &= h\ell(x_{k+\tau}, \xi_{k,k+1}) \\ &= hL(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}, \xi^{\mathcal{U}}_{k,k+1} - A(x_{k+\tau})\xi^{\mathcal{D}}_{k,k+1}) \end{aligned} \tag{3.21}$$

and

$$\begin{aligned} \ell_c^d(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}) &= \ell^d(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}) = h\ell_c(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}) \\ &= hL_c(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}) = hL_c(x_{k+\tau}, \Delta x_{k,k+1}/h) \\ &= hL(x_{k+\tau}, \Delta x_k/h, -A(x_{k+\tau})\Delta x_k/h), \end{aligned} \tag{3.22}$$

where $L(x, \dot{x}, \dot{s})$ is the Lagrangian of the continuous-time Chaplygin system. From formulae (3.21), (3.22), and (3.18), one obtains

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

Next, since we utilize the frame (3.18) just like in the continuous-time setting, the formula

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

is established with an aid of the arguments of Section 3.3.3. To keep the formulae shorter, we write the latter expression as

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

Finally,

$$\begin{aligned}\big(u[\ell^d](q_{k+\tau},\xi_{k,k+1})\big)_{\mathcal{D}} &= D_1\, \ell^d(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}) \\ &= D_1\, \ell^d_c(x_{k+\tau}, \xi^{\mathcal{D}}_{k,k+1}) = hD_1\, L_c(x_{k+\tau}, \Delta x_k/h).\end{aligned}$$

Summarizing, the dynamics of the discrete Chaplygin system reads

$$\begin{aligned}\tfrac{1}{h}\big((D_2\, L_c)_{k+\tau} - (D_2\, L_c)_{k-1+\tau}\big) &= \tau (D_1\, L_c)_{k-1+\tau} + (1-\tau)(D_1\, L_c)_{k+\tau} \\ &+ \tau \big\langle D_3\, L, \mathbf{i}_{\Delta x_{k-1}/h}\, B\big\rangle_{k-1+\tau} + (1-\tau)\big\langle D_3\, L, \mathbf{i}_{\Delta x_k/h}\, B\big\rangle_{k+\tau},\end{aligned} \tag{3.23}$$

where $(D_i\, L_c)_{k+\tau} := D_i\, L_c(x_{k+\tau}, \Delta x_k/h)$. Remarkably, the discrete Chaplygin equations (3.23) are identical to the discretization of continuous-time Chaplygin equations (3.17) viewed as forced Euler–Lagrange dynamics. For more details on this latter discretization of the Chaplygin equations see [17] and [32].

3.3.7 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 the constraint distribution $\mathcal{D}$, as discussed in Section 3.3.3. Assume that the dynamics of the Chaplygin system (3.17) is invariant with respect to the action of a commutative group G on Q/H.[5] Often such a situation is the result of the original system being invariant with respect to the semidirect product of 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 principle fiber bundle with the structure group G. Thus, locally, there exist the bundle coordinates $x = (r, g)$

[5] The general noncommutative setting is not studied in this chapter.

on Q/H. Let $\dim G = m$.

Under certain assumptions (see e.g. [26] and [52]), the dynamics has an m-dimensional manifold of relative equilibria, which are the solutions of (3.17) that in the bundle coordinates (r, g) read

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

As established in Karapetyan [26], some of these relative equilibria may be *partially asymptotically stable*. Karapetyan justifies stability using the center manifold stability analysis techniques, which, for nonholonomic systems under consideration, reduces to verifying that the nonzero spectrum of linearization at the relative equilibrium of interest belongs to the left half-plane.[6]

Partially asymptotically stable relative equilibria are a part of the ω-limit set of dynamics (3.17). Similarly, relative equilibria that become partially asymptotically stable after time reversal are a part of the α-limit set of dynamics (3.17).

It is important for a long-term numerical integrator to preserve the manifold of relative equilibria of (3.17) 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 [31,32], 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 limit sets.

A relative equilibrium of a discrete Chaplygin system (3.23) with commutative symmetry is a solution

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

Assume now that $\tau = 1/2$ in equations (3.23). Let $h > 0$ be the time step.

Theorem 3.4 (Lynch and Zenkov [31, 32])**.** *Discretization* (3.23) *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* (3.23) *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 proposed discrete Lagrange–d'Alembert principle ensures the necessary conditions for structural stability of the associated nonholonomic integrator.

[6]The stability analysis of relative equilibria of nonholonomic systems has a long history, starting form the results of Walker [48] and Routh [44]; see [52] for some of this history and for the energy-momentum method for nonholonomic systems.

Chapter 4

Variational Principles for Continuous-Time Hamel's Equations

In this chapter we present a variational derivation of Hamel's equations from an extension of Hamilton's principle: the Hamilton–Pontryagin principle. The result is part of a recent paper [4] by Ball, Zenkov, and Bloch.

An earlier variational derivation of Hamel's equations (2.15) is based on the formula for variations of velocity components that generalizes the variation formula for the Euler–Poincaré equations (see Marsden [33], Marsden and Ratiu [35], and Bloch et al. [10] for details). In this manner we may think of the Hamel's equations as a generalization of the Euler–Poincaré equations, which describe the mechanics of a system specified by a left-invariant Lagrangian acting on the tangent bundle of a Lie group. Bou-Rabee and Marsden [14] develop structure-preserving variational integrators for the Euler–Poincaré equations through a discretization of the Hamilton–Pontryagin principle.

It is plausible to assume that the derivation of variational integrators for Hamel's equations may be based on the Hamilton–Pontryagin principle adapted for Hamel's equations. This approach, eschewing the constrained variation approach of Bloch et al. [10] in favor of constraint-free variations, is examined below in the context of variational integrators for Chaplygin systems.

Just as in the Euler–Poincaré case, Hamel's equations contain terms whose structure at first appears to be non-variational. The presence of these terms is caused by non-vanishing Jacobi–Lie bracket of the vector fields that are used to measure the velocity components. The variational derivation of Hamel's equations we present below utilizes the Hamilton–Pontryagin principle and produces these bracket terms using *unconstrained variations*, albeit taken in a different, larger-dimensional space.

The origins of the Hamilton–Pontryagin principle may be traced back to Livens [30]; see also Pars [42]. The recent results of Yoshimura and Marsden [49–51] reveal the links between this principle, implicit Lagrangian systems, and Dirac structures. The latter are important in interconnected mechanical systems, electric circuits, electromechanical systems, and control, as discussed in e.g. van der Schaft and Maschke [46, 47], van der Schaft [45], and Bloch and Crouch [7]. As shown in Yoshimura and Marsden [50], the dynamics and the Legendre transform are the outcomes of a variational procedure when the Hamilton–Pontryagin principle is used.

4.1 The Hamilton–Pontryagin Principle

Let Q be a manifold, TQ be its tangent, and T^*Q be its cotangent bundles. Let q, (q, v), and (q, p) be local coordinates on Q, TQ, and T^*Q, respectively. Let $t \mapsto (q(t), v(t), p(t))$, $t \in [t_0, t_F]$, be a curve in the Pontryagin bundle $TQ \oplus T^*Q$. Following Yoshimura and Marsden [49–51], define the action functional on $TQ \oplus T^*Q$ by the formula

$$S_{HP} = \int_{t_0}^{t_F} \left[L\left(q(t), v(t)\right) + \langle p(t), \dot{q}(t) - v(t) \rangle \right] dt. \tag{4.1}$$

Consider the space of curves in $TQ \oplus T^*Q$ that satisfy the conditions $q(t_0) = q_0$, $q(t_F) = q_F$, with $t_0 \leq t \leq t_F$, where q_0 and q_F are two points in the configuration space Q. The variation of action (4.1) on this space of curves is computed to be

$$\delta S_{HP} = \int_{t_0}^{t_F} \left[\left(\frac{\partial L}{\partial q} - \dot{p} \right) \delta q + \left(\frac{\partial L}{\partial v} - p \right) \delta v + \left(\dot{q} - v \right) \delta p \right] dt.$$

Theorem 4.1. *The following statements are equivalent:*

1. *The curve* $(q(t), v(t), p(t))$ *is a critical point of the action functional* (4.1) *on the space of curves in* $TQ \oplus T^*Q$ *connecting* $q_0 \in Q$ *to* $q_F \in Q$ *on the interval* $[a, b]$, *with variations satisfying* $\delta q(t_0) = \delta q(t_F) = 0$.

2. *The curve* $(q(t), v(t), p(t))$ *satisfies the* ***implicit Euler–Lagrange equations***

$$\frac{\partial L}{\partial q} - \dot{p} = 0, \qquad p = \frac{\partial L}{\partial v}, \qquad \dot{q} = v. \tag{4.2}$$

Equations (4.2) include the Euler–Lagrange equations, the *Legendre transform* $p = \partial_v L$, and the *second order condition* $\dot{q} = v$. We emphasize that variations δv and δp are not induced by variations δq.

4.2 Hamel's Equations and the Hamilton–Pontryagin Principle

We start by rewriting action (4.1) using the frame $u_i(q)$, $i = 1, \dots, n$, as defined in (2.9). Denote the components of $\dot{q}$, v, and p relative to the frame $u_i(q)$ and its dual by ξ, η, and μ, respectively:

$$\dot{q} = \xi^j u_j = \xi^j \psi_j^i \frac{\partial}{\partial q^i}, \tag{4.3a}$$

$$v = \eta^j u_j = \eta^j \psi_j^i \frac{\partial}{\partial q^i}, \tag{4.3b}$$

$$p = \mu_j u^j = \mu_j \phi_i^j dq^i. \tag{4.3c}$$

The action functional (4.1) becomes

$$S_{HP} = \int_{t_0}^{t_f} \left[\ell\left(q(t), \eta(t)\right) + \langle p(t), \dot{q}(t) - v(t) \rangle \right] dt, \tag{4.4}$$

where $\dot{q}$, v, and p are given by formulae (4.3).

Theorem 4.2. *The following statements are equivalent:*

1. *The curve $(q(t), \eta(t), \mu(t)), t_0 \le t \le t_F$, is a critical point of the action functional* (4.4) *on the space of curves in $TQ \oplus T^*Q$ connecting q_0 and q_F on the interval $[t_0, t_F]$, where we choose variations of the curve $(q(t), \eta(t), \mu(t))$ that satisfy $\delta q(t_0) = \delta q(t_F) = 0$.*

2. *The* ***implicit Hamel's equations***

$$u_j[\ell] - \dot{\mu}_j + \phi_{m,r}^k \psi_i^m \psi_j^r \mu_k \eta^i - \phi_{r,m}^k \psi_i^m \psi_j^r \mu_k \xi^i = 0, \tag{4.5a}$$

$$\mu = \frac{\partial \ell}{\partial \eta}, \tag{4.5b}$$

$$\xi = \eta \tag{4.5c}$$

hold. Coupled with (4.3a)*, the implicit Hamel's equations capture the dynamics for the Lagrangian $\ell(q, \xi)$.*

Proof. Taking the variation of (4.4) gives

$$\begin{aligned}
\delta S &= \int_{t_0}^{t_f} \left[\delta \ell\left(q, \eta\right) + \delta \left\langle p, \dot{q} - v \right\rangle \right] dt \\
&= \int_{t_0}^{t_f} \left[\frac{\partial \ell(q, \eta)}{\partial q^i} \delta q^i + \frac{\partial \ell(q, \eta)}{\partial \eta^i} \delta \eta^i + \langle \delta p, \dot{q} - v \rangle + \langle p, \delta \dot{q} \rangle - \langle p, \delta v \rangle \right] dt.
\end{aligned}$$

Next, we evaluate δv and obtain

$$
\begin{aligned}
\delta v = \delta\eta^i u_i + \eta^i \delta u_i &= \delta\eta^i u_i + \eta^i \psi^m_{i,s} \delta q^s \frac{\partial}{\partial q^m} \\
&= \delta\eta^i u_i + \eta^i \psi^m_{i,s} \phi^k_m \delta q^s u_k \\
&= \delta\eta^i u_i - \eta^i \psi^m_i \phi^k_{m,s} \delta q^s u_k,
\end{aligned}
\tag{4.6}
$$

where the last step is a consequence of the inverse matrix differentiation rule: $\partial(\psi)\phi = -\psi\partial(\phi)$. Therefore, the term $\langle p, \delta v\rangle$ becomes

$$
\langle p, \delta v\rangle = \mu_i \delta\eta^i - \mu_k \eta^i \psi^m_i \phi^k_{m,s} \delta q^s.
$$

Integration by parts replaces the term $\langle p, \delta\dot{q}\rangle$ with $-\langle p, \delta q\rangle$, as the term $d\langle p, \delta q\rangle/dt$ vanishes after integration. Evaluating $\dot{p}$, we obtain

$$
\begin{aligned}
\dot{p} = \dot{\mu}_j u^j + \mu_i \frac{du^i}{dt} &= \dot{\mu}_j u^j + \phi^k_{s,r} \psi^r_i \mu_k \xi^i dq^s \\
&= \dot{\mu}_j \phi^j_s dq^s + \phi^k_{s,r} \psi^r_i \mu_k \xi^i dq^s.
\end{aligned}
$$

Therefore,

$$
-\langle \dot{p}, \delta q\rangle = -\dot{\mu}_j \phi^j_s \delta q^s - \phi^k_{s,r} \psi^r_i \mu_k \xi^i \delta q^s.
$$

Using these formulae (and through a change of indices in the $-\langle \dot{p}, \delta q\rangle$ term), the variation of action (4.4) becomes

$$
\begin{aligned}
\delta S_{HP} &= \int_{t_0}^{t_F} \Bigg[\left(\frac{\partial \ell}{\partial q^s} - \dot{\mu}_j \phi^j_s - \phi^k_{s,r} \psi^r_i \mu_k \xi^i + \mu_k \eta^i \psi^m_i \phi^k_{m,s} \right) \delta q^s \\
&\qquad + \left(\frac{\partial \ell}{\partial \eta^i} \mu_i \right) \delta\eta^i + \langle \delta p, \dot{q} - v\rangle \Bigg] dt \\
&= \int_{t_0}^{t_F} \Bigg[\left(\frac{\partial \ell}{\partial q^m} \psi^m_j - \dot{\mu}_j - \phi^k_{r,m} \psi^m_i \psi^r_j \mu_k \xi^i + \phi^k_{m,r} \psi^m_i \psi^r_j \mu_k \eta^i \right) \phi^j_s \delta q^s \\
&\qquad + \left(\frac{\partial \ell}{\partial \eta^i} \mu_i \right) \delta\eta^i + \langle \delta p, \dot{q} - v\rangle \Bigg] dt.
\end{aligned}
$$

Recall that the variations δv and δp are not induced by δq. By the independence of the variation δq, $\delta\eta$, and δp, vanishing of the variation of the action functional (4.4) is equivalent to the implicit Hamel's equations (4.5).

□

Recall that structure functions are defined in index notation in (2.13). Thus substituting (4.5b) and (4.5c) in (4.5a) and utilizing (2.13) produces Hamel's equations (2.15).

We emphasize that equations (4.5) include Hamel's equations, the Legendre transform $\mu = \partial_\eta \ell$, and the second order condition $\xi = \eta$.

Remark 4.1. As we have seen from the above development, in the case of regular Lagrangian mechanical systems, the Hamilton–Pontryagin principle unifies the Lagrangian and Hamiltonian formulations of mechanics. This is illustrative in the case of variational integrators, as seen in Bou-Rabee and Marsden [14], where integrators derived from a discretized Hamilton–Pontryagin principle give rise to discrete analogues of the reconstruction condition (i.e. identifying ξ with $\dot{q}$) and Legendre transform. Bou-Rabee and Marsden further describe the implementation of an s-stage Runge–Kutta discretization of the kinematic constraints in the discretized Hamilton–Pontryagin principle, yielding higher order symplectic partitioned Runge–Kutta methods [14].

4.3 The Hamilton–Pontryagin Principle and Discrete Chaplygin Equations

In this section we present an alternate discretization of the Chaplygin equations using the Hamilton–Pontryagin principle. The derivations presented below will illustrate both the usefulness of the Hamilton–Pontryagin perspective and of frame selection in non-holonomically constrained systems.

4.3.1 A Discrete Variational Principle on the Pontryagin Bundle

We would now like to describe a discrete path through the Pontryagin bundle and a corresponding discrete action mapping such discrete paths to the field of real numbers. Note that in the arbitrary path through the Pontryagin bundle, $v(t)$ and $p(t)$ do not necessarily correspond to the dynamical velocity and momentum of the system but are arbitrarily selected smooth sections of the tangent and cotangent bundles. It is only through the variational principle that we identify $v(t)$ with the time-derivative of position $\dot{q}(t)$ and $p(t)$ with the corresponding momentum via a Legendre transform. In the following development, note that $v_{k,k+1}$ is similarly considered to be an *arbitrary* vector, and is only identified with the discrete velocity through the discrete Hamilton–Pontryagin principle.

Taking a cue from our previous discretization, we define discrete vectors and covectors at the k-th steps as vectors and covectors at the interpolation points, so that

$$v_{k,k+1} \in T_{q_{k+\tau}} Q \quad \text{and} \quad p_{k,k+1} \in T^*_{q_{k+\tau}} Q.$$

As such, our discrete path is the union of discrete positions, velocities and momenta

$$\{q_k, v_{k,k+1}, p_{k,k+1}\} = \{q_k\}_{k=0}^{N} \cup \{v_{k,k+1}\}_{k=0}^{N-1} \cup \{p_{k,k+1}\}_{k=0}^{N-1}$$

and the corresponding discrete action S^d_{HP} on the Pontryagin bundle is the mapping

$$S^d_{HP}(\{q_k, v_{k,k+1}, p_{k,k+1}\}) = h\sum_{k=0}^{N-1}\left[L(q_{k+\tau}, v_{k,k+1}) + \left\langle p_{k,k+1}, \frac{1}{h}\frac{d}{ds}\Big|_{s=\tau}(q_{k+s}) - v_{k,k+1}\right\rangle\right].$$

Variations in the discrete Pontryagin bundle case work similarly to discrete variations discussed previously, except now we consider the $2N$ arbitrarily selected perturbations of the discrete vectors and covectors as $\delta v_{k,k+1}$ and $\delta p_{k,k+1}$ at the quadrature point of each step in addition to the variations of the $(N-1)$ step points δq_k.

It is straightforward to show that the discrete variational principle

$$\delta S^d_{HP}\left(\{q_k, v_{k,k+1}, p_{k,k+1}\}\right) = 0, \tag{4.7}$$

where variations are fixed at the endpoints so that $\delta q_0 = \delta q_N = 0$, results in the usual discrete Euler–Lagrange equations, and additionally identifies $v_{k,k+1}$ with the discrete velocity $\frac{1}{h}\frac{d}{d\tau}q_{k+\tau}$ and $p_{k,k+1}$ with a discrete version of the Legendre transform.

4.3.2 The Chaplygin System and the Pontryagin Bundle

We now present a variational derivation of the equations of motion of a (commutative) Chaplygin system using the Pontryagin bundle. We refer to Section 3.3.3 for the overall development of the system.

We consider an action on the Pontryagin bundle over Q (in the presence of structures associated with a commutative Chaplygin system) to be

$$S_{HP}(x, s, v, \eta, p, \mu) = \int_{t_0}^{t_F}\left[L(x^\alpha, s^a, v^\alpha, \eta^a) + \langle p, \dot{r} - v\rangle + \langle \mu, \dot{s} - \eta\rangle\right] dt$$

where p and v denote covector and vector sections of the Pontryagin bundle over Q/H and μ and η likewise denote covector and vector sections of the Pontryagin bundle over H, a Lie group acting on the configuration space.

Remark 4.2. In this section, we will constrain ourselves to the case when H is a *commutative* group, as in Chapter 3. While the development in the noncommutative case is robust, it is also significantly more complicated without offering additional insight into the usefulness of frames in the treatment of nonholonomically constrained systems. The commutative case will also fit better into the overall development of this dissertation.

In the noncommutative case left trivialized velocities are defined as $\xi = s^{-1}\dot{s} \in \mathfrak{h}$. In the commutative case, ξ is trivially associated with $\dot{s}$, so that $\dot{s} = \xi \in \mathfrak{h}$.

The Chaplygin systems have two distinguishing features. First, the Lagrangian L is invariant with respect to the group action of H. As mentioned in the remark above, the group velocity $\dot{s}$ is naturally identified with a vector in $\mathfrak{h}$, and $\{\partial_{s_a}\}$ and $\{ds^a\}$ will correspond to ordered bases of the Lie algebra $\mathfrak{h}$ and its dual space $\mathfrak{h}^*$ respectively.

Second, the system is subject to velocity constraints given by the connection

$$\mathcal{A}(x) : T_x(Q/H) \times \mathfrak{h} \to \mathfrak{h}.$$

Specifically, the constraint is such that when the connection acts on vectors in the tangent bundle over the entire configuration space, the result is zero so that $\mathcal{A}(x)(\dot{x}, \dot{s}) = 0$. The connection itself is defined to be of the form $\mathcal{A}(x)(\dot{x}, \dot{s}) = \dot{s} + A(x)\dot{x}$. Thus, the constraint reads $\dot{s} = -A(x)\dot{x}$, and the connection form may be described in terms of components with respect to the coordinate induced frame and coframe as a $p \times (n-p)$ matrix $A(x)$:

$$\mathcal{A}^a(x) = ds^a + A^a_\alpha(x)dx^\alpha.$$

Now, according to the Lagrange–d'Alembert principle, we do not impose our constraints on the system before we take variations in order to find suitable paths through the Pontryagin bundle. Rather we impose constraints *after* after taking variations. However, the constraints do come into play in the context of horizontal variations, as discussed below. This means that for horizontal variations, $(\delta x, \delta s)$ become

$$(\delta x, \delta s) \mapsto (\delta x, -A(x)\delta x). \tag{4.8}$$

To summarize, the action $\mathcal{S}_{HP}$ defined for paths over the Pontryagin bundle corresponding to our Chaplygin system is defined as

$$\mathcal{S}_{HP}(x, s, v, \eta, p, \mu) = \int_{t_0}^{t_F} \left[L(x^\alpha, v^\alpha, \eta^a) + \langle p, \dot{x} - v\rangle + \langle \mu, \dot{s} - \eta\rangle \right] dt.$$

In this setting it is fairly straightforward to show that the variational principle over the Pontryagin bundle,

$$\delta\mathcal{S}_{HP}(x, s, v, \eta, p, \mu) = 0,$$

subject to horizontal variations defined in equation (4.8), reduces to the condition

$$\begin{aligned}\int_{t_0}^{t_F} \Bigg[&\left(\frac{\partial L}{\partial x^\alpha} - \dot{p}_\alpha + \dot{\mu}_a A^a_\alpha(x) - [\mu, \dot{s}]^*_a A^a_\alpha(x) \right) \delta x^\alpha \\ &+ \left(\frac{\partial L}{\partial v^\alpha} - p_\alpha \right) \delta v^\alpha + \left(\frac{\partial L}{\partial \eta^a} - \mu_a \right) \delta \eta^a \\ &+ \delta p_\alpha \left(\dot{x}^\alpha - v^\alpha \right) + \delta \mu_a \left(\dot{s}^a - \eta^a \right) \Bigg] dt = 0.\end{aligned}$$

Linear independence of the various displacement vectors and covectors allow us to look at each piece of the above equation independently. The differential equations describing the evolution of the Chapylgin system in terms of the *unconstrained Lagrangian*, given in coordinate form with respect to the basis $\frac{\partial}{\partial x^\alpha}$ are

$$\frac{\partial L}{\partial x^\alpha} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^\alpha} + \frac{d}{dt}\frac{\partial L}{\partial \dot{s}^a} A^a_\alpha(x) = 0.$$

Now we would like to impose our constraint, namely $\dot{s} = -A(x)\dot{x}$. Notice this allows us to define a new constrained Lagrangian $L^c : T(Q/H) \to \mathbb{R}$ such that

$$L^c(x^\alpha, \dot{x}^\alpha) = L(x^\alpha, \dot{x}^\alpha, -A^a_\alpha(x)\dot{x}^\alpha),$$

which we will differentiate as follows in order to substitute into our differential equation

$$\begin{aligned}\frac{\partial L^c}{\partial x^\alpha} &= \frac{\partial L}{\partial x^\alpha} - \frac{\partial L}{\partial \dot{s}^a} A^a_{\beta,\alpha}\dot{x}^\beta \\ \frac{d}{dt}\frac{\partial L^c}{\partial \dot{x}^\alpha} &= \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^\alpha} - \frac{d}{dt}\frac{\partial L}{\partial \dot{s}^a} A^a_\alpha(x) - \frac{\partial L}{\partial \dot{s}^a} A^a_{\alpha,\beta}(x)\dot{x}^\beta.\end{aligned} \tag{4.9}$$

Thus our differential equations describing the evolution of the Chaplygin system in terms of the *constrained Lagrangian are*

$$\begin{aligned}\frac{\partial L^c}{\partial x^\alpha} - \frac{d}{dt}\frac{\partial L^c}{\partial \dot{x}^\alpha} &= \frac{\partial L}{\partial \dot{s}^a}\dot{x}^\beta \left(A^a_{\alpha,\beta}(x) - A^a_{\beta,\alpha}(x) \right) \\ &= \frac{\partial L}{\partial \dot{s}^a}\dot{x}^\beta B^a_{\alpha,\beta}(x)\dot{x}^\beta.\end{aligned} \tag{4.10}$$

4.3.3 The Chaplygin System and Frame Selection

In the previous section, we derived the differential equations of motion of a Chaplygin system through the Hamilton–Pontryagin principle. To describe vectors and covectors in this action we used components defined with respect to the coordinate induced frame and coframe spanning the Pontryagin bundle, locally

$$\operatorname{span}\left\{\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial s^a}, dx^\alpha, ds^a\right\} = (T_x(Q/H) \times \mathfrak{h}) \oplus (T_x^*(Q/H) \times \mathfrak{h}^*)\,.$$

We now investigate a frame based on the connection $\mathcal{A}(x)$ that clarifies the imposition of the constraint, consistent with work by Maruskin et al. [38]. Recall that our constraint may be written (in the local coordinate-induced frame) as

$$\mathcal{A}(x)(v,\eta) = (v^\alpha A^a_\alpha(x) + \eta^a)\frac{\partial}{\partial s^a} = 0,$$

so that a vector on the principal bundle (v, η) satisfies the constraints if and only if $\eta = -A(x)v$.

Suppose we (locally) select a new frame over each fiber such that $\operatorname{span}\{u_\alpha(x), u_a(x)\} = T_x(Q/H) \times \mathfrak{h}$ and we define the frame as an x-dependent linear transformation of the locally induced coordinate frame as

$$u_\alpha(x) = \frac{\partial}{\partial x^\alpha} - A^a_\alpha(x)e_a \quad \text{ and } \quad u_a(x) = \frac{\partial}{\partial s^a}. \tag{4.11}$$

It is fairly straightforward (using the natural pairing $\langle u^\alpha + u^a, u_\beta + u_b\rangle = 0$) to show that the corresponding dual spanning the covector bundle is

$$u^\alpha(x) = dx^\alpha \quad \text{ and } \quad u^a(x) = ds^a + A^a_\alpha(x)dx^\alpha. \tag{4.12}$$

Notice that this frame (and its dual) are a specific choice of the more general class of non-coordinate frames that we might select in Hamel's formalism. For the sake of convenience we shall refer to (4.11) and (4.12) as the Chaplygin frame and coframe.

We would next like to examine the components of a vector written with respect to the Chaplygin frame as they relate to the components of the same vector in the coordinate induced frame. Suppose that our vector (v, η) has coordinate-induced frame components (v^α, η^a) as usual, but the components of the same vector with respect to (4.11) are (w^α, ζ^a). That is,

$$(v,\eta) = v^\alpha\frac{\partial}{\partial x^\alpha} + \eta^a\frac{\partial}{\partial s^a} = w^\alpha u_\alpha(x) + \zeta^a u_a(x).$$

Expanding the Chaplygin frame out in terms of the coordinate induced frame gives us the equation

$$w^\alpha\frac{\partial}{\partial x^\alpha} + (\zeta^a - A^a_\alpha(x)w^\alpha)\frac{\partial}{\partial s^a} = v^\alpha\frac{\partial}{\partial x^\alpha} + \eta^a\frac{\partial}{\partial s^a}$$

thus, we see that components of vectors over the orbit space are unaffected by the change of frame and components of vectors in the Lie algebra undergo a linear transformation, so that

$$\zeta^a u_a(x) = (\eta^a + A^a_\alpha(x) v^\alpha) u_a(x).$$

Through a similar calculation, we can show that the components of covectors in the dual Lie algebra remain unchanged when we introduce the Chaplygin coframe, but components of the covectors in the cotangent bundle over the orbit space undergo a linear transformation so that the components in the Chaplygin coframe will be

$$w_\alpha u^\alpha(x) = (p_\alpha - \mu_a A^a_\alpha(x)) u^\alpha(x), \quad \text{or equivalently,} \quad p_\alpha = w_\alpha + \mu_a A^a_\alpha(x).$$

Notice that because of our choice of Chaplygin frame, the constraint condition on a vector $(\dot{x}, \dot{s})$ (namely $\dot{s} = -A(x)\dot{x}$) is identical to the condition that the ξ^a components of the vector $(\dot{x}, \dot{s})$ with respect to the Chaplygin frame are zero. In fact, the vector $\xi^a u_a(x) = \dot{s}^a \frac{\partial}{\partial s^a} + A^a_\alpha(x)\dot{x}^\alpha \frac{\partial}{\partial s^a}$ corresponds exactly to the term Ω defined in Kobilarov, Marsden, and Sukhatme [27] as the ***fixed angular velocity***, and $\xi^{\mathcal{U}}$ in the development in Chapter 3. This makes implementation of our constraint almost trivial:

$$\xi^a = 0 \qquad \Longleftrightarrow \qquad \dot{s}^a = -A^a_\alpha(x)\dot{x}^\alpha.$$

In particular, it is easy to see that the constrained Lagrangian ℓ^c is representable by first examining the Lagrangian ℓ written in terms of our Chaplygin frame:

$$L(x^\alpha, v^\alpha, \eta^a) = L(x^\alpha, v^\alpha, \zeta^a - A^a_\alpha(x) v^\alpha) = \ell(x^\alpha, v^\alpha, \zeta^a)$$

and when we impose the constraint, $\zeta^a = 0$ allows us to write $\ell(x^\alpha, v^\alpha, 0) = \ell^c(x^\alpha, v^\alpha)$. In fact, this is equivalent to the approach of Kobilarov et al. dealing with systems with nonholonomic constraints and symmetries: $\zeta^a u_a(x)$ corresponds to their fixed angular velocity term Ω. Since we don't wish to impose the constraint until after we carry through the Hamilton–Pontryagin principle, our action function on the Pontryagin bundle may be written in terms of the Chaplygin frame (including the ζ^a components) as

$$\begin{aligned}
\mathcal{S}_{HP}(x, v, p, \eta, \mu) &= \int_{t_0}^{t_F} [\ell(x^\alpha, v^\alpha, \zeta^a) + \langle p, \dot{x} - v\rangle + \langle \mu, \dot{s} - \eta\rangle]\, dt \\
&= \int_{t_0}^{t_F} [\ell(x^\alpha, v^\alpha, \zeta^a) + (p_\alpha - \mu_a A^a_\alpha(x))(\dot{x}^\alpha - v^\alpha) + \mu_a(\dot{s}^a + A^a_\alpha(x)\dot{x}^\alpha - \zeta^a)]\, dt \\
&= \int_{t_0}^{t_F} [\ell(x^\alpha, v^\alpha, \zeta^a) + p_\alpha(\dot{x}^\alpha - v^\alpha) + \mu_a(\dot{s}^a - \zeta^a) + \mu_a A^a_\alpha(x) v^\alpha]\, dt.
\end{aligned}$$

The Hamilton–Pontryagin principle results in the dynamical condition on the system

$$\int_{t_0}^{t_F}\Bigg[\left(\frac{\partial \ell}{\partial x^\alpha}-\dot{p}_\alpha+\dot{\mu}_a A^a_\alpha(x)-[\mu,\dot{s}]^*_a A^a_\alpha(x)+\mu_a A^a_{\beta,\alpha}(x)v^\beta\right)\delta x^\alpha$$
$$+\left(\frac{\partial \ell}{\partial v^\alpha}-p_\alpha+\mu_a A^a_\alpha(x)\right)\delta v^\alpha+\left(\frac{\partial \ell}{\partial \zeta^a}-\mu_a\right)\delta\zeta^a$$
$$+\delta p_\alpha\left(\dot{x}^\alpha-v^\alpha\right)+\delta\mu_a\left(\dot{s}^a+A^a_\alpha(x)v^\alpha-\zeta^a\right)\Bigg]dt=0.$$

Again, by linear independence of our virtual displacement vectors, we can start to identify our arbitrary paths through the tangent and cotangent bundles with the dynamical velocity and momentum of our system, so that

$$v^\alpha=\dot{x}^\alpha,\quad \zeta^a=\dot{s}^a+A^a_\alpha(x)\dot{x}^\alpha,\quad \mu_a=\frac{\partial \ell}{\partial \zeta^a},\quad \text{and}\quad p_\alpha=\frac{\partial \ell}{\partial \dot{x}^\alpha}+\mu_a A^a_\alpha(x).$$

Thus, upon substitution of these terms (including the time-derivative of p_a), and imposition of the constraint, linear independence of the δx^α displacement components again gives us the differential equations (4.10) corresponding to the Chaplygin setting. Recall that, by commutativity of H, the dual bracket term disappears:

$$\frac{d}{dt}\frac{\partial \ell^c}{\partial \dot{x}^\alpha}-\frac{\partial \ell^c}{\partial x^\alpha}=\frac{\partial \ell}{\partial \zeta^a}\dot{x}^\beta\left(A^a_{\beta,\alpha}-A^a_{\alpha,\beta}\right).$$

In this case the differential equations are given in components with respect to the vectors $u^\alpha(x)$ of the Chaplygin coframe. Of course, recall that we have established that these $(n-p)$ covectors are identical to the coordinate induced coframe, that is $u^\alpha(x)=dx^\alpha$. In other words, the Chaplygin equations are just the $(n-p)$ Hamel's equations associated with horizontal variations.

4.3.4 The Discrete Chaplygin Equations

Now that we have defined the Chaplygin frame and coframe as (locally) spanning the Pontryagin bundle built on the principal bundle Q, it will be relatively straightforward to show how to use our discretization technique (4.7) to find discrete difference equations that approximate the dynamics of our system. In particular let us describe our interpolation paramaterizations in the Orbit space Q/H and the Lie group H as

$$\tau x_{k+1}+(1-\tau)x_k=x_{k+\tau}\quad\text{and}\quad \tau s_{k+1}+(1-\tau)s_k=s_{k+\tau}.$$

Then we may identify the discrete dynamic velocities scaled by time-step h[1] for parameter value $\tau \in [0,1]$ as

$$\Delta x_{k,k+1} = \frac{1}{h}\frac{d}{d\tau}\left(x_{k+\tau}\right) = \frac{x_{k+1} - x_k}{h} \in T_{x_{k+\tau}}(Q/H),$$
$$\Delta s_{k,k+1} = \frac{1}{h}\frac{d}{d\tau}\left(s_{k+\tau}\right) = \frac{s_{k+1} - s_k}{h} \in \mathfrak{h}.$$

We then may define the discrete action of our Chaplygin system (in coordinates with respect to the Chaplygin frame and coframe) as

$$\begin{aligned} \mathcal{S}^d_{HP} = \sum_{k=0}^{N-1} & h\ell(x^\alpha_{k+\tau}, v^\alpha_{k,k+1}, \zeta^a_{k,k+1}) + p_{k,k+1;\alpha}(\Delta x^\alpha_{k,k+1} - v^\alpha_{k,k+1}) \\ & + \mu_{k,k+1;a}(\Delta s^a_{k,k+1} + A^a_\alpha(x_{k+\tau})\Delta x^\alpha_{k,k+1} - \zeta^a_{k,k+1}), \end{aligned} \tag{4.13}$$

where again, ζ^a are the components of the section (v,η) through the tangent bundle corresponding to the $u_a(x)$ terms of the Hamel frame. Now the discrete variational principle on the given action results in the condition

$$\begin{aligned} \sum_{k=0}^{N-1} & h\frac{\partial \ell_{k+\tau}}{\partial x^\alpha}\delta x^\alpha_{k+\tau} + h\frac{\partial \ell_{k+\tau}}{\partial v^\alpha}\delta v^\alpha_{k,k+1} + h\frac{\partial \ell_{k+\tau}}{\partial \zeta^a}\delta \zeta^a_{k,k+1} - p_{k,k+1;\alpha}\delta v^\alpha_{k,k+1} \\ & + \mu_{k,k+1;a}A^a_\alpha(x_{k+\tau})\delta\Delta x^\alpha_{k,k+1} - \mu_{k,k+1;a}\delta\zeta^a_{k,k+1} + \delta p_{k,k+1;\alpha}(\Delta x^\alpha_{k,k+1} - v^\alpha_{k,k+1}) \\ & + \delta\mu_{k,k+1;a}(\Delta s^a_{k,k+1} + A^a_\alpha(x_{k+\tau})\Delta x^\alpha_{k,k+1} - \zeta^a_{k,k+1}) + \mu_{k,k+1;a}A^a_{\beta,\alpha}(x_{k+\tau})\Delta x^\beta_{k,k+1}\delta x^\alpha_{k+\tau} \\ & + p_{k,k+1;\alpha}\frac{1}{h}\frac{d}{d\tau}(\delta x^\alpha_{k+\tau}) + \mu_{k,k+1;a}\frac{1}{h}\frac{d}{d\tau}(\delta s_{k+\tau})^a = 0. \end{aligned} \tag{4.14}$$

We will go through the process of grouping like terms and using the linear independence of our discrete virtual displacement vectors and covectors in the above condition in order to simplify the equation to a difference relation involving only the δx virtual displacements. As such, we will need to identify the virtual displacement of the quadrature in the symmetry group with the displacements in our orbit space according to the constraint given by the connection $\mathcal{A}(x)$. That is, we will say that

$$\delta s_{k+\tau} = -A(x_{k+\tau})\delta x_{k+\tau}. \tag{4.15}$$

Returning to our discretization (4.14), it is straightforward to group terms by their linearly independent virtual displacements to arrive at the identifications of the discrete velocities and

[1] Recall that in Chapter 3 we defined $\Delta x_k = \frac{d}{d\tau}(x_{k+\tau})$. In this chapter, for notational convenience we include the time step h in our definition of discrete velocity. In this way, $\Delta x_{k,k+1}$ is understood as a discrete analogue of dynamical velocity.

momenta as

$$
\begin{aligned}
v^{\alpha}_{k,k+1} &= \Delta x^{\alpha}_{k,k+1} = \frac{1}{h}(x^{\alpha}_{k+1} - x^{\alpha}_{k}), \\
\zeta^{a}_{k,k+1} &= \Delta s^{a}_{k,k+1} + A^{a}_{\alpha}(x_{k+\tau})\Delta x^{\alpha}_{k,k+1}, \\
\mu_{k,k+1;a} &= h\frac{\partial \ell_{k+\tau}}{\partial \zeta^{a}}, \qquad \text{and} \\
p_{k,k+1;\alpha} &= h\frac{\partial \ell_{k+\tau}}{\partial \dot{x}^{\alpha}}.
\end{aligned}
$$

Thus, we see that (4.14) reduces to

$$
\begin{aligned}
&\sum_{k=0}^{N-1} h\frac{\partial \ell_{k+\tau}}{\partial x^{\alpha}}\delta x^{\alpha}_{k+\tau} + \mu_{k,k+1;a}A^{a}_{\beta,\alpha}(x_{k+\tau})\Delta x^{\beta}_{k,k+1}\delta x^{\alpha}_{k+\tau} \\
&\quad + \frac{1}{h}\left(h\frac{\partial \ell_{k+\tau}}{\partial \dot{x}^{\alpha}} + \mu_{k,k+1;a}A^{a}_{\alpha}(x_{k+\tau})\right)\left(\delta x^{\alpha}_{k+1} - \delta x^{\alpha}_{k}\right) \\
&\quad + \frac{1}{h}\mu_{k,k+1;a}(\delta s^{a}_{k+1} - \delta s^{a}_{k}) = 0.
\end{aligned}
\tag{4.16}
$$

We rewrite the third line of (4.16) in terms of the constraint on the virtual displacements (4.15) as

$$
\left\langle \mu_{k,k+1}, \frac{1}{h}\frac{d}{d\tau}(\delta s_{k+\tau})\right\rangle = -\left\langle \mu_{k,k+1}, \frac{1}{h}\frac{d}{d\tau}\left(A(x_{k+\tau})\delta x_{k+\tau}\right)\right\rangle.
$$

Thus, carrying through with the parameter τ-derivative, and inserting our constrained Lagrangian function ℓ^{c}, our whole condition becomes (with respect to the Chaplygin frame at $u^{\alpha}(x_{k+\tau})$)

$$
\begin{aligned}
&\sum_{k=0}^{N-1} \frac{\partial \ell^{c}_{k+\tau}}{\partial x^{\alpha}}\delta x^{\alpha}_{k+\tau} + \mu_{k,k+1;a}A^{a}_{\beta,\alpha}(x_{k+\tau})\Delta x^{\beta}_{k,k+1}\delta x^{\alpha}_{k+\tau} \\
&\quad + \frac{1}{h}\left(\frac{\partial \ell^{c}_{k+\tau}}{\partial \dot{x}^{\alpha}} + \mu_{k,k+1;a}A^{a}_{\alpha}(x_{k+\tau})\right)\left(\delta x^{\alpha}_{k+1} - \delta x^{\alpha}_{k}\right) \\
&\quad - \mu_{k,k+1;a}A^{a}_{\alpha,\beta}(x_{k+\tau})\Delta x^{\beta}_{k,k+1}\delta x^{\alpha}_{k+\tau} - \frac{1}{h}\mu_{k,k+1;a}A^{a}_{\alpha}(x_{k+\tau})\left(\delta x^{\alpha}_{k+1} - \delta x^{\alpha}_{k}\right) = 0.
\end{aligned}
\tag{4.17}
$$

Then, the discrete dynamic condition describing our Chaplygin system reads

$$
\begin{aligned}
&\sum_{k=0}^{N-1}\Bigg(\frac{\partial \ell^{c}_{k+\tau}}{\partial x^{\alpha}} + \mu_{k,k+1;a}\Delta x^{\beta}_{k,k+1}\left(A^{a}_{\beta,\alpha}(x_{k+\tau}) - A^{a}_{\alpha,\beta}(x_{k+\tau})\right) \\
&\qquad + \left(\frac{1}{h}\frac{\partial \ell^{c}_{k+\tau}}{\partial \dot{x}^{\alpha}}\left(\delta x^{\alpha}_{k+1} - \delta x^{\alpha}_{k}\right)\right)\delta \Delta x^{\alpha}_{k,k+1} = 0,
\end{aligned}
$$

wherein we include the discrete constraint on velocities $\Delta s_{k,k+1} = -A(x_{k+\tau})\Delta x_{k,k+1}$ or equivalently $\zeta_{k,k+1} = 0$.

Finally, we note again that our discrete virtual displacements in the above equation take the form $\delta x_{k+\tau} = \tau\delta x_{k+1} + (1-\tau)\delta x_k$. If we change the indices of summation to account for the fact that variations at the endpoints $\delta x_0 = \delta x_N = 0$, we arrive at the discrete Chaplygin equations

$$\begin{aligned}\frac{1}{h}\left(\frac{\partial \ell^c_{k+\tau}}{\partial \dot{x}^\alpha} - \frac{\partial \ell^c_{k+\tau-1}}{\partial \dot{x}^\alpha}\right) - (1-\tau)\frac{\partial \ell^c_{k+\tau}}{\partial x^\alpha} - \tau\frac{\partial \ell^c_{k+\tau-1}}{\partial x^\alpha}& \\ = (1-\tau)\mu_{k,k+1;a}\Delta x^\beta_{k,k+1}\left(A^a_{\beta,\alpha}(x_{k+\tau}) - A^a_{\alpha,\beta}(x_{k+\tau})\right)& \\ + \tau\mu_{k-1,k;a}\Delta x^\beta_{k-1,k}\left(A^a_{\beta,\alpha}(x_{k+\tau-1}) - A^a_{\alpha,\beta}(x_{k+\tau-1})\right).&\end{aligned} \tag{4.18}$$

These equations are identical to the discrete Chaplygin equations derived in Chapter 3. The development presented in the current chapter is intended to emphasize the usefulness of both the discrete Hamilton–Pontryagin principle and Hamel's formalism to the incorporation of nonholonomic constraints. We have assumed hyperregularity of the Lagrangian so that the "discrete" Legendre transforms defined above are consistent; the discrete Hamilton–Pontryagin additionally provides a discrete perspective on a unification of Lagrangian and Hamiltonian mechanics.

Again, notice that when we write the Lagrangian with respect to the Chaplygin frame as described above, imposing the constraint becomes as easy as setting the components $\zeta^a = 0$ for $a = 1, \ldots, p$. The Lagrangian ℓ also naturally incorporates the constraint; there is no need to refer to formulas such as those in equations (4.9) to transform partial derivatives. The discrete Chaplygin equations are the $(n-p)$ discrete Hamel's equations associated with horizontal variation.

Chapter 5

Discrete Hamel's Equations as Structure Preserving Integrators

5.1 A Road Map to Discrete Structure Preservation

In 2.3.2 we showed that the update map of associated with the discrete Euler–Lagrange equations conserves a discrete symplectic two-form. This was observed (in a procedure laid out by Marsden and West [36]) to be a consequence of the discrete Lagrangian one-forms being obtained via differentiation of the discrete Lagrangian function:

$$\begin{aligned} dL_d(q_k, q_{k+1}) &= D_1 L_d(q_k, q_{k+1}) dq_k + D_2 L_d(q_k, q_{k+1}) dq_{k+1} \\ &= -\Theta^-_{L_d}(q_k, q_{k+1}) + \Theta^+_{L_d}(q_k, q_{k+1}). \end{aligned}$$

Using the property of the external derivative ($d(dL_d) = 0$),

$$\begin{aligned} 0 = -d\Theta^-_{L_d}(q_k, q_{k+1}) + d\Theta^+_{L_d}(q_k, q_{k+1}) \quad &\Rightarrow \\ d\Theta^-_{L_d}(q_k, q_{k+1}) = d\Theta^+_{L_d}(q_k, q_{k+1}) = \Omega_{L_d}(q_k, q_{k+1}).& \end{aligned}$$

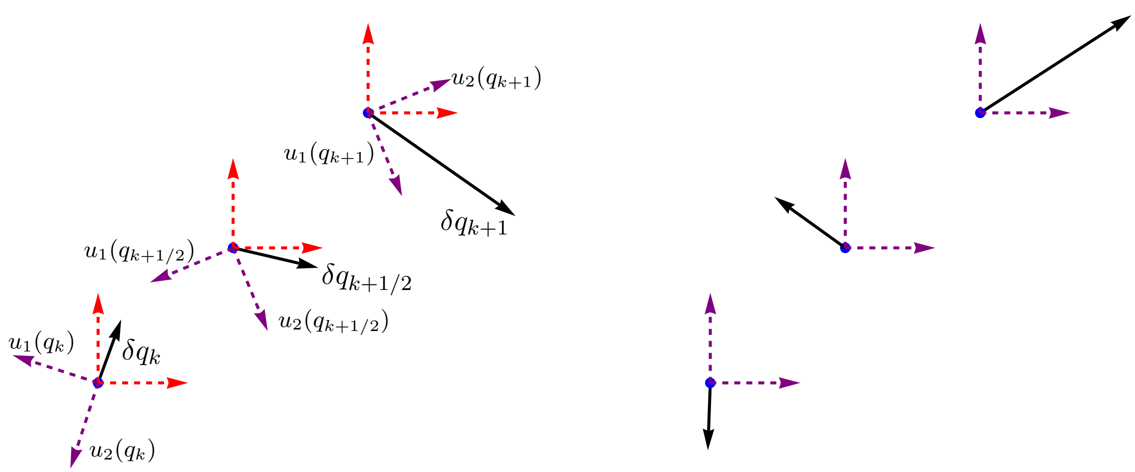
Thus, at each step specified by $(q_k, q_{k+1}) \in Q \times Q$, while there are two distinct Lagrangian one-forms, there is guaranteed to be only one symplectic two-form $\Omega_{L_d}(q_k, q_{k+1})$. The form itself is clearly conserved under the transformation F_{L_d}: points $(q_{k-1}, q_k), (q_k, q_{k+1}) \in Q \times Q$ that satisfy $F_{L_d}(q_{k-1}, q_k) = (q_k, q_{k+1})$ by definition are solutions of the discrete Euler–Lagrange equations (2.23) and hence

$$\begin{aligned} \Theta^-_{L_d}(q_k, q_{k+1}) = \Theta^+_{L_d}(q_{k-1}, q_k) \quad &\Rightarrow \\ d\Theta^-_{L_d}(q_k, q_{k+1}) = d\Theta^+_{L_d}(q_{k-1}, q_k) \quad &\Rightarrow \\ \Omega_{L_d}(q_k, q_{k+1}) = \Omega_{L_d}(q_{k-1}, q_k).& \end{aligned}$$

We have recalled this procedure to provide context to a natural question about the discrete Hamel's equations obtained in (3.6): does the discrete update map F_{ℓ_d} induced by these integrators similarly conserve a modified discrete symplectic two-form? The earlier derivation in Section 3.2 does not lend itself to the above procedure; our variational principle in that case was based on a flexible interpretation of variations themselves (see Figures 5.1, 5.2, and 5.3 for a geometric illustration). Specifically, we defined variations in terms of vectors $\zeta_k \in T_{q_k}Q$ for $k = 1, \ldots, N-1$ so that

$$\begin{aligned}
\delta s^d(q_d) &= \sum_{k=0}^{N-1} d\ell^d(q_{k+\tau}, \xi_{k,k+1})(\delta q_{k+\tau}, \delta\xi_{k,k+1}) \\
&= \sum_{k=0}^{N-1} d\ell^d(q_{k+\tau}, \xi_{k,k+1}) \left(u(q_{k+\tau}) \cdot \zeta_{k+\tau}, \tfrac{1}{h}(\zeta_{k+1} - \zeta_k) + [\xi_{k,k+1}, \zeta_{k+\tau}]_{k+\tau} \right),
\end{aligned}$$

where $\zeta_{k+\tau} = (1-\tau)\zeta_k + \tau\zeta_{k+1}$ and the transition from $(\delta q_{k+\tau}, \delta\xi_{k,k+1})$ to the vectors (ζ_k, ζ_{k+1}) is defined by equations (3.7) and (3.8).



(a) Virtual displacement vectors.

(b) Vectors measured against new frame.

Figure 5.1: In Figures 5.1a and 5.1b, the virtual displacement vectors δq_k and δq_{k+1}, and the vector $\delta q_{k+1/2}$ induced by the pushforward of the discretization mapping $Q \times Q \to TQ$ are transformed into the frame u_i. The underlying configuration space is assumed to be $Q = \mathbb{R}^2$.

(a) Vector $\zeta_{k+1/2}$ in the frame.

(b) Transforming to the coordinate basis.

Figure 5.2: In Figure 5.2a, $\zeta_{k+1/2}$ is depicted in components as the linear average of $\phi(q_k)\delta q_k$ and $\phi(q_{k+1})\delta q_{k+1}$. Furthermore, notice that in Figure 5.2b, $\zeta_{k+1/2} \neq \phi(q_{k+1/2})\delta q_{k+1/2}$.

While this principle has been shown to be equivalent to the desired discrete Hamel's equations (3.6), it is not immediately clear how to implement the above procedure of Marsden and West to show symplecticity. Specifically, it is not immediately apparent that the exterior derivative of an underlying function on $Q \times Q$ gives rise to two distinct one-forms. Nonetheless, examining (3.6), one may *expect* that the ***discrete Hamel's one-forms*** $\Theta^-_{\ell_d}$ and $\Theta^+_{\ell_d}$ over $Q \times Q$ should be

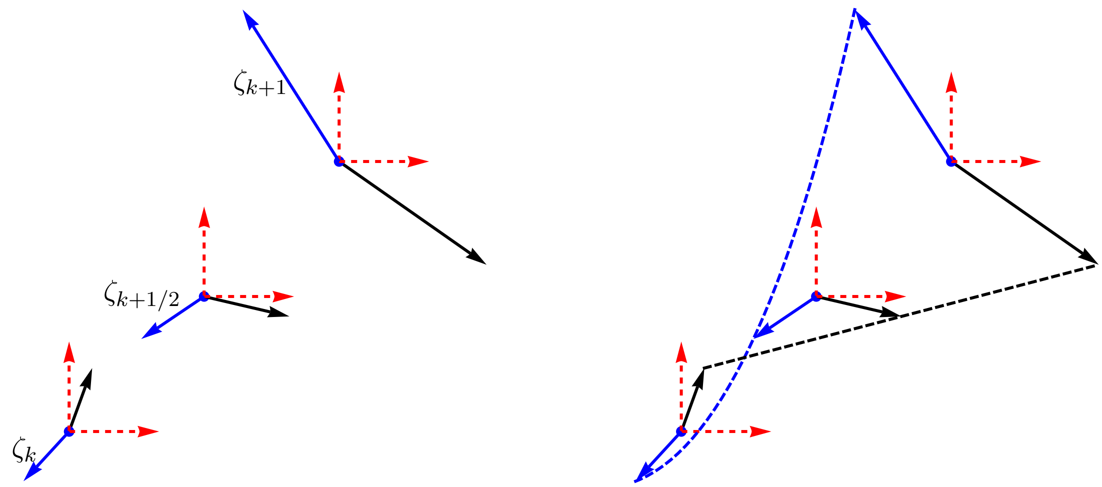
$$\begin{aligned} -\Theta^-_{\ell_d}(q_k, q_{k+1}) &= \left(-\frac{1}{h}\mu_{k,k+1} + (1-\tau)u[\ell^d]_{k+\tau} + (1-\tau)[\xi_{k,k+1}, \mu_{k,k+1}]^*_{q_{k+\tau}}\right) \cdot (u(q_k), 0) \\ \Theta^+_{\ell_d}(q_k, q_{k+1}) &= \left(\frac{1}{h}\mu_{k,k+1} + \tau u[\ell^d]_{k+\tau} + 1 - \tau[\xi_{k,k+1}, \mu_{k,k+1}]^*_{q_{k+\tau}}\right) \cdot (0, u(q_{k+1})) \end{aligned} \tag{5.1}$$

as it is then clear that the discrete Hamel's equations will be equivalent to the condition $\Theta^-_{\ell_d}(q_k, q_{k+1})_j = \Theta^+_{\ell_d}(q_{k-1}, q_k)_j$.

It is straightforward to verify that these one-forms $\Theta^\pm_{\ell_d}$ are not simply the discrete Lagrangian one-forms $\Theta^\pm_{L_d}$ written in a different frame; they are instead geometrically distinct one-forms over $Q \times Q$, i.e.

$$\Theta^\pm_{L_d}(q_k, q_{k+1}) \neq \Theta^\pm_{\ell_d}(q_k, q_{k+1}) \cdot (u(q_k), u(q_{k+1})).$$

Likewise, the one-forms $\Theta^\pm_{\ell_d}$ are not expressible in terms of exterior derivative of ℓ^d or any other



(a) Vectors $\zeta_k, \zeta_{k+1/2}, \zeta_{k+1}$ in the coordinate frame.

(b) Variation of the step at (q_k, q_{k+1}).

Figure 5.3: The vectors $\zeta_k, \zeta_{k+1/2}, \zeta_{k+1}$ defined in Figure 5.2 are now written with respect to the coordinate basis. In Figure 5.3b, we see an interpretation of the two displaced steps. Vectors $\zeta_k, \zeta_{k+1/2}, \zeta_{k+1}$ correspond to the interpretation of variations presented in Chapter 3. In the current chapter we will interpret variations in a manner consistent with the vectors $\delta q_k, \delta q_{k+1/2}, \delta q_{k+1}$ and compensate with the diffeomorphism Φ defined below. The problem can be summarized up in the inequality:

$$\phi_i^j(q_{k+1/2})(\delta q_k^i + \delta q_{k+1}^i)/2 \neq \phi_i^j(q_k)\delta q_k^i/2 + \phi_i^j(q_{k+1})\delta q_{k+1}^i/2.$$

discretization of the Lagrangian presented so far. This is just as well: if the converse were the case then we might come to the conclusion that the discrete Hamel's equations are algebraically equivalent to the discrete Euler–Lagrange equations and would therefore not present any new or useful insight (beyond some algebraic manipulation) into the mechanical systems we seek to simulate.

Instead, the discrete Hamel's equations *are distinct* from the discrete Euler–Lagrange equations. However, if we seek to prove that the resulting update maps conserve a modified discrete symplectic two-form over $Q \times Q$ following Marsden and West [36], we must first show that there exists a hypothetical function $\mathcal{L}_d : Q \times Q \to \mathbb{R}$ such that $d\mathcal{L}_d = -\Theta^-_{\ell_d} + \Theta^+_{\ell_d}$.In the next section we will make use of the Poincaré lemma to prove the existence of such a function in the case where Q is a vector space.

5.2 Structure Preservation

As we have throughout this dissertation, we define a Lagrangian mechanical system and Hamel's equations as in 2.2.2, and a discrete Lagrangian function $L_d : Q \times Q \to \mathbb{R}$ as in equation (2.21). Furthermore, recall that we restrict ourselves to the case where Q is a vector space. Again, in order to prove that the discrete update map generated by solutions of the discrete Hamel's equations conserves a symplectic two-form over $Q \times Q$, we seek a function $\mathcal{L}_d : Q \times Q \to \mathbb{R}$ such that

$$d\mathcal{L}_d = -\Theta^-_{\ell_d} + \Theta^+_{\ell_d}, \tag{5.2}$$

where $\Theta^\pm_{\ell_d}$ are defined in (5.1) as the discrete Hamel's one-forms.

Recall that we have defined $(q_{k+\tau}, v_{k,k+1}) \in TQ$ in Chapter 3 in terms of the discretization mapping $Q \times Q \to TQ$ as

$$(q_{k+\tau}, v_{k,k+1}) = \left((1-\tau)q_k + \tau q_{k+1}, \tfrac{1}{h}\tfrac{d}{ds}\Big|_{s=\tau} ((1-s)q_k + sq_{k+1})\right), \tag{5.3}$$

and furthermore we use the notation $\xi_{k,k+1}$ to refer to the components of the vector $v_{k,k+1} \in T_{q_{k+\tau}}Q$ measured against the frame, i.e. $\xi_{k,k+1} \cdot u(q_{k+\tau}) = v_{k,k+1}$, or in index notation:

$$\xi^j_{k,k+1} = \tfrac{1}{h}\tfrac{d}{ds}\Big|_{s=\tau} \left((1-s)q^i_k + sq^i_{k+1}\right) \phi^j_i(q_{k+\tau}).$$

5.2.1 A Modified Discrete Variational Principle

We start with the following observation: we may define a diffeomorphism $\Psi : T(Q\times Q) \to T(TQ)$ so that $hdL \circ \Psi_{(q_k,q_{k+1})} = -\Theta^-_{\ell_d}(q_k, q_{k+1}) + \Theta^+_{\ell_d}(q_k, q_{k+1}) \in T^*(Q \times Q)$.

Theorem 5.1. *Define the mapping $\Psi : T(Q \times Q) \to T(TQ)$ (in index notation) to be*

$$\begin{aligned}
&\Psi((q_k, q_{k+1}), (\delta q_k, \delta q_{k+1})) \\
&\quad = \Bigg((q_{k+\tau}, v_{k,k+1}), \\
&\qquad \bigg(\left((1-\tau)\delta q^j_k \phi^m_j(q_k) + \tau\delta q^j_{k+1}\phi^m_j(q_{k+1})\right) \psi^i_m(q_{k+\tau}) \frac{\partial}{\partial q^i}\Big|_{q_{k+\tau}}, \\
&\qquad \frac{1}{h}\frac{d}{ds}\Big|_{s=\tau} \left(\left((1-s)\delta q^j_k \phi^m_j(q_k) + s\delta q^j_{k+1}\phi^m_j(q_{k+1})\right) \psi^i_m(q_{k+s})\right) \frac{\partial}{\partial v^i}\Big|_{v_{k,k+1}} \bigg)\Bigg)
\end{aligned} \tag{5.4}$$

for arbitrary $((q_k, q_{k+1}), (\delta q_k, \delta q_{k+1})) \in T(Q \times Q)$. Furthermore, recall that we have defined ψ (and its inverse ϕ) in 2.2.2 to represent a frame field comprised of C^1 sections of FQ. Then:

1. *Ψ is a C^1-diffeomorphism when $\tau = 1/2$, and*

2. $hdL \circ \Psi = -\Theta^-_{\ell_d} + \Theta^+_{\ell_d}$.

Proof. First, we note that Ψ acting on the spaces $Q \times Q \to TQ$ is nothing but the discretization mapping (5.3), which is differentiable and an isomorphism when Q is a vector space.

Next, we carry out the parameter derivative in (5.4), and note that $\Psi_{(q_k,q_{k+1})}$ acts on vectors $(\delta q_k, \delta q_{k+1})$ linearly, so that in index notation:

$$\begin{aligned}
&\Psi_{(q_k,q_{k+1})}(\delta q^i_k, \delta q^i_{k+1}) = \\
&\qquad \Bigg(\left((1-\tau)\psi^j_m(q_{k+\tau})\phi^m_i(q_k)\right)\delta q^i_k + \left(\tau\psi^j_m(q_{k+\tau})\phi^m_i(q_{k+1})\right)\delta q^i_{k+1}, \\
&\qquad\quad \left(-\frac{1}{h}\psi^j_m(q_{k+\tau})\phi^m_i(q_k) + (1-\tau)\psi^j_{m,a}(q_{k+\tau})v^a_{k,k+1}\phi^m_i(q_k)\right)\delta q^i_k \\
&\qquad + \left(\frac{1}{h}\psi^j_m(q_{k+\tau})\phi^m_i(q_{k+1}) + \tau\psi^j_{m,a}(q_{k+\tau})v^a_{k,k+1}\phi^m_i(q_{k+1})\right)\delta q^i_{k+1}\Bigg).
\end{aligned}$$

Thus, as the mapping $\Psi_{(q_k,q_{k+1})}$ is a linear mapping between vector spaces $T_{(q_k,q_{k+1})}(Q \times Q)$ and $T_{(q_{k+\tau},v_{k,k+1})}(TQ)$, the operation is expressible in matrix notation as

$$\left[\Psi_{(q_k,q_{k+1})}\right]\begin{bmatrix}\delta q_k \\ \delta q_{k+1}\end{bmatrix} \in T_{(q_{k+\tau},v_{k,k+1})}(TQ),$$

where the matrix $[\Psi_{(q_k,q_{k+1})}]$ is written in block form as

$$\begin{aligned}
&\left[\Psi_{(q_k,q_{k+1})}\right] = \\
&\begin{bmatrix}
(1-\tau)\psi(q_{k+\tau})\phi(q_k) & \tau\psi(q_{k+\tau})\phi(q_{k+1}) \\
\left(-\frac{1}{h}\psi(q_{k+\tau}) + (1-\tau)\psi_{,a}(q_{k+\tau})v^a_{k,k+1}\right)\phi(q_k) & \left(\frac{1}{h}\psi(q_{k+\tau}) + \tau\psi_{,a}(q_{k+\tau})v^a_{k,k+1}\right)\phi(q_{k+1})
\end{bmatrix}.
\end{aligned}$$

A block matrix $\left(\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right)$ is invertible if A and its Schur complement $D - CA^{-1}B$ are invertible. Clearly, "A" in the above matrix expression of Ψ is nonsingular, since ψ and ϕ are necessarily in $GL(n)$ for all $q \in Q$. Furthermore, $CA^{-1}B$ is

$$\begin{aligned}
CA^{-1}B &= \left(-\tfrac{1}{h}\psi(q_{k+\tau}) + (1-\tau)\psi_{,a}(q_{k+\tau})v^a_{k,k+1}\right)\phi(q_k)\cdot\left(\frac{\psi(q_k)\phi(q_{k+\tau})}{1-\tau}\right)\cdot\tau\psi(q_{k+\tau})\phi(q_{k+1}) \\
&= \frac{\tau}{1-\tau}\left(-\tfrac{1}{h}\psi(q_{k+\tau}) + (1-\tau)\psi_{,a}(q_{k+\tau})v^a_{k,k+1}\right)\phi(q_{k+1}),
\end{aligned}$$

and hence the Schur complement of A is the matrix

$$\begin{aligned}
D - CA^{-1}B &= \left(\frac{1}{h}\left(1 + \frac{\tau}{1-\tau}\right)\psi(q_{k+\tau}) + (\tau - (1-\tau))\,\psi_{,a}(q_{k+\tau})v^a_{k,k+1}\right)\phi(q_{k+1}) \\
&= \left(\frac{1}{h}\left(\frac{1}{1-\tau}\right)\psi(q_{k+\tau}) + (2\tau - 1)\,\psi_{,a}(q_{k+\tau})v^a_{k,k+1}\right)\phi(q_{k+1}).
\end{aligned}$$

If we select $\tau = 1/2$ (the choice of parameter that leads to the midpoint approximation), the Schur complement reduces to

$$D - CA^{-1}B = \frac{2}{h}\psi(q_{k+1/2})\phi(q_{k+1})$$

and is clearly nonsingular. Therefore, the matrix representation of $\Psi_{(q_k, q_{k+1})}$ is invertible, hence the mapping is a (at least C^1) diffeomorphism. Thus, we see that $\Psi : T(Q \times Q) \to T(TQ)$ is a C^1-diffeomorphism as long as $\tau = 1/2$.

The second result of the theorem is proved by direct computation. The exterior derivative of the Lagrangian function $L : TQ \to \mathbb{R}$ returns a one-form on TQ, i.e. $dL \in T^*(TQ)$, or $dL : T(TQ) \to \mathbb{R}$. This one-form is expressible in coordinates as

$$dL(q,v) = \frac{\partial L(q,v)}{\partial q^i}dq^i + \frac{\partial L(q,v)}{\partial v^i}dv^i,$$

and if we recall that in Hamel's formalism we have expressed the Lagrangian as $\ell(q^i, \xi^j) = \ell(q^i, \dot{q}^i\phi^j_i(q)) = L(q^i, \dot{q}^i)$, we see that the partial derivatives constituting the components of dL above may be written as

$$\begin{aligned}
\frac{\partial L(q,\dot{q})}{\partial q^i} &= \frac{\partial \ell(q,\xi)}{\partial q^i} + \frac{\partial \ell(q,\xi)}{\partial \xi^a}\xi^b\psi^j_b(q)\phi^a_{j,i}(q) \text{ and} \\
\frac{\partial L(q,\dot{q})}{\partial \dot{q}^i} &= \frac{\partial \ell(q,\xi)}{\partial \xi^j}\phi^j_i(q).
\end{aligned}$$

Therefore, dL evaluated at the point $(q_{k+\tau}, v_{k,k+1}) \in TQ$ is also expressible as

$$dL_{(q_{k+\tau}, v_{k,k+1})} = \left(\frac{\partial \ell_{k+\tau}}{\partial q^i} + \frac{\partial \ell_{k+\tau}}{\partial \xi^a}\xi^b\psi^j_b(q_{k+\tau})\phi^a_{j,i}(q_{k+\tau})\right)dq^i + \left(\frac{\partial \ell_{k+\tau}}{\partial \xi^j}\phi^j_i(q_{k+\tau})\right)dv^i$$

where we recall that we have defined the shorthand notation $\ell_{k+\tau} = \ell(q_{k+\tau}, \xi_{k,k+1})$. Then, applying equation (5.4), the one-form $dL \circ \Psi \in T^*(Q \times Q)$ acts on vectors $(\delta q_k, \delta q_{k+1}) \in$

$T_{(q_k,q_{k+1})}(Q \times Q)$ so that

$$
\begin{aligned}
&dL \circ \Psi_{(q_k,q_{k+1})}(\delta q_k, \delta q_{k+1}) = \\
&\left(\frac{\partial \ell_{k+\tau}}{\partial q^i} + \frac{\partial \ell_{k+\tau}}{\partial \xi^m} \xi^r \psi^a_r(q_{k+\tau}) \phi^m_{a,i}(q_{k+\tau}) \right) \psi^i_b(q_{k+\tau}) \left((1-\tau)\delta q^j_k \phi^b_j(q_k) + \tau \delta q^j_{k+1} \phi^b_j(q_{k+1}) \right) \\
&+ \left(\frac{\partial \ell_{k+\tau}}{\partial \xi^m} \phi^m_i(q_{k+\tau}) \right) \left(-\frac{1}{h} \psi^i_b(q_{k+\tau}) + (1-\tau) \psi^i_{b,a}(q_{k+\tau}) \xi^r_{k,k+1} \psi^a_r(q_{k+\tau}) \right) \phi^b_j(q_k) \delta q^j_k \\
&+ \left(\frac{\partial \ell_{k+\tau}}{\partial \xi^m} \phi^m_i(q_{k+\tau}) \right) \left(\frac{1}{h} \psi^i_b(q_{k+\tau}) + \tau \psi^i_{b,a}(q_{k+\tau}) \xi^r_{k,k+1} \psi^a_r(q_{k+\tau}) \right) \phi^b_j(q_{k+1}) \delta q^j_{k+1}.
\end{aligned}
$$

Considering that $\partial\psi \cdot \phi = -\psi \cdot \partial\phi$ because $\phi = \psi^{-1}$, and remembering that the structure functions are defined in terms of ψ and ϕ as

$$
c^m_{ij}(q) = \psi^a_i(q)\psi^b_j(q)(\phi^m_{a,b}(q) - \phi^m_{b,a}(q)),
$$

it takes only a small amount of index gymnastics to verify that the above equation may be split up further so that:

$$
\begin{aligned}
&h dL \circ \Psi_{(q_k,q_{k+1})}(\delta q_k, \delta q_{k+1}) = \\
&\left((1-\tau) \left(\frac{\partial \ell^d_{k+\tau}}{\partial q^i} \psi^i_j(q_{k+\tau}) + \frac{\partial \ell^d_{k+\tau}}{\partial \xi^m} \xi^i_{k,k+1} c^m_{ij}(q_{k+\tau}) \right) - \frac{1}{h} \frac{\partial \ell^d_{k+\tau}}{\partial \xi^j} \right) \phi^j_b(q_k) \delta q^b_k \\
&+ \left(\tau \left(\frac{\partial \ell^d_{k+\tau}}{\partial q^i} \psi^i_j(q_{k+\tau}) + \frac{\partial \ell^d_{k+\tau}}{\partial \xi^m} \xi^i_{k,k+1} c^m_{ij}(q_{k+\tau}) \right) + \frac{1}{h} \frac{\partial \ell^d_{k+\tau}}{\partial \xi^j} \right) \phi^j_b(q_{k+1}) \delta q^b_{k+1},
\end{aligned}
$$

where we have also scaled by the time-step h (remember that $\ell^d = h\ell$). Recall that we have identified the directional derivative operator at $q_{k+\tau}$ with respect to the frame as

$$
u_j[\ell^d]_{k+\tau} = \frac{\partial \ell^d_{k+\tau}}{\partial q^i} \psi^i_j(q_{k+\tau}),
$$

the discrete momentum as

$$
\mu_{k,k+1;j} = \frac{\partial \ell^d_{k+\tau}}{\partial \xi^j},
$$

and the dual bracket operating on discrete velocity and momentum as

$$
[\xi_{k,k+1}, \mu_{k,k+1}]^*_{q_{k+\tau};j} u^j(q_{k+\tau}) = \mu_{k,k+1;m} \xi^i_{k,k+1} c^m_{ij}(q_{k+\tau}) u^j(q_{k+\tau}) \in T^*_{q_{k+\tau}} Q.
$$

Furthermore, notice that $\phi^j_m(q_k)\delta q^m_k$ are just the components of the vector δq_k measured against the frame on the fiber $T_{q_k}Q$. Then the coordinate expression for the one-form $h dL \circ \Psi$ over $Q \times Q$

is written as

$$
\begin{aligned}
hdL \circ \Psi(q_k, q_{k+1}) &= \left((1-\tau) \left(u[\ell^d]_{k+\tau} + [\xi_{k,k+1}, \mu_{k,k+1}]^*_{q_{k+\tau}} \right) - \frac{1}{h}\mu_{k,k+1} \right) \cdot u(q_k) \\
&\quad + \left(\tau \left(u[\ell^d]_{k+\tau} + [\xi_{k,k+1}, \mu_{k,k+1}]^*_{q_{k+\tau}} \right) + \frac{1}{h}\mu_{k,k+1} \right) \cdot u(q_{k+1}) \\
&= -\Theta^-_{\ell_d}(q_k, q_{k+1}) + \Theta^+_{\ell_d}(q_k, q_{k+1})
\end{aligned}
$$

so that we see $hdL\circ\Psi$ is exactly the difference of the Hamel's one-forms, i.e. $hdL\circ\Psi = \Theta^+_{\ell_d} - \Theta^-_{\ell_d}$, as expressed in equation (5.1).

□

The above theorem also shows us that the discrete Hamel's equations are equivalent to a modification of the discrete Hamilton's principle.

Corollary 5.2. *The following statements are equivalent.*

1. *The discrete trajectory $q_d \in \mathcal{C}_d(Q)$ satisfies the modified variational principle*

$$
\sum_{k=0}^{N-1} dL \circ \Psi_{(q_k, q_{k+1})}(\delta q_k, \delta q_{k+1}) = 0 \tag{5.5}
$$

   for arbitrary virtual displacements defined as in equation (2.22) *where the variation is fixed at the endpoints so that $\delta q_0 = \delta q_N = 0$.*

2. *The discrete trajectory satisfies the discrete Hamel's equations* (3.6) *for $k = 1, \ldots, N-1$.*

Proof. The proof is straightforward following the method of Marsden and West [36] and recounted in earlier 2.3. Specifically, note that the modified variational principle is written in index notation as

$$
\sum_{k=0}^{N-1} -\Theta^-_{\ell_d}(q_k, q_{k+1})_j \phi^j_i(q_k)\delta q^i_k + \Theta^+_{\ell_d}(q_k, q_{k+1})_j \phi^j_i(q_{k+1})\delta q^i_{k+1} = 0
$$

as a consequence of the previous theorem. We may manipulate the summation index so that the principle states

$$
\sum_{k=0}^{N-1} -\Theta^-_{\ell_d}(q_k, q_{k+1})_j \phi^j_i(q_k)\delta q^i_k + \sum_{k=1}^{N} \Theta^+_{\ell_d}(q_{k-1}, q_k)_j \phi^j_i(q_k)\delta q^i_k = 0
$$

and because variation at the endpoints is fixed we find the principle is equivalent to the statement

$$\sum_{k=1}^{N-1}\left(-\Theta^-_{\ell_d}(q_k,q_{k+1})_j+\Theta^+_{\ell_d}(q_{k-1},q_k)_j\right)\phi^j_i(q_k)\delta q^i_k=0.$$

Finally, we note that because ϕ is an invertible matrix, for arbitrary displacements δq_k the discrete trajectory q_d equivalently satisfies the discrete Hamel's equations. These equations may be expressed in terms of the one forms $\Theta^-_{\ell_d}, \Theta^+_{\ell_d} \in T^*(Q\times Q)$ as

$$\pi^-\Theta^-_{\ell_d}(q_k,q_{k+1})=\pi^+\Theta^+_{\ell_d}(q_{k-1},q_k),$$

where $\pi^\pm$ are the two natural projections mapping $T^*(Q\times Q)\to T^*Q$, defined so that

$$\begin{aligned}\pi^-(dq_k,dq_{k+1})_{(q_k,q_{k+1})}&=dq_k\in T^*_{q_k}Q \text{ and}\\ \pi^+(dq_k,dq_{k+1})_{(q_k,q_{k+1})}&=dq_{k+1}\in T^*_{q_{k+1}}Q\end{aligned}$$

□

Remark 5.1. It is straightforward to demonstrate that the discrete Hamel's one-forms $\Theta^\pm_{\ell_d}$, as defined above, reduce to the discrete Lagrangian one-forms $\Theta^\pm_{L_d}$ when the transformations ψ and ϕ describing the frame are the identity. This corresponds to the trivial case when the frame *is* the coordinate-induced frame, where in the continuous setting the Hamel's equations reduce to the Euler–Lagrange equations. In the discrete setting, this implies that the discrete Hamel's equations reduce to the discrete Euler–Lagrange equations in the case of a trivial frame.

5.2.2 A Closed One-Form and the Poincaré Lemma

Now that we have shown that $hdL\circ\Psi$, a one-form over $Q\times Q$ induces through a variational principle the discrete Hamel's equations in much the same way that the dL_d induces the discrete Euler–Lagrange equations, our next step will be to determine whether or not the form is closed. We begin with a statement that illustrates how the exterior derivative behaves under composition.

Proposition 5.3. *Suppose M and N are n-dimensional differentiable manifolds, $f : TN \to \mathbb{R}$ such that f is differentiable, and $\Psi : TM \to TN$ where Ψ is a local C^1-diffeomorphism. Then the mapping is $f\circ\Psi : TM \to \mathbb{R}$ is (locally) differentiable. For arbitrary $(\alpha,\beta)\in T_{(q,v)}M$, in local coordinates*

$$d(f\circ\Psi)_{(q,v)}(\alpha,\beta)=\langle df_{\Psi(q,v)},(\Psi_*)_{(q,v)}(\alpha,\beta)\rangle \tag{5.6}$$

where $\langle\cdot,\cdot\rangle$ is the natural pairing between the vector space $T_{\Psi(q,v)}(TN)$ and its dual $T^*_{\Psi(q,v)}(TN)$.

Remark 5.2. This result is an easily verifiable consequence of the chain rule and the push-forward in differential geometry. f in the above proposition can be thought of as a C^1 one-form on the manifold N, likewise $f \circ \Psi$ as the same on M. The exterior derivative of $f \circ \Psi$ can be thought of as a mapping $d(f \circ \Psi) : T(TM) \to T\mathbb{R} \cong \mathbb{R}^2$ such that

$$\begin{aligned} d(f \circ \Psi)_{(q,v)} &: T_{(q,v)}(TM) \to \mathbb{R} \\ d(f \circ \Psi)_{(q,v)} &\in T^*_{(q,v)}(TM). \end{aligned}$$

The relation described above leads directly to the observation that the one-form $hdL \circ \Psi$ that induces the discrete Hamel's equations in our modified discrete variational principle (5.5) is closed.

Theorem 5.4. *The one-form $hdL \circ \Psi$ defined over $Q \times Q$ is closed.*

Proof. The verification of this theorem is made trivial by the above proposition. dL is a C^1 one-form over $Q \times Q$, so by (5.6),

$$\begin{aligned} &d(hdL \circ \Psi)_{((q_k,q_{k+1}),(v_k,v_{k+1}))}((\alpha_k, \alpha_{k+1}), (\beta_k, \beta_{k+1})) = \\ &\qquad \left\langle hd(dL)_{\Psi((q_k,q_{k+1}),(v_k,v_{k+1}))}, (\Psi_*)_{((q_k,q_{k+1}),(v_k,v_{k+1}))}((\alpha_k, \alpha_{k+1}), (\beta_k, \beta_{k+1}))\right\rangle, \end{aligned}$$

and because $d(dL) = 0$, $d(hdL \circ \Psi) = 0$ for arbitrary vectors in $T_{((q_k,q_{k+1}),(v_k,v_{k+1}))}(T(Q \times Q))$. □

Now, recall that one of our stated goals is to find a function $\mathcal{L}_d : Q \times Q \to \mathbb{R}$ satisfying equation (5.2). Because we have specified that Q (and hence $Q \times Q$) is a vector space, by the Poincaré lemma we may conclude that since $hdL \circ \Psi = -\Theta^-_{\ell_d} + \Theta^+_{\ell_d}$ is *closed* by theorem 5.4, it is also exact and consequently there must exist a function $\mathcal{L}_d : Q \times Q \to \mathbb{R}$ such that

$$d\mathcal{L}_d = hdL \circ \Psi = -\Theta^-_{\ell_d} + \Theta^+_{\ell_d}. \tag{5.7}$$

Furthermore, this allows us to reformulate our modified variational principle (5.5) into a more familiar form as a principle of critical action, where the ***modified discrete action*** is defined as

$$s_d(q_d) = \sum_{k=0}^{N-1} \mathcal{L}_d(q_k, q_{k+1}) \tag{5.8}$$

and the principle itself may be stated as follows.

Corollary 5.5. *The following statements are equivalent:*

1. *The discrete trajectory $q_d \in \mathcal{C}_d(Q)$ is critical with respect to the modified discrete action*

$$s_d(q_d) = \sum_{k=0}^{N-1} \mathcal{L}_d(q_k, q_{k+1})$$

where q_d is fixed at the endpoints $q_0, q_N \in Q$.

2. *The discrete trajectory $q_d = \{q_k\}_{k=0}^{N-1}$ satisfies the discrete Hamel's equations* (3.6) *for $k = 1, \ldots, N-1$.*

Proof. Variation of the action results in

$$\begin{aligned}\delta s_d(q_d) &= \sum_{k=0}^{N-1} d\mathcal{L}_d(q_k, q_{k+1})(\delta q_k, \delta q_{k+1}) \\ &= \sum_{k=0}^{N-1} \left(-\Theta^-_{\ell_d}(q_k, q_{k+1}) + \Theta^+_{\ell_d}(q_k, q_{k+1})\right)(\delta q_k, \delta q_{k+1})\end{aligned}$$

by the relation specified in equation (5.7). Next, manipulation of the summation index reveals

$$\begin{aligned}\delta s_d(q_d) &= \sum_{k=0}^{N-1} \left(-\Theta^-_{\ell_d}(q_k, q_{k+1})\right)(\delta q_k, \delta q_{k+1}) + \sum_{k=1}^{N} \left(\Theta^+_{\ell_d}(q_{k-1}, q_k)\right)(\delta q_{k-1}, \delta q_k) \\ &= \sum_{k=1}^{N-1} \left[-\Theta^-_{\ell_d}(q_k, q_{k+1})\delta q_k + \Theta^+_{\ell_d}(q_{k-1}, q_k)\delta q_k\right] - \Theta^-_{\ell_d}(q_0, q_1)\delta q_0 + \Theta^+_{\ell_d}(q_{N-1}, q_N)\end{aligned}$$

and since the trajectory is fixed at the endpoints, i.e. $\delta q_0 = \delta q_N = 0$, then the trajectory is critical with respect to the action (that is, $\delta s_d(q_d) = 0$) if and only if

$$\pi^- \Theta^-_{\ell_d}(q_k, q_{k+1}) = \pi^+ \Theta^+_{\ell_d}(q_{k-1}, q_k)$$

for $k = 1, \ldots, N-1$, which, again, is exactly the statement of the discrete Hamel's equations (3.6). □

Remark 5.3. The application of the Poincaré lemma above depends upon Q being a vector space. If Q were a general differentiable manifold, we could assert at least local existence of $\mathcal{L}_d$ as long as the local coordinate chart is a mapping to an open contractible subset of $\mathbb{R}^n$, i.e. the manifold is locally contractible. The following arguments demonstrating conservation of a symplectic form might be thus extended to the case were Q is a general differentiable manifold by emphasizing that symplecticity would only be a *local* property. Nonetheless, for our purposes and for applications it suffices to prove conservation of a symplectic two-form in the case where Q is a vector space.

Remark 5.4. In Section 2.3 we defined regular (and hyperregular) discrete Lagrangian functions $L_d : Q \times Q \to \mathbb{R}$ as functions whose discrete Legendre transforms/fiber derivatives are local (or global) isomorphisms $\mathbb{F}^\pm L_d : Q \times Q \to T^*Q$. Consequently, the discrete Euler–Lagrange equations are said to be locally or globally well defined when the underlying discrete Lagrangian is, respectively, regular or hyperregular.

Similarly, the discrete Hamel's equations are well defined when the function $\mathcal{L}_d$ is hyperregular. The discrete fiber derivatives and Legendre transforms can be developed in an identical manner to the exposition in Section 2.3. However, for a Lagrangian function L, the relationship between regularity/hyperregularity of the discrete functions L_d and $\mathcal{L}_d$ is not quite clear for an arbitrary choice of non-coordinate frame. In other words, it is an open question as to whether regularity of L_d implies regularity of $\mathcal{L}_d$. Of course, in applications a frame will be chosen so that the discrete Hamel's equations are a *simplification* of the discrete Euler–Lagrange equations, so it is not too restrictive to assume that the function $\mathcal{L}_d$ is hyperregular, as we shall do for the remainder of this paper.

5.2.3 Structure Conservation

Corollary 5.6. *The discrete update map F_{ℓ_d} induced by the discrete Hamel's equations* (3.6) *conserves a discrete symplectic two-form $\Omega_{\ell_d} = d\Theta^-_{\ell_d} = d\Theta^+_{\ell_d}$ that we refer to as the* **discrete Hamel's symplectic two-form**.

Proof. We have already shown that the one-form $d\mathcal{L}_d = hdL \circ \Psi$ is expressible as the difference of the discrete Hamel's one-forms so that $hdL \circ \Psi = -\Theta^-_{\ell_d} + \Theta^+_{\ell_d}$. As a consequence of theorem 5.4,

$$
\begin{aligned}
& d(dL \circ \Psi) = 0 = -d\Theta^-_{\ell_d} + d\Theta^+_{\ell_d} \Rightarrow \\
& d\Theta^-_{\ell_d}(q_k, q_{k+1}) = d\Theta^+_{\ell_d}(q_k, q_{k+1}) = \Omega_{\ell_d}(q_k, q_{k+1})
\end{aligned}
$$

so that at each step specified by (q_k, q_{k+1}) for $k = 0, \ldots, N-1$, there is a unique two-form $\Omega_{\ell_d}(q_k, q_{k+1})$.

Next, as in Section 2.3.2, we observe that through iteration of F_{ℓ_d}, a discrete trajectory $q_d = \{q_k\}_{k=0}^N$ is uniquely identified by initial conditions $(q_0, q_1) \in Q \times Q$. We may similarly define a restricted discrete modified action $\hat{s}_d : Q \times Q \to \mathbb{R}$ as the modified discrete action s_d (5.8) restricted to the trajectory induced through iteration of the initial conditions (q_0, q_1) by F_{ℓ_d}:

$$
\hat{s}_d(q_0, q_1) = s_d(\{q_k\}_{k=0}^{N-1}), \quad (q_k, q_{k+1}) = F_{\ell_d}(q_{k-1}, q_k), \quad \forall k = 1, \ldots, N-1.
$$

Then variation of the restricted modified discrete action returns

$$
\begin{aligned}
d\hat{s}_d(q_0, q_1) \cdot (\delta q_0, \delta q_1) = & - \Theta^-_{\ell_d}(q_0, q_1) \cdot (\delta q_0, \delta q_1) \\
& + \Theta^+_{\ell_d}(F^{N-1}_{\ell_d}(q_0, q_1)) \cdot (F^{N-1}_{\ell_d})_{*(q_0,q_1)}(\delta q_0, \delta q_1) \\
= & \left((F^{N-1}_{\ell_d})^* \Theta^+_{\ell_d} - \Theta^-_{\ell_d} \right) ((q_0, q_1), (\delta q_0, \delta q_1)) \, ,
\end{aligned}
$$

where again we are left with only the terms that would otherwise cancel out when we consider variation at the endpoints fixed, because when the trajectory is restricted to solutions of the discrete Hamel's equations, $\Theta^-_{\ell_d}(F_{\ell_d}(q_k, q_{k+1})) = \Theta^+_{\ell_d}(q_k, q_{k+1})$. In fact, this implies that $d\Theta^+_{\ell_d}(q_{k-1}, q_k) = d\Theta^-_{\ell_d}(q_k, q_{k+1})$ when we restrict the discrete Hamel's one-forms to the solution space. Finally, because $d^2\hat{s} = 0$ and by compatibility of the pullback with the exterior derivative, we see that the above results imply conservation of the two-form Ω_{ℓ_d}

$$(F^{N-1}_{\ell_d})^*(\Omega_{\ell_d}) = \Omega_{\ell_d}$$

where $\Omega_{\ell_d} = d\Theta^+_{\ell_d} = -d\Theta^-_{\ell_d}$ is defined as the ***discrete Hamel symplectic form*** on $Q\times Q$. Just as in Section 2.3.2, the process holds for any contiguous sequence of substeps, and thus a single step induced by the update map F_{ℓ_d} also conserves the discrete symplectic form: $F^*_{\ell_d}(\Omega_{\ell_d}) = \Omega_{\ell_d}$. The update map $F_{\ell_d} : Q \times Q \to Q \times Q$ is therefore a discrete symplectic map. Symplecticity (i.e. skew symmetry, isotropy, and nondegeneracy) is easily verified because the two-form is expressible in coordinates in terms of the function $\mathcal{L}_d$ as

$$\Omega_{\ell_d} = \frac{\partial^2 \mathcal{L}_d(q_k, q_{k+1})}{\partial q^i_k \partial q^j_{k+1}} dq^i_k \wedge dq^j_{k+1}.$$

It is important to remember that Ω_{ℓ_d} is a two-form over $Q \times Q$, and hence is thought of as a mapping $T(Q \times Q) \times T(Q \times Q) \to \mathbb{R}$. Isotropy is demonstrated when we verify that, in fact:

$$(\Omega_{\ell_d})_{(q_k,q_{k+1})}\left((v_k, v_{k+1}), (v_k, v_{k+1})\right) = \frac{\partial^2 \mathcal{L}_d(q_k, q_{k+1})}{\partial q^i_k \partial q^j_{k+1}} \left(v^i_k v^j_{k+1} - v^i_{k+1} v^j_k\right) = 0$$

by commutativity of mixed partial derivatives. Skew-symmetry is a trivial consequence of isotropy, and nondegeneracy is verifiable because $\mathcal{L}_d$ has been assumed to be hypperregular.

□

The discrete Hamel's equations satisfy a discrete version of Noether's theorem 2.3 in much the same way as the discrete Euler–Lagrange equations. Again, we consider the group action $\Phi : G \times Q \to Q$ extended to $Q \times Q$ as $\Phi^{Q\times Q} : G \times (Q \times Q) \to Q \times Q$ in (2.27), and infinitesimal generator $\xi_{Q\times Q} : Q \times Q \to T(Q \times Q)$ given in (2.28). The ***discrete Hamel's momentum maps*** $J^{\pm}_{\ell_d} : Q \times Q \to \mathfrak{g}^*$ are defined such that for $\xi \in \mathfrak{g}$

$$\begin{aligned} J^-_{\ell_d}(q_k, q_{k+1}) \cdot \xi &= \langle -D_1\mathcal{L}_d(q_k, q_{k+1}), \xi_Q(q_k)\rangle = \Theta^-_{\ell_d} \cdot \xi_{Q\times Q}(q_k, q_{k+1}), \text{ and} \\ J^+_{\ell_d}(q_k, q_{k+1}) \cdot \xi &= \langle D_2\mathcal{L}_d(q_k, q_{k+1}), \xi_Q(q_{k+1})\rangle = \Theta^+_{\ell_d} \cdot \xi_{Q\times Q}(q_k, q_{k+1}) \end{aligned} \tag{5.9}$$

This proof of the following theorem is easily demonstrated by replacing all instances of the discrete Lagrangian L_d with $\mathcal{L}_d$ in Section 2.3.3.

Theorem 5.7 (Discrete Noether's Theorem for Hamel's Formalism)**.** *Suppose $\mathcal{L}_d : Q \times Q \to \mathbb{R}$ is a discrete function such that $hdL \circ \Psi = d\mathcal{L}_d$, approximating a Lagrangian mechanical system, and furthermore suppose that $\mathcal{L}_d$ is invariant with respect to the action Φ of a Lie group G lifted to $Q \times Q$ as in* (2.27)*. Then:*

1. $\Phi_g^{Q \times Q} : Q \times Q \to Q \times Q$ *is a* special *discrete symplectic map. Also, the discrete Lagrangian momentum maps will be equivalent so that $J_{\ell_d}^+ = J_{\ell_d}^- = J_{\ell_d}$ and can thus be referred to as one map $J_{\ell_d} : Q \times Q \to \mathfrak{g}^*$.*
2. *The update map F_{ℓ_d} conserves the discrete Lagrangian momentum map J_{ℓ_d}, i.e. $J_{\ell_d} \circ F_{\ell_d} = J_{\ell_d}$.*

Remark 5.5. We note that it is unnecessary for the actual derivation of the discrete Hamel's equations to explicitly compute $\mathcal{L}_d$. Thus, it may ultimately be useful to examine how invariance of L_d might relate to invariance on $\mathcal{L}_d$. We leave this for future development.

Chapter 6

Conclusions and Future Work

This dissertation has introduced variational integrators motivated by Hamel's formalism for the equations of motion in mechanics. Recall that Hamel's equations are a generalization of the Euler–Poincaré equations, and also of the Euler–Lagrange equations. In much the same way, from a very broad perspective the discrete Hamel's equations introduced in this dissertation might be thought of as a generalization of the discrete Euler–Poincaré numerical integrators derived by Bou-Rabee and Marsden [14], and also of the discrete Euler–Lagrange equations reviewed in Marsden and West [36].

Notably, it has been demonstrated that the discrete Hamel's equations, as variational integrators, are structure-preserving in that they conserve a discrete symplectic two-form and, in the case of symmetry, conserve discrete momentum. In systems where the Hamel's equations are a simplification over the Euler–Lagrange equations, preliminary results indicate that similar simplifications will carry over to the discrete setting.

While the results presented in this dissertation are largely theoretical, it should be reiterated that the discrete Hamel's equations have practical implications. It may be noted that, when an appropriate frame is selected, simulations designed using the discrete Hamel's equations can preserve manifolds of relative equilibria of certain nonholonomic systems, and thus are of interest in the construction of nonholonomic integrators.

The results presented in this dissertation offer a variety of future avenues of research. Such possibilities include:

- It may be of particular interest to more closely examine the relationship between the variational integrators developed in this dissertation and recent integrators developed by Bou-Rabee and Marsden [14] on Lie groups, especially linking their retraction selection with the idea of frame selection in this dissertation. This may also serve as a first step towards extension of the discrete Hamel's formalism to manifolds.

- Extensions of the currently presented Hamel's formalism to the development and study of energy conserving integrators may be of interest.
- Future developments may also include an extension of Hamel's formalism to higher order approximations of mechanical systems.

REFERENCES

[1] R Abraham and J E Marsden. *Foundations of Mechanics*. AMS Chelsea Publishing, 2nd edition, 2008.

[2] R Abraham, J E Marsden, and T Ratiu. *Manifolds, Tensor Analysis, and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2nd edition, 1988.

[3] V I Arnold. *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2nd edition, 1989.

[4] K Ball, D V Zenkov, and A M Bloch. Variational structures for Hamel's equations and stabilization. *Proc. IFAC*, pages 178–183, 2012.

[5] K R Ball and D V Zenkov. Hamel's formalism and variational integrators. *Preprint*, 2013.

[6] A M Bloch. *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Appl. Math.* Springer-Verlag, New York, 2003.

[7] A M Bloch and P E Crouch. Representations of Dirac structures on vector spaces and nonlinear LC circuits. In *Differential Geometry and Control*, volume 64 of *Proc. Sympos. Pure Math.*, pages 103–117. AMS, Providence, Rhode Island, 1997.

[8] A M Bloch, P S Krishnaprasad, J E Marsden, and R Murray. Nonholonomic mechanical systems with symmetry. *Arch. Rational Mech. Anal.*, 136:21–99, 1996.

[9] A M Bloch, J E Marsden, and D V Zenkov. Quasivelocities and stabilization of relative equilibria of underactuated nonholonomic systems. *Proc. CDC*, 48:3335–3340, 2009.

[10] A M Bloch, J E Marsden, and D V Zenkov. Quasivelocities and symmetries in nonholonomic systems. *Dynamical Systems: An International Journal*, 24(2):187–222, 2009.

[11] Anthony Bloch, PS Krishnaprasad, Jerrold E Marsden, and Tudor S Ratiu. The euler-poincaré equations and double bracket dissipation. *Communications in mathematical physics*, 175(1):1–42, 1996.

[12] A I Bobenko and Yu B Suris. Discrete Lagrangian reduction, discrete Euler–Poincaré equations, and semidirect products. *Letters in Mathematical Physics*, 49:79–93, 1999.

[13] A I Bobenko and Yu B Suris. Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top. *Communications in Mathematical Physics*, 204(1):147–188, 1999.

[14] N Bou-Rabee and J E Marsden. Hamilton–Pontryagin integrators on Lie groups part I: Introduction and structure-preserving properties. *Foundations of Computational Mathematics*, 9(2):197–219, 2009.

[15] S A Chaplygin. On the motion of a heavy body of revolution on a horizontal plane. *Physics Section of the Imperial Society of Friends of Physics, Anthropology, and Ethnographics*, 9:10–16, 1897. (Russian).

[16] N G Chetaev. On Gauss' principle. *Izv. Fiz-Mat. Obsc. Kazan. Univ. Ser. 3*, 6:68–71, 1932–1933.

[17] J Cortés and S Martínez. Nonholonomic integrators. *Nonlinearity*, 14:1365–1392, 2001.

[18] L Euler. Decouverte d'un nouveau principe de Mecanique. *Mémoires de l'académie des sciences de Berlin*, 6:185–217, 1752.

[19] Yu N Fedorov and D V Zenkov. Discrete nonholonomic LL systems on Lie groups. *Nonlinearity*, 18:2211–2241, 2005.

[20] Yu N Fedorov and D V Zenkov. Dynamics of the discrete chaplygin sleigh. *Disc. Cont. Dyn. Syst.*, (extended volume):258–267, 2005.

[21] E Hairer, C Lubich, and G Wanner. *Geometric Numerical Integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin/Heidelberg, 2nd edition, 2006.

[22] G Hamel. Die Lagrange–Eulersche Gleichungen der Mechanik. *Z. Math. Phys.*, 50:1–57, 1904.

[23] W R Hamilton. On a general method in dynamics, part I. *Phil. Trans. Roy. Soc. Lond.*, pages 247–308, 1834.

[24] W R Hamilton. On a general method in dynamics, part II. *Phil. Trans. Roy. Soc. Lond.*, pages 95–144, 1835.

[25] D D Holm, J E Marsden, and T S Ratiu. The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.*, 137:1–81, 1998.

[26] A V Karapetyan. On the problem of steady motions of nonholonomic systems. *J. Appl. Math. Mech.*, 44:418–426, 1980.

[27] M Kobilarov, J E Marsden, and G S Sukhatme. Geometric discretization of nonholonomic systems with symmetries. *Discrete and Continuous Dynamical Systems–Series S*, 3(1):61–84, 2010.

[28] VV Kozlov. The problem of realizing constraints in dynamics. *Journal of Applied Mathematics and Mechanics*, 56(4):594–600, 1992.

[29] J L Lagrange. *Mécanique Analytique.* Chez la Veuve Desaint, 1788.

[30] G H Livens. On Hamilton's principle and the modified function in analytical dynamics. *Proc. Roy. Soc. Edinburgh*, 39:113–119, 1919.

[31] C Lynch and D Zenkov. Stability of stationary motions of discrete-time nonholonomic systems. In V V Kozlov, S N Vassilyev, A V Karapetyan, N N Krasovskiy, V N Tkhai, and F L Chernousko, editors, *Problems of Analytical Mechanics and Stability Theory. Collection of Papers Dedicated to the Memory of Academician Valentin V. Rumyantsev*, pages 259–271. Fizmatlit, Moscow, 2009. (Russian).

[32] C Lynch and D V Zenkov. Stability of relative equilibria of discrete nonholonomic systems with semidirect symmetry. preprint, 2010.

[33] J E Marsden. *Lectures on Mechanics*, volume 174 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1992.

[34] J E Marsden, S Pekarsky, and S Shkoller. Discrete Euler–Poincaré and Lie–Poisson equations. *Nonlinearity*, 12:1647–1662, 1999.

[35] J E Marsden and T S Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Appl. Math.* Springer-Verlag, New York, 2nd edition, 1999.

[36] J E Marsden and M West. Discrete mechanics and variational integrators. *Acta Numerica*, 10:357–514, 2001.

[37] Jerrold E Marsden and Jürgen Scheurle. The reduced euler-lagrange equations. *Fields Institute Comm*, 1:139–164, 1993.

[38] J M Maruskin, A M Bloch, J E Marsden, and D V Zenkov. A fiber bundle approach to the transpositional relations in nonholonomic mechanics. *J. Nonlinear Sci.*, 22(4):431–461, 2012.

[39] J Moser and A Veselov. Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Communications*, 139(2):217–243, 1991.

[40] Ju I Neimark and N A Fufaev. *Dynamics of Nonholonomic Systems*, volume 33 of *Translations of Mathematical Monographs*. AMS, Providence, Rhode Island, 1972.

[41] I Newton. *Philosophi Naturalis Principia Mathematica.* Londini Societatis Regiae ac Typis, Josephi and Streater, 1687.

[42] L A Pars. *A Treatise on Analytical Dynamics.* Wiley, New York, 1965.

[43] H Poincaré. Sur une forme nouvelle des équations de la mécanique. *CR Acad. Sci.*, 132:369–371, 1901.

[44] E J Routh. *Treatise on the Dynamics of a System of Rigid Bodies.* MacMillan, London, 1860.

[45] A J van der Schaft. Implicit Hamiltonian systems with symmetry. *Reports on Mathematical Physics*, 41:203–221, 1998.

[46] A J van der Schaft and B M Maschke. On the Hamiltonian formulation of nonholonomic mechanical systems. *Reports on Mathematical Physics*, 34:225–233, 1994.

[47] A J van der Schaft and B M Maschke. The Hamiltonian formulation of energy conserving physical systems with external ports. *Archiv für Elektronik und Übertragungstechnik*, 49:362–371, 1995.

[48] G T Walker. On a dynamical top. *Quart. J. Pure Appl. Math.*, 28:175–184, 1896.

[49] H Yoshimura and J E Marsden. Dirac structures in Lagrangian mechanics. Part I: Implicit Lagrangian systems. *Journal of Geometry and Physics*, 57:133–156, 2006.

[50] H Yoshimura and J E Marsden. Dirac structures in Lagrangian mechanics. Part II: Variational structures. *Journal of Geometry and Physics*, 57:209–250, 2006.

[51] H Yoshimura and J E Marsden. Reduction of Dirac structures and the Hamilton–Pontryagin principle. *Reports on Math. Phys.*, 60:381–426, 2007.

[52] D V Zenkov, A M Bloch, and J E Marsden. The energy-momentum method for the stability of nonholonomic systems. *Dynamics and Stability of Systems*, 13:128–166, 1998.