J. Nonlinear Sci. (2006): pp. OF1–OF46
DOI: 10.1007/s00332-005-0698-1

Journal of
Nonlinear
Science

Integrators for Nonholonomic Mechanical Systems

R. McLachlan and M. Perlmutter
IFS, Massey University, Palmerston North, New Zealand 5301
e-mail: {R.McLachlan,M.Perlmutter}@massey.ac.nz

Received March 8, 2005; revised manuscript accepted January 3, 2006
Online publication June 21, 2006
Communicated by C. Rowley

Summary. We study a discrete analog of the Lagrange-d'Alembert principle of nonholonomic mechanics and give conditions for it to define a map and to be reversible. In specific cases it can generate linearly implicit, semi-implicit, or implicit numerical integrators for nonholonomic systems which, in several examples, exhibit superior preservation of the dynamics. We also study discrete nonholonomic systems on Lie groups and their reduction theory, and explore the properties of the exact discrete flow of a nonholonomic system.

1. Introduction

1.1. Nonholonomic Dynamical Systems

Hamiltonian systems have an enormous wealth of distinguishing features. They preserve energy, symplecticity, and phase space volume. They can preserve momentum in the presence of symmetries. They have a well-developed theory of integrability, near-integrability, and symmetry reduction. Moreover, all of these properties have discrete-time analogs which can give numerical integrators extraordinarily good long-time stability and robustness.

When the system is subject to a constraint, the whole picture can change drastically. The best-understood situation is that of holonomic, or position-only constraints. The position constraint $g(q) = 0$ ($q \in Q$, the configuration manifold of the system), along with its implied velocity constraint $Tg(q)\dot{q} = 0$, leads to a Hamiltonian system on the submanifold defines by these constraints. Numerical integrators that preserve these constraints and the associated structures (symplecticity, symmetry, etc.) are well known and widely used in applications such as molecular dynamics, where the constraints may be molecular bond lengths or angles [18].

By contrast, the situation for nonholonomic constraints is less well understood. These are velocity constraints that do not arise as the derivative of any position constraint. The resulting mechanical systems will not preserve the symplectic form. They may or may not preserve energy, phase space volume, or momentum, and their integrability and reduction theories are completely different from the Hamiltonian case [13], [16], [20]. They arise most commonly in systems with rolling contact (such as wheeled vehicles or the famous rattleback or Celtic stone, which spins stably in one direction only) or sliding contact (such as ice skates, whose velocity must be in the direction of the skate). Other nonholonomic dynamical systems that have been studied as models include the car with trailers, the rolling penny, and the rolling ball on a turntable. The behavior of these systems is often quite unexpected.

Surveys of nonholonomic systems can be found in Bloch et al. [4], Bloch [2], and Vershik and Gershkovich [29]. A major thrust of present research is to give a complete description of the Hamiltonian [1], [8], [20], [21], [22] and Lagrangian [3], [9] geometry of nonholonomic systems. Specific topics include a fiber bundle and connection formulation; the evolution of the momentum in systems with symmetry; the possible conservation of volume on the constraint manifold; the appearance of dissipative behavior; extreme cases of symmetries in which the constraints are transverse (Chaplygin case) or tangent (Lie group case) to the symmetries; and reduction and control. For example, in one recent approach, following Cartan's study of nonintegrable distributions Koiller et al. [16], [15] study nonholonomic systems as geodesics of a nonholonomic connection, involving the Levi-Civita connection and orthogonal projection onto the nonintegrable distribution. Local invariants for general distributions are obtained following Cartan's method of equivalence. Eventually, one wants to understand better the relationship between the algebraic properties of the distribution and the consequences for the dynamics such as integrability and behavior under reduction.

Like Hamiltonian dynamics, nonholonomic dynamics can be subtle, and many systems can be studied only via long numerical simulations. The development of special-purpose numerical integrators for nonholonomic mechanical systems is hindered by the lack of a clear theory describing the structural features of their dynamics. However, numerical integrators derived from discrete variational principles have proved very reliable both in situations where the class of dynamics is well understood (e.g. in Hamiltonian ODEs [26] and PDEs [7]) and where it is not as well understood (e.g. in systems with collisions or with multiple timescales [28]). They therefore seem suitable to apply here, because the nonholonomic equations of motion come from the Lagrange-d'Alembert principle, which (in the way it includes the forces due to the constraints) is *not* a standard variational principle. In this paper we further investigate, along lines introduced by Cortés [14], integrators satisfying a discrete analog of the Lagrange-d'Alembert principle.

We only consider the case in which the velocity constraints are linear, i.e., take the form $A(q)\dot{q} = 0$, and the Lagrangian is regular. Such nonholonomic systems do at least preserve energy. If, in addition, the Lagrangian is of simple mechanical (kinetic minus potential) type, then the system is also reversible, a feature which is known to control dynamics in a way very reminiscent to symplecticity. One of our discrete Lagrange-d'Alembert (DLA) integrators, eq. (4.18) below, is second-order accurate, time-symmetric, reversible, and requires only one force evaluation per time step. It is the nonholonomic analog of the widely used integrators SHAKE and RATTLE [18] for holonomic constraints. It performs markedly well in our numerical tests.

More precisely, such linear velocity constraints define a distribution C on the configuration manifold Q. If C is integrable, then the constraint is said to be holonomic, and Q foliates into leaves on each of which one has standard Hamiltonian dynamics; if C is nonintegrable, then the constraint is said to be nonholonomic. The term *holonomic* (= integrable) was introduced by Hertz in 1894, who also distinguished between the geometry of *straightest* paths (what we are calling nonholonomic dynamics, governed by the Lagrange-d'Alembert principle), and the geometry of *shortest* paths, or geodesics in (Q, C), which are governed by a true variational principle (sometimes called vakanomic mechanics). The study of nonholonomic manifolds (Q, C) has many links to foliation theory, control theory, thermodynamics, and quantum theory, and partial differential equations [29].

The equations of motion for nonholonomic systems can be determined from the Lagrange-d'Alembert principle. A discrete analog, the *discrete Lagrange-d'Alembert (DLA) principle*, was introduced by Cortés [14]. (See also [11], [12], [19].) In this paper we further develop the theory of this principle, apply it to develop some practical integrators, and test these on nonholonomic systems showing a range of behavior. We also apply the principle to nonholonomic systems on Lie groups, showing that it can produce an exact analog of the continuous theory.

1.2. Survey of the Paper

The theory of (continuous time) nonholonomic systems is reviewed in Section 2 and the discrete Lagrange-d'Alembert (DLA) principle introduced and studied in Section 3. The DLA principle requires on $Q \times Q$ a discrete Lagrangian, a discrete constraint distribution which is a submanifold of $Q \times Q$ of dimension $n + k$ that contains the diagonal, and, in addition, a continuous distribution on Q. The principle seeks a discrete curve that is a critical point of the discrete action sum subject to variations that lie in a continuous distribution. One then applies the discrete analog of the tangent to the curve lying in the continuous distribution, namely, that the discrete tangent lift must lie in the discrete constraint distribution in $Q \times Q$. This determines the discrete Lagrange-d'Alembert (DLA) equations of motion. In Proposition 3 we give conditions for the DLA principle to define an (in general, implicit) integrator; however, since the flow is not symplectic, we do not expect the integrator to preserve a corresponding symplectic form on $Q \times Q$ as in the discrete Euler-Lagrange [24] equations. Instead we study its reversibility properties.

We consider nonholonomic systems that admit reversing symmetries and demand that the integrators for such systems also preserve an analogous reversing symmetry. A reversing symmetry on TQ is just an involution of the tangent bundle taking tangent vectors to their negatives so that a nonholonomic system admits a reversing symmetry when the Lagrangian function is invariant under this $\mathbb{Z}^2$-action. (Note that the distribution is automatically invariant.) The flow then inherits the property that $q(t)$ is a solution if and only if $R \circ q(-t)$ is a solution where R is the reversing map. The discrete analog of the $\mathbb{Z}^2$-action is just transposition in $Q \times Q$. We show in Proposition 4 that the discrete flow is reversible when the discrete Lagrangian and discrete constraint distribution are invariant under this transposition.

We consider several examples of DLA integrators in Section 4. We study the basic geometric properties of a typical class of integrators determined by a ("finite difference") map $\varphi\colon N_0(\Delta) \to TQ$ where $N_0(\Delta)$ is a neighborhood of the diagonal $\Delta \subset Q \times Q$. Symmetry properties of the DLA equations are then inherited from symmetry properties of the continuous equations provided φ is equivariant with respect to the corresponding actions. For example, when φ is transposition equivariant, we retain reversibility of the discrete integrator from reversibility of the continuous flow. The finite difference map φ can be constructed naturally from the geodesic flow corresponding to a Riemannian structure on Q. On $\mathbb{R}^n$ this leads to the two most basic integrators, one first-order, linearly implicit, and nonreversible, and one second-order, reversible, and implicit. The most practical method, which is second-order, reversible, and implicit only in the Lagrange multipliers, is constructed in a slightly different way using two different finite difference maps.

In Section 5 we apply these integrators to three nonholonomic systems. The first system has $Q = \mathbb{R}^3$, being the lowest dimensional Euclidean space with a nonintegrable distribution, and harmonic oscillator Lagrangian. The dynamics are integrable and a reversible DLA integrator captures their qualitative properties precisely. By contrast, neither a nonreversible DLA integrator nor a reversible non-DLA integrator preserve the same qualitative properties. The second system also has $Q = \mathbb{R}^3$, but is nonintegrable. A reversible DLA integrator preserves quasiperiodic and chaotic orbits over long times, with performance reminiscent of a symplectic integrator. The third system has $Q = \mathbb{R}^7$ and is fully chaotic. A reversible DLA integrator is still markedly better than standard methods but does show some energy drift.

In Section 6, we study integrators for symmetric nonholonomic systems on Lie groups G. Under reduction, the continuous equations of motion project to a vector field on the Lie algebra $\mathfrak{g}$ obtained by projecting the unconstrained (Euler-Poincaré) vector field on $\mathfrak{g}$ to the velocity constraint distribution with respect to the kinetic energy metric [16]. On the other hand, the discrete reduction of Euler-Lagrange equations for Lie groups leads to the discrete Euler-Poincaré (DEP) equations [5], [23]. Here, we obtain for nonholonomic systems a reduced discrete flow on G depending on a choice of symmetric discrete constraint distribution and discrete Lagrangian. The reduced equations then depend explicitly on the subspace of $\mathfrak{g}$ corresponding to the original nonintegrable distribution on G. We also obtain an explicit reconstruction principle to obtain the unreduced discrete flow on $G \times G$ from the reduced flow on G. These equations naturally generalize the DEP equations in that if we take the distribution to be TG (and corresponding discrete constraint distribution to be all of $G \times G$), we recover the DEP equations.

(The systems considered here are complementary to the so-called generalized Chaplygin systems in which the symmetry directions and constraint distribution have zero intersection and together span the tangent bundle, the reduction and discrete reduction of which were considered in [8] and [14]. In our case, the constraint distribution is contained in the symmetry distribution.)

Section 7 is somewhat independent of the rest of the paper. It concerns the exact flow of the Lagrange-d'Alembert equations. By fixing the time step, this can be regarded as a kind of integrator, and one is interested in its properties not only as part of the fundamental study of nonholonomic mechanics but so that they can be mimicked by actual, practical integrators. In Theorem 5 we obtain a set of equations, the *exact discrete Lagrange-d'Alembert* (EDLA) equations, satisfied by the flow. This involves the construction of an *exact* discrete constraint submanifold of $Q \times Q$ and its fibers over the two natural projections and their intersections, which, provided the Lagrangian is reversible, we prove are submanifolds with natural dynamical interpretations. Let $\tau\colon TQ \to Q$ be the tangent projection and let $\psi^t\colon C \to C$ be the flow of the Lagrange-d'Alembert equations. Given a time h solution (q_0, q_1) to the Lagrange-d'Alembert equations, that is $q_1 = \tau \circ \psi^h(v_{q_0})$ for some $v_{q_0} \in C$, we obtain an implicit equation satisfied by $q_2 = \tau \circ \psi^{2h}(v_{q_0})$

There are some crucial differences between the EDLA and DLA equations. The EDLA equations do not determine q_2 from (q_0, q_1); however, given a pair (q_0, q_2) we prove that the EDLA equations determine a q_1 with the property that the discrete time h flow contains the orbit sequence (q_0, q_1, q_2). Finally, we define a discrete Legendre transformation so that the momentum for the pair (q_0, q_1) matches the momentum for (q_1, q_2) where q_1 is the locally unique solution of the EDLA equations.

2. Review of Smooth Theory

2.1. The Lagrange-d'Alembert Principle

In this section we review the fundamental principle of nonholonomic mechanics, the Lagrange-d'Alembert principle, and the associated Lagrange-d'Alembert equations. The data for a constrained mechanical system is (Q, L, C), where Q is the configuration space, C is a k-dimensional distribution on Q, which can be thought of as a subbundle of TQ, and $L\colon TQ \to \mathbb{R}$ is the Lagrangian. Fixing initial and final points $q(0), q(T) \in Q$, and fixing a time interval $I = [0, T]$, we then consider the space of all smooth maps from I to Q joining $q(0)$ to $q(T)$. Denote this space by $\Omega(q(0); q(T))$ and denote a point in this space by $q(t)$. As for unconstrained mechanics, we consider the action functional on this space given by

$$S(q(t)) = \int_0^T L(q(t), \dot{q}(t))\, dt. \tag{2.1}$$

We then look for critical points $q(t)$ of the action functional with respect to variations that lie in the constraint distribution. This determines a family of curves. We then choose the unique curve that also satisfies the condition

$$\dot{q}(t) \in C_{q(t)}, \tag{2.2}$$

for all $t \in I$. The constraint distribution is described by the intersection of the kernels of $n - k$ one-forms in general position. That is,

$$C_q = \bigcap_{j=1}^{n-k} \ker A_j(q), \tag{2.3}$$

where $A_j \in \Lambda^1(Q)$ for $j \in \{1, \ldots, n-k\}$.

In local coordinates q^i, $i \in \{1, \ldots, n\}$, we represent the one-forms as $\sum_{i=1}^n A_{ji} dq^i$, $l \in \{1, \ldots, n-k\}$. If we introduce Lagrange multipliers, $\lambda_1, \ldots, \lambda_{n-k}$, the condition that $q(t)$ is critical with respect to variations lying in the distribution is equivalent to the equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \sum_{j=1}^{n-k} \lambda_j A_{ji}. \tag{2.4}$$

These are $2n$ differential equations with $2n + n - k$ variables. The remaining $n - k$ equations are obtained by imposing the constraint condition on the curve $q(t)$,

$$\sum_i A_{ji}(q(t))\dot{q}^i(t) = 0, \tag{2.5}$$

for $j \in \{1, \ldots, n-k\}$. Equations (2.4) and (2.5) constitute the *Lagrange-d'Alembert equations*.

2.2. Time Reversibility of Lagrange-d'Alembert Flow

An important property of nonholonomic mechanics for Lagrangians that are of the form kinetic minus potential energy is that they have a time-reversing symmetry. This is the result of a $\mathbb{Z}_2$ symmetry of the system. The $\mathbb{Z}^2$-action is generated by the involutive diffeomorphism $R\colon TQ \to TQ$ given by

$$R(v_q) = -v_q. \tag{2.6}$$

Notice that this diffeomorphism of TQ is not the tangent lift of any diffeomorphism on Q since it covers the identity diffeomorphism, but is not itself the identity. A nonholonomic system is symmetric with respect to this $\mathbb{Z}^2$-action provided the Lagrangian function is R-invariant, since the velocity distribution, being a vector bundle over Q, is automatically R-invariant. For example, any nonholonomic system whose Lagrangian is of the form kinetic minus potential energy is symmetric with respect to this action. The dynamical consequences of this $\mathbb{Z}^2$-action are contained in the following proposition, the proof of which we omit.

Proposition 1. *Let (Q, L, C) be a constrained mechanical system. Suppose the Lagrangian $L\colon TQ \to \mathbb{R}$ is R-invariant, $L \circ R = L$, where R is the involutive diffeomorphism of TQ. Let t be such that the flow $\psi_t\colon C \to C$ exists. We then have*

$$R \circ \psi_t = \psi_{-t} \circ R. \tag{2.7}$$

Time-reversing symmetry will have important consequences for the discretization of nonholonomic systems, and will be a key property that determines the success of the integrator to behave well with respect to energy conservation. From the viewpoint of geometric integration it is expected that if the continuous system has time-reversing symmetry, then, as this is a geometric property of the flow, we should require that the integrator also admit such a symmetry. In a subsequent section we will obtain, from the discrete variational principle, a discrete time-reversing symmetry for the nonholonomic integrator.

2.3. Formulation on T^*Q

We briefly formulate the constrained mechanical system on the cotangent bundle and remark that this system fails to exhibit, in the case that the constraints are nonintegrable, the geometric invariance properties of a Hamiltonian system, namely that the flow is symplectic. Starting with the data (Q, L, C) and making the assumption that the Legendre transformation $\mathbb{F}L\colon\ TQ \to T^*Q$ is invertible, we can form both the subbundle D of T^*Q,

$$D := \mathbb{F}L(C), \tag{2.8}$$

and the Hamiltonian function $H \in C^\infty(T^*Q)$,

$$H := (\mathbf{i}(Z)\mathbf{d}L - L) \circ \mathbb{F}L^{-1},$$

where Z is the *Liouville* vector field on TQ given in coordinates (x, v) by $Z = \sum v_i \frac{\partial}{\partial v_i}$, the unique linear radial vector field tangent to the fibers of TQ. We can now construct the so-called projection bundle W [20], [21], which is a subbundle of $T(T^*Q)$ complementary to TD. To construct W, one first considers the annihilator C^0 of C which is a subbundle of T^*Q, and then forms the pull back bundle with respect to the projection $\pi\colon\ T^*Q \to Q$. The pull back π^*C^0 is a subbundle of $T^*(T^*Q)$. Finally, one takes

$$W := \tilde{\omega}^{-1}(\pi^* C^0),$$

where $\tilde{\omega}^{-1}\colon\ T^*(T^*Q) \to T(T^*Q)$ is induced from the canonical symplectic form ω on T^*Q. W is then a subbundle of $T(T^*Q)$.

Definition 1. L is *normal* provided the matrix of second partial derivatives of the Lagrangian restricted to the fibers of TQ is positive definite, that is, provided the matrix

$$\left(\frac{\partial^2 L_q}{\partial v^i \partial v^j}(v)\right), \quad 1 \le i, j \le n, \tag{2.9}$$

where $L_q = L|_{T_q Q}$, is positive definite.

In this case, we have the following theorem.

Theorem 1. *Let (Q, L, C) be a nonholonomic mechanical system with a Lagrangian that is both regular and normal. We then have*

$$TD \oplus W = T(T^*Q), \tag{2.10}$$

and consequently the restriction of the Hamiltonian vector field X_H to the constraint subbundle $D \subset T^*Q$ splits as a direct sum of two vector fields tangent to D and W, respectively. That is,

$$X_H|_D = X_D + X_W. \tag{2.11}$$

The vector field X_D tangent to D encodes the dynamics of the nonholonomic system and is related via the Legendre transformation to the Lagrange-d'Alembert flow on TQ.

Namely, a curve $c\colon I \to Q$ is a solution of the Lagrange-d'Alembert equations for (Q, L, C) (equations (2.4) and (2.5)) if and only if the curve $\tilde{c}\colon I \to T^*Q$ defined by

$$\tilde{c}(t) = \mathbb{F}L \circ \frac{d}{dt}c(t) \tag{2.12}$$

is an integral curve of the vector field X_D on D.

Proof. The proof is given in [20]. □

3. The Discrete Lagrange-d'Alembert Principle

3.1. Review of Discrete Euler-Lagrange Theory

Before embarking on discretizing the Lagrange-d'Alembert principle in order to obtain a discrete analog of the continuous equations of motion (2.4) and (2.5), we recall the schemes of Moser and Veselov [27] and Marsden et al. [24], [23]. These discretize unconstrained mechanical systems by approximating curves on Q with sequences of points and use a corresponding discrete variational principle to pick out an action-minimizing sequence.

Fix $q_0, q_N \in Q$ for some integer N and fix a time step h. Consider the space of sequences $\Omega(q_0; q_N) \simeq Q^{N-1}$, where each element is a discrete path joining q_0 to q_N. Denote an element in $\Omega(q_0; q_N)$ by $[q]$ which we will alternatively write as $q_0, q_1, \ldots, q_{N-1}, q_N$, adjoining the end points. We think of each pair (q_i, q_{i+1}), $i \in \{0, \ldots, N\}$ as an evaluation of the curve at times ih and $(i+1)h$, respectively. The discretization depends on the choice of a discrete Lagrangian, $L_d\colon Q \times Q \to \mathbb{R}$, from which we form the discrete action sum,

$$S_d([q]) = \sum_{i=0}^{N-1} L_d(q_i, q_{i+1}). \tag{3.1}$$

We compute the critical point of this action sum with respect to arbitrary variations of the discrete curve. That is, we attempt to solve the equation

$$T_{[q]}S_d \cdot \delta[q] = 0, \tag{3.2}$$

where $\delta[q] \in T_{(q_1,\ldots,q_{N-1})}Q^{N-1}$ is an arbitrary variation of the discrete curve $[q]$. This is equivalent to the $N - 1$ *discrete Euler-Lagrange* (DEL) equations,

$$D_2L_d(q_{i-1}, q_i) + D_1L_d(q_i, q_{i+1}) = 0, \tag{3.3}$$

where i ranges from 1 to $N - 1$. Moreover, under mild conditions on L_d, equations (3.3) can be solved sequentially. The first equation ($i = 1$) determines q_2 from q_0 and q_1, which is fed into the second equation ($i = 2$) to determine q_3 from q_1 and q_2, and so on.

Equation (3.3) is a discrete version of the Euler-Lagrange equations on TQ and it determines, provided L_d is sufficiently regular, a diffeomorphism, or discrete flow, $F_{L_d}\colon Q \times Q \to Q \times Q$ retaining the second-order nature of the Euler-Lagrange equations in that $\pi_1 \circ F_{L_d} = \pi_2$ where $\pi_i\colon Q \times Q \to Q$ are projections on the ith factors. Consequently, projecting the solution curve in $Q \times Q$ to a curve in Q, one can then take the lift of this curve to $Q \times Q$ and recover the original curve. This is the discrete analog of the flow being second order.

One can obtain a momentum-matching interpretation of the DEL equations by introducing the discrete Legendre transformations $\mathbb{F}^+ L_d\colon Q \times Q \to T^*Q$ and $\mathbb{F}^- L_d\colon Q \times Q \to T^*Q$ given by

$$\mathbb{F}^+ L_d(q_0, q_1) = (q_1, D_2 L_d(q_0, q_1)) \in T^*_{q_1} Q, \tag{3.4}$$

which is a bundle map with respect to $\pi_2\colon Q \times Q \to Q$, and

$$\mathbb{F}^- L_d(q_0, q_1) = (q_0, -D_1 L_d(q_0, q_1)) \in T^*_{q_0} Q, \tag{3.5}$$

which is a bundle map with respect to $\pi_1\colon Q \times Q \to Q$. There are now two important observations about the DEL equations (3.3). First, denote by $p^+(q_0, q_1)$ and $p^-(q_0, q_1)$ the momenta of a point in $Q \times Q$ defined through the discrete Legendre transformations. We can interpret (3.3) as the equation demanding that the momentum on the initial point matches the momentum on the updated point by the discrete flow map, F_{L_d}. That is,

$$p^+(q_{i-1}, q_i) = p^-(q_i, q_{i+1}) \tag{3.6}$$

for each i. There is, thus, a well-defined momentum value for each discrete time i. Notice that we can write $p^+(q_i, q_{i+1}) = \mathbb{F}^+ L_d(F_{L_d}(q_{i-1}, q_i))$. Second, let us recall that using the discrete Legendre transformations we can recast the discrete flow on $Q \times Q$ as a discrete flow on T^*Q by $(q_i, p_i) \mapsto (q_{i+1}, p_{i+1})$, where $p_i = \mathbb{F}^- L_d(q_i, q_{i+1})$ and $p_{i+1} = \mathbb{F}^+ L_d(q_i, q_{i+1})$. This gives the interpretation of L_d as a generating function for the symplectic transformation on T^*Q.

The link to the Euler-Lagrange equations is made more explicit by way of the *exact* discrete Lagrangian L_d^e [24], the action integral of the solution to the Euler-Lagrange equations joining the points q_0 and q_1 in time h. Its discrete flow is the time-h evaluation of the Euler-Lagrange flow on TQ. More precisely, we have

Theorem 2 (Exact discrete Lagrangian correspondence theorem). *Let $L_d^e\colon Q \times Q \to \mathbb{R}$ be the exact discrete Lagrangian. The DEL equations (3.3) then produce a discrete flow on $Q \times Q$ which is equal to the discrete time evaluation of the actual solution to the Euler-Lagrange equations.*

As a consequence of this theorem, one sees that a choice of discrete Lagrangian is really an approximation of the action integral $\int_0^h L(q(t), \dot{q}(t))\, dt$ where $(q(t), \dot{q}(t))$ is a solution of the Euler-Lagrange equations with $q(0) = q_0$ and $q(h) = q_1$.

3.2. The Discrete Lagrange-d'Alembert Principle

We start by defining a discrete nonholonomic system which will, given sufficient regularity, determine a discrete second-order flow on a submanifold of $Q \times Q$ and generalize the discrete Euler-Lagrange equations (3.3). We will later develop methods to define a discrete nonholonomic system from a given continuous one.

Definition 2. A discrete nonholonomic system is given by the quadruple (Q, L_d, C_d, A_d) where

1. C_d is a submanifold of $Q \times Q$ of dimension $n + k$ with the additional property that
$$\Delta = \{(q, q) \mid q \in Q\} \subset C_d.$$
We call C_d the discrete constraint distribution.
2. A_d denotes a set of $n - k$ independent one-forms, $A_1, \ldots, A_{n-k}$, on Q.
3. $L_d\colon\ Q \times Q \to \mathbb{R}$ is the discrete Lagrangian.

If we are modeling a smooth nonholonomic system (Q, L, C), we often find the following condition linking the tangent space of the discrete manifold along the diagonal with the continuous constraint distribution: for all $q_0 \in Q$,

$$0 \times v_{q_0} \in T_{(q_0,q_0)}C_d \iff v_{q_0} \in C_{q_0}. \tag{3.7}$$

If we are given a smooth nonholonomic system, the requirement (3.7) links the discrete constraint submanifold to the continuous distribution, although it does not uniquely specify it. It is satisfied by all the examples we shall consider.

Example 1. *If the continuous constraint is integrable, then the natural discrete constraint distribution is the submanifold of $Q \times Q$ given by*

$$\bigcup_l \mathcal{F}_l \times \mathcal{F}_l, \tag{3.8}$$

the disjoint union of the direct product of the leaves $\mathcal{F}_l$ of the distribution on Q. A dimension count of this submanifold gives $k + k + \dim(\text{leafspace}) = k + k + (n - k) = n + k$, as required. This discrete constraint distribution is reversible (see Section 3.3).

With this discrete data, a discrete Lagrange-d'Alembert principle was proposed in [14] which leads to a set of discrete equations on $Q \times Q$, which, assuming enough regularity, leads to a second-order diffeomorphism $F_{L_d}\colon\ Q \times Q \to Q \times Q$.

Definition 3 (Discrete Lagrange-d'Alembert (DLA) Principle). A discrete curve $[q]$ satisfies the discrete Lagrange-d'Alembert principle provided that it is a critical point of the discrete action sum $S_d([q]) = \sum_{i=0}^{N-1} L_d(q_i, q_{i+1})$ with respect to variations $\delta[q]$, vanishing at the end points q_0 and q_N, that lie in $\cap_{j=i}^{n-k} \ker A_j(q)$, i.e., for each $A_j \in A_d$, $A_j(\delta[q]) = 0$, and which also satisfies $(q_i, q_{i+1}) \in C_d$ for $i \in \{0, \ldots, N - 1\}$.

It is shown in [14] that this principle leads to the following.

Proposition 2. *The DLA principle leads to the following set of equations. For each $i \in \{0, \ldots, N - 1\}$,*

$$D_2 L_d(q_i, q_{i+1}) + D_1 L_d(q_{i+1}, q_{i+2}) = \sum_{j=1}^{n-k} \lambda_j A_j(q_{i+1}), \tag{3.9}$$

$$(q_i, q_{i+1}) \in C_d. \tag{3.10}$$

We refer to (3.9), (3.10) as the DLA equations. We next look more closely at equations (3.9) and (3.10) to formulate a regularity condition guaranteeing the existence of a unique discrete flow map from the discrete Lagrange-d'Alembert principle which is second order and therefore satisfies

$$F_{L_d}(q_{i-1}, q_i) = (q_i, q_{i+1}), \tag{3.11}$$

where q_{i+1} satisfies (3.9) and (3.10) provided that $(q_{i-1}, q_i) \in C_d$. This regularity condition is equivalent to the one formulated in [14], but is often easier to check. It is useful to first make the following definition, which follows naturally from equation (3.9).

Definition 4. For each $(q_0, q_1) \in Q \times Q$, we define the map $\psi_{(q_0,q_1)}\colon Q \to C^*_{q_1}$ by

$$\psi_{(q_0,q_1)}(q_2) = \iota^*_{q_1}(D_2 L_d(q_0, q_1) + D_1 L_d(q_1, q_2)), \tag{3.12}$$

where $C^*_{q_1}$ is the dual space of $C_{q_1} \subset T_{q_1}Q$ and where $\iota^*_{q_1}\colon T^*_{q_1}Q \to C^*_{q_1}$ is the projection map dual to the linear inclusion $\iota_{q_1}\colon C_{q_1} \hookrightarrow T_{q_1}Q$.

Proposition 3. *Let $(q_0, q_1) \in C_d$. Let $\pi_1\colon Q \times Q \to Q$ be projection on the first factor. Suppose $\pi_1|_{C_d}\colon C_d \to Q$ is a submersion. The forward discrete flow map F_{L_d} is then guaranteed to exist locally uniquely provided for each $q_2 \in \psi^{-1}_{(q_0,q_1)}(0) \cap (\pi_1|_{C_d})^{-1}(q_1)$, for each nonzero $v_{q_2} \in T_{q_2}C_d(q_1)$,*

$$\left\langle D_2 D_1 L_d(q_1, q_2) \cdot v_{q_2}, v_{q_1} \right\rangle \neq 0, \tag{3.13}$$

for all $v_{q_1} \in C_{q_1}$. When this condition holds for all $q_1 \in Q$, the discrete Lagrange-d'Alembert equations produce a uniquely defined diffeomorphism $F_{L_d}\colon C_d \to C_d$.

Proof. Consider the first discrete Lagrange-d'Alembert equation,

$$D_2 L_d(q_0, q_1) + D_1 L_d(q_1, q_2) = \sum_{j=1}^{n-k} \lambda_j A_j(q_1). \tag{3.14}$$

Since the $\lambda_j(q_1)$ are arbitrary, and the $A_j(q_1)$ are a basis for $\ker \iota^*_{q_1} = C^0_{q_1} \subset T^*_{q_1}Q$, the solutions of this equation can be written equivalently as the set $\psi^{-1}_{(q_0,q_1)}(0)$. By the regular value theorem, this set is a submanifold if 0 is a regular value of $\psi_{(q_0,q_1)}$. Since the image of $\psi_{(q_0,q_1)}$ is a linear space, we have $T_{q_2}\psi_{(q_0,q_1)}\colon T_{q_2}Q \to C^*_{q_1}$. This map is given by

$$T_{q_2}\psi_{(q_0,q_1)} = T_{q_2}(\iota^*_{q_1}(D_1 L_d(q_1, q_2))) = \iota^*_{q_1} \circ D_2 D_1 L_d(q_1, q_2). \tag{3.15}$$

Notice that $D_2 D_1 L_d(q_1, q_2) \cdot v_{q_2}$ is an element of $T^*_{q_1}Q$ so that the map $v_{q_2} \mapsto \iota^*_{q_1} \circ D_2 D_1 L_d(q_1, q_2) \cdot v_{q_2}$ is well defined. 0 is a regular value when for all $q_2 \in \psi^{-1}_{(q_0,q_1)}(0)$ this map is surjective on $C^*_{q_1}$. In this case we see that $\psi^{-1}_{(q_0,q_1)}(0)$ is a submanifold of dimension $n - \dim(C^*_{q_1}) = n - k$. On the other hand, consider the set

$$C_d(q_1) = \{q \mid (q_1, q) \in C_d\} = \pi_2 \circ (\pi_1|_{C_d})^{-1}(q_1) \simeq (\pi_1|_{C_d})^{-1}(q_1), \tag{3.16}$$

where π_1 and π_2 are the projections from $Q \times Q$ to Q. Since $\pi_1|_{C_d}$ is assumed to be a submersion, this set is a submanifold of dimension $n + k - n = k$. The discrete flow is guaranteed to exist and be well defined when these two submanifolds of complementary dimension intersect transversely. We then have

$$\begin{aligned}
&\psi^{-1}_{(q_0,q_1)}(0) \;\cap\; (\pi_1|_{C_d})^{-1}(q_1) \text{ transversely at } q_2 \\
&\qquad\iff \ker T\psi_{(q_0,q_1)}(q_2) \cap \ker T_{q_2}\pi_1|_{C_d} = 0 \\
&\qquad\iff \text{for all } v_{q_2} \in \ker T_{q_2}\psi_{(q_0,q_1)},\; T_{(q_1,q_2)}\pi_1 \cdot v_{q_2} \neq 0 \\
&\qquad\iff \text{for all } v_{q_2} \in T_{q_2}\pi_1^{-1}(q_1),\; \langle D_2 D_1 L_d(q_1,q_2)\cdot v_{q_2}, v_{q_1}\rangle \neq 0 \\
&\qquad\qquad \text{for all } v_{q_1} \in C_{q_1}.
\end{aligned}$$

Therefore, provided this holds, the discrete flow will map (q_0, q_1) to the point (q_1, q_2) where q_2 is the locally unique point of intersection of the two transverse submanifolds. □

3.3. Discrete Reversibility

Since reversibility is a key geometric property of the continuous Lagrange-d'Alembert equations, we desire that our integrator maintains a corresponding discrete analog of reversibility. In the following we formulate discrete reversibility and verify a natural condition on L_d and C_d guaranteeing that the discrete Lagrange-d'Alembert equations that they generate are discrete reversible.

Definition 5. Let $R_d\colon\ Q \times Q \to Q \times Q$ denote the diffeomorphism

$$R_d(q_0, q_1) = (q_1, q_0),$$

which is the natural discrete counterpart to $R\colon\ TQ \to TQ$ in Proposition 1. C_d is reversible when $R_d(C_d) = C_d$. L_d is reversible when $L_d \circ R_d = L_d$.

Proposition 4. *Suppose that L_d and C_d are reversible and satisfy the regularity condition of Proposition 3 so that the discrete flow F_{L_d} is a well-defined diffeomorphism of C_d. Then the discrete flow F_{L_d} determined by the discrete Lagrange-d'Alembert equations (3.9) and (3.10) is discrete reversible, that is,*

$$F_{L_d} \circ R_d \circ F_{L_d} = R_d. \tag{3.17}$$

Proof. Starting with $(q_{i-1}, q_i) \in C_d$, let $(q_i, q_{i+1}) = F_{L_d}(q_{i-1}, q_i)$. We need to show that

$$F_{L_d}(q_{i+1}, q_i) = (q_i, q_{i-1}),$$

so that

$$F_{L_d}(R_d(F_{L_d}(q_{i-1}, q_i))) = F_{L_d}(q_{i+1}, q_i) = (q_i, q_{i-1}) = R_d(q_{i-1}, q_i),$$

from which equation (3.17) follows. Now, from the reversibility of C_d, we know that since $(q_i, q_{i+1}) = F_{L_d}(q_{i-1}, q_i) \in C_d$, it follows that $(q_{i+1}, q_i) \in C_d$. Because of the regularity assumption on L_d and C_d, we know from Proposition 3 that given $(q_{i+1}, q_i) \in C_d$ there exists a unique $\bar{q}_{i-1}$ that satisfies $F_{L_d}(q_{i+1}, q_i) = (q_i, \bar{q}_{i-1})$. We will show that in fact $\bar{q}_{i-1} = q_{i-1}$. Since L_d is R_d-invariant, we have

$$D(L_d \circ R_d) = DL_d = D_1 L_d + D_2 L_d. \tag{3.18}$$

On the other hand,

$$\begin{aligned} D(L_d \circ R_d)(q_0, q_1) \cdot (v_{q_0}, v_{q_1}) &= DL_d(q_1, q_0) \circ T_{(q_0,q_1)} R_d \cdot (v_{q_0}, v_{q_1}) \\ &= DL_d(q_1, q_0) \cdot (v_{q_1}, v_{q_0}) \\ &= D_1 L_d(q_1, q_0) \cdot v_{q_1} + D_2 L_d(q_1, q_0) \cdot v_{q_0}. \end{aligned}$$

Comparing with equation (3.18), we conclude

$$D_1 L_d(q_1, q_0) = D_2 L_d(q_0, q_1) \quad \text{and} \quad D_1 L_d(q_0, q_1) = D_2 L_d(q_1, q_0). \tag{3.19}$$

We know that q_{i+1} satisfies the equations

$$D_1 L_d(q_i, q_{i+1}) + D_2 L_d(q_{i-1}, q_i) = \sum_{j=1}^{n-k} \lambda_j A_j(q_i),$$
$$(q_i, q_{i+1}) \in C_d.$$

However, using equations (3.19), we can rewrite this first equation as

$$D_2 L_d(q_{i+1}, q_i) + D_1 L_d(q_i, q_{i-1}) = \sum_{j=1}^{n-k} \lambda_j A_j(q_i). \tag{3.20}$$

On the other hand, $\bar{q}_{i-1}$ satisfies

$$D_1 L_d(q_i, \bar{q}_{i-1}) + D_2 L_d(q_{i+1}, q_i) = \sum_{j=1}^{n-k} \bar{\lambda}_j A_j(q_i),$$

together with $(q_i, \bar{q}_{i-1}) \in C_d$. Comparing with equation (3.20), we see that q_{i-1} satisfies this equation with $\bar{\lambda}_j = \lambda_j$. Furthermore, $(q_i, q_{i-1}) \in C_d$ since $(q_{i-1}, q_i) \in C_d$. Therefore, by uniqueness $\bar{q}_{i-1} = q_{i-1}$, from which the proposition follows. $\square$

4. Construction of Integrators via Finite Difference Maps

As we have defined it, a discrete nonholonomic system on Q requires L_d, C_d, and A_d to be specified. However, when discretizing a given continuous system (Q, L, C), it is always possible to choose A_d to be the collection of one-forms that determines C. In many cases L_d and C_d can be specified through a *finite difference map* φ.

Definition 6. A *finite difference map* φ is a diffeomorphism $\varphi\colon N_0(\Delta) \to T_0 Q$, where $N_0(\Delta)$ is a neighborhood of the diagonal Δ in $Q \times Q$ and $T_0 Q$ denotes a neighborhood of the zero section of TQ, which satisfies the following:

1. $\varphi(\Delta)$ is the zero section of TQ;
2. $\tau \circ \varphi(N_0(\Delta)) = Q$; and
3. On the diagonal,
$$\tau \circ \varphi|_\Delta = \pi_1|_\Delta = \pi_2|_\Delta,$$
where π_i are the projections from $Q \times Q$ to Q.

Eq. (4.8) gives a simple example of a finite difference map.

A finite difference map φ is *not* in general a bundle map with respect to either π_i. Nevertheless, φ determines a natural foliation of $N_0(\Delta)$ corresponding to the foliation of T_0Q by the fibers of the tangent projection $\tau\colon TQ \to Q$ restricted to the neighborhood of the zero section T_0Q. We write $\tau_0 := \tau|_{T_0Q}$. We then have the following definition.

Definition 7 (Vertical foliation of $N_0(\Delta)$). For $(q_0, q_1) \in N_0(\Delta)$,

$$\mathcal{L}_{(q_0,q_1)} := \varphi^{-1}_{\tau\circ\varphi(q_0,q_1)}(T_{\varphi(q_0,q_1)}Q \cap T_0Q), \tag{4.1}$$

where we use the notation $\varphi_q^{-1} := \varphi^{-1}|_{T_qQ\cap T_0Q}$.

Notice that since φ and hence φ^{-1} are diffeomorphisms, the leaves $\mathcal{L}_{(q_0,q_1)}$ are n-dimensional submanifolds of $N_0(\Delta)$. By construction, the leaves get mapped by φ to a fixed fiber in T_0Q and therefore all the points in a given leaf correspond to tangent vectors with the same base point in T_0Q. This is why we call $\mathcal{L}$ the vertical foliation of $N_0(\Delta)$. Curves in $N_0(\Delta)$ that lie on a fixed leaf of $\mathcal{L}$ correspond under the map φ to curves in the fiber of $\tau_0\colon T_0Q \to Q$ and their derivatives correspond to vertical tangent vectors lying in the kernel of the tangent map $T\tau_0\colon T(T_0Q) \to TQ$.

We collect some of the properties of the foliation in the following:

Proposition 5. *The leaves $\mathcal{L}$ determine a smooth foliation of $N_0(\Delta)$ and*

(i) *For $(q_0, q_1) \in N_0(\Delta)$, $\mathcal{L}_{(q_0,q_1)}$ intersects Δ in the unique point $(\bar q, \bar q)$ where $\bar q := \tau_0 \circ \varphi(q_0, q_1)$.*
(ii) *This intersection is transverse.*
(iii) *The tangent spaces to the leaves are given by*
$$T_{(q_0,q_1)}\mathcal{L}_{(q_0,q_1)} = T_{\varphi(q_0,q_1)}\varphi_{\bar q}^{-1}\left(V_{\varphi(q_0,q_1)}(TTQ)\right),$$
where $V(TTQ)$ is the vertical subbundle of TTQ given by ker $T\tau$.

Proof. First, since φ^{-1} is a diffeomorphism, $\varphi_q^{-1}\colon T_qQ\cap T_0Q \to N_0(\Delta)$ is smooth and invertible on its range and $T\varphi_q^{-1}$ has constant rank. It follows that φ_q^{-1} is a diffeomorphism so that each leaf $\mathcal{L}_{(q_0,q_1)}$ is a smooth n-dimensional submanifold. The leaves are disjoint since for different q's, the spaces $T_qQ \cap T_0Q$ are disjoint and φ^{-1} is a diffeomorphism. Furthermore, each point $(q_0, q_1) \in N_0(\Delta)$ lies on the unique (again, since φ is a diffeomorphism) leaf $\mathcal{L}_{(q_0,q_1)}$ given in the definition. Since φ_q^{-1} is a diffeomorphism, it sends only 0_q to the diagonal, proving (i). To prove that the intersection is transverse, i.e., that $T_{(\bar{q},\bar{q})}\mathcal{L}_{(q_0,q_1)} \cap T_{(\bar{q},\bar{q})}\Delta = 0$, suppose there exists $(\bar{q}(t), \bar{q}(t))$ a curve in $\mathcal{L}_{(q_0,q_1)}$ with $\bar{q}(0) = \bar{q}$. Then, for all t, we have $\varphi(\bar{q}(t), \bar{q}(t)) = 0 \in T_{\bar{q}}Q$. However, since φ is a diffeomorphism, we must have $\bar{q}(t) = \bar{q}$ for all t, which implies

$$\left.\frac{d}{dt}\right|_{t=0} (\bar{q}(t), \bar{q}(t)) = 0,$$

from which (ii) follows. By definition of the leaves, it is clear that every tangent vector to $(q_0, q_1) \in \mathcal{L}_{(q_0,q_1)}$ is given by

$$\left\{ \left.\frac{d}{dt}\right|_{t=0} \varphi_{\bar{q}}^{-1}(\varphi(q_0, q_1) + tv_{\bar{q}}) \colon\ v_{\bar{q}} \in T_{\bar{q}}Q \right\} \simeq T_{(q_0,q_1)}\mathcal{L}_{(q_0,q_1)}.$$

Finally, since $\varphi_{\bar{q}}^{-1}$ is a diffeomorphism, its derivative is injective, from which (iii) follows. □

Using a finite difference map φ, we can construct a continuous constraint distribution C_d from the continuous distribution C as demonstrated in the following proposition.

Proposition 6. *Given a diffeomorphism* $\varphi\colon N_0(\Delta) \to T_0Q$, *define the map* $F\colon N_0(\Delta) \to \mathbb{R}^{n-k}$ *by*

$$F = A_1 \circ \varphi \times \cdots \times A_{n-k} \circ \varphi =: \mathcal{A} \circ \varphi. \tag{4.2}$$

Then, F *is a submersion, so in particular,* 0 *is a regular value of* F, *and consequently* $C_d := F^{-1}(0)$ *is a well-defined discrete constraint submanifold.*

Proof. Since the A_j are one-forms on Q, as maps from TQ into $\mathbb{R}$, they have the property that they are linear maps on each fiber and, furthermore, for each $q \in Q$ they are linearly independent elements of the vector space T_q^*Q. Denote by $\mathbb{F}D_{v_q}A_j$ the fiber derivative of A_j at the point v_q. Linearity on the fiber then gives, for each $q \in Q$,

$$\mathbb{F}D_{v_q}A_j\left(\mathbb{L}(v_q, w_q)\right) := \left.\frac{d}{dt}\right|_{t=0} A_j(v_q + tw_q) = A_j(q)(w_q), \tag{4.3}$$

where $\mathbb{L}\colon TQ \otimes_Q TQ \to TTQ$ is the vertical lift operator which is a bundle map covering the identity and taking values in the vertical subbundle $V(TQ)$ of TTQ given by $V := \ker T\tau$ with τ the tangent projection. In fact it is easy to check that $\mathbb{L}$ is a bundle isomorphism $\mathbb{L}\colon TQ \otimes_Q TQ \simeq V(TQ)$.

Furthermore, since

$$T_{(q_0,q_1)}F = T_{\varphi(q_0,q_1)}A_1 \circ T_{(q_0,q_1)}\varphi \times \cdots \times T_{\varphi(q_0,q_1)}A_{n-k} \circ T_{(q_0,q_1)}\varphi, \tag{4.4}$$

we can use the fact that φ is a diffeomorphism from $N_0(\Delta)$ onto a neighborhood of the zero section of TQ to produce a subspace of $T_{(q_0,q_1)}(Q \times Q)$ that gets mapped by $T\varphi$ isomorphically to the vertical subbundle, $V(TQ)$. Of course, by the previous proposition, this subspace of $T_{(q_0,q_1)}(Q \times Q)$ is simply $T_{(q_0,q_1)}\mathcal{L}_{(q_0,q_1)}$. Finally, using the fact that the A_j are linear on the fibers of TQ (equation (4.3)), we see that $T(A_j|_{T_{\bar{q}}Q}) = A_j(\bar{q})$ so that, by equation (4.4), and the linear independence of the A_j, the image of $V_{(q_0,q_1)}$ under TF is all of $\mathbb{R}^{n-k}$, as required. □

The next proposition describes the tangent space of the discrete constraint submanifold along the diagonal Δ. Before proceeding, it is useful to remark that the discrete submanifold C_d admits a regular foliation induced by the $\mathcal{L}$ leaves as follows. Define for each $q \in Q, \mathcal{C}_q := \varphi_q^{-1}(C_q)$. Since φ_q is a diffeomorphism, the $\mathcal{C}_q$ are smooth submanifolds of $N_0(\Delta)$ of dimension k. Furthermore, they are contained in C_d since clearly for any $X \in \mathcal{C}_q, \mathcal{A} \circ \varphi(X) = 0$ by construction. It is also clear that $\bigcup_{q\in Q} \mathcal{C}_q = \mathcal{L} \cap C_d$ since $\mathcal{L}_{(q_0,q_1)} \cap C_d = \varphi_{\bar{q}}^{-1}(C_{\bar{q}})$. These leaves form the vertical foliation of C_d. We now have the following:

Proposition 7. *Let $C_d = F^{-1}(0)$, where F is given by the previous proposition. We then have the following properties of C_d.*

(i) *$T_{(q_0,q_0)}C_d$ contains a vertical subspace given by $T\varphi_{q_0}^{-1}\left(\mathbb{L}(0_{q_0}, C_{q_0})\right)$.*

(ii) *The tangent vector $0 \times v_{q_0} \in T_{(q_0,q_0)}N_0(\Delta)$ lies in $T_{(q_0,q_0)}C_d$ if and only if the following condition is satisfied. For the map defined by $\varphi_{q_0}(q) := \varphi(q_0, q)$, in any local trivialization $\varphi_{q_0}^{\text{loc}}\colon U_{q_0} \to \bar{U}_{q_0} \times \mathbb{R}^n$,*

$$T\pi_2 \circ D\varphi_{q_0}^{\text{loc}} \cdot v_{q_0} \in \ker A_j, \tag{4.5}$$

for each $j \in \{1, \dots, n-k\}$, where $\pi_2\colon \bar{U}_{q_0} \times \mathbb{R}^n \to \mathbb{R}^n$ is projection on the second factor.

Proof. Part (i) is clear since for each $v_{q_0} \in C_{q_0}$ we can construct the curve $\varphi_{q_0}^{-1}(tv_{q_0})$ which lies in $\mathcal{L}_{(q_0,q_0} \cap C_d$. Taking the derivative of this curve, we obtain $\frac{d}{dt}\big|_{t=0}\varphi_{q_0}^{-1}(tv_{q_0}) = T\varphi_{q_0}^{-1}\left(\mathbb{L}(0_{q_0}, v_{q_0})\right)$. Since $T\varphi_{q_0}^{-1}$ and $\mathbb{L}$ are isomorphisms, it follows that $T\varphi_{q_0}^{-1}\left(\mathbb{L}(0_{q_0}, C_{q_0})\right)$ is a k-dimensional subspace of $T_{(q_0,q_0)}C_d$.

To obtain the second condition, we compute in local coordinates as follows. Since the image of a point $(q_0, q_0) \in \Delta$ under the map φ is the zero vector in $T_{q_0}Q$, we take a chart domain around 0_{q_0} with local coordinates $(q^1, \dots, q^n, \frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n})$. In these coordinates we express the one-forms as $A_j = \Sigma_{l=1}^n A_{jl} dq^l$. Since φ_{q_0} maps a neighborhood U_{q_0} of q_0 smoothly into a neighborhood of $0_{q_0} \in T_0Q$, we have the following local coordinate expression,

$$\varphi_{q_0}(q) = (g_1(q), \dots g_n(q), f_1(q) \dots f_n(q)),$$

for smooth functions $g_i\colon U_{q_0} \to \mathbb{R}$ and $f_i\colon U_{q_0} \to \mathbb{R}$.

Now, given a curve $(q_0, q_0(t)) \in C_d$ through (q_0, q_0) tangent to $0 \times v_{q_0}$, we have $\mathcal{A} \circ \varphi(q_0, q_0(t)) = 0$ so that $0 \times v_{q_0} \in T_{(q_0,q_0)}C_d$ if and only if $0 \times v_{q_0} \in \ker T_{(q_0,q_0)}(\mathcal{A} \circ \varphi)$. Using the local coordinate expressions for φ_{q_0} and $\mathcal{A}$ we have, denoting $\bar{x}(t) := (g_1(q(t)), \dots, g_n(q(t)))$ (the local coordinate expression for $\tau \circ \varphi_{q_0}(q(t))$),

$$\begin{aligned}
\mathcal{A} \circ \varphi(q_0, q_0(t)) &= \langle \mathcal{A}(\bar{x}(t)), \varphi_{q_0}(q(t))\rangle \\
&= (\langle A_1(\bar{x}(t)), \varphi_{q_0}(q(t))\rangle, \dots, \langle A_{n-k}(\bar{x}(t)), \varphi_{q_0}(q(t))\rangle) \\
&= \left(\sum_{l=1}^{n} A_{1l}(\bar{x}(t)) f_l(q(t)), \dots, \sum_{l=1}^{n} A_{(n-k)l}(\bar{x}(t)) f_l(q(t))\right),
\end{aligned}$$

so that $0 \times v_{q_0} \in T_{(q_0,q_0)}C_d$ if and only if $\frac{d}{dt}\big|_{t=0} \sum_{l=1}^{n} A_{jl}(\bar{x}(t)) f_l(q(t)) = 0$ for each $j \in \{1, \ldots n-k\}$. However,

$$\begin{aligned} \frac{d}{dt}\bigg|_{t=0} \sum_{l=1}^{n} A_{jl}(\bar{x}(t)) f_l(q(t)) &= \sum_{l=1}^{n} (\mathbf{d}A_{jl}(q_0) \cdot \dot{\bar{x}}(0) f_l(q(0)) + A_{jl}(q_0)\mathbf{d} f_l(q_0) \cdot v_{q_0}) \\ &= \sum_{l=1}^{n} A_{jl}(q_0)\mathbf{d} f_l(q_0) v_{q_0}, \end{aligned}$$

because $f_l(q_0) = 0$. Therefore, $0 \times v_{q_0} \in T_{(q_0,q_0)}C_d$ if and only if $\sum_{l=1}^{n} A_{jl}(q_0)\mathbf{d} f_l(q_0) v_{q_0} = 0$ for each $j \in \{1, \ldots n-k\}$, which is equivalent to (4.5), proving (ii). □

A class of DLA integrators is now given by taking $A_d = A$, $L_d = L \circ \varphi$, and $C_d = (\mathcal{A} \circ \varphi)^{-1}(0)$ as in Proposition 6. Furthermore, if φ is reversible (i.e., equivariant with respect to the $\mathbb{Z}_2$-actions on $Q \times Q$ and TQ, i.e., $\varphi(R_d(q_1, q_2)) = R(\varphi(q_1, q_2))$ or $\varphi(q_2, q_1) = -\varphi(q_1, q_2)$), then by Proposition 4 the discrete nonholonomic system (Q, L_d, C_d, A_d) is also reversible.

Finite difference maps from geodesic flow. So far we have not specified how to construct finite difference maps φ. Here are two possibilities in the case that Q has a Riemannian metric. The first is nonreversible and the second is reversible. In the following section when $Q = \mathbb{R}^n$ with the Euclidean metric, these will provide our most basic DLA integrators.

Let ψ^t denote the geodesic flow on TQ. Let $N_0(\Delta)$ be a neighborhood of the diagonal such that each pair $(q_0, q_1) \in N_0(\Delta)$ is nonconjugate (such a neighborhood is always constructible when Q is compact, for example).

The nonreversible φ: $N_0(\Delta) \to T_0Q$ is defined by

$$\varphi(q_0, q_1) := v_{q_0}, \tag{4.6}$$

where v_{q_0} is the unique (since $(q_0, q_1) \in N_0(\Delta)$) velocity so that $\tau \circ \psi^h(v_{q_0}) = q_1$, while the reversible φ: $N_0(\Delta) \to T_0Q$ is defined by

$$\varphi(q_0, q_1) := \psi^{h/2}(v_{q_0}). \tag{4.7}$$

Since the geodesic flow is reversible, setting $v_{q_1} := \psi^h(v_{q_0})$, we have $v_{q_0} = \psi^{-h}(v_{q_1}) = R\left(\psi^h(-v_{q_1})\right)$, so that $\tau \circ \psi^h(-v_{q_1}) = q_0$ and therefore

$$\begin{aligned} \varphi(q_1, q_0) &= \psi^{h/2}(-v_{q_1}) = \psi^{h/2}(R \circ \psi^h(v_{q_0})) = \psi^{h/2} \circ \psi^{-h}(-v_{q_0}) \\ &= \psi^{-h/2}(-v_{q_0}) = \psi^{-h/2}(R(v_{q_0})) = R(\psi^{h/2}(v_{q_0})) = -\varphi(q_0, q_1), \end{aligned}$$

as required.

The φ in (4.6), however, is *not* in general reversible since $\varphi(q_1, q_0)$ is a tangent vector over the point q_1 and is not related to $\varphi(q_0, q_1)$, and we cannot conclude that $A_j(\varphi(q_1, q_0)) = 0$ when $A_j(\varphi(q_0, q_1)) = 0$.

An argument in favor of constructing integrators using finite difference maps is that when C is integrable, the "exact" choice of C_d given in Example 1 arises in this way for a particular choice of φ.

Proposition 8. *Let $C = T\mathcal{F}$ where $\mathcal{F}$ is a foliation of Q, that is, C is integrable. Suppose that $\mathcal{F}$ is geodesible, i.e., there exists a metric on Q such that each leaf $\mathcal{F}_l$ of $\mathcal{F}$ is a totally geodesic submanifold of Q. Then there exists a finite difference map φ: $N_0(\Delta) \to T_0 Q$ such that the constraint submanifold C_d determined by φ is $C_d = \bigcup_l \mathcal{F}_l \times \mathcal{F}_l$.*

Proof. Let $\varphi(q_0, q_1) = v_{q_0}$ as in equation (4.6). Then we have that $(q_0, q_1) \in C_d = (\mathcal{A} \circ \varphi)^{-1}(0)$ iff $\mathcal{A}\varphi(q_0, q_1) = 0$ iff $\mathcal{A}v_q = 0$ iff $v_q \in T_q\mathcal{F}$ iff q_0 and q_1 lie in the same leaf of $\mathcal{F}$. □

(Such a metric can always be defined locally by choosing coordinates such that the foliation takes the form $x_i = \text{const.}$, $i = 1, \ldots, \dim \mathcal{F}$, together with the Euclidean metric in these coordinates. In these coordinates, $\varphi(q_0, q_1) = (q_0, (q_1 - q_0)/h)$.)

The proposition can be extended to cover constraint distributions that are merely partially integrable, i.e. $C \subset T\mathcal{F}$, with φ defining a C_d contained in $\bigcup_l \mathcal{F}_l \times \mathcal{F}_l$, so that at least those constraints that are holonomic are preserved.

When C is integrable (i.e., when the constraints are holonomic), the DLA equations together with $A_d = A$ and the "exact" C_d of Proposition 8 are identical to the constrained discrete Euler-Lagrange equations, the constraint being that the orbit stays on its initial leaf. The discrete Euler-Lagrange equations generate symplectic integrators such as SHAKE and RATTLE [18], [24].

This result is mainly of theoretical interest, for it is equivalent to (partially) integrating the constraints. However, in two simple cases this can be achieved automatically by standard finite difference maps.

Proposition 9. *Either of the finite difference maps (4.6), (4.7) with a Euclidean metric preserves holonomic linear constraints exactly. The midpoint finite difference map (4.7) with a Euclidean metric (equation (4.13) below) preserves quadratic constraints exactly.*

Proof. The linear case is an instance of Proposition 8, for the leaves of a linear constraint are totally geodesic with respect to the Euclidean metric. For the quadratic case, first note that (4.7) is equivariant under all affine changes of coordinates. Apply such a change of coordinates to bring the constraint into the form $\sum \lambda_i (q^i)^2 = \text{const}$, where $\lambda_i \in \{0, 1, -1\}$. In these coordinates the constraint one-form is $\sum \lambda_i q^i dq^i$ and the equation $(q_0, q_1) \in C_d$ reads

$$\sum \lambda_i \left(\frac{q_0^i + q_1^i}{2} \right) \left(\frac{q_1^i - q_0^i}{h} \right) = 0,$$

or

$$\sum \lambda_i (q_1^i)^2 = \sum \lambda_i (q_0^i)^2,$$

that is, the constraint is preserved. If other constraints are present, C_d is merely contained in (not equal to) $\bigcup_l \mathcal{F}_l \times \mathcal{F}_l$, but the quadratic constraint is still preserved. □

We now give three specific integrators constructed (in one way or another) using finite difference maps.

First-order nonreversible integrator, linearly implicit. Let $Q = \mathbb{R}^n$ together with the Euclidean metric. The map φ defined in (4.6) is then

$$\varphi(q_0, q_1) = (q_0, (q_1 - q_0)/h) \in T_{q_0}Q. \tag{4.8}$$

Define $C_d = F^{-1}(0)$ with F defined as in equation (4.2). By Proposition 6 we know that C_d is a well-defined submanifold. Let us next consider the regularity condition of Proposition 3. First note that

$$\begin{aligned}(\pi_1|_{C_d})^{-1}(q_1) &= \left\{(q_1, q_2) \mid (q_1, (q_2 - q_1)/h) \in C_{q_1}\right\}\\ &= \{(q_1, q_1 + hv_{q_1})\colon\ v_{q_1} \in C_{q_1}\},\end{aligned}$$

so that $T_{q_2}((\pi_1|_{C_d})^{-1}(q_1)) = 0 \times C_{q_1}$. Therefore, by Proposition 3, the requirement that F_{L_d} is well defined is that

$$\text{Hess } L_d(q_1, q_2)|_{C_{q_1} \times C_{q_2}} \tag{4.9}$$

is nondegenerate.

Suppose our Lagrangian is of the form $L = T - V$ where T is the kinetic energy associated with the Euclidean metric on $\mathbb{R}^n$ and V is the potential energy. We identify TQ and T^*Q with $Q \times Q$ using the Euclidean metric, and the constraint one-forms A_d with the matrix $A \in \mathbb{R}^{n-k,n}$.

With

$$L_d(q_0, q_1) = L \circ \varphi = \frac{1}{2}\left\|\frac{q_1 - q_0}{h}\right\|_2^2 - V(q_0), \tag{4.10}$$

the integrator (3.9), (3.10) is given by

$$\begin{aligned}\frac{q_{i+1} - 2q_i + q_{i-1}}{h^2} + \nabla V(q_i) &= A(q_i)^{\mathrm{T}}\lambda_i,\\ A(q_i)(q_{i+1} - q_i) &= 0,\end{aligned} \tag{4.11}$$

where we now label the discrete time i. Since, in this case, $Q \times Q \cong TQ \cong \mathbb{R}^{2n}$, a convenient formulation in velocity variables is given by defining the velocity $v_i := (q_{i+1} - q_i)/h$, in terms of which the method (4.11) is

$$\begin{aligned}q_{i+1} &= q_i + hv_i,\\ v_{i+1} &= v_i + h(-\nabla V(q_{i+1}) + A(q_{i+1})^{\mathrm{T}}\lambda_{i+1}),\\ A(q_{i+1})v_{i+1} &= 0,\end{aligned} \tag{4.12}$$

where the initial condition should satisfy the constraint $A(q_0)v_0 = 0$. Note that the Lagrange multipliers λ_{i+1} can be determined by solving the linear system

$$A(q_{i+1})A(q_{i+1})^{\mathrm{T}}\lambda_{i+1} = A(q_{i+1})\left(\nabla V(q_{i+1}) - \frac{v_i}{h}\right),$$

that is, the method is linearly implicit.

Second-order reversible integrator, implicit. A second-order method can be constructed using the reversible finite difference map (4.7), i.e.,

$$\varphi(q_0, q_1) = ((q_0 + q_1)/2, (q_1 - q_0)/h). \tag{4.13}$$

The DLA equations are

$$\begin{aligned} \frac{q_{i+1} - 2q_i + q_{i-1}}{h^2} + \frac{1}{2}\left(\nabla V\left(\frac{q_{i-1} + q_i}{2}\right) + \nabla V\left(\frac{q_i + q_{i+1}}{2}\right)\right) &= A(q_i)^{\mathrm{T}}\lambda_i, \\ A\left(\frac{q_i + q_{i+1}}{2}\right)(q_{i+1} - q_i) &= 0. \end{aligned} \tag{4.14}$$

To get a velocity formulation, it is convenient to define

$$\begin{aligned} \bar{q}_i &:= (q_i + q_{i-1})/2, \\ v_i &= (q_i - q_{i-1})/h, \end{aligned} \tag{4.15}$$

so that the constraint has the simple form $A(\bar{q}_i)v_i = 0$. The method is $(\bar{q}_i, v_i) \mapsto (\bar{q}_{i+1}, v_{i+1})$ where

$$\begin{aligned} \bar{q}_{i+1/2} &= \bar{q}_i + \frac{1}{2}hv_i: \\ v_{i+1} &= v_i + h\left(\frac{1}{2}\left(\nabla V(\bar{q}_i) + \nabla V(\bar{q}_{i+1})\right) + A(\bar{q}_{i+1/2})^{\mathrm{T}}\lambda_i\right), \\ \bar{q}_{i+1} &= \bar{q}_{i+1/2} + \frac{1}{2}hv_{i+1}, \\ A(\bar{q}_{i+1})v_{i+1} &= 0, \end{aligned} \tag{4.16}$$

where the initial condition should satisfy the constraint $A(\bar{q}_0)v_0 = 0$. The method is fully implicit and reduces in the absence of constraints to the trapezoidal rule. As is well known, the trapezoidal rule is equivalent under the discrete Legendre transform $p_i = v_i + \frac{h}{2}\nabla V(q_i)$ to the (symplectic) midpoint rule in the variables (q, p). One can make the same transformation here if desired.

Second-order reversible, semi-implicit. Another way to construct a second-order reversible method is to compose the first-order method (4.12), $F_{L_d}(h)$, with its adjoint $F_{L_d}^{-1}(-h)$. The adjoint is

$$\begin{aligned} v_{i+1} &= v_i + h(-\nabla V(q_{i+1}) + A(q_i)^{\mathrm{T}}\lambda_i), \\ q_{i+1} &= q_i + hv_{i+1}, \\ A(q_{i+1})v_{i+1} &= 0. \end{aligned} \tag{4.17}$$

In the composition $F_{L_d}^{-1}(-h/2) \circ F_{L_d}(h/2)$ of the two methods, the two velocity updates can be merged to obtain the method

$$
\begin{aligned}
q_{i+1/2} &= q_i + \frac{1}{2} h v_i: \\
v_{i+1} &= v_i + h(-\nabla V(q_{i+1/2}) + A(q_{i+1/2})^{\mathrm{T}} \lambda_{i+1}), \\
q_{i+1} &= q_{i+1/2} + \frac{1}{2} h v_{i+1}, \\
A(q_{i+1}) v_{i+1} &= 0.
\end{aligned} \tag{4.18}
$$

There are $n - k$ constraints, so one needs to solve a set of $n - k$ equations for λ_{i+1}, which are in general nonlinear. However, the force $\nabla V(q_{i+1/2})$ is only evaluated once per time step, so the method is semi-implicit. We shall see in Section 5 that this method, the nonholonomic analog of the popular SHAKE and RATTLE methods for holonomic constraints [18], performs extremely well in numerical tests.

We now show that the integrator (4.18) satisfies the DLA principle. However, it is not constructed from a finite difference map as were the previous examples (4.11) and (4.14). It is constructed from two finite difference maps, one for L_d and one for C_d. For L_d we use the nonreversible finite difference map (4.8); although L_d (4.10) is nonreversible, the discrete action associated with it is in fact second-order and reversible. For C_d we use the reversible finite difference map (4.13), that is,

$$
C_d = \left\{ (q_0, q_1) \colon A_j \left(\frac{q_0 + q_1}{2} \right) (q_1 - q_0) = 0 \right\}. \tag{4.19}
$$

Note that by an immediate application of Proposition 7, $0 \times v_{q_0} \in T_{(q_0, q_0)} C_d$ for $v_{q_0} \in \ker A_j(q_0)$ for each j.

The method on $Q \times Q$ is then given by the first equation of (4.11) together with the constraint (4.19). The velocity formulation in variables $(\bar{q}_i, v_i)$ of equation (4.15) is exactly given by (4.18) with $q_i = \bar{q}_i$. That is, the composite method (4.18) also satisfies the discrete Lagrange-d'Alembert principle.

Although methods constructed using the finite difference map (4.7) are quite general, yielding reversible, second-order integrators for any Lagrangian on which the geodesics on Q can be computed, for most simple mechanical systems the method (4.18) is far more efficient, since it requires only one force evaluation per time step.

5. Numerical Results

5.1. The Contact Oscillator

We now explore the behavior of the semi-implicit, reversible, DLA method (4.18) in some examples to see how well it preserves the qualitative features of their dynamics. We will consider three systems of increasing complexity: in this section, an integrable system on $T^*\mathbb{R}^3$ (the "contact oscillator"), comparing a reversible DLA method with a nonreversible DLA method and a reversible, non-DLA method; in Section 5.2, a non-integrable system on $T^*\mathbb{R}^3$ (a nonlinear perturbation of the contact oscillator) showing both regular and chaotic orbits, comparing a reversible DLA method with a standard, nongeometric integrator (MATLAB's `ode15s`); and in Section 5.3, a fully chaotic system on $T^*\mathbb{R}^7$, comparing a reversible DLA method with another standard method for differential-algebraic equations (DASSL). In all cases, the reversible DLA method (4.18) is the best, both for efficiency and for qualitative preservation of the dynamics.

The simplest nonholonomic systems are those with a single constraint. There are no nonintegrable distributions on $\mathbb{R}^2$, so the simplest case is to take $Q = \mathbb{R}^3$. Every nonintegrable one-form can be put in the form $dx + y\,dz$ in local coordinates (x, y, z); we therefore consider this constraint on $\mathbb{R}^3$. That is, we take $A(x, y, z) = (1, 0, y)$. The free particle with this constraint was studied in [1], but its orbits are unbounded. To get a simple system with manifestly bounded orbits, we take the harmonic oscillator Lagrangian

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \frac{1}{2}(x^2 + y^2 + z^2). \tag{5.1}$$

We call the corresponding nonholonomic system

$$\begin{aligned} \ddot{x} + x &= \lambda, \\ \ddot{y} + y &= 0, \\ \ddot{z} + z &= \lambda y, \\ \dot{x} + y\dot{z} &= 0, \end{aligned}$$

the *nonholonomic oscillator*.

We now show that all orbits of the nonholonomic oscillator are quasiperiodic with at most two frequencies. Most orbits form a three-parameter family of two-tori in the five-dimensional constraint manifold C. First, the evolution of y is unconstrained and we have $y(t) = a\sin t + b\cos t$. We choose the origin of time so that $b = 0$ and $y = a\sin t$. Differentiating the constraint gives $\ddot{x} + y\ddot{z} + \dot{y}\dot{z} = 0$, and substituting the equations of motion gives $\lambda - x + y(\lambda y - z) + \dot{y}\dot{z} = 0$, i.e.,

$$\lambda = \frac{x + yz - \dot{y}\dot{z}}{1 + y^2} = \frac{x + za\sin t - \dot{z}a\cos t}{1 + a^2\sin^2 t}.$$

Introducing the velocities $v_x = \dot{x}$, $v_z = \dot{z}$ and substituting for λ gives a system of four nonautonomous ODEs which are linear. The equation for $\dot{v}_x$ can be eliminated in favour of the constraint $v_x = -v_z a\sin t$, leaving the three equations

$$\begin{pmatrix} \dot{x} \\ \dot{z} \\ \dot{v}_z \end{pmatrix} = \begin{pmatrix} 0 & 0 & -a\sin t \\ 0 & 0 & 1 \\ \frac{a\sin t}{1+a^2\sin^2 t} & \frac{-1}{1+a^2\sin^2 t} & \frac{-a^2\sin t\cos t}{1+a^2\sin^2 t} \end{pmatrix} \begin{pmatrix} x \\ z \\ v_z \end{pmatrix}.$$

In terms of the new variable $\tilde{v}_z := (1 + a^2\sin^2 t)^{1/2} v_z$, these become

$$\begin{pmatrix} \dot{x} \\ \dot{z} \\ \dot{\tilde{v}}_z \end{pmatrix} = \begin{pmatrix} 0 & 0 & \alpha(t) \\ 0 & 0 & \beta(t) \\ -\alpha(t) & -\beta(t) & 0 \end{pmatrix} \begin{pmatrix} x \\ z \\ \tilde{v}_z \end{pmatrix}, \tag{5.2}$$

where

$$\alpha(t) = -a(1 + a^2 \sin^2 t)^{-1/2} \sin t, \qquad \beta(t) = (1 + a^2 \sin^2 t)^{-1/2}.$$

The matrix of coefficients in (5.2) is 2π-periodic and antisymmetric, so for each a there is an orthogonal matrix $\Omega(a) \in \mathrm{SO}(3)$ such that the time-2π flow of (5.2) is given by $(x, z, \tilde{v}_z)^{\mathrm{T}} \mapsto \Omega(a)(x, z, \tilde{v}_z)^{\mathrm{T}}$. This map is simply a rotation whose angle and axis depend on a. The orbits of the nonholonomic oscillator can therefore be classified as follows: (i) a two-parameter family of periodic orbits with period 2π and parameters x and energy (these have $a = 0$); (ii) a two-parameter family of periodic orbits with period 2π and parameters $a \neq 0$ and energy, with $(x, z, \tilde{v}_z)$ lying on the axis of rotation of $\Omega(a)$; and (iii) a three-parameter family of quasiperiodic orbits with quasiperiods 2π and $2\pi/\gamma$, where γ is the angle of rotation of $\Omega(a)$, the parameters being a, energy, and the latitude of $(x, z, \tilde{v}_z)$ with respect to the axis of rotation of $\Omega(a)$.

How well do the DLA integrators preserve this integrable structure? If the constraint were not present, we would have three harmonic oscillators and any DLA integrator would be integrable.

We first consider the reversible semi-implicit method (4.18). The $(y, v_y)_i$ variables are unconstrained and hence obey the standard leapfrog method for the harmonic oscillator,

$$\begin{pmatrix} y \\ v_y \end{pmatrix} \mapsto \begin{pmatrix} 1 & h/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -h & 1 \end{pmatrix} \begin{pmatrix} 1 & h/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ v_y \end{pmatrix} =: M(h) \begin{pmatrix} y \\ v_y \end{pmatrix},$$

where the eigenvalues of M (for $0 < h < 2$) are $e^{\pm i\theta}$, with $\theta = 2\sin^{-1}\frac{h}{2}$. Thus $(y, v_y)_i$ is given by an explicit periodic function with period $2\pi/\theta$, evaluated at integer times. For this system, equations (4.18) can be solved explicitly for λ, so the method is explicit in this case. Eliminating λ and further eliminating v_x using the constraint gives a linear map $(x, z, v_z)^{\mathrm{T}}_{i+1} = R(i, a)(x, z, v_z)^{\mathrm{T}}_i$ where $R(t, a) \in \mathbb{R}^{3\times 3}$ is periodic in t with period $2\pi/\theta$. At this point it is hard to make further analytic progress because $R(t, a)$ is fairly complicated. Instead, we have iterated this reduced map numerically, with the the following results.

For each $a > 0$ there is a time step $h^* > 0$ such that for all $h < h^*$, all orbits of the integrator are quasiperiodic and the continuous and discrete flows are conjugate. Specifically, choosing $h = 2\sin(\pi/N)$ for some integer N, the dynamics of (y, v_y) are periodic with period N and the matrix $\prod_{i=0}^{N-1} R(i)$ of the time-N map has eigenvalues $1, \exp(\pm i\alpha)$ for some α depending on a and h. (The invariant spheres of the continuous system are deformed to become invariant ellipses, just as the invariant circles of the standard harmonic oscillator become invariant ellipses under leapfrog.) The critical time step h^* is equal to 2 for $a = 0$ and is approximately π/a as $a \to \infty$.

That is, the dynamics of the integrator (4.18) is integrable and is conjugate to that of the exact flow of the nonholonomic oscillator, just as one has for the leapfrog method applied to the harmonic oscillator.

Now consider the first-order nonreversible DLA method (4.10). The dynamics of (y, v_y) are identical, and the map reduces to a three-dimensional linear nonautonomous map, as before. However, this map is not conservative. We find numerically that the eigenvalues of the time-N map $\prod_{i=0}^{N-1} R(i)$ are all less than 1 in modulus and $\lim_{i\to\infty}(x, z, v_z)_i = (0, 0, 0)$. The qualitative dynamics is not preserved. Preserving reversibility is crucial to capturing the dynamics.

However, the DLA equations themselves also play a role. We consider integrating the nonholonomic oscillator with a *non*-DLA method which is nevertheless reversible and also has the same dynamics in (y, v_y), namely the method

$$\begin{aligned} \frac{q_{i+1} - 2q_i + q_{i-1}}{h^2} + \nabla V(q_i) &= A(q_i)^{\mathrm{T}}\lambda_i, \\ A(q_i)(q_{i+1} - q_{i-1}) &= 0. \end{aligned} \tag{5.3}$$

The constraint is enforced by a reversible, second-order approximation of the true constraints, but not one associated with a discrete constraint manifold in the sense used above. Introducing the velocity variables $v_i = (q_{i+1} - q_i)/h$, the method can be written in the form

$$\begin{aligned} q_{i+1} &= q_i + hv_i, \\ v_{i+1} &= v_i + h(-\nabla V(q_i) + A(q_i)^{\mathrm{T}}\lambda_i), \\ A(q_i)(v_i + v_{i+1}) &= 0, \end{aligned}$$

so that

$$A(q_i)A(q_i)^{\mathrm{T}}\lambda_i = A(q_i)\left(\nabla V(q_i) - \frac{2}{h}p_i\right).$$

However, it is no longer clear how to constrain the initial conditions. We find numerically that the linear time-N map $(x, z, v_x, v_z)_0 \mapsto (x, z, v_x, v_z)_N$ has eigenvalues 1, 1, and $\exp(\pm i\alpha)$ for some α depending on a and h. The dynamics are still a rotation, but a rotation in $\mathbb{R}^4$ instead of $\mathbb{R}^3$. The extra eigenvalue 1 indicates that there is an invariant three-dimensional subspace (corresponding to constraining the initial condition) on which the map is a rotation; but this subspace depends on a (that is, on the initial condition) and on h. Therefore, just maintaining reversibility is insufficient to get qualitatively correct dynamics.

5.2. A Nonintegrable System on $T^*\mathbb{R}^3$

We now consider a nonlinear perturbation to the contact oscillator, modifying the Lagrangian from (5.1) to

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \frac{1}{2}(x^2 + y^2 + z^2 + \varepsilon x^2 z^2). \tag{5.4}$$

The constraint is still $\dot{x} + y\dot{z} = 0$. The perturbation has been chosen so that the y equation is still $\ddot{y} + y = 0$. Orbits are confined to the three-manifold defined by the intersection of the constraint surface, the energy surface, and $y^2 + v_y^2 =$const. On this manifold we define a Poincaré section by $v_y = 0$, $\dot{v}_y > 0$, so that the Poincaré map is simply the time-2π flow of the system. The section is topologically a two-sphere in (x, z, v_z)-space; we plot (z, v_z) when $x > 0$, which shows one half of the sphere.

Figure 1 shows the phase portraits for four values of ε as calculated by the explicit, reversible, DLA method (4.18) with 40 time steps (and hence 40 force evaluations) per

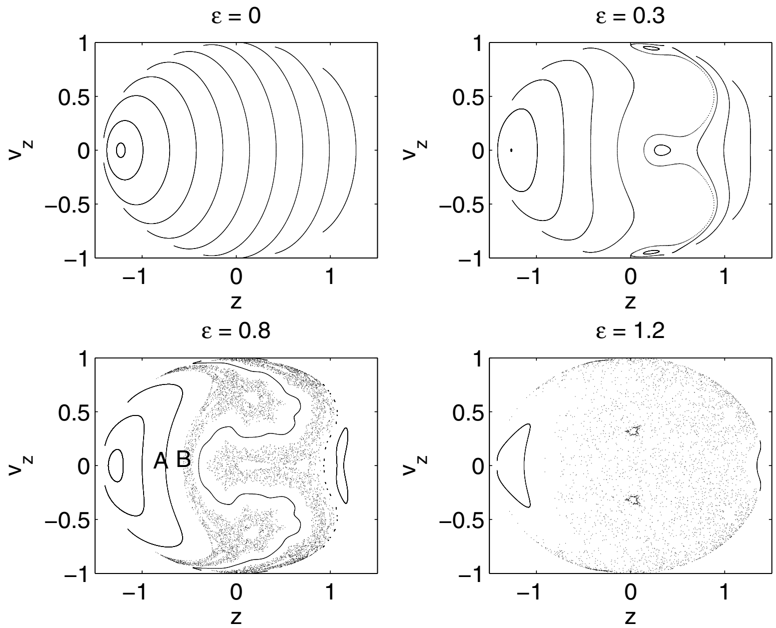

Fig. 1. Phase portraits of a three-degree-of-freedom system with a singular nonholonomic constraint (Lagrangian in (5.4) on energy level $H = 1.5$ and $y^2 + v_y^2 = 1$). 8000 iterates of the Poincaré map (with 40 time steps per iterate) are shown for each of several initial conditions and for four values of ε; for $\varepsilon = 0$, the system reduces to the (integrable) contact oscillator, (5.1).

period. We show 8000 iterates for each of several initial conditions. $\varepsilon = 0$ is the contact oscillator; the rotation described in Section 5.1 is evident in the Poincaré section. As ε increases, the invariant circles progressively break up and are replaced by chaotic bands. The integrator evidently has a mixture of quasiperiodic and chaotic orbits; the phase portraits of Figure 1 are all qualitatively correct, the main numerical error being an $\mathcal{O}(h^2)$ shift in the positions of the orbits. By contrast, a standard, non-geometric integrator for differential-algebraic equations (MATLAB's `ode15s`) does not preserve the phase portraits. Consider the quasiperiodic orbit marked A in Figure 1. Setting the absolute tolerance to 10^{-6} in `ode15s`, so that 160 time steps and more than 300 function evaluations are needed per period, the orbit still drifts away from its correct location in less than 400 iterations (see Figure 2). The same behavior is seen in long enough runs for any tolerance.

On quasiperiodic orbits, the energy error under (4.18) is bounded. A longer simulation (50000 iterations of the Poincaré map) of the chaotic orbit marked B in Figure 1 is shown in Figures 3 (40 time steps per period) and 4 (80 time steps per period). Apart from the orbit segment visiting different parts of the entire orbit, which is to be expected in any realization of a chaotic orbit, only a few small changes in the fine structure of the orbit are visible, despite the large time steps. The energy errors scale as $\mathcal{O}(h^2)$; they appear to oscillate, but a small amount of drift is possibly also present. We will examine this drift more in the next sample.

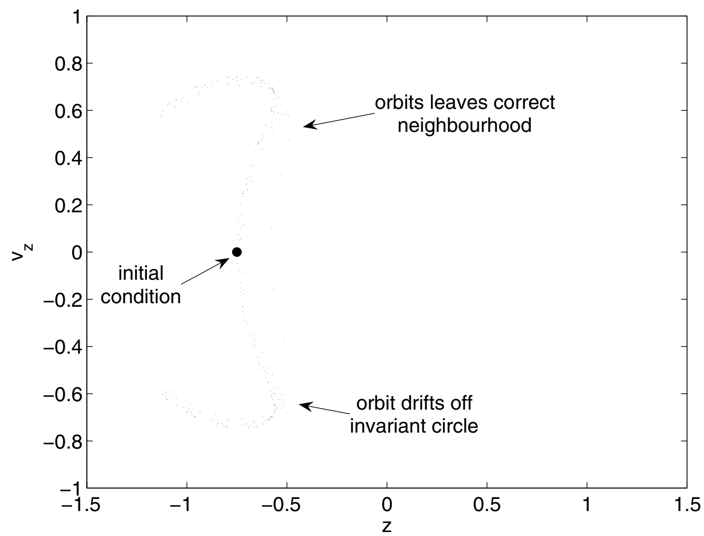

Fig. 2. The orbit marked A in Figure 1, but integrated with MATLAB's `ode15s`. The tolerance is set to 10^{-6}; the orbit rapidly drifts away from its correct location. 400 iterates of the Poincaré map are shown, compared to 8000 in Figure 1.

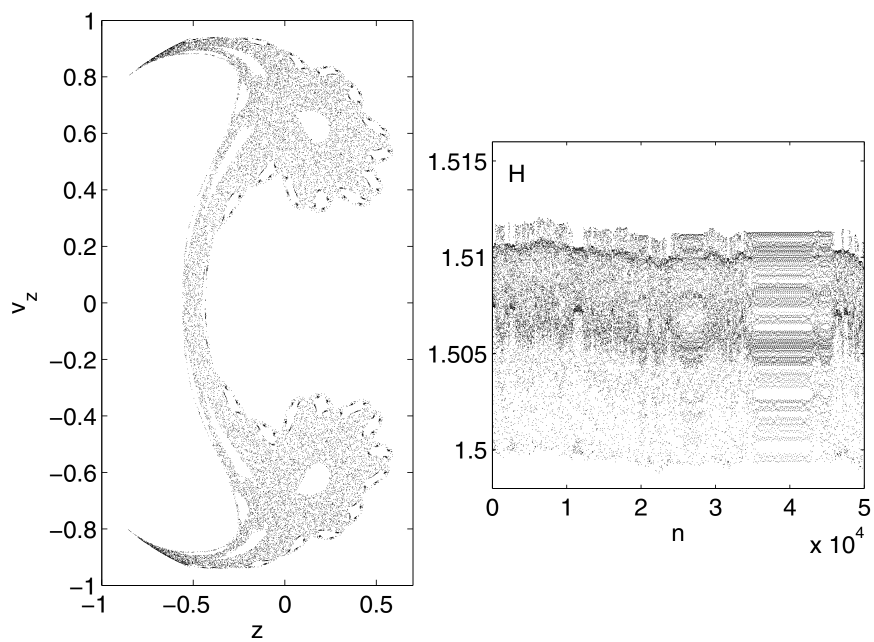

Fig. 3. Left: 50,000 iterations of the chaotic orbit marked B in Figure 1, with 40 times steps per period of the reversible DLA method (4.18). Right: Energy for this simulation (the initial energy is 1.5).

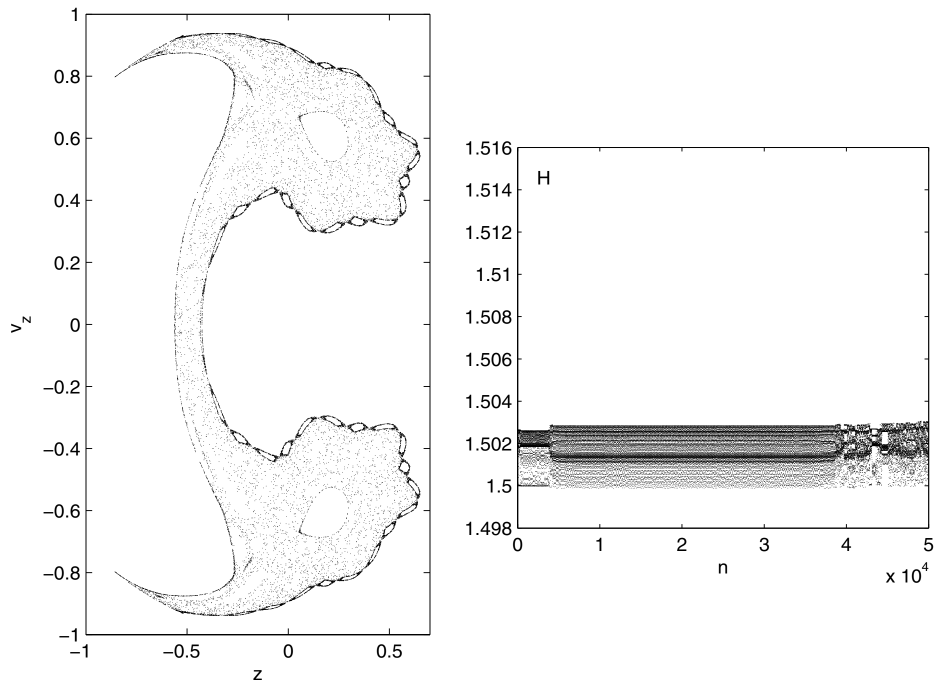

Fig. 4. As for Figure 3, but 80 time steps per period.

5.3. A Fully Chaotic System

For general, nonintegrable systems, it is harder to assess the merits of a variational vs. a standard integrator, because even the key geometric features of nonholonomic dynamics itself are not known. That is, what feature should one check in a comparison? We have chosen to monitor the energy in a nonholonomic system that conserves energy.

We consider the configuration space $\mathbb{R}^{2n+1}$ with coordinates $q = (x, y_1, \ldots, y_n, z_1, \ldots, z_n)$, and Lagrangian

$$L = T - V, \qquad T = \frac{1}{2}\|\dot{q}\|_2^2, \qquad V = \frac{1}{2}\left(\|q\|_2^2 + z_1^2 z_2^2 + \sum_i y_i^2 z_i^2\right), \tag{5.5}$$

and a single nonholonomic constraint

$$\dot{x} + \sum_{i=1}^{n} y_i \dot{z}_i = 0. \tag{5.6}$$

As before, $\dot{H} = 0$ (where $H = T + V$) and the system is reversible.

The energy behavior for the reversible DLA method (4.18) and another standard package for differential-algebraic systems, DASSL, are compared in Figure 5. We find that the energy error for the DLA method is not bounded (as in symplectic integrators), but nor does it display the secular, $\mathcal{O}(t)$ drift of standard integrators such as DASSL.

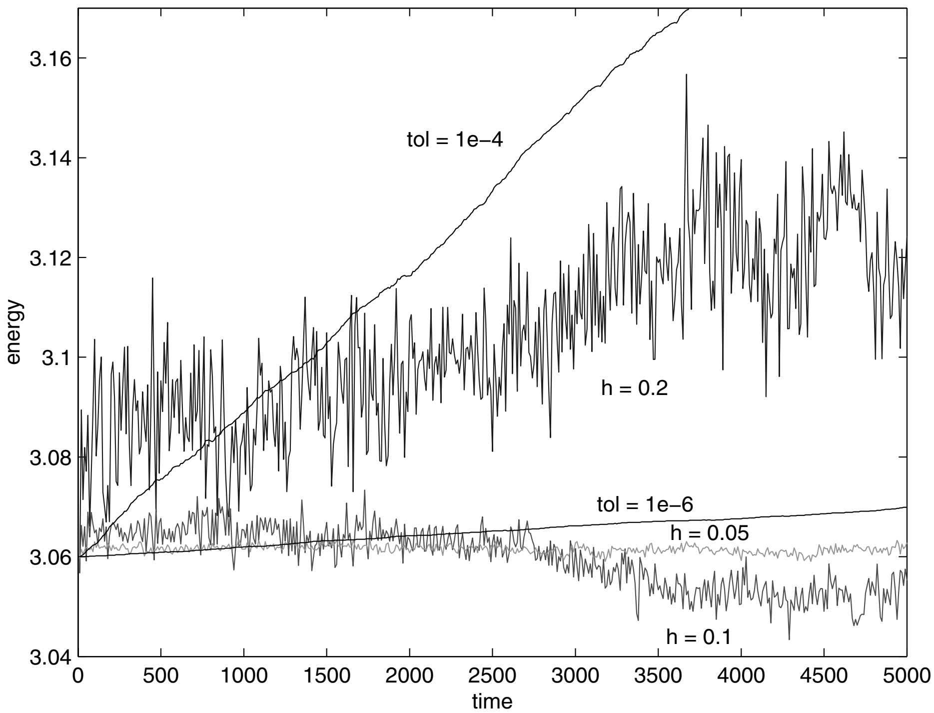

Fig. 5. Energy behavior for two methods compared for the chaotic system (5.5). Two runs are shown for the standard method DASSL (with error tolerance `tol` set to 10^{-4} and 10^{-6}, respectively) and three for the reversible DLA method (4.18) (with $h = 0.2$, 0.1, and 0.05). At tolerance 10^{-6}, DASSL used an average time step of $h = 0.026$ and an average of 2.6 function evaluations per time step.

Instead it follows a random walk, so that the energy error after time t is $\mathcal{O}(\sqrt{t})$—in this example, it is

$$|H(t) - H(0)| \sim 0.01h^2\sqrt{t}.$$

The energy errors for 100 sample trajectories are shown in Figure 6. The diffusion rate depends strongly on the energy, because of the quartic nonlinearities in H. In fact, as $H \to 0$, the system becomes integrable and no energy drift is seen; the diffusion rate is found to vary approximately as $\|q\|^4$. When the dynamics are ergodic on a symmetric set, this behavior can be explained using some ideas from ergodic theory [25]. The key point for getting this nice behavior is that the integrator is reversible.

To sum up,

1. Nonreversible integrators (both standard packages and DLA integrators) are dissipative and do not show the correct long-time dynamics of the example systems.
2. A reversible but non-DLA integrator, although not dissipative, does not capture the dynamics of the contact oscillator as well as the reversible DLA integrator.
3. The reversible DLA integrator, being explicit with a single force evaluation in these examples, performs extremely well. On the six-dimensional, integrable contact oscillator (5.1), its behavior is essentially perfect, its dynamics being integrable and

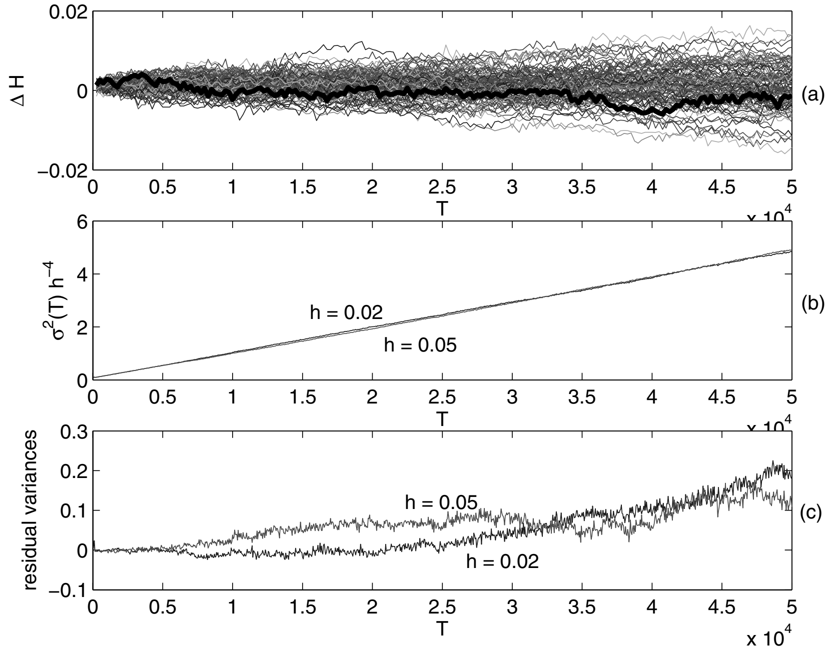

Fig. 6. Energy drift on $H(0) = 3.06$ for the reversible integrator (4.18) applied to the fully chaotic system (5.5). (a) The observed energy error for 100 different initial conditions with time step $h = 0.05$ integrated up to time $T = 50000$. One typical simulation is singled out in bold. (b) The variance $\sigma^2(T)$ of the energy errors for 10,000 different initial conditions integrated up to time $T = 50000$ for two different time steps, scaled by their expected h^4 dependence on the time step. The growth is roughly linear in time and the two results are almost indistinguishable. (c) The residual variance from (b) with the expected linear trend subtracted. Only a relatively small nonlinear behavior remains.

conjugate to the exact flow, a situation reminiscent of the leapfrog method applied to the *unconstrained* harmonic oscillator, even though the harmonic oscillator is linear and the contact oscillator is nonlinear. On the six-dimensional, nonintegrable system (5.4), it preserves quasiperiodic and chaotic orbits over long times, with performance reminiscent of a symplectic integrator. On the fourteen-dimensional, chaotic system (5.5), its performance is not quite as good as a symplectic integrator, because the energy error now drifts in a random walk instead of being bounded.

6. Nonholonomic Integrators on Lie Groups

6.1. Discrete reduction

A nonholonomic system (Q, L, C) admits a symmetry provided there is a Lie group action G on Q (and corresponding lift to TQ) leaving both L and C invariant. The nonholonomic flow is then G-equivariant, and provided the action is sufficiently regular (e.g. free and proper), the flow will reduce to the quotient subbundle C/G. The reduced equations arise from a reduced Lagrange d'Alembert principle [9]. Analogously, the discrete nonholonomic data (Q, L_d, C_d, A_d) is symmetric with respect to the action of G on Q (and its corresponding diagonal lifted action to $Q \times Q$) provided L_d, C_d, and A_d are G-invariant. In this case, the discrete flow map $F_{L_d}\colon C_d \to C_d$ determined by the DLA principle is G-equivariant [14].

We now specialize to the situation where $Q = G$. In this section we consider a discretization of the following nonholonomic system with symmetry on TG. Let $\mathfrak{k} \subset \mathfrak{g}$ be a k-dimensional subspace of the Lie algebra $\mathfrak{g}$ of G, which we assume to be *not* closed under the Lie bracket. Define the *left*-invariant distribution, $C \subset TG$ by $C_g = g \cdot \mathfrak{k}$. Clearly C is nonintegrable precisely because $\mathfrak{k}$ is not a Lie subalgebra of $\mathfrak{g}$. Finally we are also given a G-invariant Lagrangian, $L\colon TG \to \mathbb{R}$, of the type $L = T - V$. The Lagrange-d'Alembert equations for this system, being G-equivariant, will induce a flow on the quotient space $C/G \simeq k$. In fact these reduced equations can be realized as a reduced Lagrange-d'Alembert principle as defined in [16]. We wish to consider the corresponding discrete system. The equations we will determine are the nonholonomic version of the discrete Euler–Poincaré equations as found in [6], [23]. The reduced equations will agree with these discrete Euler–Poincaré equations when we take the case of no constraints.

Throughout this section we use concatenated notation for the corresponding left and right tangent and cotangent lifted actions. Thus, for $v_h \in T_hG$, $\alpha_h \in T_h^*G$, and $g \in G$, we have

$$g \cdot v_h := T_h L_g \cdot v_h,$$

$$v_h \cdot g := T_h R_g \cdot v_h,$$

$$g \cdot \alpha_h := T_{gh}^* L_{g^{-1}} \alpha_h \in T_{gh}^* G,$$

and

$$\alpha_h \cdot g := T_{gh}^* R_{g^{-1}} \alpha_h \in T_{hg}^* G.$$

With this convention, we can now formulate the following:

Theorem 3. *Let (G, L, C) be as specified in the previous paragraph, i.e., L is a G-invariant Lagrangian, and C is a k-dimensional, left G-invariant distribution on G. Let $\pi\colon TG \to \mathfrak{g}$ denote the quotient by the left G-action and define $l\colon \mathfrak{g} \to \mathbb{R}$ by $L = l \circ \pi$. Let $L_d\colon G \times G \to \mathbb{R}$ be a discrete Lagrangian and C_d be a discrete constraint submanifold. Assume that L_d and C_d are both G-invariant with respect to the diagonal action of G on $G \times G$. The discrete Lagrange-d'Alembert equations of Proposition 2 determine the equations on the reduced space G given by (where $\omega_0, \omega_1 \in G$)*

$$\begin{cases} \omega_0^{-1} \cdot \mathbf{d} l_d(\omega_0) - \mathbf{d} l_d(\omega_1) \cdot \omega_1^{-1} = \sum_{j=1}^{n-k} \lambda_j \mu_j, \\ \qquad\qquad\qquad \omega_1 \in c_d, \end{cases} \tag{6.1}$$

where $\mu_1, \ldots, \mu_{n-k}$ are a basis of the space $\mathfrak{k}^0$, (i.e., they satisfy $\bigcap_{j=1}^{n-k} \ker \mu_j = \mathfrak{k}$) and c_d is the submanifold of G determined by $c_d = \pi_d(C_d)$ where $\pi_d\colon G \times G \to G$, the quotient by the diagonal left G-action is given by $\pi_d(g_1, g_2) := g_1^{-1} g_2$ and $l_d\colon G \to \mathbb{R}$ is defined by $L_d = l_d \circ \pi_d$.

Proof. Let A_j be the left-invariant one-form on G determined by $\mu_j \in \mathfrak{g}^*$, i.e. $A_j(g)(v_g) := \langle \mu_j, g^{-1} \cdot v_g \rangle$. Applying the discrete Lagrange-d'Alembert principle of Proposition 2 to the data (Q, L_d, C_d, A_d), we obtain the following equations determining a discrete flow F_{L_d} on $C_d \subset G \times G$:

$$D_2 L_d(g_0, g_1) + D_1 L_d(g_1, g_2) = \sum_{j=1}^{n-k} \lambda_j(g_1) A_j(g_1), \tag{6.2}$$

and then imposing the condition on the set of g_2 satisfying the above equation that

$$(g_1, g_2) \in C_d \iff g_1^{-1} g_2 \in c_d \subset G. \tag{6.3}$$

We next write the first equation in terms of l_d. Let $g(t)$ be a curve through g_1 with $\frac{d}{dt}\big|_{t=0} g(t) = v_{g_1}$. We have

$$\begin{aligned} D_2 L_d(g_0, g_1) \cdot v_{g_1} &= \frac{d}{dt}\bigg|_{t=0} l_d \circ \pi_d(g_0, g(t)) = \frac{d}{dt}\bigg|_{t=0} l_d(g_0^{-1} g(t)) \\ &= \mathbf{d} l_d(g_0^{-1} g_1) \cdot (g_0^{-1} \cdot v_{g_1}), \end{aligned}$$

and similarly

$$\begin{aligned} D_1 L_d(g_1, g_2) \cdot v_{g_1} &= \frac{d}{dt}\bigg|_{t=0} l_d(g_1(t)^{-1} g_2) \\ &= \mathbf{d} l_d(g_1^{-1} g_2) \cdot ((-g_1^{-1} \cdot v_{g_1} \cdot g_1^{-1}) \cdot g_2), \end{aligned}$$

using the fact that $\frac{d}{dt}\big|_{t=0} g_1(t)^{-1} = -g_1^{-1} \cdot v_{g_1} \cdot g_1^{-1} \in T_{g_1^{-1}} G$. Next, given a sequence (g_k, g_{k+1}) of points in $G \times G$, we construct the projected sequence in G given by $\omega_k := \pi(g_k, g_{k+1}) = g_k^{-1} g_{k+1}$. Notice that given the image sequence ω_k and an initial point g_0 in G, we can reconstruct the unique sequence in $G \times G$ by solving $g_0^{-1} g_1 = \omega_0$ for g_1 and then iterating down the chain. We can then rewrite each of the above expressions as follows:

$$\begin{aligned} \langle \mathbf{d} l_d(g_0^{-1} g_1), (g_0^{-1} \cdot v_{g_1}) \rangle &= \langle \mathbf{d} l_d(\omega_0), g_0^{-1} g_1 g_1^{-1} \cdot v_{g_1} \rangle \\ &= \langle \omega_0^{-1} \cdot \mathbf{d} l_d(\omega_0), g_1^{-1} \cdot v_{g_1} \rangle = \langle g_1 \cdot \omega_0^{-1} \cdot \mathbf{d} l_d(\omega_0), v_{g_1} \rangle. \end{aligned}$$

Similarly, for the second term we obtain

$$\langle \mathbf{d} l_d(\omega_1), (-g_1^{-1} \cdot v_{g_1} \cdot g_1^{-1}) \cdot g_2 \rangle = \langle -g_1 \cdot \mathbf{d} l_d(\omega_1) \cdot \omega_1^{-1}, v_{g_1} \rangle.$$

Notice that the right-hand side of equation (6.2) can be expressed in terms of the μ_i since by left G-invariance of the one-forms A_j we have

$$\sum_{j=1}^{n-k} \lambda_j(g_1) A_j(g_1) = g_1 \cdot \sum_{j=1}^{n-k} \lambda_j \mu_j.$$

Consequently we see that equation (6.2) reads, for all v_{g_1},

$$\langle g_1 \cdot \omega_0^{-1} \cdot \mathbf{d}l_d(\omega_0) - g_1 \cdot \mathbf{d}l_d(\omega_1) \cdot \omega_1^{-1}, v_{g_1} \rangle = \left\langle g_1 \cdot \sum_{j=1}^{n-k} \lambda_j \mu_j, v_{g_1} \right\rangle,$$

which holds if and only if equation (6.1) holds. That is, g_2 solves equation (6.2) if and only if ω_1 solves equation (6.1). Finally, we recall that $(g_1, g_2) \in C_d$ if and only if $\omega_1 \in c_d$, since by G-invariance of C_d, $(g_1, g_2) \in C_d$ if and only if $(e, g_1^{-1} g_2) \in C_d$ if and only if $g_1^{-1} g_2 \in c_d$ by definition of c_d. □

We now have the following concerning the relationship between the solutions to equations (6.2) and (6.3) and the reduced equations of the preceding theorem.

Corollary 4 (Reconstruction). *The sequence $g_0, g_1, \ldots, g_k$ is a solution to the discrete Lagrange-d'Alembert equations with initial condition (g_0, g_1) (6.2) if and only if the sequence $\omega_0, \omega_1, \omega_2, \ldots$ is a solution to the reduced Lagrange-d'Alembert equations (6.1). In other words, the solutions are π-related. Second, given a solution to the reduced equations with initial condition ω_0, and given an initial point $g_0 \in G$, there exists a unique g_1 such that $(g_0, g_1) \in C_d$, $\pi(g_0, g_1) = \omega_0$. Furthermore, the reconstructed solution to the equations (6.2) with initial condition (g_0, g_1) are given by*

$$g_1 = g_0 \omega_0, \qquad g_2 = g_1 \omega_1, \ldots, \qquad g_k = g_{k-1} \omega_{k-1}. \tag{6.4}$$

Proof. The first statement is an immediate consequence of the proof of the previous theorem. Next, since $\omega_0 \in c_d$, we have $(e, \omega_0) \in C_d$ and then by G-invariance we have that $(g_0, g_0\omega_0) \in C_d$. From this we can read off $g_1 = g_0 \omega_0$. Uniqueness is clear. We need to show that the g_i sequence thus constructed solves equations (6.2). It is clear that they satisfy $(g_i, g_{i+1}) \in C_d$. We argue inductively. From the proof of the previous theorem, starting with $i = 0$, we have

$$\begin{aligned} D_2 L_d(g_0, g_1) + D_1 L_d(g_1, g_2) &= g_1 \cdot \omega_0^{-1} \cdot \mathbf{d}l_d(\omega_0) - g_1 \cdot \mathbf{d}l_d(\omega_1) \cdot \omega_1^{-1} \\ &= g_1 \cdot \sum_{j=1}^{n-k} \lambda_j \mu_j = \lambda^T A(g_1), \end{aligned}$$

so that g_2 solves equation (6.2). For the induction step, we can apply the same argument to the point (g_i, g_{i+1}) where $g_{i+1} = g_i \omega_i$, which shows that g_{i+1} solves equation (6.2), concluding the proof of the second statement. □

6.2. Rigid Body with a Nonholonomic Constraint

For an example of a dynamical system on a Lie group with a nonholonomic constraint we consider the rigid body. Let $G = \mathrm{SO}(n)$ with Lie algebra $\mathfrak{g} = \mathfrak{so}(n)$. Let $\mathfrak{k}$ be a subspace of $T_e G$ defined by the vanishing of the one-forms μ_i. The constraint is that the velocities should lie in the left-invariant distribution $G\mathfrak{k}$, which is nonintegrable when it is not a subalgebra. This is determined at each point by the vanishing of the left-invariant one-forms A_j defined by $A_j(g) = g \cdot \mu_j$. Equivalently, the angular velocity $v := g^{-1}\dot{g} \in \mathfrak{k}$.

Consider the Lagrangian $L = \int \frac{1}{2}\mathrm{tr}(\dot{g}J\dot{g}^{\mathrm{T}})\,dt$ of the free rigid body with symmetric moments of inertia J. The variation of L is

$$\begin{aligned}\delta L &= \int \mathrm{tr}\left(\dot{g}J\delta\dot{g}^{\mathrm{T}}\right)\,dt = -\int \mathrm{tr}\left(\ddot{g}J\delta g^{\mathrm{T}}\right)\,dt \\ &= -\int \mathrm{tr}\left(g^{-1}\ddot{g}J(g^{-1}\delta g)^{\mathrm{T}}\right)\,dt.\end{aligned}$$

The Lagrange-d'Alembert principle now states that $\delta L = 0$ for all variations $g^{-1}\delta g \in \mathfrak{k}$.

The equation of motion is that the component of $\mathrm{skew}(g^{-1}\ddot{g}J)$ that lies in $\mathfrak{k}$ should vanish, where $\mathrm{skew}(A) = A - A^{\mathrm{T}}$. Differentiating $v = g^{-1}\dot{g}$ gives $\dot{v} = g^{-1}\ddot{g} - g^{-1}\dot{g}g^{-1}\dot{g} = g^{-1}\ddot{g} - v^2$, so the equation of motion is that the component of

$$\mathrm{skew}((\dot{v} + v^2)J)$$

that lies in $\mathfrak{k}$ should vanish. In terms of the angular momentum m from the Legendre transform $\mathbb{F}_e L\colon T_eG \to T_e^*G$, $v \mapsto m := Jv + vJ$, the equations of motion are

$$\begin{aligned}\dot{g} &= gv, \\ \dot{m} &= [m, v] + \sum_{j=1}^{n-k}\lambda_j\mu_j, \\ \left\langle v, \mu_j\right\rangle &= 0, \\ m &= Jv + vJ.\end{aligned}$$

The last three equations form a reduced nonholonomic system on $\mathfrak{g}^*$. They are an instance of Theorem 3.2 in [16], which states that the angular momentum component of such a nonholonomic system is the projection of the unconstrained Hamiltonian vector field on $\mathfrak{g}^*$.

We now apply the discrete Lagrange-d'Alembert principle to construct a discrete version of this system. We need to define L_d, A_d, and C_d. For the discrete Lagrangian we choose the (reversible) Moser-Veselov Lagrangian

$$L_d = \mathrm{tr}(g_0 J g_1^{-1}),$$

and we choose $A_d = A$ as in the continuous system. We will specify C_d below.

The reduced Lagrangian is $l_d(\omega) = \mathrm{tr}(\omega J)$, and to express the the reduced DLA equations (6.1) we need to evaluate the following derivative of l_d. For all $\xi \in \mathfrak{g}$ we have

$$\begin{aligned}\left\langle\omega^{-1}\mathbf{d}l_d(\omega), \xi\right\rangle &= \langle\mathbf{d}l_d(\omega), \omega\cdot\xi\rangle = \left\langle\mathbf{d}l_d(\omega), \left.\frac{d}{dt}\right|_{t=0}\omega\cdot\exp t\xi\right\rangle \\ &= \left.\frac{d}{dt}\right|_{t=0} l_d(\omega\cdot\exp t\xi) = \left.\frac{d}{dt}\right|_{t=0}\mathrm{tr}(\omega\exp(t\xi)J) \\ &= \mathrm{tr}(\omega\xi J) = \mathrm{tr}(J\omega\xi) = \langle\!\langle J\omega - \omega^{\mathrm{T}}J, \xi\rangle\!\rangle, \end{aligned} \tag{6.5}$$

where $\langle\!\langle , \rangle\!\rangle$ is the Killing inner product on $\mathfrak{g}$. Similarly,

$$\left\langle \mathrm{d}l_d(\omega)\cdot\omega^{-1}, \xi\right\rangle = \langle\!\langle \omega J - J\omega^{\mathrm{T}}, \xi\rangle\!\rangle.$$

Thus the reduced DLA equations (6.1) become

$$(J\omega_0 - \omega_0^{\mathrm{T}}J) - (\omega_1 J - J\omega_1^{\mathrm{T}}) = \sum_{j=1}^{n-k}\lambda_j\mu_j.$$

Just as in the general theory for unconstrained systems [6], this equation can, like the continuous equation, be reduced to $\mathfrak{g}^*$, which will suggest a natural choice for the discrete constraint C_d. The discrete angular momentum M is defined by

$$hM = \mathrm{d}l_d(\omega)\cdot\omega^{-1} = \omega J - J\omega^{\mathrm{T}}, \tag{6.6}$$

which suggests the reduced discrete constraint that M satisfies the same constraint as m. That constraint is

$$m \in \mathfrak{m} := \mathbb{F}_e L\cdot\mathfrak{k} = \{Jv + vJ : v\in\mathfrak{k}\}.$$

(The reduced velocity constraint submanifold c_d is then defined via (6.6), and C_d in turn via $\omega = g_0^{-1}g_1$.)

The reduced integrator on $\mathfrak{m}\subset\mathfrak{g}^*$ is then determined by $M_0\mapsto M_1$, as follows: first determine $\omega_0\in\mathrm{SO}(n)$ such that

$$hM_0 = \omega_0 J - J\omega_0^{\mathrm{T}}, \tag{6.7}$$

and then determine M_1 by

$$M_1 = \omega_0^{\mathrm{T}}M_0\omega_0 + \sum_{j=1}^{n-k}\lambda_j\mu_j, \tag{6.8}$$

where the (scaled) Lagrange multipliers λ_j are determined by the constraint $M_1\in\mathfrak{m}$.

The position update is provided by $g_1 = g_0\omega_0$. The integrator is defined for the same values of h as the unconstrained Moser-Veselov rigid body.

Note that the absence of constraints, the integrator (6.7,6.8) is identical to the original Moser-Veselov discrete rigid body (except that they used $\omega_0 = g_1^{-1}g_0$). With constraints, the integrator is defined by first calculating the unconstrained update and then projecting it to $\mathfrak{m}$, which is in striking analogy with the continuous reduced equations, for which the vector field undergoes the same projection to $\mathfrak{m}$. This is permitted, although not forced, by the discrete Lagrange-d'Alembert principle; a different choice of C_d would still allow a reduction to $\mathfrak{g}^*$, but not necessarily to a linear subspace of it.

7. Properties of the Exact Discrete Flow

In this section, we formulate the exact discrete Lagrange-d'Alembert (EDLA) equations, which are satisfied by the exact discrete flow. This approach is motivated by the correspondence between the Euler-Lagrange and discrete Euler-Lagrange equations for unconstrained systems, discussed in Section 3.1. However, we shall see that in the nonholonomic case the EDLA equations, unlike the discrete Euler-Lagrange equations, do not determine a discrete flow on the exact discrete constraint submanifold, although they are deterministic in the sense that if we fix q_0 and q_2, they determine q_1 such that the projection of the discrete flow contains the orbit sequence (q_0, q_1, q_2).

7.1. The Exact Discrete Constraint Manifold

In order to formulate the equations, it is necessary to define from the Lagrangian L and the vector bundle C, for a fixed time step h, an "integrated" object, C_d^e, which is the submanifold of $Q \times Q$ consisting of all pairs of points connectible by the Lagrange-d'Alembert flow in time h. This object is interesting in its own right. As it is a submanifold of the pair groupoid $Q \times Q$ (see [30] for a treatment of Lagrangian mechanics on groupoids), it is natural to study the source and target maps restricted to C_d^e. We show that if the Lagrangian is reversible, then the restricted maps are surjective submersions. The fibers, and in particular the intersection of their images under the opposing maps, are, assuming constant rank intersection, smooth submanifolds which have a natural interpretation as the set of all points that join q_0 to q_2 under the composition of two different flows corresponding to a v_{q_0} and a $v_{q_1} \neq \psi^h(v_{q_0})$. Imposing the exact discrete equation picks out the q_1 with the property that there is a velocity over q_0 whose flow passes through q_1 and then reaches q_2 at time $2h$.

Finally, it is possible to define a discrete Legendre transformation with the property that for the exact discrete flow on C_d^e, there is a well-defined conjugate momentum value for each time step whose value is simply the conjugate momentum of the smooth solution evaluated at the discrete times.

We begin by defining the exact discrete constraint distribution and the exact Lagrangian for nonholonomic systems.

Nonconjugate solutions. First we define the notion of a nonconjugate solution $q(t)$ for a nonholonomic system joining q_0 to q_1 in time h. This will generalize the definition of nonconjugate solutions for unconstrained Lagrangian mechanics. Let $\psi^t\colon C \to C$ denote the flow of the nonholonomic system. Notice that there must exist $v_{q_0} \in C_{q_0}$ such that $q(t) = \tau \circ \psi^t(v_{q_0})$. Furthermore, $q(t)$ is nonconjugate provided there is a neighborhood U of v_{q_0} in C such that for all $v_q \in U$, the map,

$$\mathbb{F}D_{v_q}(\tau \circ \psi^h)\colon C_q \to T_{\tau\circ\psi^h(v_q)}Q, \tag{7.1}$$

where $\mathbb{F}D$ indicates the fiber derivative, is injective. If we take a sufficiently small neighborhood $U_{(q_0,q_1)}$ of (q_0, q_1) in C_d^e, and a point $(\bar{q}_0, \bar{q}_1) \in U_{(q_0,q_1)}$, this definition allows us to conclude the existence of a unique $v_{\bar{q}_0}$ such that $\tau \circ \psi^h(v_{\bar{q}_0}) = \bar{q}_1$. Furthermore, suppose we take the unconstrained limit $C = TQ$. Then this definition agrees with the condition that $q(t)$ is a nonconjugate solution joining q_0 to q_1, and allows us to conclude the existence of neighborhoods U_{q_0}, U_{q_1} of q_0 and q_1 respectively such that for any pair $(q_0', q_1') \in U_{q_0} \times U_{q_1}$, there exists a solution $q'(t)$ joining q_0' and q_1' in time h. Finally, it is possible to show that injectivity of $\mathbb{F}D_{v_q}$ is equivalent to the requirement that each solution of the following Jacobi equation with nonzero initial data does not vanish along the curve $q(t) = \tau \circ \psi^t(v_{q_0})$. For each i,

$$\frac{d}{dt}\left(\frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_i}\dot{U}_j + \frac{\partial^2 L}{\partial q_j \partial \dot{q}_i}U_j\right) - \frac{\partial^2 L}{\partial \dot{q}_j \partial q_i}U_j = \sum_{l=1}^{n-k} \frac{\partial \lambda_l}{\partial q_j} U_j A_{li} + \lambda_l \frac{\partial A_{li}}{\partial q_j} U_j, \tag{7.2}$$

with summation over j understood. Notice this agrees with the usual Jacobi equation when the constraints vanish, $A_i = 0$.

Notice that if the constraint is integrable, this condition will fail, but then one can apply this definition restricted to the integral submanifolds of the distribution.

For the following two definitions, fix a small time step h.

Definition 8. C_d^e, the *exact discrete constraint distribution*, is the subset of $Q \times Q$ consisting of pairs (q_0, q_1) such that there exists a C^2 curve $q(t)$ joining q_0 to q_1 in time h and the curve $(q(t), \dot{q}(t))$ satisfies the Lagrange-d'Alembert equations. (q_0, q_1) is called *nonconjugate* provided the solution curve $q(t)$ is nonconjugate.

Proposition 10. *In a neighborhood of a nonconjugate point (q_0, q_1) there exist smooth coordinates realizing C_d^e locally as a submanifold of $Q \times Q$ of dimension $n + k$.*

Proof. Consider the vector bundle $C \to Q$. Pick a neighborhood U_0 of the point q_1 small enough so that the bundle trivializes over U_0, i.e., $C|_{U_0} \simeq U_0 \times C_{q_0}$. Let us denote by Φ the trivializing bundle map. By nonconjugacy of (q_0, q_1), we know that $\tau \circ \psi^h$, restricted to a neighborhood $\mathcal{B}_{q_0}$ of v_{q_0} in the fiber over q_0 is injective. As this is an open condition in q_0, it follows that there exists a neighborhood $U_0' \subset U_0$ so that for each $q_0' \in U_0'$, there is a corresponding neighborhood $\mathcal{B}_{q_0'}$ of in the fiber of C over q_0' so that $\tau \circ \psi^h$ restricted to that neighborhood is injective. Shrinking the base neighborhood of the bundle if necessary in the trivialization, we can then find a fixed neighborhood V_0 of v_{q_0} in the model vector space C_{q_0} so that $\tau \circ \psi^h \circ \Phi$ is injective on each fiber of $U_0' \times V_0$. It follows that $U_0' \times V_0$ is a coordinate domain for C_d^e with smooth coordinate map

$$U_0' \times V_0 \ni (q_0', v') \mapsto \tau \circ \psi^h \circ \Phi(q_0', v'). \tag{7.3}$$

By construction this takes the point (q_0, v_{q_0}) to $(q_0, q_1) \in C_d^e$. Finally, it is clear from construction that $\dim U_0' \times V_0 = n + k$. □

We can now define the exact discrete Lagrangian for a nonholonomic system as follows.

Definition 9. The *exact discrete Lagrangian* is the map $L_d^e\colon C_d^e \to \mathbb{R}$ defined by

$$L_d^e(q_0, q_1) = \int_0^h L(q(t), \dot{q}(t))\, dt, \tag{7.4}$$

where $q(t)$ is the unique curve joining q_0 to q_1 in time h given in the definition of C_d^e.

If there are no constraints, this reduces to the exact discrete Lagrangian of unconstrained mechanics. If the constraints are integrable, notice that this exact discrete Lagrangian is defined on the submanifold of $Q \times Q$ given by $\bigcup_l \mathcal{F}_l \times \mathcal{F}_l$ where the $\mathcal{F}_l$ are the leaves of the distribution.

In the next proposition we show that C_d^e satisfies the requirements of Definition 2 for a discrete constraint distribution modeled on a distribution C (equation (3.7)).

Proposition 11. *Fix $q_0 \in Q$. We then have the following:*

(i) $(q_0, q_0) \in C_d^e$;
(ii) $0 \times C_{q_0} \subset T_{(q_0,q_0)} C_d^e$;
(iii) *If L is reversible, then we also have $C_{q_0} \times 0 \subset T_{(q_0,q_0)} C_d^e$.*

Proof. It is obvious that (i) holds since $q(t) = q_0$ satisfies the Lagrange-d'Alembert equations corresponding to initial data $(q_0, v_{q_0} = 0)$. To prove the (ii) we first show that the Lagrange-d'Alembert solutions admit the following scaling action of $\mathbb{R}^*$: $q(t)$ is a solution of Lagrange-d'Alembert if and only if $q(ct)$ is a solution for each nonzero real number c. It is easy to obtain this fact from the Lagrange-d'Alembert principle analogous to the proof of Proposition 1. For fixed c and fixed T consider the isomorphism on the space of C^2 curves with fixed endpoints

$$\tilde{\varphi}_c \colon \Omega([0,T]; q_0, q_1) \to \Omega\left(\left[O, \frac{1}{c}T\right]; q_0, q_1\right), \tag{7.5}$$

given by $q(t) \mapsto q(ct)$. This map induces an isomorphism on the tangent spaces taking a variation $\delta q(t)$ to the variation $\delta q_c(t) = \delta q(ct)$. It follows that $q(t)$ is a critical point of the action integral with respect to $\delta q(t) \in C_{q(t)}$ and satisfies $\dot{q}(t) \in C_{q(t)}$ if and only if $q(ct)$ is critical with respect to $\delta q_c(t)$ and satisfies $\dot{\varphi(q)}(t) \in C_{q(t)}$.

Next, fix $v_{q_0} \in C_{q_0}$ and consider the curve in C_d^e through the point (q_0, q_0) given by $\epsilon \mapsto (q_0, \tau \circ \psi^h(\epsilon \cdot \frac{1}{h} v_{q_0}))$. Let $q(t)$ denote the solution to the Lagrange-d'Alembert equations with initial data $(q_0, \frac{1}{h} v_{q_0})$. Then by the scaling action invariance of solutions to Lagrange-d'Alembert we have that $q(t\epsilon)$ is a solution with initial data $(q_0, \epsilon \cdot \frac{1}{h} v_{q_0})$. Therefore, $q(h\epsilon) = \tau \circ \psi^h(\epsilon \cdot \frac{1}{h} v_{q_0})$ and

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} q(h\epsilon) = h\dot{q}(0) = v_{q_0}, \tag{7.6}$$

as required. Finally for (iii), note that if $(q_0, q_1) \in C_d^e$ so that there exists v_{q_0} such that $\tau \circ \psi^h(v_{q_0}) = q_1$, then if $v_{q_1} = \psi^h(v_{q_0})$, we have, using reversibility of the Lagrange-d'Alembert flow,

$$\psi^h(R(v_{q_1})) = R(\psi^{-h} v_{q_1}) = R(v_{q_0}),$$

so that $\tau \circ \psi^h(-v_{q_1}) = q_0$ proving that $(q_1, q_0) \in C_d^e$, that is, C_d^e is reversible. Therefore $(q_0, q_0(\epsilon)) \in C_d^e$ implies $(q_0(\epsilon), q_0) \in C_d^e$ so that, after taking the derivative in ϵ, we get that $C_{q_0} \times 0 \subset T_{(q_0,q_0)} C_d^e$, as required. □

Reversibility of C_d^e. Reversibility of the Lagrange-d'Alembert flow translates to C_d^e being closed under the inverse operator on the pair groupoid $Q \times Q$. That is, $(q_0, q_1) \in C_d^e \iff (q_1, q_0) \in C_d^e$. To see this, given $(q_0, q_1) \in C_d^e$, there exists $v_{q_0} \in C_{q_0}$ such that $\tau \circ \psi^h(v_{q_0}) = q_1$. Let $v_{q_1} = \psi^h(v_{q_0})$. Again using the fact that $\tau \circ \psi^{-t} = \tau \circ (R \circ \psi^t \circ R) = \tau \circ \psi^t \circ R$, we have

$$\tau(\psi^h(R(v_{q_1}))) = \tau \circ \psi^{-h}(v_{q_1}) = \tau \circ \psi^{-h}(\psi^h(v_{q_0})) = \tau(v_{q_0}) = q_0,$$

proving that $\tau \circ \psi^h(-v_{q_1}) = q_0$, which shows that $(q_1, q_0) \in C_d^e$.

A closer look at C_d^e.

$$\begin{array}{c} C_d^e \\ \mathbf{s}\big\downarrow\big\downarrow\mathbf{t}. \\ Q \end{array}$$

Borrowing terminology from the theory of groupoids, we call **s** and **t** source and target maps, respectively. They are simply given by $\mathbf{s} = \pi_1|_{C_d^e}$ and $\mathbf{t} = \pi_2|_{C_d^e}$. We then have the interpretation that the s-fiber over a point q consists of the points in C_d^e which correspond to time-h solutions that originate at q. The reachable set is then $\mathbf{t} \circ \mathbf{s}^{-1}(q)$ which is a k-dimensional submanifold of Q corresponding to the endpoints of time h trajectories of the Lagrange-d'Alembert flow. It is worth proving that in the case that the Lagrangian is reversible, the maps **s** and **t** are surjective submersions (**s** is always a surjective submersion). Indeed, to see that **s** is a submersion, take $v_{q_0} \in T_{q_0}Q$ arbitrary. Let $q_0(t)$ be a curve such that $\frac{d}{dt}\big|_{t=0} q_0(t) = v_{q_0}$. Take an arbitrary section $V_{q_0}(t)$ over $q_0(t)$ in the bundle $C \to Q$. Then $t \mapsto \left(q_0(t), \tau \circ \psi^h(V_{q_0}(t))\right)$ is a curve in C_d^e and by construction

$$T\mathbf{s} \cdot \frac{d}{dt}\bigg|_{t=0} (q_0(t), \tau \circ \psi^h(V_{q_0}(t))) = \frac{d}{dt}\bigg|_{t=0} q_0(t) = v_{q_0},$$

as required. Now assume the flow is reversible. By reversibility it is clear that ψ^{-t} exists and is defined for exactly the time that the forward flow is defined since $\psi^{-t} = R \circ \psi^t \circ R$. Now, fix $q \in Q$ and $v_q \in C_q$. Since $\psi^{-h}(v_q) = R \circ \psi^h(-v_q)$, we have $\psi^h(\psi^{-h}(v_q)) = v_q$ and therefore $(\tau \circ \psi^{-h}(v_q), q) \in C_d^e$, so that **t** is surjective. To see that **t** is submersive, fix $v_q \in T_qQ$ and let $q(t)$ be a curve such that $\frac{d}{dt}\big|_{t=0} q(t) = v_q$. As before, let $V_{q(t)}$ be a section of $C \to Q$ over $q(t)$. Define the curve $q_0(t) := \tau \circ R\left(\psi^h(-V_{q(t)})\right)$ and a section over $q_0(t)$ by $V_{q_0(t)} := R\left(\psi^{-h}(V_{q(t)})\right)$. Then, since $\psi^h\left(V_{q_0(t)}\right) = \psi^h\left(\psi^{-h}(V_{q(t)})\right) = V_{q(t)}$, it follows that $(q_0(t), q(t)) \in C_d^e$, and by construction we then have

$$T_{(q_0,q)}\mathbf{t} \cdot \frac{d}{dt}\bigg|_{t=0} (q_0(t), q(t)) = v_q,$$

proving that **t** is a surjective submersion.

Geometry of the fibers. First, the fact that **s** and **t** are submersions proves that the fibers are smooth submanifolds of dimension k. We then consider the projections of these fibers onto the opposing legs of the two submersions, i.e., $\mathbf{t} \circ \mathbf{s}^{-1}(q_0)$ and $\mathbf{s} \circ \mathbf{t}^{-1}(q_2)$. It is clear that these are submanifolds of Q of dimension k.

Proposition 12. *Given a reversible Lagrangian, the following properties of C_d^e hold true:*

(i) *Given $(q_0, q_1) \in C_d^e$ and $(q_1, q_2) \in C_d^e$, both nonconjugate as in Definition 8, the tangent spaces of the fibers of* **s** *and* **t** *are expressed in terms of the flow as follows:*

$$V_{q_0}^{q_1} := T_{q_1}\left(\mathbf{t}\circ\mathbf{s}^{-1}(q_0)\right) = \mathbb{F}D_{v_{q_0}}(\tau\circ\psi^h)(C_{q_0}), \tag{7.7}$$

and

$$V_{q_2}^{q_1} := T_{q_1}\left(\mathbf{s}\circ\mathbf{t}^{-1}(q_2)\right) = \mathbb{F}D_{v_{q_2}}(\tau\circ\psi^{-h})(C_{q_2}), \tag{7.8}$$

where $v_{q_0} \in C_{q_0}$ is such that $\tau\circ\psi^h(v_{q_0}) = q_1$ and v_{q_2} satisfies $\tau\circ\psi^{-h}(v_{q_2}) = q_1$.

(ii) *Given $(q_0, q_2) \in C_d^e$ nonconjugate, the submanifolds $\mathbf{t}(\mathbf{s}^{-1}(q_0))$ and $\mathbf{s}(\mathbf{t}^{-1}(q_2))$ have nonempty intersection. Let $q \in \mathbf{t}(\mathbf{s}^{-1}(q_0)) \cap \mathbf{s}(\mathbf{t}^{-1}(q_2))$. We then have*

$$\dim T_q\left(\mathbf{t}(\mathbf{s}^{-1}(q_0))\right) \cap T_q\left(\mathbf{s}(\mathbf{t}^{-1}(q_2))\right) \geq 1. \tag{7.9}$$

(iii) *Let $(q_0, q_2) \in C_d^e$ be nonconjugate. For each $q_1 \in \mathbf{t}(\mathbf{s}^{-1}(q_0)) \cap \mathbf{s}(\mathbf{t}^{-1}(q_2))$, suppose the subspace $V_{q_0,q_2}^{q_1} := V_{q_0}^{q_1} \cap V_{q_2}^{q_1}$ has constant dimension (i.e. independent of q_1). Then the intersection set $\mathbf{t}(\mathbf{s}^{-1}(q_0)) \cap \mathbf{s}(\mathbf{t}^{-1}(q_2))$ is a submanifold of dimension* $\dim V_{q_0 q_2}^{q_1}$.

Proof. To prove (i), we simply observe that $q(t)$ is a curve through $\mathbf{t}(\mathbf{s}^{-1}(q_0))$ if and only if $(q_0, q(t)) \in C_d^e$ if and only if $q(t) = \tau\circ\psi^h(v_{q_0}(t))$ for some curve $v_{q_0}(t) \in C_{q_0}$ passing through v_{q_0} at $t = 0$ where v_{q_0} satisfies $\tau\circ\psi^h(v_{q_0}) = q_1$. Therefore every tangent vector $v_{q_1} \in T_{q_1}(\mathbf{t}\circ\mathbf{s}^{-1}(q_0))$ is the derivative of a curve of the form $\tau\circ\psi^h(v_{q_0}(t))$ and therefore, since

$$\left.\frac{d}{dt}\right|_{t=0} \tau\circ\psi^h(v_{q_0}(t)) = \langle \mathbb{F}D_{v_{q_0}}(\tau\circ\psi^h), w_{q_0}\rangle,$$

where $w_{q_0} := \frac{d}{dt}\big|_{t=0} v_{q_0}(t) \in C_{q_0}$, equation (7.7) follows. Note that by the nonconjugacy assumption of the point (q_0, q_1), it follows that $\dim V_{q_0}^{q_1} = k$. Similarly, by reversibility we see that every $q(t)$ smooth curve passing through q_1 and satisfying $(q(t), q_2) \in C_d^e$ if and only if there exists a curve $v_{q_2}(t) \in C_{q_2}$ such that $q(t) = \tau\circ\psi^{-h}(v_{q_2}(t))$ from which equation (7.8) follows by differentiating.

Both facts in (ii) will follow from the scaling symmetry of the Lagrange-d'Alembert equations and the reversibility of L. Note that the scaling symmetry tells us that for each nonnegative real number s, $\tau\circ\psi^h(s v_{q_0}) = \tau\circ\psi^{sh}(v_{q_0})$. Now, supposing $(q_0, q_2) \in C_d^e$, there must exist a $v_{q_0} \in C_{q_0}$ such that $\tau\circ\psi^h(v_{q_0}) = q_2$. It follows by scaling that $\tau\circ\psi^{2h}(\frac{1}{2}v_{q_0}) = q_2$ so that we can take $q_1 := \tau\circ\psi^h(\frac{1}{2}v_{q_0})$ and $C_{q_1} \ni v_{q_1} := \psi^h(v_{q_0})$. Now consider the projected integral curve $[0, 2h] \ni t \mapsto \tau\circ\psi^t(v_{q_0})$. Again by the scaling action, each point $(q_0, \tau\circ\psi^t(v_{q_0}))$ lies in C_d^e since we can write, for each t, $\psi^t(v_{q_0}) = \psi^{sh}(v_{q_0}) = \psi^h(s v_{q_0})$ for $s := t/h$. Let $v_{q_2} := \psi^{2h} v_{q_0}$. Furthermore, we have $(\tau\circ\psi^t(v_{q_0}), q_2) \in C_d^e$ since $\tau\circ\psi^{2h-t}\psi^t(v_{q_0}) = q_2$ and by reversibility, $\psi^{2h-t}(R(v_{q_2})) = R\circ\psi^{-(2h-t)}v_{q_2}$. However, $\tau\circ R\circ\psi^{-(2h-t)}v_{q_2} = \tau\circ\psi^t(v_{q_0})$ by construction so that $(q_2, \tau\circ\psi^t(v_{q_0})) \in C_d^e$. Finally, by discrete reversibility of C_d^e it follows that $(\tau\circ\psi^t(v_{q_0}), q_2) \in C_d^e$.

For (iii) notice that the dimension of the space $V^{q_1}_{q_0q_2}$ is locally constant for all $q_1 \in \mathbf{t}(\mathbf{s}^{-1}(q_0)) \cap \mathbf{s}(\mathbf{t}^{-1}(q_2))$ by smoothness of the maps $q_1 \mapsto V^{q_1}_{q_0}$ and $q_1 \mapsto V^{q_1}_{q_2}$. Assuming the dimension is constant over the entire set, we can apply the constant rank theorem to conclude that $\mathbf{t}(\mathbf{s}^{-1}(q_0)) \cap \mathbf{s}(\mathbf{t}^{-1}(q_2))$ is a submanifold of Q. □

7.2. The Exact Discrete Equations

We next determine, for reversible Lagrangians, the "exact discrete Lagrange-d'Alembert" (EDLA) equations (7.12) that are satisfied by the sequence $\{q_n = q(nh)\}$ where h is a fixed time step and $q(t)$ is a solution to Lagrange-d'Alembert.

Theorem 5 (Exact Discrete Lagrange-d'Alembert). *Assume L is a reversible Lagrangian. Let $(q_0, q_1) \in C^e_d$ be nonconjugate. Let $q_2 = \tau \circ \psi^h(v_{q_1})$ so that $(q_1, q_2) \in C^e_d$, and assume that the time step h is small enough that (q_1, q_2) is also nonconjugate. Define the function $q_{0,1}\colon C^e_d \times [0, h] \to Q$ by $q_{0,1}(\bar{q}_0, \bar{q}_1, t, h) = \bar{q}(t)$ where $\bar{q}(t)$ is the time t evaluation of the curve joining $\bar{q}_0$ to $\bar{q}_1$ in time h whose tangent lift is a solution to the Lagrange-d'Alembert equations. We define the following covectors, $F^+(q_0, q_1) \in (V^{q_1}_{q_0})^*$ given by*

$$F^+(q_0, q_1)(v_{q_1}) = \int_0^h \left\langle \lambda^T(q_{0,1}(q_0, q_1, t, h)) A(q_{0,1}(q_0, q_1, t, h)), \frac{\partial q_{0,1}}{\partial q_1} \cdot v_{q_1} \right\rangle dt, \tag{7.10}$$

and $F^-(q_1, q_2) \in (V^{q_1}_{q_2})^*$ given by

$$F^-(q_1, q_2)(v_{q_1}) = \int_0^h \left\langle \lambda^T(q_{0,1}(q_1, q_2, t, h)) A(q_{0,1}(q_1, q_2, t, h)), \frac{\partial q_{0,1}}{\partial q_0} \cdot v_{q_1} \right\rangle dt, \tag{7.11}$$

where we use the shorthand notation

$$\lambda^T(q) A(q) := \sum_{j=1}^{n-k} \lambda_j(q) A_j(q) \in T^*_q Q.$$

Then the following equation is satisfied on the subspace $V^{q_1}_{(q_0,q_2)}$:

$$D_1 L^e_d(q_1, q_2) + D_2 L^e_d(q_0, q_1) = -F^+(q_0, q_1) - F^-(q_1, q_2). \tag{7.12}$$

Proof. First observe that given $v_{q_1} \in V^{q_1}_{q_0,q_2}$ there exist curves $(q_0, q_1(\epsilon)) \in C^e_d$ and $(\bar{q}_1(\epsilon), q_2) \in C^e_d$ such that $\frac{d}{d\epsilon}\big|_{\epsilon=0} q_1(\epsilon) = \frac{d}{d\epsilon}\big|_{\epsilon=0} \bar{q}_1(\epsilon) = v_{q_1}$. Next, computing the derivative of the function $q_{0,1}$ at the point $(q_0, q_1) \in C^e_d$, we obtain

$$\frac{\partial q_{0,1}}{\partial q_1}(q_0, q_1) \cdot v_{q_1} = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} q_{0,1}(q_0, q_1(\epsilon), t, h).$$

As a function of t, this is a vector field along the solution curve $q(t)$ joining q_0 to q_1. Furthermore this vector field vanishes at the initial endpoint q_0 since, for each ϵ,

$q_{0,1}(q_0, q_1(\epsilon), t, h)$ is a solution curve which originates at q_0. Similarly $\frac{\partial q_{0,1}}{\partial q_0}(q_1, q_2) \cdot v_{q_1}$ is a vector field along the solution curve joining q_1 to q_2 with the property that it vanishes at the endpoint q_2. Starting from the definition $L_d^e(q_0, q_1) = \int_0^h L(q_{0,1}(q_0, q_1, t, h)\, dt$, we compute, for fixed $v_{q_1} \in V_{q_1}^{q_0,q_2}$, $D_2 L_d^e(q_0, q_1) \cdot v_{q_1}$ as follows:

$$\begin{aligned} D_2 L_d^e(q_0, q_1) \cdot v_{q_1} &= \int_0^h \frac{\partial L}{\partial q} \frac{\partial q_{0,1}}{\partial q_1} \cdot v_{q_1} + \frac{\partial L}{\partial \dot{q}} \frac{\partial \dot{q}_{0,1}}{\partial q_1} \cdot v_{q_1}\, dt \\ &= \int_0^h \frac{\partial L}{\partial q} \frac{\partial q_{0,1}}{\partial q_1} \cdot v_{q_1} + \frac{\partial L}{\partial \dot{q}} \frac{d}{dt} \frac{\partial q_{0,1}}{\partial q_1} \cdot v_{q_1}\, dt \qquad (7.13) \\ &= \int_0^h \frac{\partial L}{\partial q} \frac{\partial q_{0,1}}{\partial q_1} \cdot v_{q_1} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \frac{\partial q_{0,1}}{\partial q_1} \cdot v_{q_1}\, dt + \frac{\partial L}{\partial \dot{q}} \frac{\partial q_{0,1}}{\partial q_1} \cdot v_{q_1} \Big|_{t=0}^{t=h}, \end{aligned}$$

where we use the equality of mixed partial derivatives in the second equality and the third is from integration by parts. Recall that $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q}$ is the Euler-Lagrange operator, denoted $D_{\mathrm{EL}} L$: $\ddot{Q} \to T^*Q$, which is a bundle map from second derivatives of curves, $\ddot{Q}$, to T^*Q, so that the integrand of the last equation is just

$$\int_0^h \left\langle -D_{\mathrm{EL}} L(\ddot{q}_{0,1}(q_0, q_1)), \frac{\partial q_{0,1}}{\partial q_1}(q_0, q_1) \cdot v_{q_1} \right\rangle dt.$$

The integrand is well defined since $D_{\mathrm{EL}} L(\ddot{q}_{0,1}(q_0, q_1))$ is a covector field along $q_{0,1}(q_0, q_1, t, h)$ and it is paired with a vector field along the same curve. Since the curve $t \mapsto q_{0,1}(q_0, q_1, t, h)$ is a solution to the Lagrange-d'Alembert equations, we have

$$D_{\mathrm{EL}} L(\ddot{q}_{0,1}(q_0, q_1)) = \lambda^{\mathrm{T}}(q_{0,1}(t)) A(q_{0,1}(t)),$$

and therefore, equation (7.13) becomes

$$D_2 L_d^e(q_0, q_1) \cdot v_{q_1} = -F^+(q_0, q_1) \cdot v_{q_1} + \frac{\partial L}{\partial \dot{q}} \frac{\partial q_{0,1}}{\partial q_1}(q_0, q_1) \cdot v_{q_1} \Big|_{t=0}^{t=h}. \qquad (7.14)$$

In a completely analogous way, one can compute

$$\begin{aligned} D_1 L_d^e(q_1, q_2) \cdot v_{q_1} &= \int_0^h \left\langle -D_{\mathrm{EL}}(\ddot{q}_{0,1}(q_1, q_2)), \frac{\partial q_{0,1}}{\partial q_0}(q_1, q_2) \cdot v_{q_1} \right\rangle dt \\ &\quad + \frac{\partial L}{\partial \dot{q}} \frac{\partial q_{0,1}}{\partial q_0}(q_1, q_2) \cdot v_{q_1} \Big|_{t=0}^{t=h} \\ &= -F^-(q_1, q_2) \cdot v_{q_1} + \frac{\partial L}{\partial \dot{q}} \frac{\partial q_{0,1}}{\partial q_0}(q_1, q_2) \cdot v_{q_1} \Big|_{t=0}^{t=h}. \qquad (7.15) \end{aligned}$$

We now consider the boundary terms in equations (7.14) and (7.15). Recall that the vector field $\frac{\partial q_{0,1}}{\partial q_1}(q_0, q_1) \cdot v_{q_1}$ vanishes at $q_{0,1}(q_0, q_1, t = 0, h) = q_0$ and, evaluated at $q_{0,1}(q_0, q_1, t = h, h) = q_1$, is equal to v_{q_1} by construction. Similarly, the vector field

$\frac{\partial q_{0,1}}{\partial q_0}(q_1, q_2) \cdot v_{q_1}$ vanishes at $t = h$ and is equal to v_{q_1} at $t = 0$. We then have

$$\begin{aligned} D_2 L_d^e(q_0, q_1) \cdot v_{q_1} &= -F^+(q_0, q_1) \cdot v_{q_1} + \frac{\partial L}{\partial \dot q} \frac{\partial q_{0,1}}{\partial q_1}(q_0, q_1) \cdot v_{q_1} \Big|_{t=0}^{t=h} \\ &= -F^+(q_0, q_1) \cdot v_{q_1} \\ &\quad + \left\langle \frac{\partial L}{\partial \dot q}(\dot q_{0,1}(q_0, q_1, t, h)), \frac{\partial q_{0,1}}{\partial q_1}(q_0, q_1) \cdot v_{q_1} \right\rangle \Big|_{t=0}^{t=h} \\ &= -F^+(q_0, q_1) \cdot v_{q_1} + \mathbb{F}L(\dot q_{0,1}(q_0, q_1, t = h, h)) \cdot v_{q_1}, \end{aligned} \tag{7.16}$$

and that

$$\begin{aligned} D_1 L_d^e(q_1, q_2) \cdot v_{q_1} &= -F^-(q_1, q_2) \cdot v_{q_1} + \frac{\partial L}{\partial \dot q} \frac{\partial q_{0,1}}{\partial q_0}(q_1, q_2) \cdot v_{q_1} \Big|_{t=0}^{t=h} \\ &= -F^-(q_1, q_2) \cdot v_{q_1} \\ &\quad + \left\langle \frac{\partial L}{\partial \dot q}(\dot q_{0,1}(q_1, q_2, t, h)), \frac{\partial q_{0,1}}{\partial q_0}(q_1, q_2) \cdot v_{q_1} \right\rangle \Big|_{t=0}^{t=h} \\ &= -F^-(q_1, q_2) \cdot v_{q_1} - \mathbb{F}L(\dot q_{0,1}(q_1, q_2, t = 0, h)) \cdot v_{q_1}. \end{aligned} \tag{7.17}$$

Finally observe that $\mathbb{F}L(\dot q_{0,1}(q_0, q_1, t = h, h)) = \mathbb{F}L(\dot q_{0,1}(q_1, q_2, t = 0, h))$, and therefore adding these equations for $D_1 L_d^e$ and $D_2 L_d^e$ yields

$$D_2 L_d^e(q_0, q_1) + D_1 L_d^e(q_1, q_2) = -F^+(q_0, q_1) - F^-(q_1, q_2),$$

proving the theorem. □

Observe that if there are no constraints, the tangent spaces $V_{q_1}^{q_0}$, $V_{q_1}^{q_2}$ become the full tangent space $T_{q_1} Q$ provided the time step is small enough and the point (q_0, q_1) is nonconjugate for the Euler-Lagrange flow. In this case the right-hand side of equation (7.12) vanishes and we therefore recover Theorem 2.

Another interpretation of the EDLA equations (7.12) arises if we are given a nonconjugate pair $(q_0, q_2) \in C_d^e$ and would like to determine the discrete orbit (q_0, q_1, q_2). Naturally we require $q_1 \in \mathbf{t}(\mathbf{s}^{-1}(q_0)) \cap \mathbf{s}(\mathbf{t}^{-1}(q_2))$ (i.e., $(q_0, q_1) \in C_d^e$ and $(q_1, q_2) \in C_d^e$), which is a submanifold of dimension $\dim V_{q_0 q_2}^{q_1}$. Then (7.12) puts $\dim V_{q_0 q_2}^{q_1}$ further conditions on q_1 and thus determines q_1 locally uniquely.

However, unlike the DLA equations (3.9), (3.10), eq. (7.12) is not a suitable principle for an *initial* value problem. Given $(q_0, q_1) \in C_d^e$, we would like to find q_2 such that (q_0, q_1, q_2) forms a discrete orbit. We can require that $(q_1, q_2) \in C_d^e$, a submanifold of dimension k, but (7.12)) represents $\dim V_{q_0 q_2}^{q_1}$ further equations. $V_{q_0 q_2}^{q_1}$ is the intersection of two k-planes in an n-dimensional vector space; hence, $\max(0, 2k - n) \le V_{q_0 q_2}^{q_1} \le \dim k$ and there are not in general enough equations to determine a unique q_2.

Momentum interpretation. Analagous to the theory of variational integrators for the Euler-Lagrange equations, we have a momentum interpretation of the EDLA equations. However, to get a consistent time-i value of momentum, we need to define the discrete

Legendre transformations differently than in the Euler-Lagrange theory (equations (3.4) and (3.5)). We need to take into account the impulses due to the constraint forces.

Definition 10. The discrete Legendre transformations in the presence of nonholonomic constraints are given by

$$\mathbb{F}^+ L_d\colon (q_0, q_1) \mapsto (q_1, p_1) = (q_1, D_2 L_d(q_0, q_1) + F^+(q_0, q_1)) \in (V_{q_0}^{q_1})^*, \quad (7.18)$$

and

$$\mathbb{F}^- L_d\colon (q_0, q_1) \mapsto (q_0, p_0) = (q_0, -D_1 L_d(q_0, q_1) - F^-(q_0, q_1)) \in (V_{q_1}^{q_0})^*. \quad (7.19)$$

Using these fiber-preserving maps, we have the following consequence of Theorem 5.

Corollary 6. *Let $\{q_i\}$ denote a discrete time solution to the Lagrange-d'Alembert equations which therefore satisfies the EDLA equations of the previous theorem. Let $p^+_{q_{i-1}q_i} := \mathbb{F}^+ L^e_d(q_{i-1}, q_i)$ and $p^-_{q_i q_{i+1}} := \mathbb{F}^- L^e_d(q_i, q_{i+1})$. Then*

(i) $p^+_{q_{i-1}q_i} = p^-_{q_i q_{i+1}}$.
(ii) $p^+_{q_{i-1}q_i} = \mathbb{F}L(\dot{q}_{0,1}(q_{i-1}, q_i, t = h, h))$.

In other words, there is a well-defined conjugate momentum value for each time step, and this momentum value, defined through the discrete Legendre transformation, agrees with the continuous momentum of the actual flow (see equation (2.12)), evaluated at the discrete times.

Proof. (i) is an immediate consequence of Theorem 5, since it is merely rewriting equation (7.12) in terms of the discrete momenta. For (ii), we have, from the proof of the previous theorem, that

$$\begin{aligned} p^+_{q_{i-1}q_i} &= D_2 L^e_d(q_{i-1}, q_i) + F^+(q_{i-1}, q_i) \\ &= \mathbb{F}L(\dot{q}_{0,1}(q_{i-1}, q_i, t = h, h)). \end{aligned}$$

However, this last term is simply the fiber derivative of the Lagrangian applied to the tangent vector of the solution curve through q_i proving the claim. Notice that when there are no constraints, the forces vanish and we recover the momentum matching condition of the discrete Euler-Lagrange equations. □

Failure to be a subgroupoid. The set C^e_d is *not* a subgroupoid of the pair groupoid $Q \times Q$, because it fails to be closed under multiplication. For, if (q_0, q_1) and (q_1, q_2) are elements in C^e_d, $(q_0, q_2) = (q_0, q_1) \cdot (q_1, q_2)$ will belong to C^e_d if and only if (q_0, q_1, q_2) is on a single dynamical orbit. In that case $(q_0, q_2) \in C^e_d$ since, by the scaling symmetry of the Lagrange-d'Alembert flow, we can choose $v'_{q_0} = 2v_{q_0}$ so that $\tau \circ \psi^h(v'_{q_0}) = q_2$. Equivalently we can multiply (q_0, q_1) with (q_1, q_2) in C^e_d if and only if q_1 satisfies the EDLA equations if and only if the discrete momentum of (q_0, q_1) matches the discrete momentum of (q_1, q_2).

Finally we point out that if we consider the image of $T\mathbf{t}|_\Delta$, along ker $T\mathbf{s}|_\Delta$, we simply obtain the bundle $C \to Q$. This follows since, by Proposition 11, we have

$0 \times C_{q_0} \subset T_{(q_0,q_0)} C_d^e$. Now clearly $0 \times C_{q_0} \subset \ker T_{(q_0,q_0)}\mathbf{s}$. On the other hand since $\mathbf{s}$ is a submersion and $\dim C_d^e = n + k$, it follows that $0 \times C_{q_0}$ is the entire kernel. We have shown that

$$\ker T\mathbf{s}|_\Delta \simeq C. \tag{7.20}$$

This is not an algebroid over Q (as it would be if C_d^e were a subgroupoid) since the bracket of sections, corresponding to the Jacobi bracket of vector fields with values in the distribution, is not closed as C is not integrable.

In the case that C is an integrable distribution, it is an algebroid over Q. Its corresponding C_d^e is in this case just $\bigcup_l \mathcal{F}_l \times \mathcal{F}_l$ a subgroupoid of $Q \times Q$ consisting of pairs (q_0, q_1) with the property that q_0 and q_1 are on the same leaf of the foliation of Q determined by C. In this case we are able to multiply elements of C_d^e since the dynamics restricted to each leaf is just Euler-Lagrange.

8. Conclusions

We have studied a discrete principle that is general enough to include practical integrators and appears to capture the correct dynamics, at least in some cases. However, much work remains to be done to clarify the nature of discrete nonholonomic mechanics and to pinpoint the "correct" discrete analog of the Lagrange-d'Alembert principle. We therefore close with some open questions.

The discrete systems depend on two pieces of constraint data, namely C_d and A_d. Can or should these be reduced to one, presumably C_d alone?

Should the definition of the constraint one-forms A_d be generalized to allow dependence on all of q_{i-1}, q_i, and q_{i+1} (instead of just q_i), i.e., $A_d\colon\ Q \times Q \times Q \to T^*Q$? Such A_d can still lead to reversible integrators on $Q \times Q$.

Nonholonomic systems can preserve a volume form on the constraint surface C [17]. Can DLA integrators preserve a corresponding volume form on C_d?

There are many unconstrained Hamiltonian systems with a well-understood and wide range of long-time behavior which can serve as tests for integrators. What are the equivalent nonholonomic systems?

Does the exact flow of a nonholonomic system obey the DLA equations for an appropriate choice of L_d, C_d, and A_d?

What extra property is needed to single out the q_2 of the exact flow amongst those q_2 satisfying (7.16)?

Given the discrete system (Q, L_d, A_d, C_d), does there exist in the sense of backward error analysis a Lagrangian, L^*, such that the discrete flow F_{L_d} is, up to appropriate order in the time step, the discrete time-evaluated flow of the nonholonomic system (Q, L^*, C) where C is the nonintegrable distribution determined by A_d?

Acknowledgment

We are grateful to the Marsden Fund of the Royal Society of New Zealand for financial support. We would also like to thank the referees for their useful comments.

References

[1] L. Bates and J. Śniatycki, Nonholonomic reduction, *Rep. Math. Phys.* **32** (1993), no. 1, 99–115.

[2] A. M. Bloch, *Nonholonomic Mechanics and Control*, Springer-Verlag, New York, 2003.

[3] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, R. Murray, Nonholonomic mechanical systems with symmetry, *Arch. Ratl. Mech. Anal.* **136** (1996), no. 1, 21–99.

[4] A. M. Bloch, J. E. Marsden, and D. V. Zenkov, Nonholonomic dynamics, *Notices Amer. Math. Soc.* **52** (2005), 324–333.

[5] A. I. Bobenko and Y. B. Suris, Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top, *Commun. Math. Phys.* **204** (1998), 147–188.

[6] A. I. Bobenko and Yu. B. Suris, Discrete Lagrangian reduction, discrete Euler-Poincaré equations, and semidirect products, *Lett. Math. Phys.* **49** (1999), 79–93.

[7] T. Bridges and S. Reich, Numerical methods for Hamiltonian PDEs, *J. Phys. A*, to appear.

[8] F. Cantrijn, J. Cortés, M. de León, and D. Martin de Diego, On the geometry of generalized Chaplygin systems, *Math. Proc. Cambridge Philos. Soc.* **132** (2002), no. 2, 323–351.

[9] H. Cendra, J. E. Marsden, and T. S. Ratiu, Geometric mechanics, Lagrangian reduction, and nonholonomic systems. *Mathematics Unlimited—2001 and Beyond*, 221–273, Springer, Berlin, 2001.

[10] Chaplygin, S. A. On a ball's rolling on a horizontal plane, *Regul. Chaotic Dynam.* **7** (2002), no. 2, 131–148, translation of *Mat. Sb.* **24** (1903), 139–168.

[11] J. Cortés, *Geometric, Control, and Numerical Aspects of Nonholonomic Systems*, Lecture Notes in Mathematics vol. 1793, Springer-Verlag, New York, 2002.

[12] J. Cortés, Energy-conserving nonholonomic integrators, *Disc. Contin. Dynam. Syst.*, suppl., (2003), 189–199.

[13] J. Cortés and M. de León, Reduction and reconstruction of the dynamics of nonholonomic systems, *J. Phys. A* **32** (1999), 8615–8645.

[14] J. Cortés and S. Martinez, Non-holonomic integrators, *Nonlinearity* **14** (2001), no. 5, 1365–1392.

[15] K. Ehlers, J. Koiller, R. Montgomery, and Pedro M. Rios, Nonholonomic systems via moving frames: Cartan equivalence and Chaplygin Hamiltonization, in *The Breadth of Symplectic and Poisson Geometry*, pp. 75–120, Progr. Math., 232, Birkhäuser, Boston, 2005.

[16] J. Koiller, Reduction of some classical nonholonomic systems with symmetry, *Arch. Ratl. Mech. Anal.* **118** (1992), no. 2, 113–148.

[17] I. Kupka and W. Oliva, The non-holonomic mechanics, Special issue in celebration of Jack Hale's 70th birthday, *J. Diff. Eq.* **169** (2001), no. 1, 169–189.

[18] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, Cambridge University Press, Cambridge, 2004.

[19] M. de León, D. Martin de Diego, and A. Santamaria-Merino, Geometric integrators and nonholonomic mechanics, *J. Math. Phys.* **45** (2004), 1042–1062.

[20] C.-M. Marle, Reduction of constrained mechanical systems and stability of relative equilibria, *Commun. Math. Phys.* **174** (1995), no. 2, 295–318.

[21] C.-M. Marle, Various approaches to conservative and nonconservative nonholonomic systems, *Pacific Institute of Mathematical Sciences Workshop on Nonholonomic Constraints in Dynamics (Calgary, AB, 1997)*, *Rep. Math. Phys.* **42** (1998), no. 1–2, 211–229.

[22] C.-M. Marle, On symmetries and constants of motion in Hamiltonian systems with nonholonomic constraints. Classical and quantum integrability (Warsaw, 2001), 223–242, Banach Center Publ., 59, Polish Acad. Sci., Warsaw, 2003.

[23] J. E. Marsden, S. Pekarsky, and S. Shkoller, Discrete Euler-Poincaré and Lie-Poisson Equations, *Nonlinearity* **12** (1999), no. 6, 1647–1662.

[24] J. E. Marsden and M. West, Discrete mechanics and variational integrators, *Acta Numerica* **10** (2001), 1–159.

[25] R. I. McLachlan and M. Perlmutter, Energy drift in reversible time integration, *J. Phys. A* **37**, no. 45 (2004), L593–L598.

[26] R. I. McLachlan and G. R. W. Quispel, Geometric integrators for ODEs, *J. Phys. A*, to appear.

[27] J. Moser and A. Veselov, Discrete versions of some classical integrable systems and factorization of matrix polynomials, *Commun. Math. Phys.* **139** (1991), no. 2, 217–243.
[28] A. Pandolfi, C. Kane, J. E. Marsden, and M. Ortiz, Time-discretized variational formulation of nonsmooth frictional contact, *Int. J. Num. Meth. Eng.* **53** (2002), 1801–1829.
[29] A. M. Vershik and V. Ya. Gershkovich, Nonholonomic dynamical systems, geometry of distributions and variational problems, in *Dynamical Systems VII: Integrable Systems, Nonholonomic Dynamical Systems*, pp. 1–81, Encyclopaedia of Mathematical Sciences 16, Springer-Verlag, Berlin, 1994.
[30] A. Weinstein, Lagrangian mechanics and groupoids, Mechanics day (Waterloo, ON, 1992), *Fields Inst. Commun.* **7** (1996), 207–231.