GEOMETRIC INTEGRATORS FOR HIGHER-ORDER VARIATIONAL SYSTEMS AND THEIR APPLICATION TO OPTIMAL CONTROL

LEONARDO COLOMBO, SEBASTIÁN FERRARO, AND DAVID MARTÍN DE DIEGO

Abstract. Numerical methods that preserve geometric invariants of the system, such as energy, momentum or the symplectic form, are called geometric integrators. In this paper we present a method to construct symplectic-momentum integrators for higher-order Lagrangian systems. Given a regular higher-order Lagrangian $L\colon T^{(k)}Q \to \mathbb{R}$ with $k \geq 1$, the resulting discrete equations define a generally implicit numerical integrator algorithm on $T^{(k-1)}Q \times T^{(k-1)}Q$ that approximates the flow of the higher-order Euler–Lagrange equations for L. The algorithm equations are called higher-order discrete Euler–Lagrange equations and constitute a variational integrator for higher-order mechanical systems. The general idea for those variational integrators is to directly discretize Hamilton's principle rather than the equations of motion in a way that preserves the invariants of the original system, notably the symplectic form and, via a discrete version of Noether's theorem, the momentum map.

We construct an exact discrete Lagrangian L_d^e using the locally unique solution of the higher-order Euler–Lagrange equations for L with boundary conditions. By taking the discrete Lagrangian as an approximation of L_d^e, we obtain variational integrators for higher-order mechanical systems. We apply our techniques to optimal control problems since, given a cost function, the optimal control problem is understood as a second-order variational problem.

Contents

1. Introduction

This paper is concerned with the design of geometric integrators for higher-order variational systems. The study of higher-order variational systems has regularly attracted a lot of attention

Mathematics Subject Classification (2010): 70G45, 70Hxx, 49J15.

Keywords and phrases: variational integrators, higher-order mechanics, optimal control, discrete variational calculus.

from the applied and theoretical points of view (see [14] and references therein). But recently there is a renewed interest in these systems due to new and relevant applications in optimal control for robotics or aeronautics, or the study of air traffic control and computational anatomy ([7, 9, 13, 16, 17, 18, 20, 24, 27]).

A continuous higher-order system is modeled by a Lagrangian on a higher-order tangent bundle $T^{(k)}Q$, that is, a function $L\colon T^{(k)}Q \to \mathbb{R}$. The corresponding Euler–Lagrange equations are a system of implicit $2k$-order differential equations. Of course the explicit integration of most of these Lagrangian systems is too complicated to integrate directly or even it is generically not possible. In these cases, it is necessary to discretize the equations taking approximations at several points in time over the interval of integration.

Among the different numerical integrators that one can derive for continuous higher-order systems, one of the most successful ideas is to discretize first the variational principle (instead of the equations of motion) and to derive the numerical method applying discrete calculus of variations [25, 31, 32]. The advantage of this procedure is that automatically we have preservation of some of the geometric structures involved, like symplectic forms or preservation of momentum, moreover, a good behavior of the associated energy. These methods have their roots in the optimal control literature in the 1960s [21].

In previous approaches (see for example [3, 10, 11]), the theory of discrete variational mechanics for higher-order systems was derived using a discrete Lagrangian $L_d\colon Q^{k+1} \to \mathbb{R}$ where Q^{k+1} is the cartesian product of $k+1$ copies of the configuration manifold Q. There, $k+1$ points are used to approximate the positions and the higher-order velocities (such as the standard velocities, accelerations, jerks...) and to represent in this way elements of the higher-order tangent bundle $T^{(k)}Q$.

We will see in this paper that the most natural approach is to take a discrete Lagrangian $L_d\colon T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$ since actually the discrete variational calculus is not based on the discretization of the Lagrangian itself, but on the discretization of the associated action. We will see that a suitable approximation of the action

$$\int_0^h L(q, \dot{q}, \dots, q^{(k)})\, dt$$

is given by a Lagrangian of the form $L_d\colon T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$. Moreover, we will derive a particular choice of discrete Lagrangian which gives an exact correspondence between discrete and continuous systems, the exact discrete Lagrangian. For instance, if we take the Lagrangian $L(q,\dot{q},\ddot{q}) = \frac{1}{2}\ddot{q}^2$, the corresponding exact discrete Lagrangian $L_d^e\colon TQ \times TQ \to \mathbb{R}$ is

$$\begin{aligned} L_d^e(q_0, v_0, q_h, v_h) &= \int_0^h L(q(t), \dot{q}(t), \ddot{q}(t))\, dt \\ &= \frac{6}{h^3}(q_0 - q_h)^2 + \frac{6}{h^2}(q_0 - q_h)(v_0 + v_h) + \frac{2}{h}(v_0^2 + v_0 v_h + v_h^2) \end{aligned}$$

where $q(t)$ is the unique solution of the Euler–Lagrange equations for L verifying $q(0) = q_0$, $\dot{q}(0) = v_0$, $q(h) = q_h$, $\dot{q}(h) = v_h$ for h small enough (see Section 2).

Observe from the previous example that now this theory of variational integrators for higher-order systems is even simpler, since it fits directly into the standard discrete mechanics theory for a discrete Lagrangian of the form $L_d\colon M \times M \to \mathbb{R}$ where $M = T^{(k-1)}Q$. We will show that if the original Lagrangian is regular then so is the exact discrete Lagrangian, in the sense of [25]. Moreover, in the corresponding applications, for instance in optimal control theory or splines theory, typically we are dealing with initial and final boundary conditions which are not necessary discretized, in contrast to previously proposed methods [5, 22, 23].

The paper is structured as follows. In Section 2, we show that a regular higher-order Lagrangian system has a unique solution for given nearby endpoint conditions using a direct variational proof of existence and uniqueness of the local boundary value problem, which employs a regularization procedure. In Section 3 we introduce the notion of exact discrete Lagrangian for higher-order systems and we design the construction of variational integrators for higher-order Lagrangian systems taking approximations of the exact discrete Lagrangian. We obtain the discrete Euler–Lagrange equations for a discrete Lagrangian defined in the cartesian product of two copies of $T^{(k-1)}Q$. Section 4 is devoted to the study of the relation between the discrete and continuous dynamics. We show the relation between the discrete Legendre transformations and the continuous one and we also show that the exact discrete Lagrangian associated with a higher-order regular Lagrangian is also regular. Finally, in Section 5, we apply our techniques to study optimal control problems for fully actuated mechanical systems.

2. Existence and uniqueness of solutions for the boundary value problem

2.1. Higher-order tangent bundles. First we recall some basic facts about the higher-order tangent bundle theory. For more details see [12] and [14].

Let Q be a differentiable manifold. We introduce the following equivalence relation in the set $C^k(I,Q)$ of k-differentiable curves from the interval $I \subseteq \mathbb{R}$ to Q, where $0 \in I$. By definition, two curves γ_1 and γ_2 belonging to $C^k(I,Q)$ have contact of order k at $q_0 = \gamma_1(0) = \gamma_2(0)$ if there is a local chart (φ, U) of Q such that $q_0 \in U$ and

$$\frac{d^s}{dt^s}\left(\varphi \circ \gamma_1(t)\right)\Big|_{t=0} = \frac{d^s}{dt^s}\left(\varphi \circ \gamma_2(t)\right)\Big|_{t=0},$$

for all $s = 0, \ldots, k$. The equivalence class of a curve γ will be denoted by $[\gamma]_0^{(k)}$. The set of equivalence classes will be denoted by $T^{(k)}Q$ and it is not hard to show that it has a natural structure of differentiable manifold. Moreover, $\tau_Q^k \colon T^{(k)}Q \to Q$ where $\tau_Q^k\big([\gamma]_0^{(k)}\big) = \gamma(0)$ is a fiber bundle called the *tangent bundle of order* k of Q. Clearly, $T^{(1)}Q = TQ$.

From a local chart $q^{(0)} = (q^i)$ on a neighborhood U of Q with $i = 1, \ldots, n = \dim Q$, it is possible to induce local coordinates $(q^{(0)}, q^{(1)}, \ldots, q^{(k)})$ on $T^{(k)}U = (\tau_Q^k)^{-1}(U) \equiv U \times (\mathbb{R}^n)^k$. Sometimes we will resort to the usual notation $q^{(0)} \equiv (q^i)$, $q^{(1)} \equiv (\dot{q}^i)$ and $q^{(2)} \equiv (\ddot{q}^i)$.

There is a canonical embedding $j_k \colon T^{(k)}Q \to TT^{(k-1)}Q$ defined as $j_k([\gamma]_0^{(k)}) = [\gamma^{(k-1)}]_0^{(1)}$, where $\gamma^{(k-1)}$ is the lift of the curve γ to $T^{(k-1)}Q$; that is, the curve $\gamma^{(k-1)} \colon I \to T^{(k-1)}Q$ is given by $\gamma^{(k-1)}(t) = [\gamma_t]_0^{(k-1)}$ where $\gamma_t(s) = \gamma(t+s)$. In local coordinates,

$$j_k(q^{(0)}, q^{(1)}, q^{(2)}, \ldots q^{(k)}) = (q^{(0)}, q^{(1)}, \ldots, q^{(k-1)}; q^{(1)}, q^{(2)}, \ldots, q^{(k)}) .$$

2.2. Hamilton's principle and considerations about the existence and uniqueness of solutions. Let $L \colon T^{(k)}Q \to \mathbb{R}$ be a Lagrangian of order $k \geq 1$, of class C^{k+1}. Since our result will be local, we assume from now on that Q is an open subset of $\mathbb{R}^n$. Take coordinates $(q^{(0)}, q^{(1)}, \ldots, q^{(k)})$ on $T^{(k)}Q \equiv Q \times (\mathbb{R}^n)^k$ as before. We suppose that L is regular in the sense that the Hessian matrix

$$\left(\frac{\partial^2 L}{\partial q^{(k)i}\partial q^{(k)j}}\right)$$

is a regular matrix. Let also $h > 0$ be given. We can formulate Hamilton's principle as follows.

Variational Principle 1. *Find a C^k curve $q \colon [0,h] \to Q$ such that it is a critical point of the action*

$$S_h = \int_0^h L\left(q(t), \dot{q}(t), \ldots, q^{(k)}(t)\right) dt$$

among those curves whose first $k-1$ derivatives are fixed at the endpoints, that is, with given values for $q(0), \dot{q}(0), \ldots, q^{(k-1)}(0)$ and $q(h), \dot{q}(h), \ldots, q^{(k-1)}(h)$. △

Hamilton's principle is a constrained problem in the Banach space $C^k([0,h], \mathbb{R}^n)$. Now if $q(t)$ is a solution to this problem that is not only C^k but C^{2k}, then it satisfies the well-known k^{th}-order Euler–Lagrange equations[1]

$$\sum_{j=0}^{k} (-1)^j \frac{d^j}{dt^j}\frac{\partial L}{\partial q^{(j)}} = 0. \tag{1}$$

For a regular Lagrangian, (1) can be written as an explicit $2k$-order ordinary differential equation. Existence and uniqueness of solutions for the *initial* value problem can be guaranteed using basic

[1]For $k = 1$, recall writing $\delta\dot{q} = (\dot{\delta q})$ when deriving the Euler–Lagrange equations, assuming that q is C^2.

ODE theory. Doing the same for for the boundary value problem of finding a solution $q(t)$ of (1) with given values for $q(0), \dot{q}(0), \ldots, q^{(k-1)}(0)$ and $q(h), \dot{q}(h), \ldots, q^{(k-1)}(h)$ requires different techniques. For instance, in [2, ch. 9] it is shown that there exists a unique solution to an explicit $2k$-order ODE with this kind of boundary conditions, for small enough h and close enough boundary values. See also [15, Appendix A] for results on the existence, uniqueness and smooth dependence on parameters of solutions of ODEs.

In principle, however, there could exist solutions to Hamilton's variational principle that are C^k but not C^{2k}, and thus do not satisfy (1). Therefore, uniqueness of solutions to the variational principle cannot yet be guaranteed. One possibility for avoiding this situation is stating Hamilton's principle in the (smaller) C^{2k} context from the beginning. In this section we proceed differently, acknowledging the fact the variational principle makes sense in the C^k setting. We prove local existence and uniqueness of C^k solutions to Hamilton's principle from a direct variational point of view. We will see that these solutions turn out to be automatically C^{2k}, so they satisfy Euler–Lagrange equations *a posteriori*.

Our argument for the existence and uniqueness of solutions will involve a regularization procedure which follows closely the proof by Patrick [29] for first-order Lagrangians; the formulas, of course, reduce to those in [29] for order 1, but we introduce an additional modification using orthonormal polynomials. See also [8, 19] for discussions on the regularity of extremals for variational problems.

2.3. **Non-regularity of Hamilton's principle.** We want to determine whether there exists a unique solution curve to Hamilton's principle, given endpoint conditions that are close enough. The main obstacle for a straightforward affirmative answer is that the local boundary value problem as stated above is nonregular at $h = 0$. That is, the constraint function $g\colon C^k([0,h],Q) \to (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k$

$$g\colon q(\cdot) \mapsto \left(q(0), \dot{q}(0), \ldots, q^{(k-1)}(0); q(h), \dot{q}(h), \ldots, q^{(k-1)}(h)\right)$$

maps into the diagonal of $T^{(k-1)}Q \times T^{(k-1)}Q$ for $h = 0$ and is not therefore a submersion. For $h \neq 0$, the constraint function is a submersion.

The approach consists in replacing this problem by an equivalent one that is regular at $h = 0$, and show that locally there is a unique solution to the regularized problem.

2.4. **Regularization.** First we replace the space of curves on Q in the variational problem by the space of curves on $T^{(k)}Q$, and include additional constraints. Denote an arbitrary curve by

$$\left(q(t) = q^{[0]}(t), q^{[1]}(t), \ldots, q^{[k]}(t)\right) \in T^{(k)}Q \equiv Q \times (\mathbb{R}^n)^k,$$

$t \in [0,h]$. Here we have modified our notation for coordinates on $T^{(k)}Q$, using superscripts in square brackets to make a distinction with the actual derivatives of $q(t)$.

Variational Principle 2. *Find a curve $(q^{[0]}(t), q^{[1]}(t), \ldots, q^{[k]}(t))$ on $T^{(k)}Q$, with $q^{[l]} \in C^{k-l}([0,h], \mathbb{R}^n)$, $l = 0, \ldots, k$, such that it is a critical point of*

$$S_h = \int_0^h L\left(q^{[0]}(t), q^{[1]}(t), \ldots, q^{[k]}(t)\right) dt$$

subject to the constraints

$$q^{[j+1]}(t) = \frac{dq^{[j]}}{dt}(t), \quad q^{[j]}(0) = q_1^{[j]}, \quad q^{[j]}(h) = q_2^{[j]}, \quad j = 0, \ldots, k-1,$$

where $(q_i^{[0]}, q_i^{[1]}, \ldots, q_i^{[k-1]})$, $i = 1, 2$, are given points in $T^{(k-1)}Q$. △

Now reparameterize the curve by defining

$$Q^{[j]}(u) = q^{[j]}(hu), \quad j = 0, \ldots, k, \quad u \in [0,1].$$

For $h > 0$, the curve $(Q^{[0]}(u), \dots, Q^{[k]}(u))$ satisfies an equivalent variational problem as follows. Since h is a constant for each instance of the problem, we can use

$$\frac{1}{h} \int_0^h L\left(q^{[0]}(t), q^{[1]}(t), \dots, q^{[k]}(t)\right) dt = \int_0^1 L\left(Q^{[0]}(u), \dots, Q^{[k]}(u)\right) du$$

as an objective function. The first set of constraints becomes

$$0 = \frac{dq^{[j]}}{dt}(t) - q^{[j+1]}(t) = \left(\frac{1}{h}\frac{dQ^{[j]}}{du}(u) - Q^{[j+1]}(u)\right)_{u=t/h}$$

where $j = 0, \dots, k-1$.

The reparametrized variational principle is the following.

Variational Principle 3. *Find a curve $(Q^{[0]}(u), \dots, Q^{[k]}(u))$ on $T^{(k)}Q$, $Q^{[l]} \in C^{k-l}([0,1], \mathbb{R}^n)$, $l = 0, \dots, k$, that is a critical point of*

$$S = \int_0^1 L\left(Q^{[0]}(u), \dots, Q^{[k]}(u)\right) du,$$

subject to the constraints

$$\frac{dQ^{[j]}}{du}(u) = hQ^{[j+1]}(u), \tag{2}$$

$$Q^{[j]}(0) = q_1^{[j]}, \tag{3}$$

$$Q^{[j]}(1) = q_2^{[j]}, \tag{4}$$

where $j = 0, \dots, k-1$, and $(q_i^{[0]}, q_i^{[1]}, \dots, q_i^{[k-1]})$, $i = 1, 2$, are given points in $T^{(k-1)}Q$. △

The objective S does not depend on h, and the constraints are smooth through $h = 0$.

Remark 2.1. For $h = 0$, the constraints (2) imply that $Q^{[0]}(u)$, ..., $Q^{[k-1]}(u)$ remain constant, which restricts the possible values of the endpoint conditions in order to have a compatible set of constraints. More precisely, $q_1^{[j]} = q_2^{[j]}$ for $j = 0, \dots, k-1$; otherwise there would be no curves satisfying the constraints. This kind of restriction also appears in the original variational principle 1. Moreover, the problem becomes the unconstrained problem of finding a curve $Q^{[k]}(u) \in C^0([0,1], \mathbb{R}^n)$ such that it is a critical point of

$$\int_0^1 L\left(q^{[0]}, \dots, q^{[k-1]}, Q^{[k]}(u)\right) du.$$

This means

$$\frac{\partial L}{\partial q^{[k]}}\left(q^{[0]}, q^{[1]}, \dots, q^{[k-1]}, Q^{[k]}(u)\right) = 0.$$

Differentiating with respect to u, and using the fact that the Lagrangian is regular, we obtain that $Q^{[k]}(u)$ is constant. ◊

In preparation for the next step for regularization, let us solve the constraints (2) to get

$$Q^{[j]}(u) = Q^{[j]}(0) + h \int_0^u Q^{[j+1]}(s)\, ds, \quad j = 0, \dots, k-1.$$

This means that the functions $Q^{[j]}(u)$, $j = 0, \dots, k-1$, can be expressed in terms of $Q^{[j]}(0)$, ..., $Q^{[k-1]}(0)$, the function $Q^{[k]}(u)$ and h. For example, for $k = 2$ we have

$$\begin{aligned}
Q^{[1]}(u) &= Q^{[1]}(0) + h \int_0^u Q^{[2]}(s)\, ds, \\
Q^{[0]}(u) &= Q^{[0]}(0) + h \int_0^u Q^{[1]}(s)\, ds \\
&= Q^{[0]}(0) + huQ^{[1]}(0) + h^2 \int_0^u \int_0^s Q^{[2]}(\tau)\, d\tau\, ds \\
&= Q^{[0]}(0) + huQ^{[1]}(0) + h^2 \int_0^u (u - \tau) Q^{[2]}(\tau)\, d\tau.
\end{aligned}$$

For a general k, and for $j = 0, \dots, k-1$, an iterated change of order of integration yields

$$Q^{[j]}(u) = Q^{[j]}(0) + \sum_{i=1}^{k-j-1} \frac{h^i u^i}{i!} Q^{[j+i]}(0) + h^{k-j} \int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds. \tag{5}$$

If the upper bound of summation is less than the lower bound, the sum is understood to be 0.

Note that taking $u = 1$, the final endpoint data $(q_2^{[0]}, \dots, q_2^{[k-1]})$ can now be written as

$$q_2^{[j]} = Q^{[j]}(1) = q_1^{[j]} + \sum_{i=1}^{k-j-1} \frac{h^i}{i!} q_1^{[j+i]} + h^{k-j} \int_0^1 \frac{(1-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds, \tag{6}$$

so we define

$$z^{[j]} = \int_0^1 \frac{(1-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds = \frac{1}{h^{k-j}} \left(q_2^{[j]} - \sum_{i=0}^{k-j-1} \frac{h^i}{i!} q_1^{[j+i]} \right). \tag{7}$$

We will discuss the case $h = 0$ in Remark 2.2.

Now replace the curves and endpoint data by just $Q^{[k]}(u)$, $(q_1^{[0]}, \dots, q_1^{[k-1]})$, and $(z^{[0]}, \dots, z^{[k-1]})$, to get a new variational principle.

Variational Principle 4. *Given h, $(q_1^{[0]}, \dots, q_1^{[k-1]})$ and $(z^{[0]}, \dots, z^{[k-1]})$, find a continuous curve $Q^{[k]} \colon [0,1] \to \mathbb{R}^n$ that is a critical point of*

$$\mathcal{S} = \int_0^1 L\left(Q^{[0]}(u), \dots, Q^{[k]}(u)\right) du,$$

where $Q^{[0]}(u)$, ..., $Q^{[k-1]}(u)$ are defined as in (5) *by*

$$Q^{[j]}(u) = q_1^{[j]} + \sum_{i=1}^{k-j-1} \frac{h^i u^i}{i!} q_1^{[j+i]} + h^{k-j} \int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds, \quad j = 0, \dots, k-1$$

subject to the constraints

$$\int_0^1 \frac{(1-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds = z^{[j]}, \quad j = 0, \dots, k-1. \qquad \triangle$$

Observe that the constraint functions do not depend on h and are linear on the curve $Q^{[k]}$. This variational principle is already regular through $h = 0$, as we will see when we proceed to find the solutions later.

Remark 2.2. The data $q_1^{[0]}$, ..., $q_1^{[k-1]}$, $z^{[0]}$, ..., $z^{[k-1]}$ can be transformed into the endpoint conditions for the variational principle 3 in a straightforward way, for any h, using (6) and (7). The converse (7) is possible only for $h \neq 0$, in principle. However, if $h = 0$ let $(Q^{[0]}(u), \dots, Q^{[k]}(u))$ a solution for the variational principle 3 with boundary conditions $(q_1^{[0]}, \dots, q_1^{[k-1]})$ and $(q_2^{[0]}, \dots, q_2^{[k-1]})$. Define $z^{[j]}$ by the constraint in (4). Since $Q^{[k]}$ is constant and $\frac{(1-s)^{k-j-1}}{(k-j-1)!} > 0$ in $(0,1)$, to different values of $Q^{[k]}$ correspond different values of $z^{[j]}$. Then $Q^{[k]}$ is a solution of 4 with boundary conditions $q_1^{[0]}$, ..., $q_1^{[k-1]}$, $z^{[0]}$, ..., $z^{[k-1]}$. $\diamondsuit$

Finally, we will introduce a modification that will enable us to carry out the computations in the next section easily. Consider the inner product on $C^0([0,1], \mathbb{R})$ given by

$$\langle f, g \rangle = \int_0^1 f(s) g(s)\, ds.$$

If $f \in C^0([0,1], \mathbb{R})$ and $V = (V_1, \dots, V_n) \in C^0([0,1], \mathbb{R}^n)$ we define the bilinear operation

$$\langle\!\langle f, V \rangle\!\rangle = \int_0^1 f(s) V(s)\, ds = (\langle f, V_0 \rangle, \dots, \langle f, V_n \rangle) \in \mathbb{R}^n.$$

Then the integrals appearing in the constraints in the variational principle 4 are $\langle\!\langle a_j^{[k]}, Q^{[k]} \rangle\!\rangle$, where $a_j^{[k]}$ are the polynomials

$$a_j^{[k]}(s) = \frac{(1-s)^{k-j-1}}{(k-j-1)!}, \quad j = 0, \dots, k-1.$$

These form a basis of the space of polynomials of degree at most $k-1$. Let us consider a basis $b_j^{[k]}(s)$, $j = 0, \dots, k-1$, of the same space of polynomials consisting of orthonormal polynomials on $[0,1]$, and let $(\gamma_j^{[k],i})$, where $i, j = 0, \dots, k-1$, be the invertible real matrix such that $a_j^{[k]}(s) = \gamma_j^{[k],i} b_i^{[k]}(s)$. For example, for $k = 2$,

$$a_0^{[2]}(s) = 1 - s, \quad a_1^{[2]}(s) = 1,$$

and we can take for instance the orthonormal basis

$$b_0^{[2]}(s) = \sqrt{3}(1-2s), \quad b_1^{[2]}(s) = 1;$$

therefore,

$$\gamma_0^{[2],0} = \tfrac{1}{2\sqrt{3}}, \quad \gamma_0^{[2],1} = \tfrac{1}{2}, \quad \gamma_1^{[2],0} = 0, \quad \gamma_1^{[2],1} = 1.$$

Using this matrix, the constraints can be rewritten as

$$z^{[j]} = \langle\!\langle a_j^{[k]}, Q^{[k]} \rangle\!\rangle = \gamma_j^{[k],i} \langle\!\langle b_i^{[k]}(s), Q^{[k]} \rangle\!\rangle,$$

for $j = 0, \dots, k-1$. This allows us to reformulate the variational principle in an equivalent way by replacing the data $(z^{[0]}, \dots, z^{[k-1]})$ and constraints $\langle\!\langle a_j^{[k]}, Q^{[k]} \rangle\!\rangle = z^{[j]}$ by new data $(w^{[0]}, \dots, w^{[k-1]})$ and constraints $\langle\!\langle b_j^{[k]}, Q^{[k]} \rangle\!\rangle = w^{[j]}$, $j = 0, \dots, k-1$. The old and new data are related by

$$\sum_{i=0}^{k-1} \gamma_j^{[k],i} w^{[i]} = z^{[j]}. \tag{8}$$

Variational Principle 5. *Given h, $(q_1^{[0]}, \dots, q_1^{[k-1]})$ and $(w^{[0]}, \dots, w^{[k-1]})$, find a continuous curve $Q^{[k]} \colon [0,1] \to \mathbb{R}^n$ that is a critical point of*

$$S_h = \int_0^1 L\left(Q^{[0]}(u), \dots, Q^{[k]}(u)\right) du,$$

where $Q^{[0]}(u)$, …, $Q^{[k-1]}(u)$ are defined by

$$Q^{[j]}(u) = q_1^{[j]} + \sum_{i=1}^{k-j-1} \frac{h^i u^i}{i!} q_1^{[j+i]} + h^{k-j} \int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds, \tag{9}$$

subject to the constraints

$$\int_0^1 b_j^{[k]}(s) Q^{[k]}(s)\, ds = w^{[j]}, \quad j = 0, \dots, k-1. \qquad \triangle$$

2.5. **Solution of the regularized problem.** Next, we will study the existence and uniqueness of solutions associated with Variational Principle 5. We will show that the boundary value problem is well posed, and that even though the variational problem is posed on the space of C^k solutions, the extremizers are C^{2k} and hence satisfy the Euler-Lagrange equations.

We start the proof by showing the C^{k+1} differentiability of the action S_h for the variational principle 5. Next, we compute the gradient of S_h in order to solve the equation "gradient of S_h perpendicular to constraint space". After introducing an orthogonal decomposition of the constraint space we obtain that S_h has a critical point on the constraint set if and only if the orthogonal projection of the gradient of S_h is 0 and hence we can find the stationary curve for the variational principle 5. Using the implicit function theorem we obtain existence and uniqueness of solutions for the variational principle 5. Finally, we reverse the regularization to obtain a unique C^{2k} solution of the original variational principle.

2.5.1. *Step 1 - C^{k+1} differentiability of S_h:* Let S_h be given as in the variational principle 5, regarded as a real-valued map defined on the Banach space $C^0([0,1],\mathbb{R}^n)$ of curves $Q^{[k]}(u)$. We can also consider its restriction to the Banach space $C^k([0,1],\mathbb{R}^n)$. We are going to use the following lemma [1].

Lemma 2.3 (Omega Lemma)**.** *Let E, F be Banach spaces, U open in E, and M a compact topological space. Let $g\colon U \to F$ be a C^r map, $r > 0$. The map*

$$\Omega_g\colon C^0(M,U) \to C^0(M,F) \quad \textit{defined by} \quad \Omega_g(f) = g \circ f$$

is also C^r, and $D\Omega_g(f)\cdot h = [(Dg)\circ f]\cdot h$.

The objective S_h is the composition of the maps

$$C^0([0,1],\mathbb{R}^n) \xrightarrow{\;i\;} C^0([0,1],T^{(k)}Q) \xrightarrow{\;\Omega_L\;} C^0([0,1],\mathbb{R}) \xrightarrow{\;\int\;} \mathbb{R}$$

where i is defined by $Q^{[k]}(u) \mapsto (Q^{[0]}(u),\dots,Q^{[k]}(u))$. Here $Q^{[0]}(u),\dots,\ Q^{[k-1]}(u)$ stand for the right-hand sides of (9). Both i and $\int$ are bounded affine and therefore C^∞. By the Omega Lemma, Ω_L is C^{k+1} because L is C^{k+1}, and therefore so is S_h.

If we regard S_h as defined on $C^k([0,1],\mathbb{R}^n)$, we should append the inclusion $C^k([0,1],\mathbb{R}^n) \hookrightarrow C^0([0,1],\mathbb{R}^n)$ to the left side of the diagram above. This inclusion is C^∞ because it is linear and bounded ($\|Q^{[k]}\|_{C^0} \le \|Q^{[k]}\|_{C^k}$ for all $Q^{[k]}$). Then S_h is C^{k+1} also as a map defined on $C^k([0,1],\mathbb{R}^n)$. In order to cover both cases, from now on l will denote 0 or k interchangeably.

2.5.2. *Step 2 - Computing the gradient of S_h:* We need a suitable notion of the gradient of S_h, in order to find where it is perpendicular to the constraint space. In order to do that, let us first compute $\mathbf{d}S_h[Q^{[k]}(u)]$, for $Q^{[k]}$ of class C^l. The functions $Q^{[0]}(u),\ \dots,\ Q^{[k-1]}(u)$ are defined by (9). Since S_h is smooth, we will compute $\mathbf{d}S_h$ using directional derivatives. For an arbitrary $\delta Q^{[k]}$ of class C^l, take a deformation $Q^{[k]}_\epsilon(u) = Q^{[k]}(u) + \epsilon\delta Q^{[k]}(u)$ of $Q^{[k]}(u)$. For $j = 0,\dots,k-1$, define the corresponding lower order curves as in (9) by

$$Q^{[j]}_\epsilon(u) = q^{[j]}_1 + \sum_{i=1}^{k-j-1}\frac{h^i u^i}{i!}q^{[j+i]}_1 + h^{k-j}\int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!}Q^{[k]}_\epsilon(s)\,ds, \tag{10}$$

so $Q^{[j]}_0(u) = Q^{[j]}(u)$ and

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} Q^{[j]}_\epsilon(u) = h^{k-j}\int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!}\delta Q^{[k]}(s)\,ds.$$

Denoting $a^{[k]}_j(u,s) = (u-s)^{k-j-1}/(k-j-1)!$ and $Q(u) = (Q^{[0]}(u),\dots,\ Q^{[k]}(u))$ for short, we have

$$\begin{aligned}
\mathbf{d}S_h[Q^{[k]}(u)]\cdot\delta Q^{[k]}(u) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\int_0^1 L\left(Q^{[0]}_\epsilon(u),\dots,Q^{[k]}_\epsilon(u)\right)du\\
&= \int_0^1\left(\sum_{j=0}^{k-1}\frac{\partial L}{\partial q^{[j]}}(Q(u))h^{k-j}\int_0^u a^{[k]}_j(u,s)\delta Q^{[k]}(s)\,ds + \frac{\partial L}{\partial q^{[k]}}(Q(u))\delta Q^{[k]}(u)\right)du\\
&= \sum_{j=0}^{k-1}\int_0^1\int_s^1 \frac{\partial L}{\partial q^{[j]}}(Q(u))h^{k-j}a^{[k]}_j(u,s)\delta Q^{[k]}(s)\,du\,ds + \int_0^1\frac{\partial L}{\partial q^{[k]}}(Q(u))\delta Q^{[k]}(u)\,du\\
&= \sum_{j=0}^{k-1}\int_0^1\int_u^1 \frac{\partial L}{\partial q^{[j]}}(Q(s))h^{k-j}a^{[k]}_j(s,u)\delta Q^{[k]}(u)\,ds\,du + \int_0^1\frac{\partial L}{\partial q^{[k]}}(Q(u))\delta Q^{[k]}(u)\,du\\
&= \int_0^1\left(\sum_{j=0}^{k-1}\int_u^1\frac{\partial L}{\partial q^{[j]}}(Q(s))h^{k-j}a^{[k]}_j(s,u)\,ds + \frac{\partial L}{\partial q^{[k]}}(Q(u))\right)\delta Q^{[k]}(u)\,du.
\end{aligned}$$

For each $u \in [0,1]$, the first factor in the integrand of the last expression is in $(\mathbb{R}^n)^*$. If $\sharp\colon (\mathbb{R}^n)^* \to \mathbb{R}^n$ denotes the index raising operator associated to the Euclidean inner product, define

$$\nabla S_h[Q^{[k]}(u)](u) := \left(\sum_{j=0}^{k-1} \int_u^1 \frac{\partial L}{\partial q^{[j]}}(Q(s)) h^{k-j} a_j^{[k]}(s,u)\, ds + \frac{\partial L}{\partial q^{[k]}}(Q(u)) \right)^\sharp .$$

Since $\partial L/\partial q^{[0]}, \ldots, \partial L/\partial q^{[k]}$ are C^k and the curve Q is C^l ($l = 0$ or $l = k$) , then $\nabla S_h[Q^{[k]}(u)]$ is $C^l([0,1], \mathbb{R}^n)$. Then we have a vector field

$$\nabla S_h \colon C^l([0,1], \mathbb{R}^n) \to C^l([0,1], \mathbb{R}^n)$$

which we call the gradient of S_h. By the Omega Lemma, ∇S_h is a C^k map.

2.5.3. *Step 3 - Orthogonal decomposition of the constraint space and critical points of S_h:* Let us now compute the tangent space to the constraint set. If we consider the inner product on $C^l([0,1], \mathbb{R}^n)$ given by

$$[\![V, W]\!] = \int_0^1 V(u) \cdot W(u)\, du,$$

then

$$\mathbf{d} S_h[Q^{[k]}(u)] \cdot \delta Q^{[k]}(u) = [\![\nabla S_h[Q^{[k]}(u)], \delta Q^{[k]}(u)]\!].$$

The constraints $g_j[Q^{[k]}(s)] := \langle\!\langle b_j^{[k]}, Q^{[k]} \rangle\!\rangle = w^{[j]}$, $j = 0, \ldots, k-1$, in the variational principle 5 are bounded and linear, and therefore C^∞, and the corresponding derivatives are the same functions g_j. Define

$$g = (g_0, \ldots, g_{k-1}) \colon C^l([0,1], \mathbb{R}^n) \to (\mathbb{R}^n)^k$$

so

$$E = \operatorname{Ker} g \subset C^l([0,1], \mathbb{R}^n)$$

is the tangent space to the constraint set. They are actually parallel since the constraints are linear. It is not difficult to show using the definitions that the space

$$E^\perp = \{ c^j b_j^{[k]} \mid c^0, \ldots, c^{k-1} \in \mathbb{R}^n \}$$

of $\mathbb{R}^n$-valued polynomials of degree at most $k-1$ is indeed the $[\![,]\!]$-orthogonal complement of E, which is then a split subspace (see the Appendix for a proof). The orthogonal projection $P \colon C^l([0,1], \mathbb{R}^n) = E \oplus E^\perp \to E$ is given by

$$P(\delta Q^{[k]}(u)) = \delta Q^{[k]}(u) - \sum_{j=0}^{k-1} \langle\!\langle b_j^{[k]}, \delta Q^{[k]} \rangle\!\rangle\, b_j^{[k]}.$$

Now S_h has a critical point on the constraint set (for any value of the constraints) if and only if the projection $P \nabla S_h$ of ∇S_h to the tangent space E of the constraint set is 0.

2.5.4. *Step 4 - Existence and uniqueness for the regularized problem:* In order to find solutions to the variational principle 5, we solve

$$P \nabla S_h(Q^{[k]}) = P \nabla S_h(Q_E^{[k]} \oplus Q_{E^\perp}^{[k]}) = 0$$

for $Q_E^{[k]}$, near

$$Q^{[k]} = 0, \quad w^{[0]} = \cdots = w^{[k-1]} = 0,$$
$$q_1^{[0]} = \bar{q}^{[0]}, \ldots, q_1^{[k-1]} = \bar{q}^{[k-1]}, \quad h = 0.$$

This can be solved using the implicit function theorem by requiring that the partial derivative of $P \nabla S_h(Q^{[k]})$ at the point $Q^{[k]} = 0$ with respect to the space E is a linear isomorphism. The variables $w^{[0]}, \ldots, w^{[k-1]}$, $q_1^{[0]}, \ldots, q_1^{[k-1]}$ and h are seen as parameters that can move in some neighborhood. Note that it is not necessary to solve for $Q_{E^\perp}^{[k]}$ since it is completely determined by $w^{[0]}, \ldots, w^{[k-1]}$ using the constraint equations in variational principle 5.

In order to compute this partial derivative, take a deformation of $Q^{[k]}=0$ of the form $Q^{[k]}_\epsilon = \epsilon\delta Q^{[k]}_E$, where $\delta Q^{[k]}_E \in E$. Recalling (10) and noting that $h=0$, we have

$$\begin{aligned}\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} P\frac{\partial L}{\partial q^{[k]}}(Q^{[0]}_\epsilon(u),\dots,Q^{[k]}_\epsilon(u)) &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} P\frac{\partial L}{\partial q^{[k]}}(\bar q^{[0]},\dots,\bar q^{[k-1]},Q^{[k]}_\epsilon(u))\\ &= P\frac{\partial^2 L}{\partial q^{[k]2}}(\bar q^{[0]},\dots,\bar q^{[k-1]},0)\delta Q^{[k]}_E(u) = \frac{\partial^2 L}{\partial q^{[k]2}}(\bar q^{[0]},\dots,\bar q^{[k-1]},0)\delta Q^{[k]}_E(u)\\ &\quad -\sum_{j=0}^{k-1}\left\langle\!\!\left\langle b^{[k]}_j, \frac{\partial^2 L}{\partial q^{[k]2}}(\bar q^{[0]},\dots,\bar q^{[k-1]},0)\delta Q^{[k]}_E\right\rangle\!\!\right\rangle b^{[k]}_j = \frac{\partial^2 L}{\partial q^{[k]2}}(\bar q^{[0]},\dots,\bar q^{[k-1]},0)\delta Q^{[k]}_E(u).\end{aligned}$$

Here the inner products vanish because $\frac{\partial^2 L}{\partial q^{[k]2}}(\bar q^{[0]},\dots,\bar q^{[k-1]},0)$ is a constant matrix (that is, it does not depend on u) and $\langle\!\langle b^{[j]},\delta Q^{[k]}_E\rangle\!\rangle = 0$ for $j=0,\dots,k-1$.

Then the derivative is precisely $\frac{\partial^2 L}{\partial q^{[k]2}}(\bar q^{[0]},\dots,\bar q^{[k-1]},0)$, seen as a linear map from E into itself, and if L is regular then it is an isomorphism.

By the implicit function theorem, there are neighborhoods $W_1 \subseteq (\mathbb{R}^n)^k\times(\mathbb{R}^n)^k\times\mathbb{R}$ (with variables $(q_1^{[0]},\dots,q_1^{[k-1]};w^{[0]},\dots,w^{[k-1]};h)$) containing $(\bar q^{[0]},\dots,\bar q^{[k-1]};0,\dots,0;0)$ and $W_2^l\subseteq C^l([0,1],\mathbb{R}^n)$ containing the constant curve $Q^{[k]}(u)=0$, and a C^k map $\psi\colon W_1\to W_2^l$ such that for each $(q_1^{[0]},\dots,q_1^{[k-1]};w^{[0]},\dots,w^{[k-1]};h)\in W_1$, the curve

$$Q^{[k]}=\psi(q_1^{[0]},\dots,q_1^{[k-1]};w^{[0]},\dots,w^{[k-1]};h)\in C^l([0,1],\mathbb{R}^n)$$

is the unique critical point in W_2^l of the variational problem 5. Thus, ψ maps initial conditions, constraint values (which encode the final endpoint conditions for the original problem) and h into C^l curves.

Let us now consider the cases $l=0$ and $l=k$ separately. Taking $l=k$, ψ has values in $W_2^k\subseteq C^k([0,1],\mathbb{R}^n)$. Taking $l=0$, ψ has values in $W_2^0\subseteq C^0([0,1],\mathbb{R}^n)$. However, since $C^k([0,1],\mathbb{R}^n)\subset C^0([0,1],\mathbb{R}^n)$, this ψ also provides the unique solution among the C^0 curves in a C^0-open neighborhood of the curve $u\mapsto 0$, say $\{Q^{[k]}(u)\mid \|Q^{[k]}\|_0<\epsilon\}$.

2.5.5. *Step 5 - Reverse of the regularization:* Let us now reverse the regularization in order to obtain a unique C^{2k} solution of the variational principle 1. Let $h\neq 0$. For $(q_1,q_2)=((q_1^{[0]},\dots,q_1^{[k-1]}),(q_2^{[0]},\dots,q_2^{[k-1]}))\in(\mathbb{R}^n)^k\times(\mathbb{R}^n)^k$ the corresponding values of $z^{[0]},\dots,z^{[k-1]}$ are given by (7) and the values of $w^{[0]},\dots,w^{[k-1]}$ can be computed from (8) using the inverse matrix of $\left(\gamma_j^{[k],i}\right)$. This defines a smooth function $(w^{[0]},\dots,w^{[k-1]})=\varpi(q_1,q_2,h)$. Note that the condition that q_1 and q_2 are close translates into the condition that $(w^{[0]},\dots,w^{[k-1]})$ is close to 0.

Let $h>0$ be such that $(\bar q^{[0]},\dots,\bar q^{[k-1]};0,\dots,0;h)\in W_1$. Define

$$\widetilde{W}_1=\{(q_1,q_2)\in(\mathbb{R}^n)^k\times(\mathbb{R}^n)^k\mid (q_1;\varpi(q_1,q_2,h);h)\in W_1\}$$

and for each $(q_1,q_2)=\left((q_1^{[0]},\dots,q_1^{[k-1]}),(q_2^{[0]},\dots,q_2^{[k-1]})\right)\in W_1$ define the curve $Q^{[0]}_{(q_1,q_2)}(u)$ according to (5) as

$$Q^{[0]}_{(q_1,q_2)}(u)=\sum_{i=0}^{k-1}\frac{h^iu^i}{i!}q_1^{[i]}+h^k\int_0^u\frac{(u-s)^{k-1}}{(k-1)!}\psi\left(q_1;\varpi(q_1,q_2,h);h\right)(s)\,ds.$$

Since ψ takes values in the C^k curves, $Q^{[0]}_{(q_1,q_2)}(u)$ is C^{2k} by the reasoning leading to equation (5).

Now reparameterize with $t=hu$ to get a C^{2k} curve

$$q^{[0]}_{(q_1,q_2)}(t)=\sum_{i=0}^{k-1}\frac{t^i}{i!}q_1^{[i]}+\left(\frac{t}{u}\right)^k\int_0^{t/h}\frac{(t/h-s)^{k-1}}{(k-1)!}\psi\left(q_1;\varpi(q_1,q_2,h);h\right)(s)\,ds$$

on Q, defined for $t\in[0,h]$. This curve is the unique solution of the variational principle 1 with endpoint conditions q_1 and q_2.

This solution is C^{2k}, and unique among the curves corresponding to $Q^{[k]}$ continuous with $\|Q^{[k]}\|_0 < \epsilon$. These are the C^k curves $q(t)$ on Q with $\|q^{(k)}\|_0 < \epsilon/h^k$, which are the C^k curves in some C^k neighborhood of the constant curve $t \mapsto \bar{q}^{[0]}$.

3. The exact discrete Lagrangian and discrete equations for second-order systems

Next, we will consider second-order Lagrangian systems, motivated by the study of optimal control problems. Let Q be a configuration manifold and let $L\colon T^{(2)}Q \to \mathbb{R}$ be a regular Lagrangian.

Definition 3.1. *Given a small enough*[2] $h > 0$, *the exact discrete lagrangian* $L_d^e\colon TQ \times TQ \to \mathbb{R}$ *is defined by*

$$L_d^e(q_0, \dot{q}_0, q_1, \dot{q}_1) = \int_0^h L(q(t), \dot{q}(t), \ddot{q}(t))dt,$$

where $q\colon [0,h] \to Q$ *is the unique solution of the Euler–Lagrange equations for the second-order Lagrangian* L,

$$\frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} + \frac{\partial L}{\partial q} = 0,$$

satisfying the boundary conditions $q(0) = q_0, q(h) = q_1, \dot{q}(0) = \dot{q}_0$ *and* $\dot{q}(h) = \dot{q}_1$.

Strictly speaking, the exact discrete Lagrangian is defined not on $TQ \times TQ$ but on a neighborhood of the diagonal. For the sake of simplicity, we will not make this distinction. Our idea is to take a discrete Lagrangian $L_d\colon TQ \times TQ \to \mathbb{R}$ as an approximation of $L_d^e\colon TQ \times TQ \to \mathbb{R}$, to construct variational integrators in the same way as in discrete mechanics (see section 4). In other words, for given $h > 0$ we define $L_d(q_0, v_0, q_1, v_1)$ as an approximation of the action integral along the exact solution curve segment $q(t)$ with boundary conditions $q(0) = q_0$, $\dot{q}(0) = v_0$, $q(h) = q_1$, and $\dot{q}(h) = v_1$. For example, we can use the formula

$$L_d(q_0, v_0, q_1, v_1) = hL\left(\kappa(q_0, v_0, q_1, v_1), \chi(q_0, v_0, q_1, v_1), \zeta(q_0, v_0, q_1, v_1)\right),$$

where κ, χ and ζ are functions of $(q_0, v_0, q_1, v_1) \in TQ \times TQ$ which approximate the configuration $q(t)$, the velocity $\dot{q}(t)$ and the acceleration $\ddot{q}(t)$, respectively, in terms of the initial and final positions and velocities. We can also, for instance, consider suitable linear combinations of discrete Lagrangians of this type, for instance, weighted averages of the type

$$L_d(q_0, v_0, q_1, v_1) = \frac{1}{2}L\left(q_0, v_0, \frac{v_1 - v_0}{h}\right) + \frac{1}{2}L\left(q_1, v_1, \frac{v_1 - v_0}{h}\right),$$

or other combinations.

For completeness, we will derive the discrete equations for the Lagrangian $L_d\colon TQ \times TQ \to \mathbb{R}$, but these results are a direct translation of Marsden and West [25] to our case.

Given the grid $\{t_k = kh \mid k = 0, \dots, N\}$, $Nh = T$, define the discrete path space $\mathcal{P}_d(TQ) := \{(q_d, v_d)\colon \{t_k\}_{k=0}^N \to TQ\}$. We will identify a discrete trajectory $(q_d, v_d) \in \mathcal{P}_d(TQ)$ with its image $(q_d, v_d) = \{(q_k, v_k)\}_{k=0}^N$ where $(q_k, v_k) := (q_d(t_k), v_d(t_k))$. The discrete action $\mathcal{A}_d\colon \mathcal{P}_d(TQ) \to \mathbb{R}$ along this sequence is calculated by summing the discrete Lagrangian evaluated at each pair of adjacent points of the discrete path, that is,

$$\mathcal{A}_d(q_d, v_d) := \sum_{k=0}^{N-1} L_d(q_k, v_k, q_{k+1}, v_{k+1}).$$

We would like to point out that the discrete path space is isomorphic to the smooth product manifold which consists on $N+1$ copies of TQ, the discrete action inherits the smoothness of the discrete Lagrangian, and the tangent space $T_{(q_d,v_d)}\mathcal{P}_d(TQ)$ at (q_d, v_d) is the set of maps $a_{(q_d,v_d)}\colon \{t_k\}_{k=0}^N \to TTQ$ such that $\tau_{TQ} \circ a_{(q_d,v_d)} = (q_d, v_d)$ where $\tau_{TQ}\colon TTQ \to TQ$ is the canonical projection.

[2] By this we mean, from now on, that there exists $h_0 > 0$ such that for all $h \in (0, h_0)$ the definition or proof holds.

Hamilton's principle seeks discrete curves $\{(q_k, v_k)\}_{k=0}^N$ that satisfy

$$\delta \sum_{k=0}^{N-1} L_d(q_k, v_k, q_{k+1}, v_{k+1}) = 0$$

for all variations $\{(\delta q_k, \delta v_k)\}_{k=0}^N$ vanishing at the endpoints. This is equivalent to the *discrete Euler–Lagrange equations*

$$D_3L_d(q_{k-1}, v_{k-1}, q_k, v_k) + D_1L_d(q_k, v_k, q_{k+1}, v_{k+1}) = 0, \tag{11a}$$

$$D_4L_d(q_{k-1}, v_{k-1}, q_k, v_k) + D_2L_d(q_k, v_k, q_{k+1}, v_{k+1}) = 0, \tag{11b}$$

for $1 \leq k \leq N-1$.

Given a solution $\{q_k^*, v_k^*\}_{k\in\mathbb{Z}}$ of equations (11) and assuming that the $2n \times 2n$ matrix

$$\begin{pmatrix} D_{13}L_d(q_k, v_k, q_{k+1}, v_{k+1}) & D_{14}L_d(q_k, v_k, q_{k+1}, v_{k+1}) \\ D_{23}L_d(q_k, v_k, q_{k+1}, v_{k+1}) & D_{24}L_d(q_k, v_k, q_{k+1}, v_{k+1}) \end{pmatrix}$$

is nonsingular, it is possible to define the (local) *discrete flow* $F_{L_d}\colon \mathcal{U}_k \subset TQ\times TQ \to TQ \times TQ$ mapping $(q_{k-1}, v_{k-1}, q_k, v_k)$ to $(q_k, v_k, q_{k+1}, v_{k+1})$ from (11) where $\mathcal{U}_k$ is a neighborhood of the point $(q_{k-1}^*, v_{k-1}^*, q_k^*, v_k^*)$. The simplecticity and momentum preservation of the discrete flow is derived in [25].

Example 3.2. Cubic splines Let $Q = \mathbb{R}^n$ and $L\colon T^{(2)}Q \equiv (\mathbb{R}^n)^3 \to \mathbb{R}$ be the second-order Lagrangian given by $L(q, \dot{q}, \ddot{q}) = \frac{1}{2}\ddot{q}^2$.

It is well known that the solutions to the corresponding Euler–Lagrange equations $q^{(4)} = 0$ are the so-called cubic splines $q(t) = at^3 + bt^2 + ct + d$, for $a, b, c, d \in \mathbb{R}^n$. We define $L_d\colon (\mathbb{R}^n \times \mathbb{R}^n) \times (\mathbb{R}^n \times \mathbb{R}^n) \to \mathbb{R}$ as follows. Write

$$q(0) = q(h) - h\dot{q}(h) + \frac{h^2}{2}\ddot{q}(h) + \mathcal{O}(h^3), \tag{12a}$$

$$q(h) = q(0) + h\dot{q}(0) + \frac{h^2}{2}\ddot{q}(0) + \mathcal{O}(h^3). \tag{12b}$$

Given sufficiently close $(q_0, v_0), (q_1, v_1) \in TQ$ we can use equations (12) to obtain approximations of the acceleration of the exact solution joining these boundary conditions at time h, which we call

$$a_0 = \frac{2}{h^2}(q_1 - q_0 - hv_0) \text{ and } a_1 = \frac{2}{h^2}(q_0 - q_1 + hv_1).$$

Then we define

$$L_d(q_0, v_0, q_1, v_1) = \frac{h}{2}\left(L(q_0, v_0, a_0) + L(q_1, v_1, a_1)\right) = \frac{(hv_1 + q_0 - q_1)^2}{h^3} + \frac{(-hv_0 - q_0 + q_1)^2}{h^3}.$$

Solving the discrete second-order Euler–Lagrange equations for this discrete Lagrangian, the evolution of the discrete trajectory is

$$q_{k+1} = q_{k-1} + 2hv_k, \tag{13a}$$

$$v_{k+1} = v_{k-1} + 4\left(v_k - \frac{q_k - q_{k-1}}{h}\right). \tag{13b}$$

In the following section we will continue this example and show some simulations.

3.1. **Discrete Legendre transforms.**

We define the *discrete Legendre transforms* $\mathbb{F}^+L_d, \mathbb{F}^-L_d\colon TQ\times TQ \to T^*TQ$ which maps the space $TQ\times TQ$ into T^*TQ. These are given by

$$\mathbb{F}^+L_d(q_0, v_0, q_1, v_1) = (q_0, v_0, -D_1L_d(q_0, v_0, q_1, v_1), -D_2L_d(q_0, v_0, q_1, v_1)),$$

$$\mathbb{F}^-L_d(q_0, v_0, q_1, v_1) = (q_1, v_1, D_3L_d(q_0, v_0, q_1, v_1), D_4L_d(q_0, v_0, q_1, v_1)).$$

If both discrete fibre derivatives are locally diffeomorphisms for nearby (q_0, v_0) and (q_1, v_1), then we say that L_d is *regular*.

Using the discrete Legendre transforms the discrete Euler–Lagrange equations (11) can be rewritten as

$$\mathbb{F}^- L_d(q_k, v_k, q_{k+1}, v_{k+1}) = \mathbb{F}^+ L_d(q_{k-1}, v_{k-1}, q_k, v_k).$$

It will be useful to note that

$$\begin{aligned}\mathbb{F}^- L_d \circ F_{L_d}(q_0, v_0, q_1, v_1) &= \mathbb{F}^- L_d(q_1, v_1, q_2, v_2)\\ &= (q_1, v_1, -D_1 L_d(q_1, v_1, q_2, v_2), -D_2 L_d(q_1, v_1, q_2, v_2))\\ &= (q_1, v_1, D_3 L_d(q_0, v_0, q_1, v_1), D_4 L_d(q_0, v_0, q_1, v_1))\\ &= \mathbb{F}^+ L_d(q_0, v_0, q_1, v_1),\end{aligned}$$

that is,

$$\mathbb{F}^+ L_d = \mathbb{F}^- L_d \circ F_{L_d}. \tag{14}$$

Remark 3.3. It is easy to extend this framework to higher-order mechanical systems. Let $L\colon T^{(\ell)}Q \to \mathbb{R}$ be a regular higher-order Lagrangian. Given a small enough $h > 0$, the *exact discrete Lagrangian* $L_d^e\colon T^{(\ell-1)}Q \times T^{(\ell-1)}Q \to \mathbb{R}$ is defined by

$$L_d^e(q_0^{(0)}, q_0^{(1)}, \ldots, q_0^{(\ell-1)}; q_1^{(0)}, q_1^{(1)}, \ldots, q_1^{(\ell-1)}) = \int_0^h L(q(t), \dot q(t), \ldots, q^{(\ell)}(t))dt,$$

where $q(t)\colon I \subset \mathbb{R} \to Q$ is the unique solution of the Euler–Lagrange equations for the higher-order Lagrangian L,

$$\sum_{j=0}^{\ell} (-1)^j \frac{d^j}{dt^j} \frac{\partial L}{\partial q^{(j)}} = 0,$$

satisfying the boundary conditions $q(0) = q_0^{(0)}, \dot q(0) = q_0^{(1)}, \ldots, q^{(\ell-1)}(0) = q_0^{(\ell-1)}, q(h) = q_1^{(0)}, \dot q(h) = q_1^{(1)}, \ldots, q^{(\ell-1)}(h) = q_1^{(\ell-1)}$.

The exact discrete Lagrangian is actually defined on a neighborhood of the diagonal of $T^{(\ell-1)}Q \times T^{(\ell-1)}Q$. We take $L_d\colon T^{(\ell-1)}Q \times T^{(\ell-1)}Q \to \mathbb{R}$ to be an approximation of L_d^e in order to construct variational integrators for higher-order mechanical systems.

Given a discrete path $\{(q_k^{(0)}, \ldots, q_k^{(\ell-1)}) \in T^{(\ell-1)}Q\}|_{k=0}^N$, the corresponding discrete action is defined as

$$\mathcal{A}_d := \sum_{k=0}^{N-1} L_d(q_k^{(0)}, \ldots, q_k^{(\ell-1)}; q_{k+1}^{(0)}, \ldots, q_{k+1}^{(\ell-1)}).$$

Hamilton's principle seeks discrete paths that satisfy $\delta\mathcal{A}_d = 0$ for all variations $\{(\delta q_k^{(0)}, \ldots, \delta q_k^{(\ell-1)})|_{k=0}^N\}$ vanishing at the endpoints $k = 0, N$. This is equivalent to the *discrete higher-order Euler–Lagrange equations for* L_d*:*

$$D_{i+\ell} L_d(q_{k-1}^{(0)}, \ldots, q_{k-1}^{(\ell-1)}; q_k^{(0)}, \ldots, q_k^{(\ell-1)}) + D_i L_d(q_k^{(0)}, \ldots, q_k^{(\ell-1)}; q_{k+1}^{(0)}, \ldots, q_{k+1}^{(\ell-1)}) = 0$$

for $i = 1, \ldots, \ell$ and $k = 1, \ldots, N-1$. ♢

4. Relationship between discrete and continuous variational systems

Let $L\colon T^{(2)}Q \to \mathbb{R}$ be a regular Lagrangian and, for small enough $h > 0$, consider the exact discrete Lagrangian defined before, that is, a function $L_d^e\colon TQ \times TQ \to \mathbb{R}$ given by

$$L_d^e(q_0, \dot q_0, q_1, \dot q_1) = \int_0^h L(q(t), \dot q(t), \ddot q(t))dt,$$

where $q\colon [0,h] \to Q$ is the unique solution of the Euler–Lagrange equations for the second-order Lagrangian L,

$$\frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot q} - \frac{d}{dt}\frac{\partial L}{\partial \dot q} + \frac{\partial L}{\partial q} = 0$$

satisfying the boundary conditions $q(0) = q_0, q(h) = q_1, \dot q(0) = \dot q_0$ and $\dot q(h) = \dot q_1$.

The Legendre transformation associated to L is defined to be the map $\mathbb{F}L\colon T^{(3)}Q \to T^*TQ$ given by (see [14])

$$\mathbb{F}L(q,\dot{q},\ddot{q},q^{(3)}) = \left(q,\dot{q},\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}},\frac{\partial L}{\partial \ddot{q}}\right).$$

We will see that there is a special relationship between the Legendre transform of a regular Lagrangian and the discrete Legendre transforms of the corresponding exact discrete Lagrangian L_d^e.

Theorem 4.1. *Let $L\colon T^{(2)}Q \to \mathbb{R}$ be a regular Lagrangian and $L_d^e\colon TQ\times TQ \to \mathbb{R}$, the corresponding exact discrete Lagrangian. Then L and L_d^e have Legendre transformations related by*

$$\begin{aligned}\mathbb{F}^- L_d^e(q(0),\dot{q}(0),q(h),\dot{q}(h)) &= \mathbb{F}L(q(0),\dot{q}(0),\ddot{q}(0),q^{(3)}(0))\\ \mathbb{F}^+ L_d^e(q(0),\dot{q}(0),q(h),\dot{q}(h)) &= \mathbb{F}L(q(h),\dot{q}(h),\ddot{q}(h),q^{(3)}(h)),\end{aligned}$$

where $q(t)$ is a solution of the second-order Euler–Lagrange equations.

Proof. We begin by computing the derivatives of L_d^e.

$$\begin{aligned}\frac{\partial L_d^e}{\partial q_0} &= \int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial q_0} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial q_0} + \frac{\partial L}{\partial \ddot{q}}\frac{\partial \ddot{q}}{\partial q_0}\right) dt\\ &= \int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial q_0} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial q_0} - \left(\frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial \dot{q}}{\partial q_0}\right) dt + \left(\frac{\partial L}{\partial \ddot{q}}\frac{\partial \dot{q}}{\partial q_0}\right)\Big|_0^h\\ &= \int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial q_0} + \left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial \dot{q}}{\partial q_0}\right) dt,\end{aligned}$$

where we have used integration by parts and the fact that

$$\frac{\partial \dot{q}}{\partial q_0}(0) = 0 \text{ and } \frac{\partial \dot{q}}{\partial q_0}(h) = 0.$$

Therefore,

$$\frac{\partial L_d^e}{\partial q_0} = \left(\left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial q}{\partial q_0}\right)\Big|_0^h + \int_0^h \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial q}{\partial q_0}\, dt.$$

Since $q(t)$ is a solution of the Euler–Lagrange equations for $L\colon T^{(2)}Q \to \mathbb{R}$, the last term is zero. Therefore,

$$\frac{\partial L_d^e}{\partial q_0} = \left(\left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial q}{\partial q_0}\right)\Big|_0^h = \left(-\frac{\partial L}{\partial \dot{q}} + \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)(q(0),\dot{q}(0),\ddot{q}(0),q^{(3)}(0)), \tag{15}$$

because

$$\frac{\partial q}{\partial q_0}(0) = \mathrm{Id} \text{ and } \frac{\partial q}{\partial q_0}(h) = 0.$$

On the other hand,

$$\begin{aligned}\frac{\partial L_d^e}{\partial \dot{q}_0} &= \int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial \dot{q}_0} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial \dot{q}_0} + \frac{\partial L}{\partial \ddot{q}}\frac{\partial \ddot{q}}{\partial \dot{q}_0}\right) dt =\\ &\int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial \dot{q}_0} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial \dot{q}_0} - \left(\frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial \dot{q}}{\partial \dot{q}_0}\right) dt + \left(\frac{\partial L}{\partial \ddot{q}}\frac{\partial \dot{q}}{\partial \dot{q}_0}\right)\Big|_0^h =\\ &\int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial \dot{q}_0} + \left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial \dot{q}}{\partial \dot{q}_0}\right) dt + \left(\frac{\partial L}{\partial \ddot{q}}\frac{\partial \dot{q}}{\partial \dot{q}_0}\right)\Big|_0^h =\\ &\int_0^h \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial q}{\partial \dot{q}_0}dt + \frac{\partial L}{\partial \ddot{q}}\frac{\partial \dot{q}}{\partial \dot{q}_0}\Big|_0^h + \left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial q}{\partial \dot{q}_0}\Big|_0^h.\end{aligned}$$

Since $q(t)$ is a solution of the Euler–Lagrange equations, the first term is zero, and using that

$$\frac{\partial \dot{q}}{\partial \dot{q}_0}(0) = \mathrm{Id}, \quad \frac{\partial \dot{q}}{\partial \dot{q}_0}(h) = 0, \quad \frac{\partial q}{\partial \dot{q}_0}(0) = 0, \text{ and } \frac{\partial q}{\partial \dot{q}_0}(h) = 0,$$

we have

$$\frac{\partial L_d^e}{\partial \dot{q}_0} = -\frac{\partial L}{\partial \ddot{q}}(q(0), \dot{q}(0), \ddot{q}(0)).$$

Therefore

$$\begin{aligned}\mathbb{F}^- L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h)) &= \Big(q(0), \dot{q}(0), -\frac{\partial L_d^e}{\partial q_0}(q(0), \dot{q}(0), q(h), \dot{q}(h)), -\frac{\partial L_d^e}{\partial \dot{q}_0}(q(0), \dot{q}(0), q(h), \dot{q}(h))\Big) \\ &= \mathbb{F}L(q(0), \dot{q}(0), \ddot{q}(0), q^{(3)}(0)).\end{aligned}$$

With similar arguments, we can also prove that

$$\frac{\partial L_d^e}{\partial q_1} = \left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)(q(h), \dot{q}(h), \ddot{q}(h), q^{(3)}(h))$$

and

$$\frac{\partial L_d^e}{\partial \dot{q}_1} = \frac{\partial L}{\partial \ddot{q}}(q(h), \dot{q}(h), \ddot{q}(h)),$$

and in consequence,

$$\mathbb{F}^+ L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h)) = \mathbb{F}L(q(h), \dot{q}(h), \ddot{q}(h), q^{(3)}(h)).$$ □

In what follows we will study the relation between the regularity of the continuous Lagrangian, given by the hessian matrix

$$\mathcal{W} = \left(\frac{\partial^2 L}{\partial \ddot{q}\, \partial \ddot{q}}\right)$$

and the regularity condition corresponding to the exact discrete Lagrangian $L_d^e \colon TQ \times TQ \to \mathbb{R}$

$$\mathcal{W}_d = \begin{pmatrix} D_{13} L_d^e & D_{14} L_d^e \\ D_{23} L_d^e & D_{24} L_d^e \end{pmatrix}.$$

For the next theorem, we restrict ourselves to Lagrangians that can be written locally as

$$L(q, \dot{q}, \ddot{q}) = \frac{1}{2} g_{ij}(q) \ddot{q}^i \ddot{q}^j + \ddot{q}^i f_i(q, \dot{q}) + V(q, \dot{q}), \tag{16}$$

where $(g_{ij}(q))$ is a regular matrix for all q. It is also possible to write this condition intrinsically by using a metric, a connection, a one-form and a function. This covers the kind of Lagrangians that appear in interpolation problems [16] and in optimal control problems with cost functionals of the form $\frac{1}{2}\int_0^T \|u\|^2 dt$, where u represents the control force applied to a system having a (first-order) Lagrangian of mechanical type (see section 5).

Theorem 4.2. *Let $L \colon T^{(2)}Q \to \mathbb{R}$ be a regular Lagrangian of the type* (16)*. For small enough $h > 0$, the corresponding exact discrete Lagrangian $L_d^e \colon TQ \times TQ \to \mathbb{R}$ is also regular.*

Proof. We will work locally. Given q_0, $\dot{q}_0$, q_1, $\dot{q}_1$, consider the curve $q(t)$ that solves the Euler–Lagrange equations with those boundary values, as in the definition of L_d^e. Using the Taylor expansions for $q(t)$ and $\dot{q}(t)$, we can write

$$\begin{aligned} q(h) &= q(0) + h\dot{q}(0) + \frac{h^2}{2}\ddot{q}(0) + \frac{h^3}{6} q^{(3)}(0) + \mathcal{O}\left(h^4\right), \\ \dot{q}(h) &= \dot{q}(0) + h\ddot{q}(0) + \frac{h^2}{2} q^{(3)}(0) + \mathcal{O}\left(h^3\right), \end{aligned}$$

for $h \to 0$. By differentiating these expressions with respect to the parameters q_0 and $\dot{q}_0$, we get two systems of equations from which we find

$$\begin{aligned} \frac{\partial \ddot{q}}{\partial q_0}(h) &= \frac{6}{h^2}\operatorname{Id} + \mathcal{O}\left(h^2\right), & \frac{\partial q^{(3)}}{\partial q_0}(h) &= \frac{12}{h^3}\operatorname{Id} + \mathcal{O}\left(h\right), \\ \frac{\partial \ddot{q}}{\partial \dot{q}_0}(h) &= \frac{2}{h}\operatorname{Id} + \mathcal{O}\left(h^2\right), & \frac{\partial q^{(3)}}{\partial \dot{q}_0}(h) &= \frac{6}{h^2}\operatorname{Id} + \mathcal{O}\left(h\right). \end{aligned}$$

Analogously,

$$\frac{\partial \ddot{q}}{\partial q_1}(0) = \frac{6}{h^2}\operatorname{Id} + \mathcal{O}\left(h^2\right), \quad \frac{\partial q^{(3)}}{\partial q_1}(0) = -\frac{12}{h^3}\operatorname{Id} + \mathcal{O}\left(h\right),$$

$$\frac{\partial \ddot{q}}{\partial \dot{q}_1}(0) = -\frac{2}{h}\operatorname{Id} + \mathcal{O}\left(h^2\right), \quad \frac{\partial q^{(3)}}{\partial \dot{q}_1}(0) = \frac{6}{h^2}\operatorname{Id} + \mathcal{O}\left(h\right).$$

Let us compute $D_{13}L_d^e$. Denote by F the right-hand side of (15), so

$$\begin{aligned}\frac{\partial L_d^e}{\partial q_0^i}(q(0),\dot{q}(0),q(h),\dot{q}(h)) &= \left(-\frac{\partial L}{\partial \dot{q}^i} + \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}^i}\right)(q(0),\dot{q}(0),\ddot{q}(0),q^{(3)}(0)) \\ &= F_i(q(0),\dot{q}(0),\ddot{q}(0),q^{(3)}(0)).\end{aligned}$$

Recall that $q(0),\dot{q}(0),\ddot{q}(0),q^{(3)}(0)$ are obtained as the initial conditions for the higher-order Euler–Lagrange equations that correspond to the boundary conditions $q(0),\dot{q}(0),q(h),\dot{q}(h)$. We have

$$F_i = -\frac{\partial L}{\partial \dot{q}^i} + \frac{\partial^2 L}{\partial q^j \partial \ddot{q}^i}\dot{q}^j + \frac{\partial^2 L}{\partial \dot{q}^j \partial \ddot{q}^i}\ddot{q}^j + \frac{\partial^2 L}{\partial \ddot{q}^j \partial \ddot{q}^i}q^{(3)j}.$$

Then

$$\begin{aligned}\frac{\partial^2 L_d^e}{\partial q_1^j \partial q_0^i} &= \frac{\partial F_i}{\partial q^k}\frac{\partial q^k}{\partial q_1^j} + \frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \dot{q}^k}{\partial q_1^j} + \frac{\partial F_i}{\partial \ddot{q}^k}\frac{\partial \ddot{q}^k}{\partial q_1^j} + \frac{\partial F_i}{\partial q^{(3)k}}\frac{\partial q^{(3)k}}{\partial q_1^j} = \frac{\partial F_i}{\partial \ddot{q}^k}\frac{\partial \ddot{q}^k}{\partial q_1^j} + \frac{\partial F_i}{\partial q^{(3)k}}\frac{\partial q^{(3)k}}{\partial q_1^j} \\ &= \left(-\frac{\partial^2 L}{\partial \ddot{q}^k \partial \dot{q}^i} + \frac{\partial^3 L}{\partial \ddot{q}^k \partial q^j \partial \ddot{q}^i}\dot{q}^j + \frac{\partial^3 L}{\partial \ddot{q}^k \partial \dot{q}^j \partial \ddot{q}^i}\ddot{q}^j + \frac{\partial^2 L}{\partial \dot{q}^k \partial \ddot{q}^i} + \frac{\partial^3 L}{\partial \ddot{q}^k \partial \ddot{q}^j \partial \ddot{q}^i}q^{(3)j}\right)\frac{\partial \ddot{q}^k}{\partial q_1^j} \\ &\quad + \frac{\partial^2 L}{\partial \ddot{q}^k \partial \ddot{q}^i}\frac{\partial q^{(3)k}}{\partial q_1^j} \\ &= \left(-\frac{\partial^2 L}{\partial \ddot{q}^k \partial \dot{q}^i} + \frac{\partial^2 L}{\partial \dot{q}^k \partial \ddot{q}^i} + \frac{d\mathcal{W}_{ik}}{dt}\right)\left(\frac{6}{h^2}\delta_j^k + \mathcal{O}(h^2)\right) + \frac{\partial^2 L}{\partial \ddot{q}^k \partial \ddot{q}^i}\left(-\frac{12}{h^3}\delta_j^k + \mathcal{O}(h)\right).\end{aligned}$$

In the expression above, the derivatives are evaluated at the arguments corresponding to time 0 for each function. It is important to note that the first factor involves $\ddot{q}(0)$ and $q^{(3)}(0)$, which can blow up for $h \to 0$, even in the simple case of cubic splines. However, for L of the type (16) we have

$$\frac{\partial^2 L}{\partial \ddot{q}^k \partial \dot{q}^i} = \frac{\partial f_k}{\partial \dot{q}^i}, \qquad \frac{\partial^2 L}{\partial \dot{q}^k \partial \ddot{q}^i} = \frac{\partial f_i}{\partial \dot{q}^k}, \qquad \frac{d\mathcal{W}_{ik}}{dt} = \frac{d}{dt}\frac{\partial^2 L}{\partial \ddot{q}^k \partial \ddot{q}^i} = \frac{d}{dt}g_{ik} = \frac{\partial g_{ik}}{\partial q^l}\dot{q}^l.$$

These expressions do not contain $\ddot{q}$ or $q^{(3)}$, so they are $\mathcal{O}(1)$ for $h \to 0$. Therefore,

$$D_{13}L_d^e(q(0),\dot{q}(0),q(h),\dot{q}(h)) = \frac{\partial^2 L_d^e}{\partial q_0 \partial q_1}(q(0),\dot{q}(0),q(h),\dot{q}(h)) = -\frac{12}{h^3}\mathcal{W} + \mathcal{O}\left(\frac{1}{h^2}\right).$$

The remaining derivatives in $\mathcal{W}_d$ can be computed without using the special form (16) of the Lagrangian.

$$\begin{aligned}D_{14}L_d^e(q(0),\dot{q}(0),q(h),\dot{q}(h)) &= \frac{\partial^2 L_d^e}{\partial q_0 \partial \dot{q}_1}(q(0),\dot{q}(0),q(h),\dot{q}(h)) = \frac{6}{h^2}\mathcal{W} + \mathcal{O}\left(\frac{1}{h}\right) \\ D_{23}L_d^e(q(0),\dot{q}(0),q(h),\dot{q}(h)) &= \frac{\partial^2 L_d^e}{\partial \dot{q}_0 \partial q_1}(q(0),\dot{q}(0),q(h),\dot{q}(h)) = \frac{6}{h^2}\mathcal{W} + \mathcal{O}\left(\frac{1}{h}\right) \\ D_{24}L_d^e(q(0),\dot{q}(0),q(h),\dot{q}(h)) &= \frac{\partial^2 L_d^e}{\partial \dot{q}_0 \partial \dot{q}_1}(q(0),\dot{q}(0),q(h),\dot{q}(h)) = -\frac{2}{h}\mathcal{W} + \mathcal{O}\left(1\right).\end{aligned}$$

Seeing $\mathcal{W}_d$ as a block matrix, a well-known result from linear algebra leads us to

$$\det \mathcal{W}_d = \left(-\frac{12}{h^4}\right)^{\dim Q}\det \mathcal{W}^2 + \mathcal{O}\left(\frac{1}{h^{4\dim Q - 1}}\right).$$

That is, for small enough h, if L is regular then L_d^e is regular. □

In what follows we denote $(TQ \times TQ)_2$ the subset of $(TQ \times TQ) \times (TQ \times TQ)$ given by

$$(TQ \times TQ)_2 := \{(q_0,\dot{q}_0,q_1,\dot{q}_1,\tilde{q}_1,\dot{\tilde{q}}_1,q_2,\dot{q}_2) \mid \bar{\pi}_2(q_0,\dot{q}_0,q_1,\dot{q}_1) = \bar{\pi}_1(\tilde{q}_1,\dot{\tilde{q}}_1,q_2,\dot{q}_2)\}.$$

If $L\colon T^{(2)}Q \to \mathbb{R}$ is a regular Lagrangian then the Euler–Lagrange equations for L gives rise a system of explicit 4-order differential equations

$$q^{(4)} = \Psi(q,\dot{q},\ddot{q},q^{(3)}).$$

Therefore, for h given, it is possible to derive the following application (see [2])

$$\Psi_L^h \colon T^{(3)}Q \to T^{(3)}Q$$

which maps $(q(0), \dot q(0), \ddot q(0), q^{(3)}(0)) \in T^{(3)}Q$ into $(q(h), \dot q(h), \ddot q(h), q^{(3)}(h)) \in T^{(3)}Q$. Therefore, from Theorem 4.1 we deduce the commutativity the diagram in Figure 1.

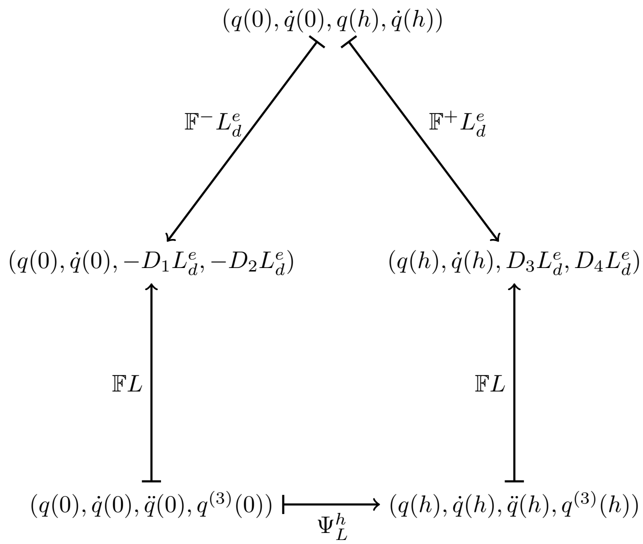

FIGURE 1. Correspondence between the discrete Legendre transforms and the continuous Hamiltonian flow.

Definition 4.3. *The discrete Hamiltonian flow is defined by* $\widetilde{F}_{L_d} \colon T^*TQ \to T^*TQ$ *as*

$$\widetilde{F}_{L_d} = \mathbb{F}^- L_d \circ F_{L_d} \circ (\mathbb{F}^- L_d)^{-1}. \tag{17}$$

Alternatively, it can also be defined as $\widetilde{F}_{L_d} = \mathbb{F}^+ L_d \circ F_{L_d} \circ (\mathbb{F}^+ L_d)^{-1}$.

Theorem 4.4. *The diagram in Figure 2 is commutative.*

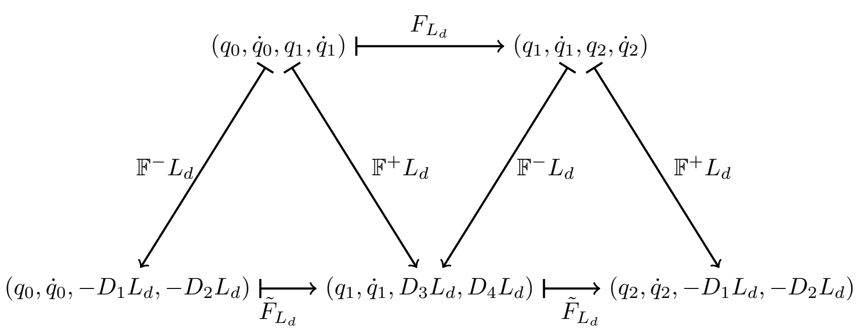

FIGURE 2. Correspondence between the discrete Lagrangian and the discrete Hamiltonian maps.

Proof. The central triangle is (14). The parallelogram on the left-hand side is commutative by (17), so the triangle on the left is commutative. The triangle on the right is the same as the triangle on the left, with shifted indices. Then parallelogram on the right-hand side is commutative, which gives the equivalence stated in the definition of the discrete Hamiltonian flow. □

Corollary 4.5. *The following definitions of the discrete Hamiltonian map are equivalent*

$$\widetilde{F}_{L_d} = \mathbb{F}^+L_d \circ F_{L_d} \circ (\mathbb{F}^+L_d)^{-1},$$
$$\widetilde{F}_{L_d} = \mathbb{F}^-L_d \circ F_{L_d} \circ (\mathbb{F}^-L_d)^{-1},$$
$$\widetilde{F}_{L_d} = \mathbb{F}^+L_d \circ (\mathbb{F}^-L_d)^{-1},$$

and have the coordinate expression $\widetilde{F}_{L_d}\colon (q_0, \dot{q}_0, p_0, \tilde{p}_0) \mapsto (q_1, \dot{q}_1, p_1, \tilde{p}_1)$*, where we use the notation*

$$\begin{aligned} p_0 &= -D_1L_d(q_0, \dot{q}_0, q_1, \dot{q}_1),\\ \tilde{p}_0 &= -D_2L_d(q_0, \dot{q}_0, q_1, \dot{q}_1),\\ p_1 &= D_3L_d(q_0, \dot{q}_0, q_1, \dot{q}_1),\\ \tilde{p}_1 &= D_4L_d(q_0, \dot{q}_0, q_1, \dot{q}_1). \end{aligned}$$

Combining Theorem (4.1) with the diagram in Figure 2 gives the commutative diagram shown in Figure 3 for the exact discrete Lagrangian.

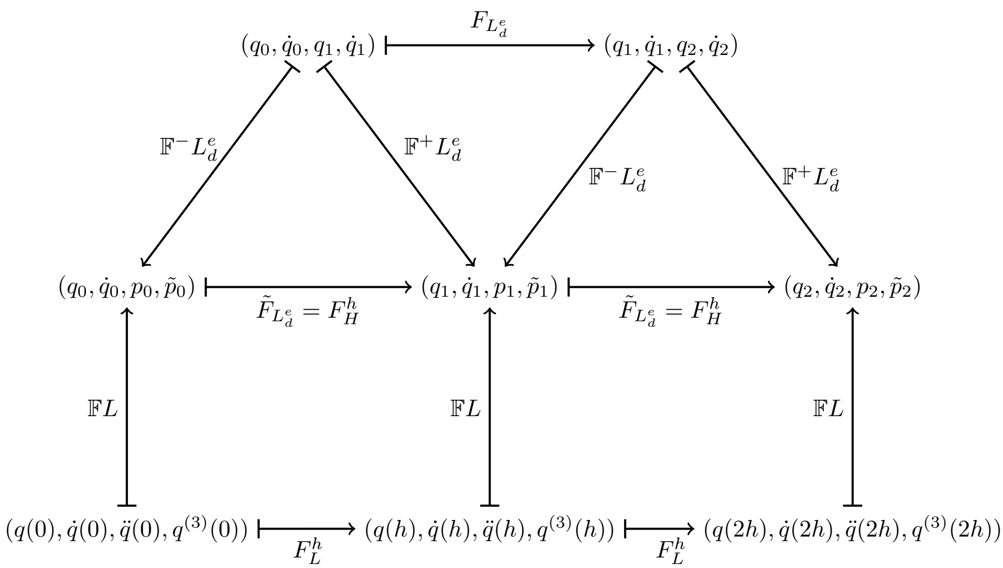

FIGURE 3. Correspondence between the exact discrete Lagrangian and the continuous Hamiltonian flow.

Here, F_H^h denotes the flow of the Hamiltonian vector field X_H associated with the Hamiltonian $H\colon T^*TQ \to \mathbb{R}$ given by $H = E_L \circ (\mathbb{F}L)^{-1}$ where $E_L\colon T^{(3)}Q \to \mathbb{R}$ denotes the energy function associated to L (see [14]).

Theorem 4.6. *Under these conditions we have that* $F_H^h = \tilde{F}_{L_d^e}$.

Example 4.7. Cubic splines (cont.) Recall that in this example $Q = \mathbb{R}^n$ and $L = \frac{1}{2}\ddot{q}^2$. Since the exact solutions for the second-order Euler–Lagrange equation for L can be found explicitly, it is easy to show that the discrete exact Lagrangian is

$$L_d^e(q_0, v_0, q_1, v_1) = \frac{6}{h^3}(q_0 - q_1)^2 + \frac{6}{h^2}(q_0 - q_1)(v_0 + v_1) + \frac{2}{h}(v_0^2 + v_0v_1 + v_1^2).$$

From the corresponding discrete second-order Euler–Lagrange equation, the evolution is

$$\begin{aligned} q_{k+1} &= 5q_{k-1} - 4q_k + 2h(v_{k-1} + 2v_k),\\ v_{k+1} &= v_{k-1} + \frac{2}{h}(q_{k-1} - 2q_k + q_{k+1}). \end{aligned}$$

It is interesting to note that both this exact method and method (13) preserve the quantity

$$\varphi(q_k, v_k, q_{k+1}, v_{k+1}) = \frac{q_{k+1} - q_k}{h} - \frac{v_k + v_{k+1}}{2}.$$

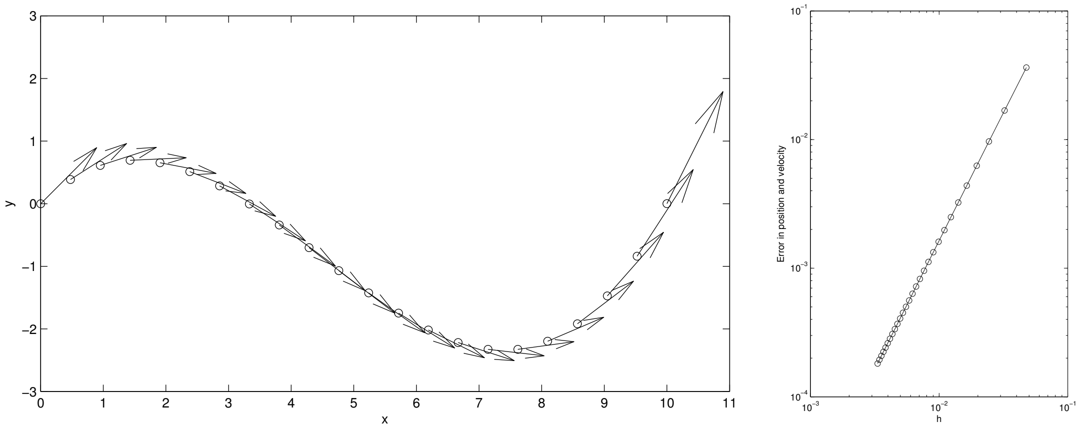

FIGURE 4. Left: simulation of the method (13) with $q_0 = (0, 0)$ $v_0 = (10, 10)$, $q_N = (10, 0)$, $v_N = (10, 20)$, $N = 21$, depicting the computed points and velocities in the xy-plane (velocities are scaled). Right: Error in position and velocity for different values of h.

4.1. **Variational error analysis.** Now we rewrite the result of Patrick [29] and Marsden and West [25] for the particular case of a Lagrangian $L_d \colon TQ \times TQ \to \mathbb{R}$.

Definition 4.8. *Let $L_d \colon TQ \times TQ \to \mathbb{R}$ be a discrete Lagrangian. We say that L_d is a discretization of order r if there exist an open subset $U_1 \subset T^{(2)}Q$ with compact closure and constants $C_1 > 0$, $h_1 > 0$ so that*

$$|L_d(q(0), \dot{q}(0), q(h), \dot{q}(h), h) - L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h), h)| \leq C_1 h^{r+1}$$

for all solutions $q(t)$ of the second-order Euler–Lagrange equations with initial conditions $(q_0, \dot{q}_0, \ddot{q}_0) \in U_1$ and for all $h \leq h_1$.

Following [25, 30], we have the next result about the order of our variational integrator.

Theorem 4.9. *If $\widetilde{F}_{L_d}$ is the evolution map of an order r discretization $L_d \colon TQ \times TQ \to \mathbb{R}$ of the exact discrete Lagrangian $L_d^e \colon TQ \times TQ \to \mathbb{R}$, then*

$$\widetilde{F}_{L_d} = \widetilde{F}_{L_d^e} + \mathcal{O}(h^{r+1}).$$

In other words, $\widetilde{F}_{L_d}$ gives an integrator of order r for $\widetilde{F}_{L_d^e} = F_H^h$.

Note that given a discrete Lagrangian $L_d \colon TQ \times TQ \to \mathbb{R}$ its order can be calculated by expanding the expressions for $L_d(q(0), \dot{q}(0), q(h), \dot{q}(h), h)$ in a Taylor series in h and comparing this to the same expansions for the exact Lagrangian. If the series agree up to r terms, then the discrete Lagrangian is of order r.

5. APPLICATION TO OPTIMAL CONTROL OF MECHANICAL SYSTEMS

In this section we will study how to apply our variational integrator to optimal control problems. We will study optimal control problems for fully actuated mechanical systems and we will show how our methods can be applied to the optimal control of a robotic leg.

In the following we will assume that all the control systems are controllable, that is, for any two points q_0 and q_f in the configuration space Q, there exists an admissible control $u(t)$ defined on some interval $[0, T]$ such that the system with initial condition q_0 reaches the point q_f at time T (see [4] and [6] for example).

5.1. **Optimal control of fully actuated systems.** Let $L\colon TQ \to \mathbb{R}$ be a regular Lagrangian and take local coordinates (q^A) on Q where $1 \leq A \leq n$. For this Lagrangian the *controlled Euler–Lagrange equations* are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^A} - \frac{\partial L}{\partial q^A} = u_A, \tag{18}$$

where $u = (u_A) \in U \subset \mathbb{R}^n$ is an open subset of $\mathbb{R}^n$, the set of control parameters.

The optimal control problem consists in finding a trajectory of the state variables and control inputs $(q^{(A)}(t), u^A(t))$ satisfying (18) given initial and final conditions $(q^A(t_0), \dot{q}^A(t_0))$, $(q^A(t_f), \dot{q}^A(t_f))$ respectively, minimizing the cost function

$$\mathcal{A} = \int_{t_0}^{t_f} C(q^A, \dot{q}^A, u_A)dt,$$

where $C\colon TQ \times U \to \mathbb{R}$.

From (18) we can rewrite the cost function as a second-order Lagrangian $\widetilde{L}\colon T^{(2)}Q \to \mathbb{R}$ given by

$$\widetilde{L}(q^A, \dot{q}^A, \ddot{q}^A) = C\left(q^A, \dot{q}^A, \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^A} - \frac{\partial L}{\partial q^A}\right)$$

replacing the controls by the Euler–Lagrange equations in the cost function (see [4] for example).

Suppose that $Q = \mathbb{R}^n$. Then we can define a discretization of the Lagrangian $\widetilde{L}\colon T^{(2)}Q \to \mathbb{R}$ by a discrete Lagrangian $\widetilde{L}_d\colon TQ \times TQ \to \mathbb{R}$,

$$\begin{aligned}\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) = {} & \frac{h}{2}\widetilde{L}\left(\frac{q_k + q_{k+1}}{2}, \frac{v_k + v_{k+1}}{2}, \frac{2}{h^2}(q_{k+1} - q_k - hv_k)\right)\\ & + \frac{h}{2}\widetilde{L}\left(\frac{q_k + q_{k+1}}{2}, \frac{v_k + v_{k+1}}{2}, \frac{2}{h^2}(q_k - q_{k+1} + hv_{k+1})\right).\end{aligned}$$

In the first term, we have computed an approximate value of the acceleration a_k by using the Taylor expansion $q_{k+1} \approx q_k + hv_k + \frac{h^2}{2}a_k$. For the second term, we have approximated a_{k+1} using $q_k \approx q_{k+1} - hv_{k+1} + \frac{h^2}{2}a_{k+1}$, as in Example 3.2.

Other natural possibilities for $\widetilde{L}_d$ are, for instance,

$$\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) = hL\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}, \frac{v_{k+1} - v_k}{h}\right)$$

or

$$\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) = \frac{1}{2}L\left(q_k, v_k, \frac{v_{k+1} - v_k}{h}\right) + \frac{1}{2}L\left(q_{k+1}, v_{k+1}, \frac{v_{k+1} - v_k}{h}\right).$$

Applying the results given in Section 3, we know that the minimizers of the cost function are obtained by solving the discrete second-order Euler–Lagrange equations

$$\begin{aligned}D_1\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) + D_3\widetilde{L}_d(q_{k-1}, v_{k-1}, q_k, v_k) = 0,\\ D_2\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) + D_4\widetilde{L}_d(q_{k-1}, v_{k-1}, q_k, v_k) = 0.\end{aligned}$$

If the matrix

$$\begin{pmatrix} D_{13}\widetilde{L}_d & D_{14}\widetilde{L}_d \\ D_{23}\widetilde{L}_d & D_{24}\widetilde{L}_d \end{pmatrix}$$

is regular, then one can define the discrete Lagrangian map to solve the optimal control problem.

Example 5.1. Two-link manipulator

We consider the optimal control of a two-link manipulator which is a classical example studied in robotics (see for example [26] and [28]). The two-link manipulator consists of two coupled (planar) rigid bodies with mass m_i, length l_i and moments of inertia with respect to the joints J_i, with $i = 1, 2$, respectively.

Let θ_1 and θ_2 be the configuration angles measured as in Figure 5. If we assume one end of the first link to be fixed in an inertial reference frame, the configuration of the system is locally

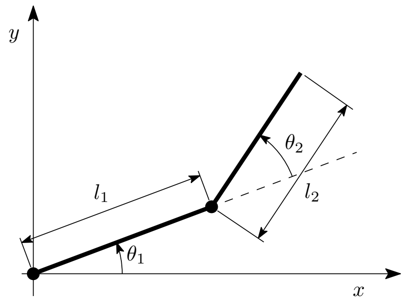

FIGURE 5. Two-link manipulator

specified by the coordinates $(\theta_1, \theta_2) \in \mathbb{S}^1 \times \mathbb{S}^1$. The Lagrangian is given by the kinetic energy of the system minus the potential energy, that is,

$$\begin{aligned} L(q,\dot q) = \frac{1}{8}(m_1+4m_2)l_1^2\dot\theta_1^2 + \frac{1}{8}m_2 l_2^2(\dot\theta_1+\dot\theta_2)^2 + \frac{1}{2}m_2 l_1 l_2\cos(\theta_2)\dot\theta_1(\dot\theta_1+\dot\theta_2) + \frac{1}{2}J_1\dot\theta_1^2 \\ + \frac{1}{2}J_2(\dot\theta_1+\dot\theta_2)^2 + g\left(\frac{1}{2}m_1 l_1\sin\theta_1 + m_2 l_1 \sin\theta_1 + \frac{1}{2}m_2 l_2(\theta_1+\theta_2)\right), \end{aligned}$$

where g is the constant gravitational acceleration.

Control torques u_1 and u_2 are applied at the base of the first link and at the joint between the two links. The equations of motion of the controlled system are

$$\begin{aligned} u_1 &= -\sin\theta_2 l_1 l_2 m_2\dot\theta_2\dot\theta_1 - \frac{1}{2}\sin\theta_2\dot\theta_2^2 l_1 l_2 m_2 + \frac{1}{2}m_2 l_2\cos(\theta_1+\theta_2)g \\ &\quad + \left(m_2 g\cos\theta_1 + \frac{1}{2}g\cos\theta_1 m_1\right)l_1 + \left(\frac{1}{4}m_2 l_2^2 + J_2 + \frac{1}{2}\cos\theta_2 l_1 l_2 m_2\right)\ddot\theta_2 \\ &\quad + \left(\cos\theta_2 l_1 l_2 m_2 + \left(\frac{m_1}{4}+m_2\right)l_1^2 + \frac{m_2 l_2^2}{4} + J_1 + J_2\right)\ddot\theta_1, \\ u_2 &= \frac{1}{2}\sin\theta_2 l_1 l_2 m_2\dot\theta_1^2 + \left(\frac{1}{4}m_2 l_2^2 + J_2 + \frac{1}{2}\cos\theta_2 l_1 l_2 m_2\right)\ddot\theta_1 \\ &\quad + \frac{1}{2}m_2 l_2\cos(\theta_1+\theta_2)g + \left(\frac{1}{4}m_2 l_2^2 + J_2\right)\ddot\theta_2. \end{aligned}$$

We look for trajectories $(\theta_1(t), \theta_2(t), u(t))$ of the state variables and control inputs for given initial and final conditions, that is, for given values of $(\theta_1(0), \theta_2(0), \dot\theta_1(0), \dot\theta_2(0))$ and $(\theta_1(T), \theta_2(T), \dot\theta_1(T), \dot\theta_2(T))$, and minimizing the cost functional

$$\mathcal{A} = \frac{1}{2}\int_0^T (u_1^2 + u_2^2)\, dt.$$

We construct the discrete Lagrangian $\widetilde{L}_d\colon T(\mathbb{S}^1\times\mathbb{S}^1)\times T(\mathbb{S}^1\times\mathbb{S}^1) \to \mathbb{R}$, discretizing the Lagrangian $\widetilde{L}\colon T^{(2)}(\mathbb{S}^1\times\mathbb{S}^1) \to \mathbb{R}$ given by

$$\begin{aligned} \widetilde{L}(\theta_1,\theta_2,\dot\theta_1,\dot\theta_2,\ddot\theta_1,\ddot\theta_2) = \frac{1}{2}\Bigg[\frac{1}{2}\sin\theta_2 l_1 l_2 m_2\dot\theta_1^2 + \left(\frac{1}{4}m_2 l_2^2 + J_2 + \frac{1}{2}\cos\theta_2 l_1 l_2 m_2\right)\ddot\theta_1 \\ + \frac{1}{2}m_2 l_2\cos(\theta_1+\theta_2)g + \left(\frac{1}{4}m_2 l_2^2 + J_2\right)\ddot\theta_2\Bigg]^2 \\ + \frac{1}{2}\Bigg[\frac{1}{2}\sin\theta_2 l_1 l_2 m_2\dot\theta_1^2 + \left(\frac{1}{4}m_2 l_2^2 + J_2 + \frac{1}{2}\cos\theta_2 l_1 l_2 m_2\right)\ddot\theta_1 \\ + \frac{1}{2}m_2 l_2\cos(\theta_1+\theta_2)g + \left(\frac{1}{4}m_2 l_2^2 + J_2\right)\ddot\theta_2\Bigg]^2 \end{aligned}$$

taking the same discretization as in equation (12) to approximate the acceleration and taking midpoint averages to approximate the position and velocity.

Figures 6 and 7 show the results from a numerical simulation of the method, taking the system from the stable mechanical equilibrium $(\theta_1(0), \theta_2(0), \dot{\theta}_1(0), \dot{\theta}_2(0)) = (-\pi/2, 0, 0, 0)$ to the unstable equilibrium $(\theta_1(T), \theta_2(T), \dot{\theta}_1(T), \dot{\theta}_2(T)) = (\pi/2, 0, 0, 0)$. We have used $T = 10$, $N = 1000$, $m_1 = 0.375$, $m_2 = 0.25$, $l_1 = 1.5$, $l_2 = 1$, $J_1 = \frac{m_1 l_1^2}{3}$, $J_2 = \frac{m_2 l_2^2}{3}$, and $g = 9.8$. In addition, the reader can find a video of the simulation in `www.youtube.com/watch?v=ZUUH0596a30`. The algorithm generates a sequence of velocities as well as positions, but we represent only the positions in the figures.

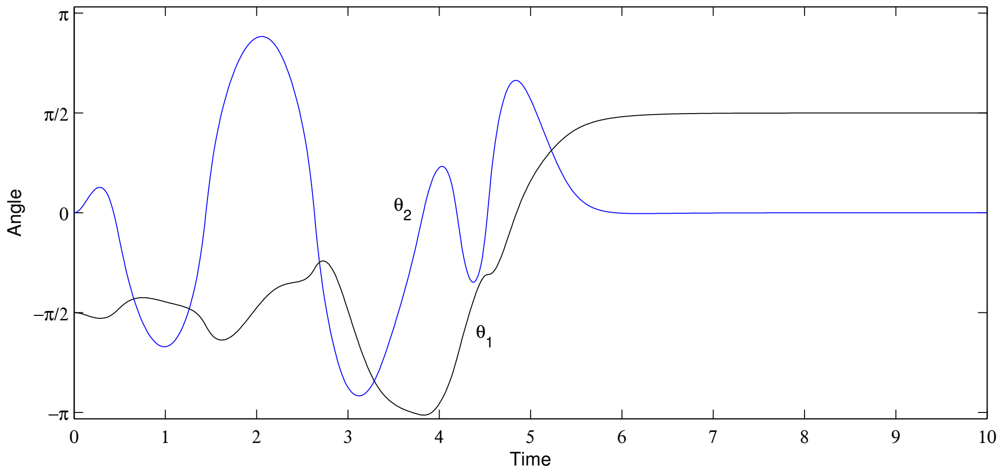

FIGURE 6. Angles θ_1 and θ_2 for the optimal control of the two-link manipulator. Initially, the two links point downwards; at $T = 10$ they point upwards.

We have also considered a different setting where the angle θ_2 is restricted to move between 0 and 170 degrees, inspired by an elbow joint. This range of motion is enforced by adding a continuous, piecewise linear function $V(\theta_2)$ to the cost function, with slope -1000 for $\theta_2 < 0°$, 0 for $0° < \theta_2 < 170°$, and 1000 for $\theta_2 > 170°$. We simulated the optimal trajectory with the same endpoint conditions and physical parameters as above, with $N = 200$. A video of the resulting motion can be found in `www.youtube.com/watch?v=0x0FHdT7emQ`.

Conclusions and future research

In this paper we design variational integrators for higher-order variational systems and their application to optimal control problems. The general idea for those variational integrators is to directly discretize Hamilton's principle rather than the equations of motion in a way that preserves the original system invariants, notably the symplectic form and, via a discrete version of Noether's theorem, the momentum map.

We show that a regular higher-order Lagrangian system has a unique solution for given nearby endpoint conditions using a direct variational proof of existence and uniqueness for the local boundary value problem using a regularization procedure assuming only C^k differentiability (instead of C^{2k} as in standard ODE theory).

We have seen that taking a discrete Lagrangian function $L_d\colon T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$ we obtain the appropriate approximation of the action $\int_0^h L(q, \dot{q}, \ldots, q^{(k)})\, dt$. Moreover, we derive a particular choice of discrete Lagrangian which gives an exact correspondence between discrete and continuous systems, the exact discrete Lagrangian. We show that if the original Lagrangian is regular then it is also the exact discrete Lagrangian and how is the relation between the discrete Legendre transformations with the continuous one.

As future research, we are interested in the construction of an exact discrete Lagrangian function for higher-order mechanical systems subject to higher-order constraints. The main point will be to show the existence and uniqueness of solutions for the boundary value problem

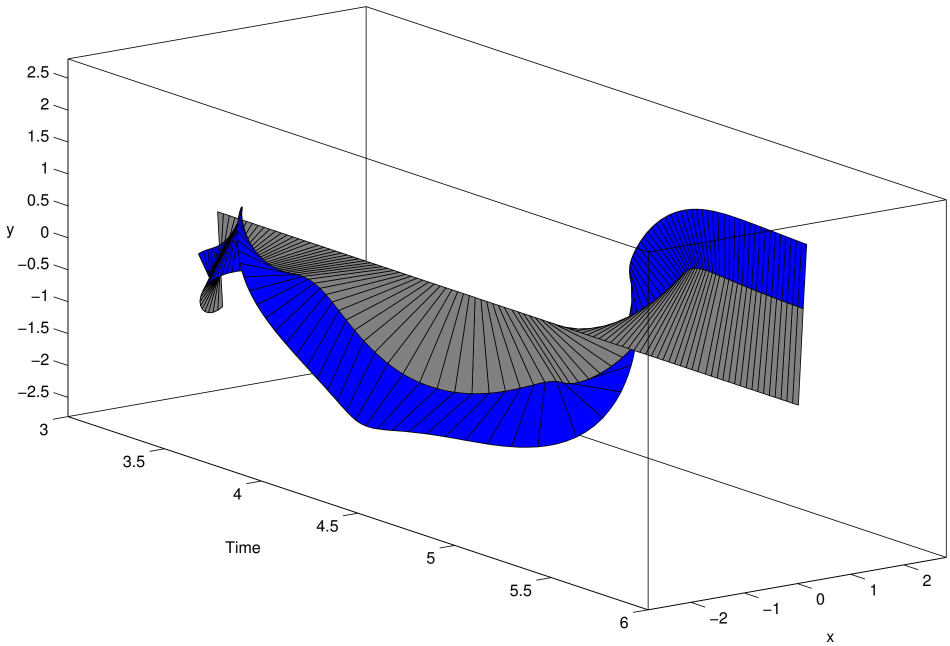

Figure 7. Evolution of the actual position of the two-link manipulator (detail for $t \in [3,6]$). Sections of this surface with the vertical plane $t = t_0$ show the two links as they are positioned at time t_0.

for higher-order systems subject to higher-order constraints. After it, one could define the exact discrete Lagrangian for constrained systems in a similar fashion that the ones shown in this work. Since optimal control problems for the class of underactuated mechanical systems can be seen as constrained higher-order variational problems, the extension of the constructions given in this work, can be useful to new developments in the field of geometric integration for optimal control problems. The case of optimal control of nonholonomic systems will be developed.

Appendix: a technical result for section 2

Let E be the kernel of g, where $g = (g_0, \dots, g_{k-1})\colon C^l([0,1], \mathbb{R}^n) \to (\mathbb{R}^n)^k$ and $g_j[\cdot] = \langle\!\langle b_j^{[k]}, \cdot \rangle\!\rangle$. In the context of section 2.5, E is the tangent space of the constraint set defined using the linear constraints g_j, and l is either 0 or k.

In this Appendix we show that the orthogonal complement of E is the space F of $\mathbb{R}^n$-valued polynomials of degree at most $k-1$,

$$F = \mathrm{span}_{\mathbb{R}^n}(b_0^{[k]}, \dots, b_{k-1}^{[k]}) = \{c^j b_j^{[k]} | c^0, \dots, c^{k-1} \in \mathbb{R}^n\},$$

where $b_j^{[k]}$, $j = 0, \dots, k-1$, is a basis of the space of real-valued polynomials of degree at most $k-1$ consisting of orthonormal polynomials on $[0,1]$.

Lemma 5.2. *$F = E^\perp$, where the orthogonal complement is taken with respect to the inner product $[\![,]\!]$ in $C^l([0,1], \mathbb{R}^n)$.*

Proof. We will prove that E and F are orthogonal (with zero intersection) and that their sum is the whole space $C^l([0,1], \mathbb{R}^n)$.

Let $e \in E$ and $c^j b_j^{[k]} \in F$.

$$[\![c^j b_j^{[k]}, e]\!] = \int_0^1 (c^j b_j^{[k]}(u)) \cdot e(u)\, du = \sum_{i=1}^n \int_0^1 c_i^j b_j^{[k]}(u) e_i(u) du$$
$$= c^j \cdot \left(\int_0^1 b_j^{[k]}(u) e_1(u), \dots, \int_0^1 b_j^{[k]}(u) e_n(u) \right)$$

$$= c^j \cdot \langle\!\langle b_j^{[k]}, e \rangle\!\rangle = c^j \cdot g_j[e] = 0,$$

since $e \in E = \operatorname{Ker} g$.

The fact that $E \cap F = \{0\}$ can be obtained either by using that the inner product is nondegenerate or directly as follows. Take $e \in E \cap F$, so $e = c^j b_j^{[k]}$. For all j', we have $0 = g_{j'}[e] = \langle\!\langle b_{j'}^{[k]}, c^j b_j^{[k]} \rangle\!\rangle = c^{j'}$, which means that $e = 0$.

Finally, take $e \in C^l([0,1], \mathbb{R}^n)$. Write

$$e = e - \sum_{j=0}^{k-1} \langle\!\langle b_j^{[k]}, e \rangle\!\rangle b_j^{[k]} + \sum_{j=0}^{k-1} \langle\!\langle b_j^{[k]}, e \rangle\!\rangle b_j^{[k]}.$$

The third term is in F. The remaining part of the right-hand side is in E since for all j',

$$\left\langle\!\!\left\langle b_{j'}, e - \sum_{j=0}^{k-1} \langle\!\langle b_j^{[k]}, e \rangle\!\rangle b_j^{[k]} \right\rangle\!\!\right\rangle = \langle\!\langle b_{j'}, e \rangle\!\rangle - \sum_{j=0}^{k-1} \delta_{j'j} \langle\!\langle b_j, e \rangle\!\rangle = 0.$$

Therefore $C^l([0,1], \mathbb{R}^n) = E + F$. From the first part of the proof, we obtain that there is an orthogonal decomposition $C^l([0,1], \mathbb{R}^n) = E \oplus F$. □

Acknowledgments

This work has been supported by MICINN (Spain) Grant MTM 2013-42870-P, ICMAT Severo Ochoa Project SEV-2011-0087 and IRSES-project "Geomech-246981". The research of S. Ferraro has been supported by CONICET Argentina (PIP 2013-2015 GI 11220120100532CO), ANPCyT Argentina (PICT 2013-1302) and SGCyT UNS. We would like to thank the referee for the helpful comments.

References

[1] R. Abraham, J. E. Marsden and T. Ratiu, *Manifolds, tensor analysis, and applications*, vol. 75 of Applied Mathematical Sciences, 2nd edition, Springer-Verlag, New York, 1988, URL http://dx.doi.org/10.1007/978-1-4612-1029-0.

[2] R. P. Agarwal, *Boundary value problems for higher order differential equations*, World Scientific Publishing Co., Inc., Teaneck, NJ, 1986, URL http://dx.doi.org/10.1142/0266.

[3] R. Benito, M. de León and D. Martín de Diego, Higher-order discrete Lagrangian mechanics, *Int. J. Geom. Methods Mod. Phys.*, **3** (2006), 421–436, URL http://dx.doi.org/10.1142/S0219887806001235.

[4] A. M. Bloch, *Nonholonomic mechanics and control*, vol. 24 of Interdisciplinary Applied Mathematics, Springer-Verlag, New York, 2003, URL http://dx.doi.org/10.1007/b97376, With the collaboration of J. Baillieul, P. Crouch and J. Marsden, With scientific input from P. S. Krishnaprasad, R. M. Murray and D. Zenkov, Systems and Control.

[5] A. M. Bloch, I. I. Hussein, M. Leok and A. K. Sanyal, Geometric structure-preserving optimal control of a rigid body, *J. Dyn. Control Syst.*, **15** (2009), 307–330, URL http://dx.doi.org/10.1007/s10883-009-9071-2.

[6] F. Bullo and A. D. Lewis, *Geometric control of mechanical systems*, vol. 49 of Texts in Applied Mathematics, Springer-Verlag, New York, 2005, URL http://dx.doi.org/10.1007/978-1-4899-7276-7, Modeling, analysis, and design for simple mechanical control systems.

[7] C. L. Burnett, D. D. Holm and D. M. Meier, Inexact trajectory planning and inverse problems in the Hamilton-Pontryagin framework, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **469** (2013), 20130249, 24, URL http://dx.doi.org/10.1098/rspa.2013.0249.

[8] G. Buttazzo, M. Giaquinta and S. Hildebrandt, *One-dimensional variational problems, an introduction*, vol. 15 of Oxford Lecture Series in Mathematics and its Applications, The Clarendon Press, Oxford University Press, New York, 1998.

[9] L. Colombo and D. Martín de Diego, Higher-order variational problems on Lie groups and optimal control applications, *J. Geom. Mech.*, **6** (2014), 451–478, URL http://dx.doi.org/10.3934/jgm.2014.6.451.

[10] L. Colombo, D. Martín de Diego and M. Zuccalli, On variational integrators for optimal control of mechanical control systems, *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Math. RACSAM*, **106** (2012), 161–171, URL http://dx.doi.org/10.1007/s13398-011-0032-8.

[11] L. Colombo, D. Martín de Diego and M. Zuccalli, Higher-order discrete variational problems with constraints, *J. Math. Phys.*, **54** (2013), 093507, 17, URL http://dx.doi.org/10.1063/1.4820817.

[12] M. Crampin, W. Sarlet and F. Cantrijn, Higher-order differential equations and higher-order Lagrangian mechanics, *Math. Proc. Cambridge Philos. Soc.*, **99** (1986), 565–587, URL http://dx.doi.org/10.1017/S0305004100064501.

[13] P. Crouch and F. Silva Leite, The dynamic interpolation problem: on Riemannian manifolds, Lie groups, and symmetric spaces, *J. Dynam. Control Systems*, **1** (1995), 177–202, URL `http://dx.doi.org/10.1007/BF02254638`.

[14] M. de León and P. R. Rodrigues, *Generalized classical mechanics and field theory*, vol. 112 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, 1985, A geometrical approach of Lagrangian and Hamiltonian formalisms involving higher order derivatives, Notes on Pure Mathematics, 102.

[15] J. Eldering, *Persistence of noncompact normally hyperbolic invariant manifolds in bounded geometry*, PhD thesis, Universiteit Utrecht, 2012.

[16] F. Gay-Balmaz, D. D. Holm, D. M. Meier, T. S. Ratiu and F.-X. Vialard, Invariant higher-order variational problems, *Comm. Math. Phys.*, **309** (2012), 413–458, URL `http://dx.doi.org/10.1007/s00220-011-1313-y`.

[17] F. Gay-Balmaz, D. D. Holm, D. M. Meier, T. S. Ratiu and F.-X. Vialard, Invariant higher-order variational problems II, *J. Nonlinear Sci.*, **22** (2012), 553–597, URL `http://dx.doi.org/10.1007/s00332-012-9137-2`.

[18] F. Gay-Balmaz, D. D. Holm and T. S. Ratiu, Higher order Lagrange-Poincaré and Hamilton-Poincaré reductions, *Bull. Braz. Math. Soc. (N.S.)*, **42** (2011), 579–606, URL `http://dx.doi.org/10.1007/s00574-011-0030-7`.

[19] M. Giaquinta and S. Hildebrandt, *Calculus of variations. I*, vol. 310 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1996.

[20] I. I. Hussein and A. M. Bloch, Dynamic interpolation on Riemannian manifolds: an application to interferometric imaging, in *Proceedings of the 2004 American Control Conference*, 2004, 685–690.

[21] B. W. Jordan and E. Polak, Theory of a class of discrete optimal control systems, *J. Electronics Control (1)*, **17** (1964), 697–711.

[22] T. Lee, M. Leok and N. H. McClamroch, Optimal attitude control of a rigid body using geometrically exact computations on SO(3), *J. Dyn. Control Syst.*, **14** (2008), 465–487, URL `http://dx.doi.org/10.1007/s10883-008-9047-7`.

[23] M. Leok and T. Shingel, Prolongation-collocation variational integrators, *IMA J. Numer. Anal.*, **32** (2012), 1194–1216, URL `http://dx.doi.org/10.1093/imanum/drr042`.

[24] L. Machado, F. Silva Leite and K. Krakowski, Higher-order smoothing splines versus least squares problems on Riemannian manifolds, *J. Dyn. Control Syst.*, **16** (2010), 121–148, URL `http://dx.doi.org/10.1007/s10883-010-9080-1`.

[25] J. E. Marsden and M. West, Discrete mechanics and variational integrators, *Acta Numer.*, **10** (2001), 357–514, URL `http://dx.doi.org/10.1017/S096249290100006X`.

[26] R. N. Murray, Z. X. Li and S. S. Sastry, *A mathematical introduction to robotic manipulation*, CRC Press, Boca Raton, FL, 1994.

[27] L. Noakes, G. Heinzinger and B. Paden, Cubic splines on curved spaces, *IMA J. Math. Control Inform.*, **6** (1989), 465–473, URL `http://dx.doi.org/10.1093/imamci/6.4.465`.

[28] S. Ober-Blöbaum, O. Junge and J. E. Marsden, Discrete mechanics and optimal control: an analysis, *ESAIM Control Optim. Calc. Var.*, **17** (2011), 322–352, URL `http://dx.doi.org/10.1051/cocv/2010012`.

[29] G. W. Patrick, Lagrangian mechanics without ordinary differential equations, *Rep. Math. Phys.*, **57** (2006), 437–443, URL `http://dx.doi.org/10.1016/S0034-4877(06)80030-3`.

[30] G. W. Patrick and C. Cuell, Error analysis of variational integrators of unconstrained Lagrangian systems, *Numer. Math.*, **113** (2009), 243–264, URL `http://dx.doi.org/10.1007/s00211-009-0245-3`.

[31] A. P. Veselov, Integrable systems with discrete time, and difference operators, *Funktsional. Anal. i Prilozhen.*, **22** (1988), 1–13, 96, URL `http://dx.doi.org/10.1007/BF01077598`.

[32] J. M. Wendlandt and J. E. Marsden, Mechanical integrators derived from a discrete variational principle, *Phys. D*, **106** (1997), 223–246, URL `http://dx.doi.org/10.1016/S0167-2789(97)00051-1`.

Leonardo Colombo: Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA

E-mail address: `ljcolomb@umich.edu`

Sebastián Ferraro: Departamento de Matemática, Universidad Nacional del Sur, Instituto de Matemática Bahía Blanca, and CONICET, Av. Alem 1253, 8000 Bahía Blanca, Argentina

E-mail address: `sferraro@uns.edu.ar`

David Martín de Diego: Instituto de Ciencias Matemáticas (CSIC-UAM-UC3M-UCM), Campus de Cantoblanco, UAM C/ Nicolas Cabrera, 15 - 28049 Madrid, Spain

E-mail address: `david.martin@icmat.es`