IMA Journal of Numerical Analysis (2012) **32** 1194–1216
doi:10.1093/imanum/drr042
Advance Access publication on November 15, 2011

Prolongation–collocation variational integrators

MELVIN LEOK AND TATIANA SHINGEL*
Department of Mathematics, University of California, San Diego, CA, USA
mleok@math.ucsd.edu *Corresponding author: tshingel@math.ucsd.edu

[Received on 10 January 2011; revised on 26 August 2011]

We introduce a novel technique for constructing higher-order variational integrators for Hamiltonian systems of ordinary differential equations. In the construction of the discrete Lagrangian we adopt Hermite interpolation polynomials and the Euler–Maclaurin quadrature formula and apply collocation to the Euler–Lagrange equation and its prolongation. Considerable attention is devoted to the order analysis of the resulting variational integrators in terms of approximation properties of the Hermite polynomials and quadrature errors. In particular, the order of the variational integrator can be computed *a priori* based on the quadrature error estimate. The analysis in the paper is straightforward compared to the order theory for Runge–Kutta methods. Finally, a performance comparison is presented on a selection of these integrators.

Keywords: Hamiltonian ODEs; Euler–Lagrange equations; Lagrangian variational integrators; variational order.

1. Introduction

One of the major themes in geometric integration is symplectic methods for solving Hamiltonian systems of ordinary differential equations (ODEs). Viewed as maps, such methods preserve the symplectic two-form underlying the dynamical evolution of the system. Variational integrators are an important class of symplectic integrators, which arise from discretizing Hamilton's principle. We refer the reader to Marsden & West (2001) for a detailed discussion of the background theory for such methods. Variational integrators are automatically symplectic and momentum-preserving. Moreover, they exhibit good energy behaviour for exponentially long times.

The construction of variational integrators combines techniques from approximation theory and numerical quadrature and is related to the Galerkin approach of converting a differential operator equation into a discrete system. Galerkin methods are semianalytic in the sense that the discrete solution is described by an element of a finite-dimensional function space, which provides an analytic expression for the numerical solution. Our goal is to develop a new class of variational integrators based on approximation spaces that are more directly amenable to variational order error analysis, and thereby obtain systematic proofs of their order of accuracy properties. To pursue this goal, we adopt the space of piecewise Hermite polynomials in the Galerkin construction. It is a well-known result in approximation theory that Hermite polynomials allow for higher-order approximation of smooth functions at relatively low cost. Recall that Lagrange polynomials are a common choice in the construction of higher-order variational integrators. However, the computation of an interpolating polynomial in Lagrange form becomes unstable when the degree of the polynomial is high. Unlike Lagrange polynomials, higher-degree Hermite polynomials produce accurate and computationally stable results.

Furthermore, the Galerkin construction based on piecewise Hermite polynomial interpolation allows us to adequately address the question of variational order error analysis. The order error analysis of

variational integrators relies on determining the order with which a discrete Lagrangian $L_{\mathrm{d}}\colon Q \times Q \to \mathbb{R}$ approximates the *exact discrete Lagrangian*,

$$L_{\mathrm{d}}^{E}(q_0, q_1; h) = \int_0^h L(q_{01}(t), \dot{q}_{01}(t))\,\mathrm{d}t, \tag{1.1}$$

where $q_{01}(t)$ is a solution curve of the Euler–Lagrange equation that satisfies the boundary conditions $q_{01}(0) = q_0$, $q_{01}(h) = q_1$. Alternatively, one can characterize the exact discrete Lagrangian in the following variational way:

$$L_{\mathrm{d}}^{E}(q_0, q_1; h) = \operatorname{ext}_{\substack{q\in C^2([0,h],Q)\\ q(0)=q_0, q(h)=q_1}} \int_0^h L(q(t), \dot{q}(t))\,\mathrm{d}t. \tag{1.2}$$

The standard way of performing the error analysis is by comparing the Taylor expansions of the exact discrete Lagrangian and the discrete Lagrangian. Consequently, the result critically depends on the extent to which the discrete trajectory is able to approximate the higher derivatives of the exact Euler–Lagrange solution curve. To enable a robust variational order error analysis for Galerkin variational integrators, we propose a novel application of the collocation approach in the setting of discrete Lagrangian mechanics, which involves prolongations of the Euler–Lagrange vector field. The proposed approach has the advantage that one is able to prove *optimal* rates of convergence of the associated variational integrators as long as sufficiently accurate quadrature formulas are used. The proof crucially depends on *a priori* estimates of the higher-derivative approximation properties of prolongation–collocation curves.

The general approach of variational integrators is conceptually related to the generating function approach introduced by Feng (1986), except that the discrete Lagrangian, which is a Type I generating function, is viewed as an approximation to Jacobi's solution of the Hamilton–Jacobi equation. In contrast, traditional approaches seek to obtain numerical solutions of the Hamilton–Jacobi equation directly. As such, the variational integrator approach has computational advantages since approximating Jacobi's solution corresponds to the numerical solution of a two-point boundary value problem, which is less computationally involved than solving the Hamilton–Jacobi partial differential equation directly.

1.1 *Outline of the paper*

We present a brief review of discrete variational mechanics and variational integrators in Section 2 and the Euler–Maclaurin quadrature formula in Section 3. In Section 4 we introduce prolongation–collocation variational integrators (PCVIs) and perform variational order error analysis in Section 5. In Section 6 we present a few numerical examples and present some conclusions and future directions in Section 7.

2. Variational integrators

Let Q be the configuration manifold of a mechanical system, with generalized coordinates q. Consider the *Lagrangian* $L\colon TQ \to \mathbb{R}$, where TQ is the tangent bundle of the configuration space Q. The tangent bundle has local coordinates (q, v). Further, let $\mathcal{C}(Q) = \mathcal{C}([0, T], Q)$ denote the space of smooth trajectories $q\colon [0, T] \to Q$ in the configuration manifold Q. The *action integral* $S\colon \mathcal{C}(Q) \to \mathbb{R}$ is defined as follows:

$$S(q) = \int_0^T L(q(t), \dot{q}(t))\,\mathrm{d}t.$$

The variational principle known as *Hamilton's principle* states that

$$\delta S = 0 \quad \text{for } \delta q(0) = \delta q(T) = 0,$$

which yields the *Euler–Lagrange equations*

$$\frac{\partial L}{\partial q}(q, \dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q})\right) = 0. \tag{2.1}$$

It is possible to rewrite equations (2.1) in terms of the generalized coordinates and momenta (q, p) on the cotangent bundle T^*Q (*phase space*). For this, we introduce the *Legendre transformation* $\mathbb{F}L\colon TQ \to T^*Q$ (see Marsden & Ratiu, 1999, Section 7.2), defined by

$$\mathbb{F}L\colon (q, \dot{q}) \mapsto \left(q, \frac{\partial L}{\partial \dot{q}}\right).$$

The *Hamiltonian* $H\colon T^*Q \to \mathbb{R}$ is given by

$$H(q, p) = \operatorname*{ext}_{\dot{q}}[p \cdot \dot{q} - L(q, \dot{q})].$$

When the Lagrangian is hyper-regular, i.e., when the Legendre transformation is a global diffeomorphism, this reduces to the familiar expression for the Hamiltonian, $H(q, p) = p \cdot \dot{q} - L(q, \dot{q})|_{p=\frac{\partial L}{\partial \dot{q}}}$. One can show that equations (2.1) are equivalent to *Hamilton's equations* (Hairer *et al.*, 2006, Theorem 1.3, p. 182),

$$\dot{p} = -\frac{\partial H}{\partial q}(p, q), \quad \dot{q} = \frac{\partial H}{\partial p}(p, q). \tag{2.2}$$

In the case of variational integrators, instead of discretizing the Euler–Lagrange equations (2.1), one discretizes Hamilton's principle. That is, one discretizes the action by introducing a *discrete Lagrangian* and replaces the action integral by an *action sum* and applies the discrete version of Hamilton's variational principle. The major advantage of this approach is that the resulting numerical algorithm automatically preserves the symplectic structure of the underlying dynamical system.

The discrete Lagrangian $L_{\mathrm{d}}(q_0, q_1, h)$ is thought of as an approximation of the action integral along the curve segment between the points $q_0 \approx q(0)$ and $q_1 \approx q(h)$. Formally, this can be expressed as

$$L_{\mathrm{d}}(q_0, q_1, h) \approx \int_0^h L(q(t), \dot{q}(t))\, \mathrm{d}t. \tag{2.3}$$

In practice, we need to assume that h is sufficiently small for this construction to be well defined, and we will assume this throughout our discussion. We omit the h-dependence and simply write $L_{\mathrm{d}}(q_0, q_1)$ when it is not essential for the exposition of the material. Given the discrete sequence of times $\{t_k = hk | k = 0, \ldots, N\}$, $h = T/N$, a discrete curve in Q is denoted by $\{q_k\}_{k=0}^N$, where $q_k \approx q(t_k)$. The discrete action sum is a function that maps the discrete trajectories $\{q_k\}_{k=0}^N$ to $\mathbb{R}$ and is given by

$$S_{\mathrm{d}}\left(\{q_k\}_{k=0}^N\right) = \sum_{k=0}^{N-1} L_{\mathrm{d}}(q_k, q_{k+1}).$$

The *discrete Hamilton principle* requires the discrete action to be stationary with respect to variations vanishing at $k = 0$ and $k = N$. From this, one derives the discrete version of the Euler–Lagrange equations, which are known as the *discrete Euler–Lagrange equations*,

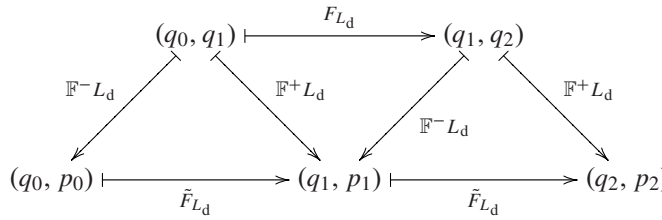
$$D_2 L_\mathrm{d}(q_{k-1}, q_k) + D_1 L_\mathrm{d}(q_k, q_{k+1}) = 0, \tag{2.4}$$

where $k = 1, 2, \ldots, N-1$. These equations implicitly define the one-step *discrete Lagrangian map* $F_{L_\mathrm{d}}\colon Q \times Q \to Q \times Q$. The discrete Legendre transforms $\mathbb{F}^{\pm} L_\mathrm{d}\colon Q \times Q \to T^*Q$ (see Marsden & West, 2001, Section 5.1) are defined by

$$\mathbb{F}^{-} L_\mathrm{d}\colon (q_0, q_1) \mapsto (q_0, -D_1 L_\mathrm{d}(q_0, q_1)),$$

$$\mathbb{F}^{+} L_\mathrm{d}\colon (q_0, q_1) \mapsto (q_1, D_2 L_\mathrm{d}(q_0, q_1)).$$

Pushing the discrete Lagrangian map F_{L_d} forward to T^*Q with the discrete Legendre transforms gives the *discrete Hamiltonian map* $\tilde{F}_{L_\mathrm{d}}\colon T^*Q \to T^*Q$ by $\tilde{F}_{L_\mathrm{d}} = \mathbb{F}^{\pm} L_\mathrm{d} \circ F_{L_\mathrm{d}} \circ (\mathbb{F}^{\pm} L_\mathrm{d})^{-1}$. The fact that the definitions of $\tilde{F}_{L_\mathrm{d}}$ are equivalent for the $+$ and $-$ case is implied by the following commutative diagram (Marsden & West, 2001, Theorem 1.5.2):



(2.5)

In coordinates, $\tilde{F}_{L_\mathrm{d}}\colon (q_0, p_0) \mapsto (q_1, p_1)$, where

$$p_0 = -D_1 L_\mathrm{d}(q_0, q_1), \quad p_1 = D_2 L_\mathrm{d}(q_0, q_1). \tag{2.6}$$

A numerical quadrature formula can be used to approximate the integral in equation (2.3). However, the functional form of the solution curve $q(t)$ is required when applying a quadrature rule and it is, in general, unknown. In practice, one can choose an interpolating function on the interval $[0, h]$ passing through q_0, q_1. Then, a quadrature rule can be applied to the integral of the Lagrangian evaluated along the interpolating function. This approach fits the general framework of Galerkin integration methods. In more detail, for the construction of Galerkin Lagrangian variational integrators, one replaces the path space $\mathcal{C}([0, T], Q)$, which is an infinite-dimensional function space, with a finite-dimensional function space, $\mathcal{C}^s([0, T], Q)$. Commonly, one uses polynomial approximations to the trajectories, letting

$$\mathcal{C}^s([0, h], Q) = \{q \in \mathcal{C}([0, h], Q) \mid q \text{ is a polynomial of degree } \leqslant s\}.$$

An approximate action $\mathcal{S}(q)\colon \mathcal{C}^s([0, h], Q) \to \mathbb{R}$ is

$$\mathcal{S}(q) = h \sum_{i=1}^{s} b_i L(q(c_i h), \dot{q}(c_i h)),$$

where $c_i \in [0, 1]$ are quadrature points, b_i are quadrature weights, $i = 1, \ldots, s$. The Galerkin discrete Lagrangian is

$$L_{\mathrm{d}}(q_0, q_1) = \operatorname*{ext}_{\substack{q \in \mathcal{C}^s([0,h], Q) \\ q(0)=q_0, q(h)=q_1}} \mathcal{S}(q). \tag{2.7}$$

In particular, for higher-order methods, one takes $q \in \mathcal{C}^s([0, h], Q)$ in the form

$$q\big(\tau h; q_0^{\nu}, h\big) = \sum_{\nu=0}^{s} q_0^{\nu} \tilde{l}_{\nu,s}(\tau),$$

where q_0^{ν}, $\nu = 1, \ldots, s-1$, are the *internal stages*, $\tilde{l}_{\nu,s}(\tau)$ are the *Lagrange basis polynomials* of degree s defined on the interval $[0, 1]$. Then, the integration scheme $(q_0, p_0) \mapsto (q_1, p_1)$ is given by

$$\begin{aligned}
-p_0 &= h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q}(c_i h) \tilde{l}_{0,s}(c_i) + \frac{1}{h} \frac{\partial L}{\partial \dot{q}}(c_i h) \dot{\tilde{l}}_{0,s}(c_i) \right], \\
0 &= h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q}(c_i h) \tilde{l}_{\nu,s}(c_i) + \frac{1}{h} \frac{\partial L}{\partial \dot{q}}(c_i h) \dot{\tilde{l}}_{\nu,s}(c_i) \right], \quad \nu = 1, \ldots, s-1, \\
p_1 &= h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q}(c_i h) \tilde{l}_{s,s}(c_i) + \frac{1}{h} \frac{\partial L}{\partial \dot{q}}(c_i h) \dot{\tilde{l}}_{s,s}(c_i) \right],
\end{aligned}$$

where we denoted $\frac{\partial L}{\partial q}(q(c_i h; q_0^{\nu}), \dot{q}(c_i h; q_0^{\nu}))$ by $\frac{\partial L}{\partial q}(c_i h)$ and similarly for $\frac{\partial L}{\partial \dot{q}}(c_i h)$. It has been established (see Marsden & West, 2001; Hairer *et al.*, 2006) that the above scheme is equivalent to a symplectic partitioned Runge–Kutta method.

There are several well-known classes of symplectic integrators that can be extended to arbitrarily high order. These include the Gauss methods, and the Lobatto IIIA–IIIB partitioned Runge–Kutta methods, which can all be viewed as collocation methods (Hairer *et al.*, 2006). In all these cases higher-order accuracy is achieved by evaluating the vector field at additional internal stages. In contrast, we will consider collocation methods with only the nodal times as collocation points for which higher-order accuracy is achieved by considering higher derivatives of the vector field. We will introduce higher-order variational integrators that have as their natural variables the position and its derivatives at only the nodal times, without the use of internal stages. This construction leads to a new class of variational integrators, named *PCVIs*. In particular, one constructs a computable discrete Lagrangian by using a Hermite–Obreschkoff method (see Hairer *et al.*, 1993, Section II.13) to obtain a numerical solution of the Euler–Lagrange boundary value problem and then applying the Euler–Maclaurin quadrature formula to approximate the action integral.

3. Quadrature

Here, we present a quadrature formula that we will use in our construction of a high-order discrete Lagrangian. The advantage of this particular rule is that it only involves function evaluations at the end points of the interval. When fast adaptive treecodes are used in conjunction with automatic differentiation techniques (Rall, 1981), it is more efficient to obtain higher-order approximations using higher-derivative information at the end points rather than evaluating the integrand at a number of internal stages.

THEOREM 3.1 (Euler–Maclaurin quadrature formula; Abramowitz & Stegun, 1972). If f is sufficiently differentiable on (a, b), then for any $m > 0$,

$$\int_a^b f(x)\,\mathrm{d}x = \frac{\theta}{2}\left[f(a) + 2\sum_{k=1}^{N-1} f(a+k\theta) + f(b)\right] - \sum_{l=1}^{m} \frac{B_{2l}}{(2l)!}\theta^{2l}\left(f^{(2l-1)}(b) - f^{(2l-1)}(a)\right)$$
$$- \frac{B_{2m+2}}{(2m+2)!}N\theta^{2m+3} f^{(2m+2)}(\xi),$$

where B_k are the Bernoulli numbers, $\theta = (b-a)/N$ and $\xi \in (a, b)$.

Let us apply Theorem 3.1 to approximate an integral $\int_0^h f(x)\,\mathrm{d}x$ in the simplest case when $N = 1$. It is easy to see that we obtain the quadrature rule

$$I(f) = \frac{h}{2}\left[f(0) + f(h)\right] - \sum_{l=1}^{m} \frac{B_{2l}}{(2l)!}h^{2l}\left(f^{(2l-1)}(h) - f^{(2l-1)}(0)\right), \tag{3.1}$$

where $I(f)$ approximates $\int_0^h f(x)\mathrm{d}x$ with an $\mathcal{O}(h^{2m+3})$ error.

4. The prolongation–collocation method

In this section we explain the construction of the discrete Lagrangian based on Hermite interpolation and the Euler–Maclaurin quadrature formula.

4.1 *Motivation for the prolongation–collocation approach*

The variational characterization of the exact discrete Lagrangian (1.2) naturally leads to the variational Galerkin discrete Lagrangian (2.7), where the infinite-dimensional function space $\mathcal{C}([0, h], Q)$ is replaced by a finite-dimensional subspace, and the integral is approximated by a quadrature formula. While this leads to a computable discrete Lagrangian, one does not necessarily obtain an optimally accurate discrete Lagrangian whose variational order is related to the best approximation properties of the chosen finite-dimensional function space.

In retrospect, the fact that the variational Galerkin approach does not readily lead to computable discrete Lagrangians with provable approximation properties is not too surprising. By construction, variational Galerkin discrete Lagrangians associated with a sequence of finite-dimensional function spaces involve extremizers of a sequence of functionals. Since the sequence of finite-dimensional function spaces converges to $\mathcal{C}([0, h], Q)$, the sequence of functionals converges to the functional that appears in the variational characterization of the exact discrete Lagrangian. However, it is unclear that the sequence of extremizers converges to the extremizer of the limiting functional since that corresponds to Γ-convergence of the sequence of functionals (Maso, 1993). The Γ-convergence of variational integrators was studied in Müller & Ortiz (2004). The issue of optimal rates of convergence of the computable discrete Lagrangians involves establishing rates of convergence of extremizers in terms of approximation rates of the finite-dimensional function spaces, which is an even more complicated process and to our knowledge has not been well understood.

As an alternative, we adopt the characterization of the exact discrete Lagrangian in terms of the Euler–Lagrange solution curve (1.1) and construct a discrete curve that approximates higher derivatives

of the Euler–Lagrange solution curve to an adequate level of accuracy. The latter is explored in detail in Section 5.

4.2 *Hermite interpolation and prolongation–collocation*

We commence by replacing $q(t)$ in (2.3) by its Hermite interpolant which is obtained by constructing a polynomial $q_d(t)$ such that values of $q(t)$ and any number of its derivatives at given points are fitted by the corresponding function values and derivatives of $q_d(t)$. In this paper we are concerned with fitting function values of $q(t)$ and its derivatives at the end points of the interval $[0, h]$. Consequently, a so-called two-point Hermite interpolant $q_d(t)$ of degree $d = 2n - 1$ can be used, which has the form

$$q_d(t) = \sum_{j=0}^{n-1} \left(q^{(j)}(0) H_{n,j}(t) + (-1)^j q^{(j)}(h) H_{n,j}(h-t) \right), \tag{4.1}$$

where

$$H_{n,j}(t) = \frac{t^j}{j!}(1 - t/h)^n \sum_{s=0}^{n-j-1} \binom{n+s-1}{s} (t/h)^s \tag{4.2}$$

are the Hermite basis functions. Note that for $n = 1$, the interpolant is a straight line joining $q(0)$ and $q(h)$. By choosing one of the simple quadrature rules to discretize the integral in (2.3) (e.g., the midpoint rule or trapezoidal rule), one obtains a class of well-known integrators that are at most second order (see Marsden & West, 2001). Therefore, the first nontrivial case of interest is $n = 2$, where we assume that the position and velocity data at the end points are available. From now on we only consider $n \geqslant 2$ when applying the Hermite interpolation formula. The detailed derivation of (4.1) can be found, for example, in Davis (1963). By construction,

$$q_d^{(r)}(0) = q^{(r)}(0), \quad q_d^{(r)}(h) = q^{(r)}(h), \quad r = 0, 1, \ldots, n-1.$$

Except for the step size h, the discrete Lagrangian $L_{\mathrm{d}}(q_0, q_1, h)$ should only depend on $q_0 \approx q(0)$, $q_1 \approx q(h)$. Therefore, letting $q_d(0) = q_0$ and $q_d(h) = q_1$, we need to approximate the higher-order derivatives of $q(t)$ by expressions that only depend on q_0, q_1. One natural approach, which is often found in the literature, is to use finite differences. In this work we propose to apply the idea of *collocation* in conjunction with the Euler–Lagrange equations (2.1). The benefits of this approach will be exemplified later when discussing the variational error analysis of the proposed class of numerical integrators (see Section 5).

The collocation approach (Hairer *et al.*, 1993) is well known in the theory of initial and boundary value problems for ODEs (Costabile & Napoli, 2007; Iserles, 2009). Roughly speaking, the technique consists of determining the unknown parameters of a parameterized curve by requiring $q_d(t)$ to satisfy the ODE at a given set of points (collocation points). To define $q_d(t)$ uniquely, one sets the number of collocation points to be equal to the number of the available degrees of freedom. In our approach we use the method of collocation in a slightly unusual manner. In particular, since the parameters in (4.1) correspond to the derivatives of the solution curve $q(t)$ at the end points of the interval $[0, h]$, we are going to use $t = 0$ and $t = h$ as collocation points for the Euler–Lagrange equations (2.1) and consider the *prolongation* (Olver, 1993) of the Euler–Lagrange equations in order to generate a sufficient number

of conditions. In other words, we increase the number of equations under consideration (not the number of collocation points) to match the number of degrees of freedom.

For example, consider the case of the quintic Hermite interpolation, i.e., set $n = 3$ in (4.1). For separable Lagrangians of the form $L(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - V(q)$, where m is the mass and $V(q)$ is the potential energy term, the Euler–Lagrange equations (2.1) become a second-order ODE of the form

$$\ddot{q}(t) = f(q(t)),$$

and its first-order prolongation can be expressed as

$$q^{(3)}(t) = f'(q(t))\dot{q}(t).$$

We set the boundary conditions $q_d(0) = q_0$ and $q_d(h) = q_1$ and the collocation conditions

$$\ddot{q}_d(0) = f(q_d(0)), \quad \ddot{q}_d(h) = f(q_d(h)),$$

$$q_d^{(3)}(0) = f'(q_d(0))\dot{q}_d(0), \quad q_d^{(3)}(h) = f'(q_d(h))\dot{q}_d(h).$$

The boundary and collocation conditions constitute the system of six equations which uniquely determines the fifth-degree polynomial $q_d(t)$ in the form of (4.1).

In general, for the Hermite polynomial of degree $2n - 1$, one would need to differentiate the Euler–Lagrange equation $n - 2$ times, thus deriving a system of $n - 1$ equations for $\ddot{q}_d(t), q_d^{(3)}(t), \ldots, q_d^{(n)}(t)$. Evaluated at 0 and h, these give $2n - 2$ collocation equations, which together with boundary conditions ($q_d(0) = q_0, q_d(h) = q_1$) constitute a sufficient number of conditions to determine the interpolant $q_d(t)$ uniquely. Note that for large n, the system of collocation conditions becomes nonlinear. Since the second- and higher-order derivatives $q_d^{(j)}(0), q_d^{(j)}(h)$ are given explicitly, the system can be recursively reduced to two implicit equations involving $\dot{q}_d(0), \dot{q}_d(h)$. In this case one would need to make use of a nonlinear root solver, such as the Newton–Raphson method, to determine $\dot{q}_d(0), \dot{q}_d(h)$.

More generally, in the case of nonseparable Lagrangians, one repeatedly implicitly differentiates the Euler–Lagrange equations to obtain a system of nonlinear equations that have to be solved for the prolongation. This is a larger system of equations than the one for separable Lagrangians, and the increased computational effort in solving this system of equations can be significant, particularly as the order of the prolongation–collocation variational integrator is increased. For simplicity of exposition and implementation, we shall restrict ourselves to the case of separable Lagrangians for the remainder of the paper.

Further, in order to discretize the integral in (2.3), we apply the Euler–Maclaurin quadrature formula (3.1). Recall that the formula involves derivatives of the integrand, in our case the Lagrangian L, with respect to the independent variable evaluated at the end points of the integration interval. The latter, however, does not require the extensive computations typically associated with the interpolating polynomial due to the use of the Hermite interpolation formula and the collocation idea explained above. In more detail, we write

$$\int_0^h L(q_d(t), \dot{q}_d(t))\,\mathrm{d}t \approx \frac{h}{2}(L(q_d(0), \dot{q}_d(0)) + L(q_d(h), \dot{q}_d(h))) - \sum_{l=1}^{m} \frac{B_{2l}}{(2l)!} h^{2l} \left(\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q_d(t), \dot{q}_d(t)) \Big|_{t=h} - \frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q_d(t), \dot{q}_d(t)) \Big|_{t=0} \right).$$

Provided that the degree of the interpolating polynomial is $2n-1$, we choose $m \leqslant \lfloor n/2 \rfloor$, where the brackets denote the greatest integer lower bound for $n/2$. So for even n, $\lfloor n/2 \rfloor = n/2$ and for odd n, $\lfloor n/2 \rfloor = (n-1)/2$. This choice of m is justified by observing that the expressions for

$$\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q_d(t), \dot{q}_d(t))\bigg|_{t=\tau}, \quad l = 1, 2, \ldots, m, \qquad \tau = 0, h,$$

include the derivatives of $q_d(t)$ up to order $2\lfloor n/2 \rfloor - 1 + 1 \leqslant n$, which satisfy the corresponding collocation conditions.

4.3 *Prolongation–collocation discrete Lagrangian*

The prolongation–collocation discrete Lagrangian is defined as

$$L_{\mathrm{d}}(q_0, q_1, h) = \frac{h}{2}(L(q_d(0), \dot{q}_d(0)) + L(q_d(h), \dot{q}_d(h))) - \sum_{l=1}^{\lfloor n/2 \rfloor} \frac{B_{2l}}{(2l)!} h^{2l} \left(\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q_d(t), \dot{q}_d(t))\bigg|_{t=h} - \frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q_d(t), \dot{q}_d(t))\bigg|_{t=0} \right), \quad (4.3)$$

where $q_d(t) \in \mathcal{C}^s(Q)$ is determined by the boundary and prolongation–collocation conditions,

$$\begin{aligned} &q_d(0) = q_0, \quad q_d(h) = q_1, \\ &\ddot{q}_d(0) = f(q_0), \quad \ddot{q}_d(h) = f(q_1), \\ &q_d^{(3)}(0) = f'(q_0)\dot{q}_d(0), \quad q_d^{(3)}(h) = f'(q_1)\dot{q}_d(h), \\ &\qquad\vdots \qquad\qquad\qquad \vdots \\ &q_d^{(n)}(0) = \frac{\mathrm{d}^{n-2}}{\mathrm{d}t^{n-2}} f(q_d(t))\bigg|_{t=0}, \quad q_d^{(n)}(h) = \frac{\mathrm{d}^{n-2}}{\mathrm{d}t^{n-2}} f(q_d(t))\bigg|_{t=h}. \end{aligned} \quad (4.4)$$

One can include fewer than $\lfloor n/2 \rfloor$ terms in the summation in (4.3). However, this will have an impact on the variational order of the corresponding integrator as will be further discussed in Section 5. The system of equations (4.4) completely defines the discrete Lagrangian (4.3). Note that we used the second-order ODE, $\ddot{q}(t) = f(q(t))$, as a prototype of the Euler–Lagrange equations for simplicity of notation only. The same idea applies for any smooth, not necessarily separable, Lagrangian function and the corresponding Euler–Lagrange equations.

Given the initial conditions (q_0, p_0), the variational integrator has the form

$$\begin{aligned} p_k &= -D_1 L_{\mathrm{d}}(q_k, q_{k+1}), \\ p_{k+1} &= D_2 L_{\mathrm{d}}(q_k, q_{k+1}), \quad k = 0, 1, \ldots, \end{aligned} \quad (4.5)$$

and defines a one-step method $(q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$. Generally, the equation $p_k = -D_1 L_{\mathrm{d}}(q_k, q_{k+1})$ together with the system of collocation conditions (4.4) can be reduced to a system of implicit equations with respect to $q_{k+1}, \dot{q}_d(t_k), \dot{q}_d(t_{k+1})$. As soon as the solution is obtained via some appropriate nonlinear root-finding method, it is inserted into the equation $p_{k+1} = D_2 L_{\mathrm{d}}(q_k, q_{k+1})$.

When the Lagrangian has a relatively simple form it makes sense to compute the expression for the discrete Lagrangian (4.3) symbolically, which can be done using the symbolic module in MATLAB or symbolic software such as Mathematica or Maple. Having computed *a priori* closed-form expressions for the right-hand side in (4.5), makes the implementation of the integrator particularly simple and fast. The reader is referred to Section 6 for some numerical examples.

4.4 *Self-adjointness of the discrete Lagrangian*

Given a one-step method, $\Phi_h\colon T^*Q \to T^*Q$, the *adjoint method* is $\Phi_h^*\colon T^*Q \to T^*Q$, which is defined by the property that

$$\Phi_h^* \circ \Phi_{-h} = \mathrm{Id}_{T^*Q},$$

that is to say that $\Phi_h^* = \Phi_{-h}^{-1}$. In the setting of discrete mechanics, it turns out that the adjoint of a discrete Hamiltonian map can be expressed in terms of the discrete Hamiltonian map associated with an adjoint discrete Lagrangian (see Marsden & West, 2001, Section 2.4). More explicitly, $\tilde{F}_{L_\mathrm{d}}^* = \tilde{F}_{L_\mathrm{d}^*}$, where $L_\mathrm{d}^*(q_0, q_1, h) = -L_\mathrm{d}(q_1, q_0, -h)$. In the following theorem, we prove that the prolongation–collocation discrete Lagrangian is self-adjoint. One consequence of self-adjointness is that the method has an even order of accuracy, and as we will see, Theorem 4.1 is compatible with the variational error analysis of Theorem 5.3.

THEOREM 4.1 If the Lagrangian L is time independent, then the prolongation–collocation discrete Lagrangian defined by (4.3) and (4.4) is self-adjoint.

Proof. Recall that the adjoint of a discrete Lagrangian is given by

$$L_\mathrm{d}^*(q_0, q_1, h) = -L_\mathrm{d}(q_1, q_0, -h).$$

Note that the prolongation–collocation discrete Lagrangian involves evaluating a quadrature approximation of the action associated with a polynomial curve $q_d(\cdot)$ that satisfies the prolongation–collocation conditions (4.4). To obtain an expression for the adjoint discrete Lagrangian, we introduce a polynomial curve $r_d(\cdot)$ that satisfies the prolongation–collocation conditions, but where the roles of q_0 and q_1 are reversed and $h \mapsto -h$. In particular, $r_d(\cdot)$ satisfies

$$\begin{aligned}
r_d(0) &= q_1, \quad r_d(-h) = q_0,\\
\ddot{r}_d(0) &= f(q_1), \quad \ddot{r}_d(-h) = f(q_0),\\
r_d^{(3)}(0) &= f'(q_1)\dot{r}_d(0), \quad r_d^{(3)}(-h) = f'(q_0)\dot{r}_d(-h),\\
&\vdots \qquad\qquad\qquad\quad \vdots\\
r_d^{(n)}(0) &= \frac{\mathrm{d}^{n-2}}{\mathrm{d}t^{n-2}} f(r_d(t))\bigg|_{t=0}, \quad r_d^{(n)}(-h) = \frac{\mathrm{d}^{n-2}}{\mathrm{d}t^{n-2}} f(r_d(t))\bigg|_{t=-h}.
\end{aligned}$$

Clearly, these are the same boundary conditions as given in (4.4), except that it is imposed on the interval $[0, h]$ for $q_d(\cdot)$ and the interval $[-h, 0]$ for $r_d(\cdot)$. This implies that the two polynomial curves $q_d(\cdot)$ and $r_d(\cdot)$ are related by a time-shift

$$q_d(t) = r_d(t - h),$$

which follows from the uniqueness of the two-point Hermite interpolating polynomial, and in the case of more general time-independent Lagrangians, on the time invariance of the Euler–Lagrange vector field. Based on this, we now compute the adjoint of the prolongation–collocation discrete Lagrangian,

$$\begin{aligned}
L_{\mathrm{d}}^{*}(q_0, q_1, h) &= -L_{\mathrm{d}}(q_1, q_0, -h) \\
&= -\frac{(-h)}{2}(L(r_d(0), \dot{r}_d(0)) + L(r_d(-h), \dot{r}_d(-h))) \\
&\quad + \sum_{l=1}^{\lfloor n/2 \rfloor} \frac{B_{2l}}{(2l)!}(-h)^{2l}\left(\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}}L(r_d(t), \dot{r}_d(t))\bigg|_{t=-h} - \frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}}L(r_d(t), \dot{r}_d(t))\bigg|_{t=0}\right) \\
&= \frac{h}{2}(L(q_d(h), \dot{q}_d(h)) + L(q_d(0), \dot{q}_d(0))) \\
&\quad + \sum_{l=1}^{\lfloor n/2 \rfloor} \frac{B_{2l}}{(2l)!}h^{2l}\left(\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}}L(q_d(t), \dot{q}_d(t))\bigg|_{t=0} - \frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}}L(q_d(t), \dot{q}_d(t))\bigg|_{t=h}\right) \\
&= \frac{h}{2}(L(q_d(0), \dot{q}_d(0)) + L(q_d(h), \dot{q}_d(h))) \\
&\quad - \sum_{l=1}^{\lfloor n/2 \rfloor} \frac{B_{2l}}{(2l)!}h^{2l}\left(\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}}L(q_d(t), \dot{q}_d(t))\bigg|_{t=h} - \frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}}L(q_d(t), \dot{q}_d(t))\bigg|_{t=0}\right) \\
&= L_{\mathrm{d}}(q_0, q_1, h).
\end{aligned}$$

Hence, the prolongation–collocation discrete Lagrangian is self-adjoint, and by Marsden & West (2001, Theorem 2.4.1), of the associated discrete Hamiltonian map is symmetric. □

5. Variational-order calculation

The construction of variational integrators naturally leads to the question of how it can be reconciled with the results from the approximation theory of function spaces and the numerical analysis of quadrature schemes. In particular, our goal is to explore the way that the approximation errors enter into the calculation of the convergence order of the corresponding integrators. Variational error analysis provides the right framework to pursue this goal.

The variational error analysis introduced in Marsden & West (2001), and refined in Patrick & Cuell (2009), is based on the idea that rather than considering how closely the numerical trajectory matches the exact flow, one can consider how the discrete Lagrangian approximates the exact discrete Lagrangian (1.1) that generates the exact flow map of the Euler–Lagrange equations. In other words, we are looking at the approximation error in

$$L_{\mathrm{d}}(q(0), q(h)) \approx \underset{\substack{q \in \mathcal{C}([0,h], Q) \\ q(0)=q_0,\, q(h)=q_1}}{\operatorname{ext}} \int_0^h L(q(t), \dot{q}(t))\,\mathrm{d}t = L_{\mathrm{d}}^{E}(q(0), q(h), h).$$

We say that a given discrete Lagrangian is of *order r* if there exist an open subset $U_v \subset TQ$ with compact closure and constants C_v and $h_v > 0$ so that

$$\left\| L_{\mathrm{d}}(q(0), q(h), h) - L_{\mathrm{d}}^{E}(q(0), q(h), h) \right\| \leqslant C_v h^{r+1} \tag{5.1}$$

for all solutions $q(t)$ of the Euler–Lagrange equations with initial condition $(q(0), \dot{q}(0)) \in U_v$ and for all $h \leqslant h_v$. Marsden & West (2001, Theorem 2.3.1) prove the equivalence of (5.1) to

(i) the discrete Hamiltonian map $\tilde{F}_{L_{\mathrm{d}}}$ being of order r;
(ii) the discrete Legendre transforms $\mathbb{F}^{\pm} L_{\mathrm{d}}$ being of order r.

In particular, the discrete Hamiltonian map is of order r if

$$\left\| \tilde{F}_{L_{\mathrm{d}}}(q(0), p(0), h) - \tilde{F}_{L_{\mathrm{d}}^{E}}(q(0), p(0), h) \right\| \leqslant \tilde{C}_v h^{r+1} \tag{5.2}$$

for all solutions $(q(t), p(t))$ of Hamilton's equations with initial condition $(q(0), p(0)) \in U_w \subset T^*Q$ and for all $h \leqslant h_w$. The order of the discrete Legendre transforms is defined analogously. Recall from diagram (2.5) that

$$\tilde{F}_{L_{\mathrm{d}}} \colon (q_0, p_0) \mapsto (q_1, p_1).$$

By construction, $\tilde{F}_{L_{\mathrm{d}}^{E}}(q(0), q(h), h)$ produces the values $(q(h), p(h))$ corresponding to the exact solution of the Hamiltonian system (2.2), whereas $\tilde{F}_{L_{\mathrm{d}}}(q(0), q(h), h)$ produces the approximate values (q_1, p_1), $q_1 \approx q(h)$, $p_1 \approx p(h)$. To summarize, the estimate (5.2) provides the local order of convergence of the discrete trajectory (q_k, p_k) to the exact flow $(q(t), p(t))$ of the Hamiltonian vector field. This order is the same as the order to which the discrete Lagrangian approximates the exact discrete Lagrangian, which we focus on.

Before we explore the inequality (5.1) for the prolongation–collocation discrete Lagrangian discussed in Section 4, we would like to establish the approximation error of $q_d^{(j)}(t)$ in comparison to $q^{(j)}(t)$ for $j = 1, 2, \ldots, n$, where $q(t)$ is the exact solution of the Euler–Lagrange equation, and $q_d(t)$ is the Hermite interpolating polynomial (4.1) of degree $2n - 1$, constructed by letting $q_d(0) = q(0)$ and $q_d(h) = q(h)$, and imposing the prolongation–collocation conditions discussed in Section 4 at the end points. Note that this can be a difficult task in general since the complexity of the collocation procedure escalates with the degree of the Hermite polynomial. However, we are only interested in the approximation order at the end points of the interval $[0, h]$ in which case the analysis is straightforward.

LEMMA 5.1 For $q(t)$, $q_d(t)$ as above, if $\dot{q}_d(\tau) = \dot{q}(\tau) + \mathcal{O}(h^p)$ for some $p > 0$ and $\tau = 0, h$, then

$$q_d^{(j)}(\tau) = q^{(j)}(\tau) + \mathcal{O}(h^p), \quad j = 3, \ldots, n.$$

Proof. Note that $\ddot{q}(\tau)$ coincides with $\ddot{q}_d(\tau)$ by construction. Therefore, we consider j starting from 3. As before, we restrict the proof to the case of a separable Lagrangian. Indeed, since the Euler–Lagrange equation is equivalent to the second-order ODE

$$\ddot{q}(t) = f(q(t)), \tag{5.3}$$

it follows that for $\tau = 0, h$,

$$\begin{aligned} q_d^{(3)}(\tau) - q^{(3)}(\tau) &= f'(q_d(\tau))\dot{q}_d(\tau) - f'(q(\tau))\dot{q}(\tau) \\ &= f'(q(\tau))\dot{q}_d(\tau) - f'(q(\tau))\dot{q}(\tau) = f'(q(\tau))(\dot{q}_d(\tau) - \dot{q}(\tau)) = \mathcal{O}(h^p), \end{aligned}$$

provided there exists a uniform bound for f'. Consecutively differentiating (5.3) and substituting the corresponding expressions for lower-order derivatives, one can see that $q^{(j)}(\tau)$ (and $q_d^{(j)}(\tau)$) can be represented as a polynomial in powers of $\dot{q}(\tau)$ (and $q_d(\tau)$, respectively) with coefficients which only depend on $q(\tau) = q_d(\tau)$,

$$q^{(4)}(\tau) = f''(q(\tau))\dot{q}(\tau)^2 + f'(q(\tau))f(q(\tau)),$$

$$q^{(5)}(\tau) = f^{(3)}(q(\tau))\dot{q}(\tau)^3 + \dot{q}(\tau)\Big(3f''(q(\tau))f(q(\tau)) + f'(q(\tau))^2\Big),$$

$$\vdots$$

As soon as f has bounded higher-order derivatives, the above formulas imply that the order of the approximation of $q^{(j)}(\tau)$ by $q_d^{(j)}(\tau)$ as a function of h is equal to the order of the approximation of $\dot{q}(\tau)$ by $\dot{q}_d(\tau)$. We consider these expressions up to the $j = n$ case, where n is determined by the number of collocation equations that were used to compute q_d. □

In the next lemma, we determine the value of p in the relation $\dot{q}_d(\tau) = \dot{q}(\tau) + \mathcal{O}(h^p)$ in terms of the degree of the polynomial $q_d(t)$.

LEMMA 5.2 Consider a polynomial $q_d(t)$ of degree $d = 2n - 1$ given by the formula

$$q_d(t) = \sum_{j=0}^{n-1} (a_{0j} H_{n,j}(t) + (-1)^j a_{1j} H_{n,j}(h - t)),$$

where $H_{n,j}(t)$ are the basis polynomial functions (4.2). By construction,

$$q_d^{(j)}(0) = a_{0j}, \quad q_d^{(j)}(h) = a_{1j}, \qquad j = 0, \ldots, n-1.$$

We let $a_{00} = q(0)$, $a_{10} = q(h)$. The coefficients a_{0j}, a_{1j}, $j = 1, 2, \ldots, n-1$, are obtained from the system of equations consisting of the Euler–Lagrange equation (2.1) and its prolongations. In particular, these are $n - 1$ differential equations evaluated on $q_d(t)$ at $t = 0$ and $t = h$. Then for $\tau = 0, h$,

$$\dot{q}_d(\tau) = \dot{q}(\tau) + \mathcal{O}(h^{2n-1}).$$

Proof. Let $q(t) \in \mathcal{C}^{2n}([0, h], Q)$ be the solution of the Euler–Lagrange equation (2.1) with boundary conditions $q(0)$ and $q(h)$. Then, $\ddot{q}(t)$ can be written in the form

$$\ddot{q}(t) = P_{2n-3}[\ddot{q}](t) + R_{n-1}[\ddot{q}](t),$$

where

$$P_{2n-3}[\ddot{q}](t) = \sum_{j=0}^{n-2} \Big(q^{(j+2)}(0) H_{n-1,j}(t) + (-1)^j q^{(j+2)}(h) H_{n-1,j}(h - t)\Big),$$

$$R_{n-1}[\ddot{q}](t) = \frac{q^{2n+2}(\xi)}{(2n-2)!} t^{n-1}(h - t)^{n-1}, \quad \xi \in (0, h),$$

and $H_{n-1}(t)$ are the basis polynomials defined by (4.2). See Davis (1963) for the proof of the above formula for sufficiently smooth functions. Note that $\ddot{q}_d(t)$ is a polynomial of degree $2n - 3$ to which the

same formula can be applied. In the latter case the remainder term is identically zero. Hence,

$$\ddot{q}_d(t) = \sum_{j=0}^{n-2} \big(q_d^{(j+2)}(0) H_{n-1,j}(t) + (-1)^j q_d^{(j+2)}(h) H_{n-1,j}(h-t)\big).$$

Subtracting $\ddot{q}_d(t)$ from $\ddot{q}(t)$ gives

$$\begin{aligned}\ddot{q}(t) - \ddot{q}_d(t) = \sum_{j=1}^{n-2} \big((q^{(j+2)}(0) - q_d^{(j+2)}(0)) H_{n-1,j}(t) + (-1)^j (q^{(j+2)}(h) \\ - q_d^{(j+2)}(h)) H_{n-1,j}(h-t)\big) + R_{n-1}[\ddot{q}](t). \end{aligned} \tag{5.4}$$

Next, we integrate the above expression from 0 to h. The left-hand side of the equation becomes

$$\int_0^h [\ddot{q}(t) - \ddot{q}_d(t)]\,\mathrm{d}t = [\dot{q}(h) - \dot{q}_d(h)] - [\dot{q}(0) - \dot{q}_d(0)].$$

Further, observe that

$$\int_0^h H_{n-1,j}(t)\,\mathrm{d}t = \int_0^h H_{n-1,j}(h-t)\,\mathrm{d}t = C_j h^{j+1}, \quad j = 1, \ldots, n-2,$$

$$\int_0^h t^{n-1}(h-t)^{n-1}\mathrm{d}t = C h^{2n-1},$$

where $C_j, C > 0$ are constants that depend on n but not on h. Now, let $\dot{q}_d(0) = \dot{q}(0) + \mathcal{O}(h^p)$ and $\dot{q}_d(h) = \dot{q}(h) + \mathcal{O}(h^p)$, where we wish to determine the value of p. It is easy to see that after integrating both sides of (5.4) and rewriting it in terms of the order conditions we arrive at

$$\mathcal{O}(h^p) = \mathcal{O}(h^{p+2}) + \mathcal{O}(h^{2n-1}).$$

Note that we used Lemma 5.1 to estimate the higher-order derivatives in (5.4). It follows immediately that $p = 2n - 1$ and the proof is finished. □

We are now ready to prove the main result of this section, which establishes the order of accuracy (as defined by (5.1)) of the prolongation–collocation discrete Lagrangian. Recall that the corresponding discrete Hamiltonian map $\tilde{F}_{L_\mathrm{d}}$ has the same order of accuracy (as defined by (5.2)) as the discrete Lagrangian L_d.

THEOREM 5.3 Assume that a Lagrangian function $L\colon TQ \to \mathbb{R}$ is sufficiently smooth and its partial derivatives are uniformly bounded. Then, the prolongation–collocation discrete Lagrangian L_d, defined by (4.3) and (4.4), and the corresponding discrete Hamiltonian map $\tilde{F}_{L_\mathrm{d}}$ has order of accuracy $2\lfloor n/2 \rfloor + 2$ for $n \geqslant 3$ and order of accuracy 2 for $n = 2$. In particular, for $n \geqslant 3$ and odd, the order is $n + 1$; for $n \geqslant 3$ and even, the order is $n + 2$.

Proof. Observe that the differentiability of the Lagrangian function L implies that it is Lipschitz continuous in each of its arguments, given that the partial derivatives are uniformly bounded. We will make use of both the differentiability and the Lipschitz continuity of L in the proof below.

We start with the simplest nontrivial case of $n = 2$, which corresponds to the space of piecewise cubic polynomials in the construction of a variational integrator. Note that it is sufficient to apply collocation to the Euler–Lagrange equation at the end points of the interval $[0, h]$ to uniquely define $q_d(t)$ satisfying the given boundary conditions. We obtain the order of approximation of the first derivative by applying Lemma 5.2. We use the simple trapezoidal rule

$$\int_a^{\mathrm{d}} f(t)\,\mathrm{d}t = \frac{b-a}{2}(f(a)+f(b)) + \mathcal{O}((b-a)^3)$$

to discretize the action integral in (2.3). Since $L(q, \dot{q})$ is Lipschitz continuous,

$$L(q_d(\tau), \dot{q}_d(\tau)) = L(q(\tau), \dot{q}(\tau)) + \mathcal{O}(h^3). \tag{5.5}$$

Therefore,

$$\begin{aligned} L_{\mathrm{d}}(q(0), q(h), h) &= \frac{h}{2}\Big(L(q_d(0), \dot{q}_d(0)) + L(q_d(h), \dot{q}_d(h))\Big) \\ &= \frac{h}{2}\Big(L(q(0), \dot{q}(0)) + L(q(h), \dot{q}(h))\Big) + \mathcal{O}(h^4) \\ &= L_{\mathrm{d}}^E(q(0), q(h)) + \mathcal{O}(h^3) + \mathcal{O}(h^4) \\ &= L_{\mathrm{d}}^E(q(0), q(h)) + \mathcal{O}(h^3). \end{aligned}$$

The combination of the trapezoidal rule and the cubic Hermite interpolation makes the analysis of the approximation of the exact discrete Lagrangian by the discrete Lagrangian elementary. As we can see, the error in approximation is determined by the error of the quadrature rule. This is due to the fact that the derivatives are approximated to sufficiently high order. The pattern persists for higher-order Hermite interpolants due to our choice of quadrature method.

It is straightforward to extend the above reasoning to the general case. Iteratively, applying Lemmas 5.1 and 5.2 to the derivatives of $L_{\mathrm{d}}(q_d(t), \dot{q}_d(t))$ with respect to time t gives the relation

$$\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q_d(t), \dot{q}_d(t))\bigg|_{t=0,h} = \frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q(t), \dot{q}(t))\bigg|_{t=0,h} + \mathcal{O}(h^{2n-1}),$$

which is analogous to (5.5). Hence,

$$\begin{aligned} L_{\mathrm{d}}(q(0), q(h), h) &= \frac{h}{2}\Big(L(q_d(0), \dot{q}_d(0)) + L(q_d(h), \dot{q}_d(h))\Big) \\ &\quad - \sum_{l=1}^{\lfloor n/2 \rfloor} \frac{B_{2l}}{(2l)!} h^{2l} \left(\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q_d(t), \dot{q}_d(t))\bigg|_{t=h} - \frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q_d(t), \dot{q}_d(t))\bigg|_{t=0} \right) \\ &= \frac{h}{2}\Big(L((0), \dot{q}(0)) + L(q(h), \dot{q}(h))\Big) + \mathcal{O}(h^{2n}) \\ &\quad - \sum_{l=1}^{\lfloor n/2 \rfloor} \frac{B_{2l}}{(2l)!} h^{2l} \left(\frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q(t), \dot{q}(t))\bigg|_{t=h} - \frac{\mathrm{d}^{2l-1}}{\mathrm{d}t^{2l-1}} L(q(t), \dot{q}(t))\bigg|_{t=0} + \mathcal{O}(h^{2n-1}) \right) \\ &= L_{\mathrm{d}}^E(q(0), q(h), h) + \mathcal{O}\big(h^{2\lfloor n/2 \rfloor + 3}\big) + \text{h.o.t.} = L_{\mathrm{d}}^E(q(0), q(h), h) + \mathcal{O}\big(h^{2\lfloor n/2 \rfloor + 3}\big). \end{aligned}$$

Thus, the prolongation–collocation discrete Lagrangian has order of accuracy $2\lfloor n/2 \rfloor + 2$ when $n \geqslant 3$, and by Marsden & West (2001, Theorem 2.3.1), the discrete Hamiltonian map has the same order of accuracy. □

Remarks. Several remarks are in order. As we have already noted, the error in approximation of the exact discrete Lagrangian by the prolongation–collocation discrete Lagrangian is determined by the order of the quadrature formula. In particular, the best order estimate is achieved if *all* the collocation conditions (4.4) are used. The collocation equations enter the terms under the summation in (4.3), which in turn determine the order of accuracy of the quadrature formula.

Secondly, if we write $n = 2k$, so that the degree of the interpolating polynomial is $d = 2n - 1 = 4k - 1$, the choices $n = 2k$ and $n = 2k + 1$ lead to the same maximal order of the quadrature, $2k + 2$. Therefore, it is preferable to use Hermite interpolating polynomials of degree $d = 4k - 1$, which minimizes the computational effort for a discrete Lagrangian for a given order of accuracy.

6. Examples

6.1 *Second-order PCVI*

Here, we derive the formulas for a simple case of the proposed variational integrator. In particular, we take the trapezoidal rule as the quadrature formula in the construction of the discrete Lagrangian that ensures second-order accuracy. Consider the Lagrangian $L(q, \dot{q}) = \frac{1}{2}\dot{q}^{\mathrm{T}} M \dot{q} - V(q)$, where M is a symmetric positive-definite mass matrix and V is a potential function. It is well known that the Euler–Lagrange equation for this Lagrangian system are

$$M\ddot{q} = -\nabla V(q), \tag{6.1}$$

which is Newton's equation of motion. We rewrite it as a system of first-order equations,

$$\dot{q} = v,$$

$$\dot{v} = -\nabla V(q).$$

We introduce discrete variables that denote the following approximations, $q_0 \approx q(0)$, $q_1 \approx q(h)$ and $v_0 \approx \dot{q}(0)$, $v_1 \approx \dot{q}(h)$. It is easy to show that the collocation conditions for the equation (6.1) applied to the Hermite cubic polynomial at the end points of the interval $[0, h]$ have the form

$$6\frac{q_1 - q_0}{h^2} - \frac{2}{h}(2v_0 + v_1) + M^{-1}\nabla V(q_0) = 0, \tag{6.2a}$$

$$6\frac{q_1 - q_0}{h^2} - \frac{2}{h}(v_0 + 2v_1) - M^{-1}\nabla V(q_1) = 0. \tag{6.2b}$$

Further, we apply the trapezoidal rule to construct the discrete Lagrangian. This will lead to a variational integrator which is second-order accurate. Henceforth,

$$L_{\mathrm{d}}(q_0, q_1) = \frac{h}{2}\left(\frac{1}{2}v_0^{\mathrm{T}} M v_0 - V(q_0) + \frac{1}{2}v_1^{\mathrm{T}} M v_1 - V(q_1)\right),$$

where v_0, v_1 are functions depending on q_0, q_1, as determined by the collocation conditions (6.2). The discrete Hamilton equations are

$$p_0 = -D_1 L_{\mathrm{d}}(q_0, q_1) = -\frac{h}{2}\left(\frac{\partial v_0}{\partial q_0} \cdot M v_0 - \nabla V(q_0) + \frac{\partial v_1}{\partial q_0} \cdot M v_1\right),$$

$$p_1 = D_2 L_{\mathrm{d}}(q_0, q_1) = \frac{h}{2}\left(\frac{\partial v_0}{\partial q_1} \cdot M v_0 + \frac{\partial v_1}{\partial q_1} \cdot M v_1 - \nabla V(q_1)\right),$$

where $\frac{\partial v_i}{\partial q_j}$ is in general a matrix of partial derivatives, and $\cdot$ denotes a matrix product. We differentiate equations (6.2) to find expressions for $\frac{\partial v_i}{\partial q_j}$ and substitute them into the discrete Hamilton equations. This, together with (6.2), leads to the system of four implicit equations with respect to the unknowns v_0, v_1, q_1 and p_1:

$$v_1 - v_0 + \frac{h}{2} M^{-1}(\nabla V(q_0) + \nabla V(q_1)) = 0, \tag{6.3a}$$

$$q_1 - q_0 - \frac{h}{2}(v_0 + v_1) + \frac{h^2}{12} M^{-1}(\nabla V(q_0) - \nabla V(q_1)) = 0, \tag{6.3b}$$

$$p_0 - \frac{1}{2} M(v_0 + v_1) - \frac{h}{2}\nabla V(q_0) + \frac{h^2}{12}\nabla(\nabla V(q_0))(2v_0 - v_1) = 0, \tag{6.3c}$$

$$p_1 - \frac{1}{2} M(v_0 + v_1) + \frac{h}{2}\nabla V(q_1) - \frac{h^2}{12}\nabla(\nabla V(q_1))(v_0 - 2v_1) = 0. \tag{6.3d}$$

In the above equations, $\nabla(\nabla V)$ denotes the derivative of ∇V, which is the Jacobian matrix for the gradient vector ∇V, or equivalently the Hessian of the potential V. We would like to point out that the above method does not belong to any existing families of symplectic integrators. In particular, it is not a partitioned Runge–Kutta method since it depends on higher derivatives of the potential V.

6.2 *Kepler problem*

We test the numerical method (6.3) on the Kepler problem, a detailed discussion of which can be found in Hairer *et al.* (2006, Section I.2). Recall that the Kepler two-body problem describes the motion of two bodies under mutual gravitational attraction. If one of the bodies is placed at the origin, the position (y_1, y_2) and the velocity (y_3, y_4) of the second body satisfy the following four-dimensional ODE:

$$\begin{aligned}
\dot{y}_1 &= y_3,\\
\dot{y}_2 &= y_4,\\
\dot{y}_3 &= -\frac{y_1}{(y_1^2 + y_2^2)^{3/2}},\\
\dot{y}_4 &= -\frac{y_2}{(y_1^2 + y_2^2)^{3/2}}.
\end{aligned} \tag{6.4}$$

The Lagrangian function is

$$L(\mathbf{y}) = \frac{1}{2}(y_3^2 + y_4^2) + \frac{1}{\sqrt{y_1^2 + y_2^2}}.$$

We use (6.3) to numerically solve (6.4) with initial conditions borrowed from Hairer *et al.* (2006, Section I.2.3):

$$y_1^0 = 1 - e, \quad y_2^0 = 0, \quad y_3^0 = 0, \quad y_4^0 = \sqrt{\frac{1+e}{1-e}},$$

where the eccentricity $e = 0.6$. With these initial conditions, the exact solution has period 2π. We can see from Fig. 1 that the second-order prolongation–collocation method (Fig. 1c) produces a more accurate elliptic orbit than the one generated by the second-order implicit midpoint rule (Fig. 1a) or the Störmer–Verlet method (SV; Fig. 1b), given the same step size h. Figure 1d shows the reference solution generated by the fourth-order symplectic Runge–Kutta method (Table 1a).

In Fig. 2 we plot the global error for the discrete Hamiltonian map for the methods MID, SV and PCVI. One can see that all three methods are second-order accurate, but that the PCVI method performs slightly better than MID and SV.

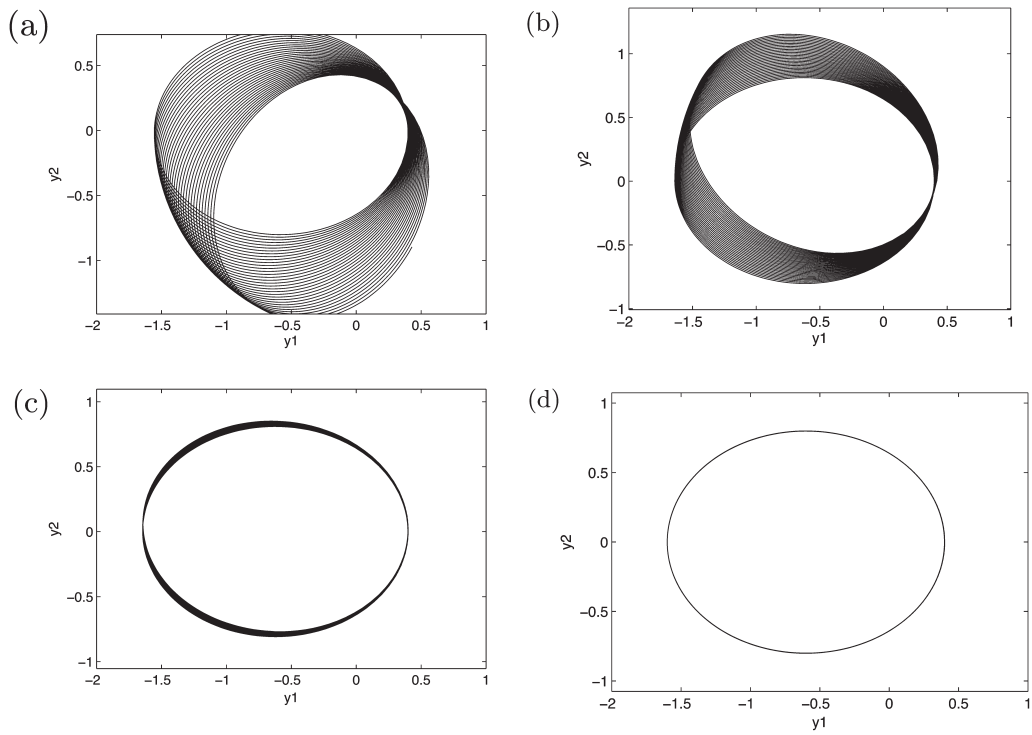


FIG. 1. The two-body Kepler problem. The step size $h = 0.05$. The first 4000 steps are shown. (a) The numerical solution using the implicit midpoint rule (MID). (b) The numerical solution using the SV method. (c) The numerical solution using the second-order PCVI. (d) The numerical solution using the fourth-order symplectic Runge–Kutta method.

TABLE 1. Butcher tableaus

$$\begin{array}{c|cc} \frac{1}{2}+\frac{1}{2\sqrt{3}} & \frac{1}{4} & \frac{1}{4}+\frac{1}{2\sqrt{3}} \\ \frac{1}{2}-\frac{1}{2\sqrt{3}} & \frac{1}{4}-\frac{1}{2\sqrt{3}} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

(a) SRK4

$$\begin{array}{c|ccc} \frac{1}{2}+\frac{1}{2}\sqrt{\frac{3}{5}} & \frac{5}{36} & \frac{2}{9}+\frac{1}{3}\sqrt{\frac{3}{5}} & \frac{5}{36}+\frac{1}{6}\sqrt{\frac{3}{5}} \\ \frac{1}{2} & \frac{5}{36}-\frac{5}{24}\sqrt{\frac{3}{5}} & \frac{2}{9} & \frac{5}{36}+\frac{5}{24}\sqrt{\frac{3}{5}} \\ \frac{1}{2}-\frac{1}{2}\sqrt{\frac{3}{5}} & \frac{5}{36}-\frac{1}{6}\sqrt{\frac{3}{5}} & \frac{2}{9}-\frac{1}{3}\sqrt{\frac{3}{5}} & \frac{5}{36} \\ \hline & \frac{5}{18} & \frac{4}{9} & \frac{5}{18} \end{array}$$

(b) SRK6

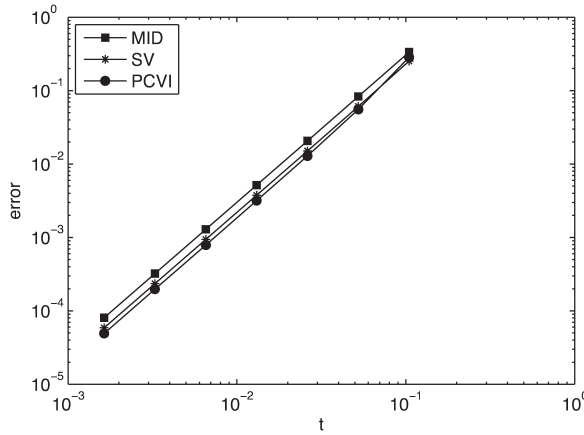


FIG. 2. The two-body Kepler problem. Global error for the discrete Hamiltonian map on $[0, \pi]$.

6.3 *Simple harmonic oscillator*

Next, we consider a harmonic oscillator system described by the equations

$$\dot{q} = p, \quad \dot{p} = -q.$$

The total energy of the system is given by the Hamiltonian $H(q, p) = \frac{1}{2}p^2 + \frac{1}{2}q^2$. For numerical tests (see Figs 3 and 4), we use the fourth- and sixth-order prolongation–collocation variational integrators (PCVI4 and PCVI6, respectively) and the symplectic Runge–Kutta methods of the corresponding order (SRK4 and SRK6, respectively). The Butcher tableaus for the Runge–Kutta methods are given in Table 1. As we can see from the plots, the standard symplectic Runge–Kutta method performs better in the sense that it encounters a slightly smaller global error. However, we would like to note the following interesting observation. Let us consider only the error in the position component, q_k, of the symplectic integrator $(q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$. It turns out that in the case of the fourth-order PCVI method, the position is globally sixth-order accurate. The error plots are given in Fig. 3b, where we can see that the position error for the symplectic Runge–Kutta method remains fourth order and is also larger in magnitude. This, however, does not contradict the variational order analysis discussed in Section 5. The theorem mentioned therein establishes the order of the integrator $(q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ in position–momentum variables but allows the integrator $(q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$ in position variables to have the

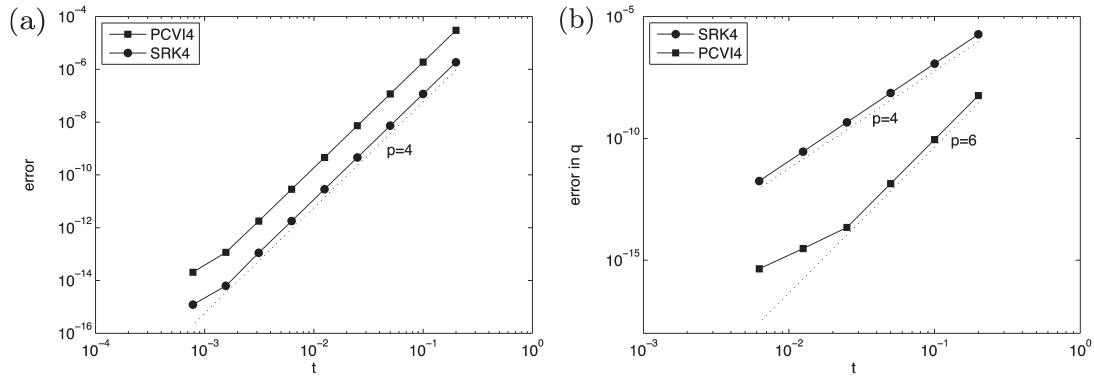


FIG. 3. Simple harmonic oscillator. (a) Global error for PCVI4 and SRK4. The dotted line is the reference line for the exact order. (b) Global error only for the position trajectory.

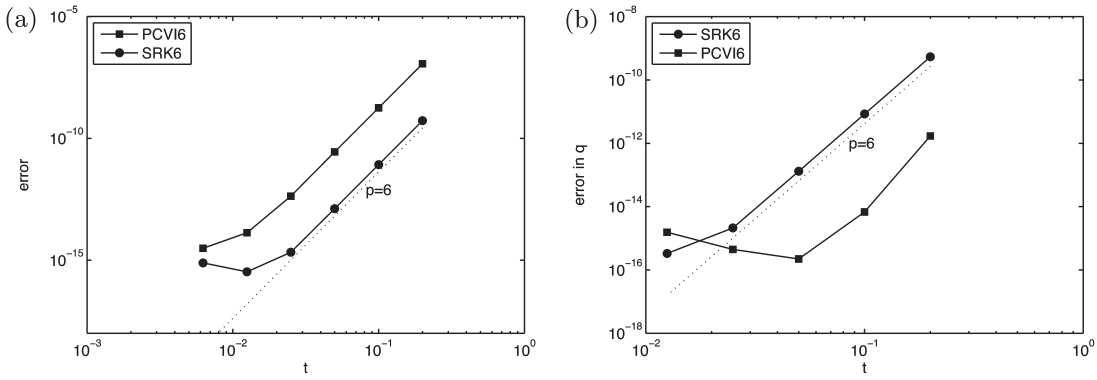


FIG. 4. Simple harmonic oscillator. (a) Global error for PCVI6 and SRK6. The dotted line is the reference line for the exact order. (b) Global error only for the position trajectory.

same or higher order. A similar pattern holds for the sixth-order PCVI (see Fig. 4b), though we do not supply a precise estimate for the order as the positional error quickly achieves machine precision, and a definitive statement would require the use of higher-precision arithmetic.

6.4 *Planar pendulum*

A planar pendulum of mass $m = 1$ with the massless rod of length $l = 1$ is a Hamiltonian system for which the equations of motion are

$$\dot{q} = p, \quad \dot{p} = -\sin q. \tag{6.5}$$

The Hamiltonian is $H(q, p) = \frac{1}{2}p^2 - \cos q$. Figure 5 displays the performance of the higher-order PCVIs constructed in Section 4 in conjunction with the pendulum equations (6.5). The global error for the corresponding numerical methods is presented in Fig. 5c and the energy error is given in Fig. 5(a and b). We see that the second-order prolongation–collocation method exhibits a relatively large energy error

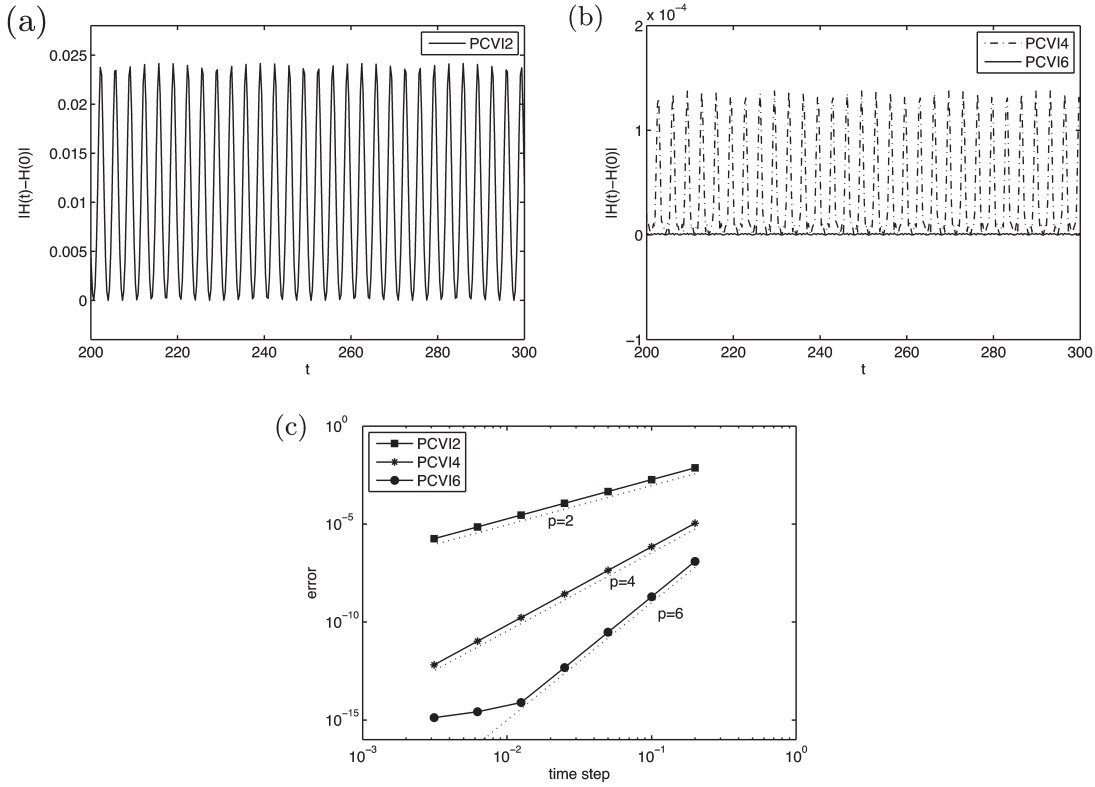


FIG. 5. Planar pendulum. (a) Energy error for PCVI2. (b) Energy error for PCVI4 and PCVI6. (c) Global error for PCVI2, PCVI4 and PCVI6. The dotted lines are the reference lines for the exact order.

and it drops dramatically for the fourth- and sixth-order methods. The energy error plots are generated using the step size $h = 0.3$.

REMARK 6.1 The codes for the numerical tests presented in this section are available for download.

7. Conclusions and future directions

In this paper we introduced a novel technique for constructing high-order variational integrators using collocation on the prolongation of the Euler–Lagrange equations. This leads to a family of arbitrarily high-order variational integrators with provable order-of-accuracy properties. The resulting methods are particularly appropriate in combination with digital feedback control since they naturally yield position and its derivatives as the output, without the need to use interpolation in order to access such data.

The resulting analytical solution, given by the piecewise defined two-point Hermite interpolants, is continuous across the nodal points only in position, and exhibits an $\mathcal{O}(h^{p+1})$ jump in velocity across nodal points, which is similar to what is achieved by other numerical methods. However, the prolongation–collocation method has the additional benefit that one also obtains a similar $\mathcal{O}(h^{p+1})$ control on the jumps in the higher-derivative terms $\ddot{q}, \ldots, q^{(n)}$, which is not generally true for other methods. It is in this sense that the PCVIs have better higher-order regularity across nodal points.

The proposed method can be viewed as a symplectic s-stage, q-derivative Runge–Kutta method (see Hairer *et al.*, 1993, Section II.13), where $s = 2$, $c_1 = 0$, $c_2 = 1$, which corresponds to the Hermite–Obreschkoff methods. It would be interesting to consider more general symplectic s-stage, q-derivative Runge–Kutta methods by using a similar construction. In this scenario, one would use a higher-derivative method to compute a numerical approximation to the Euler–Lagrange boundary value problem, and then consider a quadrature approximation of the action integral, in order to construct a computable discrete Lagrangian.

It would be desirable to explore the connection between the methods proposed in this paper and variational integrators based on global approximation techniques like splines and to extend this work to the setting of Lie groups by incorporating techniques from Lie group variational integrators (Leok, 2004) and constructive approximation techniques on Lie groups (Oswald & Shingel, 2009; Shingel, 2010). Furthermore, it would be interesting to extend the prolongation–collocation techniques to Hamiltonian variational integrators (Leok & Zhang, 2011) and Hamilton–Pontryagin variational integrators (Leok & Ohsawa, 2011) by considering prolongations of Hamilton's equations, and the implicit Euler–Lagrange equations.

Acknowledgements

We gratefully acknowledge the helpful comments and suggestions of the referees. This research was partially supported by National Science Foundation (DMS-1001521, CMMI-1029445, DMS-1065972) and NSF CAREER Award (DMS-1010687).

References

Abramowitz, M. & Stegun, I. A. (eds) (1972) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing. Dover: New York.

Costabile, F. & Napoli, A. (2007) Solving BVPs using two-point Taylor formula by a symbolic software. *J. Comput. Appl. Math.*, **210**, 136–148.

Davis, P. J. (1963) *Interpolation and Approximation*. New York: Blaisdell.

Feng, K. (1986) Difference schemes for Hamiltonian formalism and symplectic geometry. *J. Comput. Math.*, **4**, 279–289.

Hairer, E., Lubich, C. & Wanner, G. (2006) *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Berlin: Springer.

Hairer, E., Nørsett, S. P. & Wanner, G. (1993) *Solving Ordinary Differential Equations. I.* Springer Series in Computational Mathematics, Nonstiff problems, vol. 8, 2nd edn. Berlin: Springer.

Iserles, A. (2009) *A First Course in the Numerical Analysis of Differential Equations*. Cambridge, UK: Cambridge University Press.

Leok, M. (2004) Foundations of computational geometric mechanics. *Ph.D. Thesis*, California Institute of Technology, Pasadena, CA.

Leok, M. & Ohsawa, T. (2011) Variational and Geometric Structures of Discrete Dirac Mechanics. *Found. Comput. Math.,* **11**, 529–562.

Leok, M. & Zhang, J. (2011) Discrete Hamiltonian variational integrators. *IMA J. Numer. Anal.*, **31**, 1497–1532.

Marsden, J. E. & Ratiu, T. S. (1999) *Introduction to Mechanics and Symmetry*. Texts in Applied Mathematics, vol. 17, 2nd edn. New York: Springer.

Marsden, J. E. & West, M. (2001) Discrete mechanics and variational integrators. *Acta Numer.*, **10**, 357–514.

Maso, G. D. (1993) *An Introduction to Γ-Convergence*. Progress in nonlinear differential equations and their applications, vol. 8. Boston, MA: Birkhäuser Boston Inc.

MÜLLER, S. & ORTIZ, M. (2004) On the Γ-convergence of discrete dynamics and variational integrators. *J. Nonlinear Sci.*, **14**, 279–296.

OLVER, P. J. (1993) *Applications of Lie Groups to Differential Equations*. Graduate Texts in Mathematics, vol. 107, 2nd edn. New York: Springer.

OSWALD, P. & SHINGEL, T. (2009) Splitting methods for SU(N) loop approximation. *J. Approx. Theory*, **161**, 174–186.

PATRICK, G. W. & CUELL, C. (2009) Error analysis of variational integrators of unconstrained Lagrangian systems. *Numer. Math.*, **113**, 243–264.

RALL, L. B. (1981) *Automatic Differentiation: Techniques and Applications* Lecture Notes in Computer Science, vol. 120. Berlin: Springer.

SHINGEL, T. (2010) Trigonometric approximation of SO(N) loops. *Constr. Approx.*, **32**, 597–618.