Front. Math. China
DOI 10.1007/s11464-012-0190-9

RESEARCH ARTICLE

General techniques for constructing variational integrators

Melvin LEOK, Tatiana SHINGEL

Department of Mathematics, University of California, San Diego, La Jolla, CA 92093, USA

Abstract The numerical analysis of variational integrators relies on variational error analysis, which relates the order of accuracy of a variational integrator with the order of approximation of the exact discrete Lagrangian by a computable discrete Lagrangian. The exact discrete Lagrangian can either be characterized variationally, or in terms of Jacobi's solution of the Hamilton–Jacobi equation. These two characterizations lead to the Galerkin and shooting constructions for discrete Lagrangians, which depend on a choice of a numerical quadrature formula, together with either a finite-dimensional function space or a one-step method. We prove that the properties of the quadrature formula, finite-dimensional function space, and underlying one-step method determine the order of accuracy and momentum-conservation properties of the associated variational integrators. We also illustrate these systematic methods for constructing variational integrators with numerical examples.

Keywords Geometric numerical integration, geometric mechanics, symplectic integrator, variational integrator, Lagrangian mechanics
MSC 37J10, 65P10, 70H25

1 Introduction

Variational integrators are a particularly popular class of geometric numerical integrators, due to their symplectic and momentum conservation properties, and their good energy behavior. While most of the standard approaches for constructing variational integrators involve piecewise polynomial interpolants and high-order quadrature formulas, we will show that existing techniques from approximation theory, numerical quadrature, and one-step methods, can be incorporated systematically into the variational integration framework.

In this paper, we will discuss two general approaches to constructing

Received February 14, 2011; accepted December 21, 2011
Corresponding author: Melvin LEOK, E-mail: mleok@math.ucsd.edu

variational integrators. The first of this is the Galerkin variational integrator construction, which relies on a variational characterization of the exact discrete Lagrangian, and involves the choice of a numerical quadrature formula and a finite-dimensional function space. The second approach relies on the characterization of the exact discrete Lagrangian in terms of Jacobi's solution of the Hamilton–Jacobi equation, and involves the choice of a numerical quadrature formula and an underlying one-step method.

1.1 Discrete Lagrangian mechanics

Discrete Lagrangian mechanics [22] is based on a discrete analogue of Hamilton's principle, referred to as the *discrete Hamilton's principle*,

$$\delta \mathfrak{S}_d = 0,$$

where the *discrete action sum*, $\mathfrak{S}_d \colon Q^{n+1} \to \mathbb{R}$, is given by

$$\mathfrak{S}_d(q_0, q_1, \dots, q_n) = \sum_{i=0}^{n-1} L_d(q_i, q_{i+1}).$$

The *discrete Lagrangian*, $L_d \colon Q \times Q \to \mathbb{R}$, is a generating function of the symplectic flow, and is an approximation to the *exact discrete Lagrangian*,

$$L_d^E(q_0, q_1; h) = \int_0^h L(q_{01}(t), \dot{q}_{01}(t)) \mathrm{d}t, \tag{1}$$

where

$$q_{01}(0) = q_0, \quad q_{01}(h) = q_1,$$

and q_{01} satisfies the Euler–Lagrange equation in the time interval $(0, h)$. The exact discrete Lagrangian is related to the Jacobi solution of the Hamilton–Jacobi equation. Alternatively, one can characterize the exact discrete Lagrangian in the following way:

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

As we will see in the remainder of the paper, these two characterizations of the exact discrete Lagrangian will lead to two general techniques for constructing variational integrators.

The discrete variational principle then yields the *discrete Euler–Lagrange* (DEL) equation,

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \tag{3}$$

which implicitly defines the *discrete Lagrangian map*

$$F_{L_d} \colon (q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$$

for initial conditions (q_0, q_1) that are sufficiently close to the diagonal of $Q \times Q$. This is equivalent to the *implicit discrete Euler–Lagrange* (IDEL) equations,

$$p_k = -D_1 L_d(q_k, q_{k+1}), \quad p_{k+1} = D_2 L_d(q_k, q_{k+1}), \tag{4}$$

which implicitly defines the *discrete Hamiltonian map*

$$\widetilde{F}_{L_d} \colon (q_k, p_k) \mapsto (q_{k+1}, p_{k+1}),$$

where the discrete Lagrangian is the Type I generating function of the symplectic transformation.

1.2 Desirable properties of variational integrators

We summarize the many geometric structure-preserving properties of variational integrators, which account for their popularity in simulations involving mechanical systems.

Symplecticity. Given a discrete Lagrangian L_d, one obtains a discrete fiber derivative:

$$\mathbb{F}L_d \colon (q_0, q_1) \mapsto (q_0, -D_1 L_d(q_0, q_1)).$$

Variational integrators are symplectic, i.e., the pullback under $\mathbb{F}L_d$ of the canonical symplectic form Ω on the cotangent bundle T^*Q, is preserved. Pushing-forward the discrete Euler–Lagrange equations yield a symplectic-partitioned Runge–Kutta method.

Momentum conservation. Noether's theorem states that if a Lagrangian is invariant under the lifted action of a Lie group, the associated momentum is preserved by the flow. If a discrete Lagrangian is invariant under the diagonal action of a symmetry group, a discrete Noether's theorem holds, and the discrete flow preserves the discrete momentum map. For partial differential equations (PDEs) with a uniform spatial discretization, a backward error analysis implies approximate spatial momentum conservation [23].

Approximate energy conservation. While variational integrators do not exactly preserve energy, backward error analysis [1,6,7,25] shows that it preserves a modified Hamiltonian that is close to the original Hamiltonian for exponentially long times. In practice, the energy error is bounded and does not exhibit a drift. This is the temporal analogue of the approximate momentum conservation result for PDEs, as energy is the momentum map associated with time invariance.

Applicable to a large class of problems. The discrete variational approach is very general, and allows for the construction of geometric structure-preserving numerical integrators for PDEs [19], nonsmooth collisions [5], stochastic systems [2], nonholonomic systems [3], and constrained systems [20]. Furthermore, Dirac structures and mechanics allow for interconnections between Lagrangian systems, thereby providing a unified simulation framework for multi-physics systems.

Generates a large class of methods. A variational integrator can be constructed by choosing a finite-dimensional function space, and a numerical quadrature method [16]. By leveraging techniques from approximation theory, numerical analysis, and finite elements, one can construct variational integrators that are appropriate for problems that evolve on Lie groups [13,14] and homogeneous spaces [15], or exhibit multiple timescales [19,26].

1.3 Variational error analysis and discrete Noether's theorem

The variational integrator approach to constructing symplectic integrators has a few important advantages from the point of view of numerical analysis. In particular, the task of proving properties of the discrete Lagrangian map

$$F_{L_d}\colon\ Q\times Q\to Q\times Q$$

reduces to verifying certain properties of the discrete Lagrangian instead. Here, we summarize the results from [22, Theorems 1.3.3 and 2.3.1] that relate to the order of accuracy and momentum conservation properties of the variational integrator.

Discrete Noether's theorem. Given a discrete Lagrangian

$$L_d\colon\ Q\times Q\to\mathbb{R},$$

which is invariant under the diagonal action of a Lie group G on $Q\times Q$, then the discrete Lagrangian momentum map

$$J_{L_d}\colon\ Q\times Q\to\mathfrak{g}^*,$$

given by

$$J_{L_d}(q_k,q_{k+1})\cdot\xi=\langle -D_1L_d(q_k,q_{k+1}),\xi_Q(q_k)\rangle$$

is invariant under the discrete Lagrangian map, i.e.,

$$J_{L_d}\circ F_{L_d}=J_{L_d}.$$

Variational error analysis. The natural setting for analyzing the order of accuracy of a variational integrator is the variational error analysis framework introduced in [22]. In particular, [22, Theorem 2.3.1] states that if a discrete Lagrangian,

$$L_d\colon\ Q\times Q\to\mathbb{R},$$

approximates the exact discrete Lagrangian,

$$L_d^E\colon\ Q\times Q\to\mathbb{R},$$

given in (1) and (2) to order p, i.e.,

$$L_d(q_0,q_1;h)=L_d^E(q_0,q_1;h)+\mathscr{O}(h^{p+1}),$$

then the discrete Hamiltonian map,

$$\widetilde{F}_{L_d}\colon\ (q_k, p_k) \mapsto (q_{k+1}, p_{k+1}),$$

viewed as a one-step method, is order p accurate.

2 Galerkin variational integrators

The variational characterization of the exact discrete Lagrangian (2), naturally leads to the Galerkin approach to constructing variational integrators. This involves replacing the infinite-dimensional function space $C^2([0,h],Q)$ with a suitably chosen finite-dimensional subspace, and the integral with a numerical quadrature formula. In particular, it is common to choose an interpolatory approximation space, so as to satisfy the boundary conditions at time 0 and h in a straightforward manner.

We now recall the construction of higher-order Galerkin variational integrators, as originally described in [22]. Given a Lie group G, the associated *state space* is given by the tangent bundle TG. In addition, the dynamics on G is described by a *Lagrangian*, $L\colon\ TG \to \mathbb{R}$. Given a time interval $[0,h]$, the *action map*,

$$\mathfrak{S}\colon\ C^2([0,h],G) \to \mathbb{R},$$

is given by

$$\mathfrak{S}(g) := \int_0^h L(g(t), \dot{g}(t))\mathrm{d}t.$$

We approximate the action map, by numerical quadrature, to yield

$$\mathfrak{S}^s\colon\ C^2([0,h],G) \to \mathbb{R},$$

$$\mathfrak{S}^s(g) := h\sum_{i=1}^{s} b_i L(g(c_i h), \dot{g}(c_i h)),$$

where $c_i \in [0,1]$, $i = 1,\ldots,s$, are the quadrature points, and b_i are the quadrature weights.

Recall that the discrete Lagrangian should be an approximation of the form

$$L_d(g_0,g_1,h) \approx \operatorname*{ext}_{\substack{g\in C^2([0,h],G)\\ g(0)=g_0,\, g(h)=g_1}} \mathfrak{S}(g).$$

If we restrict the extremization procedure to the subspace spanned by the interpolatory function that is parameterized by $s+1$ internal points,

$$\varphi\colon\ G^{s+1} \to C^2([0,h],G),$$

we obtain the following discrete Lagrangian:

$$L_d(g_0,g_1) = \operatorname*{ext}_{\substack{g^\nu\in G\\ g^0=g_0,\, g^s=g_1}} \mathfrak{S}(\varphi(g^\nu;\cdot)) = \operatorname*{ext}_{\substack{g^\nu\in G\\ g^0=g_0,\, g^s=g_1}} h\sum_{i=1}^{s} b_i L(T\varphi(g^\nu;c_i h)).$$

2.1 Galerkin Lie group variational integrators

A particularly novel and nontrivial application of this approach is the construction of Lie group variational integrators. This is based on the underlying idea of Lie group integrators [9], which is to express the update map of the numerical scheme in terms of the exponential map

$$g_1 = g_0 \exp(\xi_{01}),$$

and thereby reduce the problem to finding an appropriate Lie algebra element $\xi_{01} \in \mathfrak{g}$, such that the update scheme has the desired order of accuracy. This is a desirable reduction, as the Lie algebra is a vector space, and as such the interpolation of elements can be easily defined. In our construction, the interpolatory method we use on the Lie group relies on interpolation at the level of the Lie algebra.

Since the momentum conservation properties of variational integrators depend on the discrete Lagrangian being invariant under the diagonal action of the symmetry group, we will introduce a construction that yields a G-invariant discrete Lagrangian whenever the continuous Lagrangian is G-invariant. This is achieved through the use of G-equivariant interpolatory functions, and in particular, natural charts on G, which we will now discuss. This then leads to the issue of reduction of higher-order Lie group integrators.

Group-equivariant interpolants. The interpolatory function is G-equivariant if

$$\varphi(gg^\nu; t) = g\varphi(g^\nu; t).$$

Proposition 1 *If the interpolatory function $\varphi(g^\nu; t)$ is G-equivariant, and the Lagrangian, $L\colon TG \to \mathbb{R}$, is G-invariant, then the discrete Lagrangian, $L_d\colon G \times G \to \mathbb{R}$, given by*

$$L_d(g_0, g_1) = \operatorname*{ext}_{\substack{g^\nu \in G \\ g^0 = g_0,\, g^s = g_1}} h \sum_{i=1}^{s} b_i L(T\varphi(g^\nu; c_i h)),$$

is G-invariant.

Proof we have

$$\begin{aligned} L_d(gg_0, gg_1) &= \operatorname*{ext}_{\substack{\widetilde{g}^\nu \in G \\ \widetilde{g}^0 = gg_0,\, \widetilde{g}^s = gg_1}} h \sum_{i=1}^{s} b_i L(T\varphi(\widetilde{g}^\nu; c_i h)) \\ &= \operatorname*{ext}_{\substack{g^\nu \in g^{-1}G \\ gg^0 = gg_0,\, gg^s = gg_1}} h \sum_{i=1}^{s} b_i L(T\varphi(gg^\nu; c_i h)) \\ &= \operatorname*{ext}_{\substack{g^\nu \in G \\ g^0 = g_0,\, g^s = g_1}} h \sum_{i=1}^{s} b_i L(TL_g \cdot T\varphi(g^\nu; c_i h)) \end{aligned}$$

$$
\begin{aligned}
&= \operatorname*{ext}_{\substack{g^\nu \in G \\ g^0 = g_0,\, g^s = g_1}} h \sum_{i=1}^{s} b_i L(T\varphi(g^\nu; c_i h)) \\
&= L_d(g_0, g_1),
\end{aligned}
$$

where we used the G-equivariance of the interpolatory function in the third equality, and the G-invariance of the Lagrangian in the forth equality. □

Remark 1 While G-equivariant interpolatory functions provide a computationally efficient method of constructing G-invariant discrete Lagrangians, we can construct a G-invariant discrete Lagrangian (when G is compact) by averaging an arbitrary discrete Lagrangian. In particular, given a discrete Lagrangian $L_d\colon\ Q \times Q \to \mathbb{R}$, the averaged discrete Lagrangian, given by

$$
\overline{L}_d(q_0, q_1) = \frac{1}{|G|} \int_{g \in G} L_d(gq_0, gq_1) \mathrm{d}g
$$

is G-equivariant. Therefore, in the case of compact symmetry groups, a G-invariant discrete Lagrangian always exists.

Natural charts. Following the construction in [21], we use the group exponential map at the identity, $\exp_e\colon\ \mathfrak{g} \to G$, to construct a G-equivariant interpolatory function, and a higher-order discrete Lagrangian. As shown in Proposition 1, this construction yields a G-invariant discrete Lagrangian if the Lagrangian itself is G-invariant.

In a finite-dimensional Lie group G, $\exp_e$ is a local diffeomorphism, and thus, there is an open neighborhood $U \subset G$ of e such that

$$
\exp_e^{-1}\colon\ U \to \mathfrak{u} \subset \mathfrak{g}.
$$

When the group acts on the left, we obtain a chart $\psi_g\colon\ L_g U \to \mathfrak{u}$ at $g \in G$ by

$$
\psi_g = \exp_e^{-1} \circ L_{g^{-1}}.
$$

We would like to construct an interpolatory function that is described by a set of control points $\{g^\nu\}_{\nu=0}^s$ in the group G at control times

$$
0 = d_0 < d_1 < d_2 < \cdots < d_{s-1} < d_s = 1.
$$

Our natural chart based at g^0 induces a set of control points $\xi^\nu = \psi_{g^0}(g^\nu)$ in the Lie algebra $\mathfrak{g}$ at the same control times. Let $\widetilde{l}_{\nu,s}(t)$ denote the Lagrange polynomials associated with the control times d_ν, which yields an interpolating polynomial at the level of the Lie algebra,

$$
\xi_d(\xi^\nu; \tau h) = \sum_{\nu=0}^{s} \xi^\nu \widetilde{l}_{\nu,s}(\tau).
$$

Applying $\psi_{g^0}^{-1}$ yields an interpolating curve in G of the form

$$\varphi(g^\nu;\tau h) = \psi_{g^0}^{-1}\bigg(\sum_{\nu=0}^{s} \psi_{g^0}(g^\nu)\widetilde{l}_{\nu,s}(\tau)\bigg),$$

where $\varphi(d_\nu h) = g^\nu$ for $\nu = 0,\dots,s$. Furthermore, this interpolant is G-equivariant, as shown in the following proposition.

Proposition 2 *The interpolatory function given by*

$$\varphi(g^\nu;\tau h) = \psi_{g^0}^{-1}\bigg(\sum_{\nu=0}^{s} \psi_{g^0}(g^\nu)\widetilde{l}_{\nu,s}(\tau)\bigg)$$

is G-equivariant.

Proof we have

$$\begin{aligned}
\varphi(gg^\nu;\tau h) &= \psi_{(gg^0)}^{-1}\bigg(\sum_{\nu=0}^{s} \psi_{gg^0}(gg^\nu)\widetilde{l}_{\nu,s}(\tau)\bigg)\\
&= L_{gg^0}\exp_e\bigg(\sum_{\nu=0}^{s} \exp_e^{-1}((gg^0)^{-1}(gg^\nu))\widetilde{l}_{\nu,s}(\tau)\bigg)\\
&= L_g L_{g^0}\exp_e\bigg(\sum_{\nu=0}^{s} \exp_e^{-1}((g^0)^{-1}g^{-1}gg^\nu)\widetilde{l}_{\nu,s}(\tau)\bigg)\\
&= L_g\psi_{g^0}^{-1}\bigg(\sum_{\nu=0}^{s} \exp_e^{-1}\circ L_{(g^0)^{-1}}(g^\nu)\widetilde{l}_{\nu,s}(\tau)\bigg)\\
&= L_g\psi_{g^0}^{-1}\bigg(\sum_{\nu=0}^{s} \psi_{g^0}(g^\nu)\widetilde{l}_{\nu,s}(\tau)\bigg)\\
&= L_g\varphi(g^\nu;\tau h).
\end{aligned}$$

□

Remark 2 In the proof that φ is G-equivariant, it was important that the base point for the chart should transform in the same way as the internal points g^ν. As such, the interpolatory function will be G-equivariant for a chart that it based at any one of the internal points g^ν that parameterize the function, but will not be G-equivariant if the chart is based at a fixed $g \in G$. Without loss of generality, we will consider the case when the chart is based at the first point g_0.

We will now consider a discrete Lagrangian based on the use of interpolation on a natural chart, which is given by

$$L_d(g_0,g_1) = \operatorname*{ext}_{\substack{g^\nu\in G\\ g^0=g_0,\, g^s=g_1}} h\sum_{i=1}^{s} b_i L(T\varphi(\{g^\nu\}_{\nu=0}^{s};c_i h)).$$

To further simplify the expression, we will express the extremal in terms of the Lie algebra elements ξ^ν associated with the ν-th control point. This relation is given by

$$\xi^\nu = \psi_{g_0}(g^\nu)\,,$$

and the interpolated curve in the algebra is given by

$$\xi(\xi^\nu;\tau h) = \sum_{\kappa=0}^{s} \xi^\kappa \widetilde{l}_{\kappa,s}(\tau),$$

which is related to the curve in the group:

$$g(g^\nu;\tau h) = g_0 \exp(\xi(\psi_{g_0}(g^\nu);\tau h)).$$

The velocity $\dot{\xi} = g^{-1}\dot{g}$ is given by

$$\dot{\xi}(\tau h) = g^{-1}\dot{g}(\tau h) = \frac{1}{h}\sum_{\kappa=0}^{s} \xi^\kappa \dot{\widetilde{l}}_{\kappa,s}(\tau).$$

Using the expression for the left-trivialized derivative of the exponential

$$T_\xi \exp = T_e L_{\exp(\xi)} \cdot \operatorname{dexp}_{\operatorname{ad}_\xi},$$

where

$$\operatorname{dexp}_w = \sum_{n=0}^{\infty} \frac{w^n}{(n+1)!},$$

we obtain the following expression for discrete Lagrangian:

$$\begin{aligned} L_d(g_0,g_1) = \operatorname*{ext}_{\substack{\xi^\nu\in\mathfrak{g},\,\xi^0=0\\ \xi^s=\psi_{g_0}(g_1)}} h\sum_{i=1}^{s} b_i L(L_{g_0}\exp(\xi(c_i h)),\\ T_{\exp(\xi(c_i h))}L_{g_0}\cdot T_e L_{\exp(\xi(c_i h))}\cdot \operatorname{dexp}_{\operatorname{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h)))\,. \end{aligned}$$

More explicitly, we can compute the conditions on the Lie algebra elements for the expression above to be extremal. This implies that

$$\begin{aligned} L_d(g_0,g_1) = h\sum_{i=1}^{s} b_i L(L_{g_0}\exp(\xi(c_i h)),\\ T_{\exp(\xi(c_i h))}L_{g_0}\cdot T_e L_{\exp(\xi(c_i h))}\cdot \operatorname{dexp}_{\operatorname{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h))) \end{aligned}$$

with $\xi^0 = 0$, $\xi^s = \psi_{g_0}(g_1)$, and the other Lie algebra elements implicitly defined by

$$\begin{aligned} 0 = h\sum_{i=1}^{s} b_i \Big[&\frac{\partial L}{\partial g}(c_i h) T_{\exp(\xi(c_i h))}L_{g_0}\cdot T_e L_{\exp(\xi(c_i h))}\cdot \operatorname{dexp}_{\operatorname{ad}_{\xi(c_i h)}} \widetilde{l}_{\nu,s}(c_i)\\ &+ \frac{1}{h}\frac{\partial L}{\partial \dot{g}}(c_i h) T^2_{\exp(\xi(c_i h))}L_{\exp(\xi(c_i h))}\cdot T^2_e L_{\exp(\xi(c_i h))}\cdot \operatorname{ddexp}_{\operatorname{ad}_{\xi(c_i h)}} \dot{\widetilde{l}}_{\nu,s}(c_i)\Big] \end{aligned}$$

for $\nu = 1, \ldots, s-1$, and where

$$\mathrm{ddexp}_w = \sum_{n=0}^{\infty} \frac{w^n}{(n+2)!}\,.$$

This expression for the higher-order discrete Lagrangian, together with the discrete Euler–Lagrange equation,

$$D_2 L_d(g_0, g_1) + D_1 L_d(g_1, g_2) = 0\,,$$

yields a *higher-order Lie group variational integrator.*

2.2 Higher-order discrete Euler–Poincaré equations

In this subsection, we will apply discrete Euler–Poincaré reduction (see, for example, [21]) to the Lie group variational integrator we derived previously, to construct a higher-order generalization of discrete Euler–Poincaré reduction.

Reduced discrete Lagrangian. We first proceed by computing an expression for the reduced discrete Lagrangian in the case when the Lagrangian is G-invariant. Recall that our discrete Lagrangian uses G-equivariant interpolation, which, when combined with the G-invariance of the Lagrangian, implies that the discrete Lagrangian is G-invariant as well. We compute the reduced discrete Lagrangian:

$$\begin{aligned}
l_d(g_0^{-1}g_1) &:= L_d(g_0, g_1)\\
&= L_d(e, g_0^{-1}g_1)\\
&= \operatorname*{ext}_{\substack{\xi^\nu \in \mathfrak{g},\, \xi^0 = 0\\ \xi^s = \log(g_0^{-1}g_1)}} h \sum_{i=1}^{s} b_i L(L_e \exp(\xi(c_i h)),\\
&\quad T_{\exp(\xi(c_i h))} L_e \cdot T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h)))\\
&= \operatorname*{ext}_{\substack{\xi^\nu \in \mathfrak{g},\, \xi^0 = 0\\ \xi^s = \log(g_0^{-1}g_1)}} h \sum_{i=1}^{s} b_i L(\exp(\xi(c_i h)),\\
&\quad T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h))),
\end{aligned}$$

where we used the fact that

$$\psi_{g_0}(g_1) = \log(g_0^{-1}g_1).$$

Setting

$$\xi^0 = 0, \quad \xi^s = \log(g_0^{-1}g_1),$$

we can solve the stationarity conditions for the other Lie algebra elements $\{\xi^\nu\}_{\nu=1}^{s-1}$ using the following implicit system of equations:

$$\begin{aligned}
0 = h \sum_{i=1}^{s} b_i \Big[&\frac{\partial L}{\partial g}(c_i h) T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}} \widetilde{l}_{\nu,s}(c_i)\\
&+ \frac{1}{h} \frac{\partial L}{\partial \dot{g}}(c_i h) T_e^2 L_{\exp(\xi(c_i h))} \cdot \mathrm{ddexp}_{\mathrm{ad}_{\xi(c_i h)}} \dot{\widetilde{l}}_{\nu,s}(c_i) \Big],
\end{aligned}$$

where $\nu = 1, \ldots, s-1$.

This expression for the reduced discrete Lagrangian is not fully satisfactory however, since it involves the Lagrangian, as opposed to the reduced Lagrangian. If we revisit the expression for the reduced discrete Lagrangian

$$\begin{aligned} &l_d(g_0^{-1}g_1) \\ &\quad = \operatorname*{ext}_{\substack{\xi^\nu \in \mathfrak{g},\, \xi^0 = 0 \\ \xi^s = \log(g_0^{-1}g_1)}} h \sum_{i=1}^{s} b_i L(\exp(\xi(c_i h)), T_e L_{\exp(\xi(c_i h))} \cdot \operatorname{dexp}_{\operatorname{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h))), \end{aligned}$$

we find that by G-invariance of the Lagrangian, each of the terms in the summation,

$$L(\exp(\xi(c_i h)), T_e L_{\exp(\xi(c_i h))} \cdot \operatorname{dexp}_{\operatorname{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h))),$$

can be replaced by

$$l(\operatorname{dexp}_{\operatorname{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h))),$$

where $l\colon\ \mathfrak{g} \to \mathbb{R}$ is the *reduced Lagrangian* given by

$$l(\eta) = L(L_{g^{-1}} g, T L_{g^{-1}} \dot{g}) = L(e, \eta),$$

where $\eta = T L_{g^{-1}} \dot{g} \in \mathfrak{g}$.

From this observation, we have an expression for the reduced discrete Lagrangian in terms of the reduced Lagrangian:

$$l_d(g_0^{-1}g_1) = \operatorname*{ext}_{\substack{\xi^\nu \in \mathfrak{g},\, \xi^0 = 0 \\ \xi^s = \log(g_0^{-1}g_1)}} h \sum_{i=1}^{s} b_i l(\operatorname{dexp}_{\operatorname{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h))).$$

As before, we set

$$\xi^0 = 0, \quad \xi^s = \log(g_0^{-1}g_1),$$

and solve the stationarity conditions for the other Lie algebra elements $\{\xi^\nu\}_{\nu=1}^{s-1}$ using the following implicit system of equations:

$$0 = h \sum_{i=1}^{s} b_i \Big[\frac{\partial l}{\partial \eta}(c_i h) \operatorname{ddexp}_{\operatorname{ad}_{\xi(c_i h)}} \tilde{\dot{l}}_{\nu,s}(c_i) \Big],$$

where $\nu = 1, \ldots, s-1$.

Discrete Euler–Poincaré equations. As shown above, we have constructed a higher-order reduced discrete Lagrangian that depends on

$$f_{k k+1} := g_k^{-1} g_{k+1}.$$

We will now recall the derivation of the discrete Euler–Poincaré equations, introduced in [21]. The variations in f_{kk+1} induced by variations in g_k, g_{k+1} are computed as follows:

$$\begin{aligned}\delta f_{kk+1} &= -g_k^{-1}\delta g_k g_k^{-1} g_{k+1} + g_k^{-1}\delta g_{k+1}\\ &= -g_k^{-1}\delta g_k f_{kk+1} + f_{kk+1} g_{k+1}^{-1}\delta g_{k+1}\\ &= TR_{f_{kk+1}}(-g_k^{-1}\delta g_k + \mathrm{Ad}_{f_{kk+1}} g_{k+1}^{-1}\delta g_{k+1}).\end{aligned}$$

Then, the variation in the discrete action sum is given by

$$\begin{aligned}\delta\mathfrak{S} &= \sum_{k=0}^{N-1} l_d'(f_{kk+1})\delta f_{kk+1}\\ &= \sum_{k=0}^{N-1} l_d'(f_{kk+1}) TR_{f_{kk+1}}(-g_k^{-1}\delta g_k + \mathrm{Ad}_{f_{kk+1}} g_{k+1}^{-1}\delta g_{k+1})\\ &= \sum_{k=1}^{N-1} [l_d'(f_{k-1k}) TR_{f_{k-1k}}\mathrm{Ad}_{f_{k-1k}} - l_d'(f_{kk+1}) TR_{f_{kk+1}}]\vartheta_k,\end{aligned}$$

with variations of the form $\vartheta_k = g_k^{-1}\delta g_k$. In computing the variation of the discrete action sum, we have collected terms involving the same variations, and used the fact that $\vartheta_0 = \vartheta_N = 0$. This yields the *discrete Euler–Poincaré equation*

$$l_d'(f_{k-1k}) TR_{f_{k-1k}}\mathrm{Ad}_{f_{k-1k}} - l_d'(f_{kk+1}) TR_{f_{kk+1}} = 0, \quad k = 1, \dots, N-1.$$

For ease of reference, we will recall the expressions from the previous discussion that define the *higher-order reduced discrete Lagrangian*:

$$l_d(f_{kk+1}) = h\sum_{i=1}^{s} b_i l(\mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}}(\dot\xi(c_i h))),$$

where

$$\xi(\xi^\nu;\tau h) = \sum_{\kappa=0}^{s} \xi^\kappa \widetilde{l}_{\kappa,s}(\tau), \quad \xi^0 = 0, \quad \xi^s = \log(f_{kk+1}),$$

and the remaining Lie algebra elements $\{\xi^\nu\}_{\nu=1}^{s-1}$ are defined implicitly by

$$0 = h\sum_{i=1}^{s} b_i \Big[\frac{\partial l}{\partial \eta}(c_i h)\mathrm{ddexp}_{\mathrm{ad}_{\xi(c_i h)}}\dot{\widetilde{l}}_{\nu,s}(c_i)\Big]$$

for $\nu = 1, \dots, s-1$, where

$$\mathrm{ddexp}_w = \sum_{n=0}^{\infty}\frac{w^n}{(n+2)!}\,.$$

When the discrete Euler–Poincaré equation is used in conjunction with the higher-order reduced discrete Lagrangian, we obtain the *higher-order Euler–Poincaré equations*.

2.3 Example: Lie group velocity Verlet method

We will now construct a Lie group analogue of the velocity Verlet method for the free rigid body. The velocity Verlet method can be derived from the context of discrete mechanics by considering the following discrete Lagrangian:

$$L_d(q_k, q_{k+1}) = \frac{h}{2}\Big[L\Big(q_k, \frac{q_{k+1}-q_k}{h}\Big) + L\Big(q_{k+1}, \frac{q_{k+1}-q_k}{h}\Big)\Big],$$

which corresponds to using a piecewise linear interpolant, and the trapezoidal rule to approximate the integral.

In the case of the free rigid body, the Lagrangian is given by

$$L(R, \dot{R}) = \frac{1}{2}\,\Omega J \Omega^{\mathrm{T}} = \frac{1}{2}\,\mathrm{tr}[S(\Omega) J_d S(\Omega)^{\mathrm{T}}],$$

where $S\colon \mathbb{R}^3 \to \mathfrak{so}(3)$ is defined by $S(x)y = x \times y$ for all $x, y \in \mathbb{R}^3$, and J_d is a modified moment of inertia that is related to the usual moment of inertia by the relations

$$J_d = \frac{1}{2}\,(\mathrm{tr}[J] I_{3\times 3} - 2J), \quad J = \mathrm{tr}[J_d] I_{3\times 3} - J_d.$$

From the kinematic relation $S(\Omega) = R^{\mathrm{T}}\dot{R}$, we have

$$S(\Omega_k) = R_k^{\mathrm{T}} \dot{R}_k \approx R_k\,\frac{R_{k+1}-R_k}{h} = \frac{1}{h}\,(F_k - I_{3\times 3}),$$

where $F_k = R_k^{\mathrm{T}} R_{k+1}$. Then, the discrete Lagrangian for the velocity Verlet method applied to the free rigid body is given by

$$\begin{aligned} L_d(R_k, R_{k+1}) &= 2 \cdot \frac{h}{2}\,\frac{1}{2}\,\frac{1}{h^2}\,\mathrm{tr}[(F_k - I_{3\times 3})^{\mathrm{T}} J_d (F_k - I_{3\times 3})] \\ &= \frac{1}{2h}\,\mathrm{tr}[(F_k - I_{3\times 3})(F_k - I_{3\times 3})^{\mathrm{T}} J_d] \\ &= \frac{1}{h}\,\mathrm{tr}[(I_{3\times 3} - F_k) J_d], \end{aligned}$$

where in the second to last equality, we used the fact that $\mathrm{tr}[AB] = \mathrm{tr}[BA]$, and in the last equality, we used the fact that F_k is an orthogonal matrix, J_d is symmetric, and $\mathrm{tr}[AB] = \mathrm{tr}[B^{\mathrm{T}} A^{\mathrm{T}}]$.

Recall that

$$\frac{\partial R^{\mathrm{T}}}{\partial R} \cdot \delta R = -R^{\mathrm{T}}(\delta R) R^{\mathrm{T}}.$$

Furthermore, the variation of R_k is given by

$$\delta R_k = R_k \eta_k,$$

where $\eta_k \in \mathfrak{so}(3)$ is a variation represented by a skew-symmetric matrix and vanishes at $k = 0$ and $k = N$. We may now compute the constrained variation of $F_k = R_k^{\mathrm{T}} R_k$, which yields

$$\delta F_k = \delta R_k^{\mathrm{T}} R_{k+1} + R_k^{\mathrm{T}} \delta R_{k+1} = \eta_k R_k^{\mathrm{T}} R_{k+1} + R_k^{\mathrm{T}} R_{k+1} \eta_{k+1} = -\eta_k F_k + F_k \eta_{k+1}.$$

Define the discrete action sum to be

$$\mathfrak{S}_d = \sum_{k=0}^{N-1} L_d(R_k, F_k).$$

Taking constrained variations of F_k yields

$$\delta \mathfrak{S}_d = \sum_{k=0}^{N-1} \frac{1}{h} \left\{ \mathrm{tr}[-\eta_{k+1} J_d F_k] + \mathrm{tr}[\eta_k F_k J_d] \right\}.$$

Using the fact that the variations η_k vanish at the endpoints, we may reindex the sum to obtain

$$\delta \mathfrak{S}_d = \sum_{k=1}^{N-1} \frac{1}{h} \mathrm{tr}[\eta_k (F_k J_d - J_d F_{k-1})].$$

The discrete Hamilton's principle states that the variation of the discrete action sum should be zero for all variations that vanish at the endpoints. Since η_k is an arbitrary skew-symmetric matrix, for the discrete action sum to be zero, it is necessary for $(F_k J_d - J_d F_{k-1})$ to be symmetric, which is to say that

$$F_{k+1} J_d - J_d F_{k+1}^{\mathrm{T}} - J_d F_k + F_k^{\mathrm{T}} J_d = 0.$$

This implicit equation for F_{k+1} in terms of F_k, together with the reconstruction equation $R_{k+1} = R_k F_k$, yields the Lie group analogue of the velocity Verlet method.

In practice, in time marching the numerical solution, we need to solve the above equation for $F_{k+1} \in \mathrm{SO}(3)$ given F_k. This equation is linear in F_{k+1}, but it is implicit due to the nonlinear constraint

$$F_{k+1}^{\mathrm{T}} F_{k+1} = I_{3\times 3}.$$

Since $J_d F_k - F_k^{\mathrm{T}} J_d$ is a skew-symmetric matrix, it may be represented as $S(g)$, where $g \in \mathbb{R}^3$, which reduces the equation to the form

$$F J_d - J_d F^{\mathrm{T}} = S(g). \tag{5}$$

We now introduce two iterative approaches to solve (5) numerically.

Exponential map. An element of a Lie group can be expressed as the exponential of an element of its Lie algebra, so $F \in \mathrm{SO}(3)$ can be expressed as

an exponential of $S(f) \in \mathfrak{so}(3)$ for some vector $f \in \mathbb{R}^3$. The exponential can be written in closed form, using Rodrigues' formula:

$$F = \exp S(f) = I_{3\times 3} + \frac{\sin\|f\|}{\|f\|} S(f) + \frac{1-\cos\|f\|}{\|f\|^2} S(f)^2. \tag{6}$$

Substituting (6) into (5), we obtain

$$S(g) = \frac{\sin\|f\|}{\|f\|} S(Jf) + \frac{1-\cos\|f\|}{\|f\|^2} S(f \times Jf).$$

Thus, (5) is converted into the equivalent vector equation $g = G(f)$, where $G\colon \mathbb{R}^3 \mapsto \mathbb{R}^3$ is given by

$$G(f) = \frac{\sin\|f\|}{\|f\|} Jf + \frac{1-\cos\|f\|}{\|f\|^2} f \times Jf.$$

We use the Newton method to solve $g = G(f)$, which gives the iteration

$$f_{i+1} = f_i + \nabla G(f_i)^{-1}(g - G(f_i)). \tag{7}$$

We iterate until $\|g - G(f_i)\| < \varepsilon$ for a small tolerance $\varepsilon > 0$. The Jacobian $\nabla G(f)$ in (7) can be expressed as

$$\begin{aligned}\nabla G(f) = {} & \frac{\cos\|f\|\,\|f\| - \sin\|f\|}{\|f\|^3} Jff^{\mathrm{T}} + \frac{\sin\|f\|}{\|f\|} J \\ & + \frac{\sin\|f\|\,\|f\| - 2(1-\cos\|f\|)}{\|f\|^4} (f \times Jf) f^{\mathrm{T}} \\ & + \frac{1-\cos\|f\|}{\|f\|^2} \{-S(Jf) + S(f)J\}.\end{aligned}$$

Cayley transformation. Similarly, given $f_c \in \mathbb{R}^3$, the Cayley transformation is a local diffeomorphism that maps $S(f_c) \in \mathfrak{so}(3)$ to $F \in \mathrm{SO}(3)$, where

$$F = \operatorname{cay} S(f_c) = (I_{3\times 3} + S(f_c))(I_{3\times 3} - S(f_c))^{-1}. \tag{8}$$

Substituting (8) into (5), we obtain a vector equation $G_c(f_c) = 0$ equivalent to (5):

$$G_c(f_c) = g + g \times f_c + (g^{\mathrm{T}} f_c) f_c - 2Jf_c = 0, \tag{9}$$

and its Jacobian $\nabla G_c(f_c)$ is written as

$$\nabla G_c(f_c) = S(g) + (g^{\mathrm{T}} f_c) I_{3\times 3} + f_c g^{\mathrm{T}} - 2J.$$

Then, (9) is solved by using Newton's iteration (7), and the rotation matrix is obtained by the Cayley transformation.

For both methods, numerical experiments show that 2 or 3 iterations are sufficient to achieve a tolerance of $\varepsilon = 10^{-15}$. Numerical iteration with the

Cayley transformation is a faster by a factor of 4-5 due to the simpler expressions in the iteration. It should be noted that since $F = \exp S(f)$ or $F = \operatorname{cay} S(f_c)$, it is automatically a rotation matrix, even when the equation $g = G(f)$ is not satisfied to machine precision. This computational approach is distinct from directly solving the implicit equation (5) with 9 variables and 6 constraints.

2.4 Numerical comparisons

The Lie Group velocity Verlet method is a second-order symplectic Lie group method, and it is natural to compare it to other second-order accurate methods which fail to preserve either the symplectic or Lie group structure.

(i) Explicit Midpoint Rule (RK): preserves neither symplectic nor Lie group properties.

(ii) Implicit Midpoint Rule (SRK): symplectic but does not preserve Lie group properties.

(iii) Crouch-Grossman (LGM): Lie group method but not symplectic.

These methods were used to simulate the dynamics of two simple dumbbell bodies acting under their mutual gravity. We consider the energy conservation properties of these integrators, and the extent to which they stay on the rotation group, as a function of step-size. Furthermore, we will explore how the computational cost scales as the step-size is varied.

Discussion of numerical results. Notice that while one typically expects a symplectic method to exhibit good energy behavior, we find that the implicit midpoint method (SRK) does more poorly than the two methods that preserve the Lie group structure—the Crouch-Grossman (LGM) and the Lie Group Velocity Verlet (LGVI) methods (see, Fig. 1). We find that preserving the Lie group structure and the symplectic structure simultaneously yields the best results in terms of energy preservation, and it is only in this instance that the integrator exhibits the bounded energy errors for long integration times that one typically associates with symplectic integrators.

Not surprisingly, the Lie group methods perform the best in terms of the orthogonality error. The increase in orthogonality error as the step-size decreases is associated to the accumulation of round-off error, since smaller step-sizes result in more matrix multiplications. This phenomena can be mitigated by a more careful implementation that utilizes compensated summation [10].

The Lie Group velocity Verlet method also exhibits the best computational efficiency, since it only requires one force evaluation per time-step, an advantage that it inherits from the vector space version of the Verlet method for separable Hamiltonian systems.

3 Variational integrators from arbitrary one-step methods

This section discusses how one can, given a p-th order accurate one-step method for an initial-value problem, and a q-th order accurate numerical quadrature

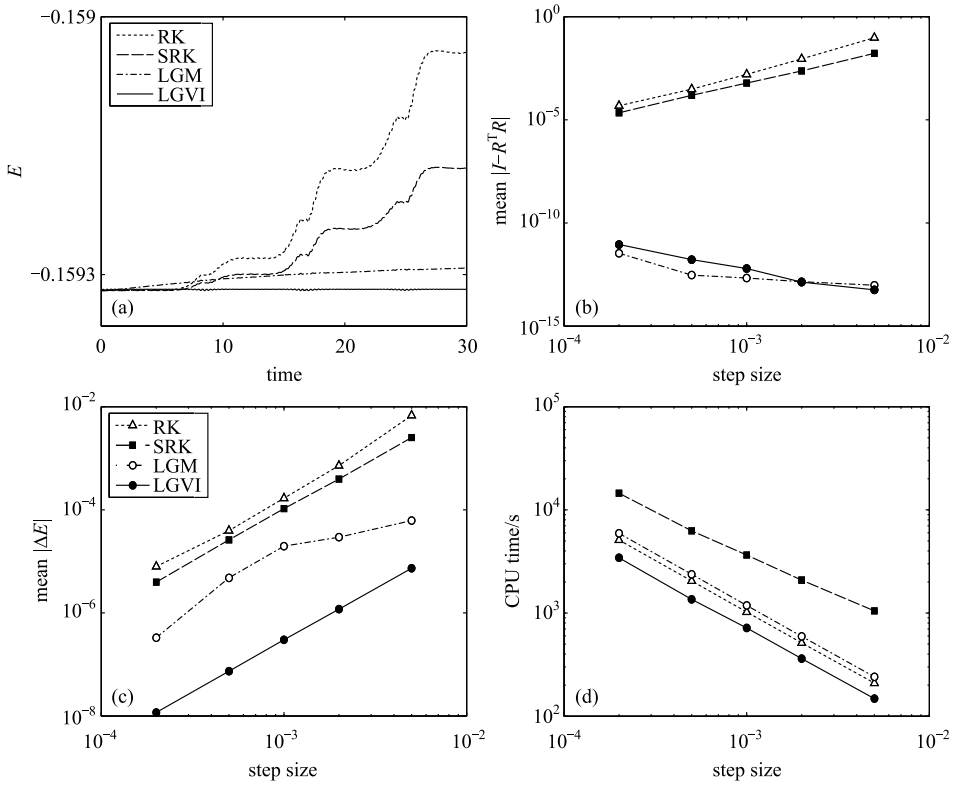

Fig. 1 Comparison of Lie Group Velocity Verlet (LGVI) with Explicit Midpoint (RK), Implicit Midpoint (SRK), and Crouch-Grossman (LGM). (a) Computed total energy for 30 seconds; (b) mean orthogonality error $\|I - R^{\mathrm{T}}R\|$ vs. step size; (c) mean total energy error $|E - E_0|$ vs. step size; (d) CPU time vs. step size.

formula, construct a variational integrator with order of accuracy $\min(p, q)$.

A related effort to develop variational integrators from arbitrary one-step methods was introduced in [24], but this was developed using a nonstandard formulation of discrete variational mechanics on phase space and curve segments [4], which is quite an abstract and involved construction. In contrast, the approach proposed in this section is far more transparent, relies on the well-understood theory of shooting methods for boundary-value problems, and is developed in the standard setting of discrete Lagrangian mechanics [22].

Outline of Approach. Notice that the characterization of the exact discrete Lagrangian associated to Jacobi's solution is expressed in terms of the action integral evaluated on a solution of a two-point boundary-value problem. A standard method of solving a boundary-value problem is to reduce it to the solution of an initial-value problem using the method of shooting.

We obtain a computable approximation to the exact discrete Lagrangian (1) in two stages: (i) apply a numerical quadrature formula to the action integral, evaluated along the exact solution of the Euler–Lagrange boundary-value problem; (ii) replace the exact solution of the Euler–Lagrange boundary-value

problem by a converged shooting solution associated with a given one-step method.

Shooting discrete Lagrangian. Given a one-step method $\Psi_h\colon TQ \to TQ$ and a numerical quadrature formula

$$\int_0^h f(x)\mathrm{d}x \approx h\sum_{i=0}^{n} b_i f(x(c_i h)),$$

with quadrature weights b_i and quadrature nodes

$$0 = c_0 < c_1 < \ldots < c_{n-1} < c_n = 1,$$

we construct the *shooting discrete Lagrangian*

$$L_d(q_0, q_1; h) = h\sum_{i=0}^{n} b_i L(q^i, v^i), \tag{10}$$

where

$$(q^{i+1}, v^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, v^i), \quad q^0 = q_0, \quad q^n = q_1. \tag{11}$$

Note that while we formally require that the endpoints are included as quadrature points, i.e., $c_0 = 0$ and $c_n = 1$, the associated weights b_0 and b_n can be zero, so this is does not constrain the type of quadrature formula we can consider.

Order of discrete Lagrangian. The order analysis of the shooting discrete Lagrangian depends critically on the global approximation properties of the shooting solution of two-point boundary-value problems.

In general, the Euler–Lagrange equation is a second-order nonlinear differential equation, and it is a standard result in the numerical analysis of the shooting method for nonlinear problems (see, for example, [11, Theorem 2.2.2]) that the approximation error in the solution of a boundary-value problem is bounded by the sum of two terms, the first of which is $\mathscr{O}(h^p)$, associated with the global error of the one-step method applied to the initial-value problem, and the second of which is related to the rate of convergence of the nonlinear solver used to determine the appropriate initial condition for the initial-value problem that would lead to the correct terminal boundary condition of the boundary-value problem. In essence, if the solution of the boundary-value problem is isolated and sufficiently regular, then a properly converged shooting method yields a solution of the two-point boundary-value problem that has error $\mathscr{O}(h^p)$, if the underlying one-step method has local truncation error $\mathscr{O}(h^{p+1})$.

This result allows us to obtain the following theorem on the order of accuracy of the shooting discrete Lagrangian, which in turn allows us to establish the order of accuracy of the associated variational integrator.

Theorem 1 *Given a p-th order accurate one-step method Ψ, a q-th order accurate quadrature formula, and a Lagrangian L that is Lipschitz continuous in*

both variables, the associated shooting discrete Lagrangian (10)-(11) *has order of accuracy* $\min(p, q)$.

Proof By [11, Theorem 2.2.2], a fully converged shooting solution $(\widetilde{q}_{01}, \widetilde{v}_{01})$, associated with a one-step method Ψ of order p, approximates the exact solution (q_{01}, v_{01}) of the Euler–Lagrange boundary-value problem with the following global error:

$$q_{01}(c_i h) = \widetilde{q}_{01}(c_i h) + \mathscr{O}(h^p), \quad v_{01}(c_i h) = \widetilde{v}_{01}(c_i h) + \mathscr{O}(h^p).$$

If the numerical quadrature formula is order q accurate, then

$$\begin{aligned}
L_d^E(q_0, q_1; h) &= \int_0^h L(q_{01}(t), v_{01}(t))\mathrm{d}t \\
&= \left[h \sum_{i=1}^m b_i L(q_{01}(c_i h), v_{01}(c_i h)) \right] + \mathscr{O}(h^{q+1}) \\
&= \left[h \sum_{i=1}^m b_i L(\widetilde{q}_{01}(c_i h) + \mathscr{O}(h^p), \widetilde{v}_{01}(c_i h) + \mathscr{O}(h^p)) \right] + \mathscr{O}(h^{q+1}) \\
&= \left[h \sum_{i=1}^m b_i L(\widetilde{q}_{01}(c_i h), \widetilde{v}_{01}(c_i h)) \right] + \mathscr{O}(h^{p+1}) + \mathscr{O}(h^{q+1}) \\
&= L_d(q_0, q_1; h) + \mathscr{O}(h^{p+1}) + \mathscr{O}(h^{q+1}) \\
&= L_d(q_0, q_1; h) + \mathscr{O}(h^{\min(p,q)+1}),
\end{aligned}$$

where we used the quadrature approximation error, the error estimates on the shooting solution, and the assumption that L is Lipschitz continuous. □

Remark 3 Notice that the numerical quadrature formula introduces an additional factor of h which gives the requisite $\mathscr{O}(h^{p+1})$ local error estimate for the discrete Lagrangian. Any consistent numerical quadrature formula would introduce this factor of h, since the integral is over the interval $[0, h]$.

Remark 4 It should be noted that the prolongation-collocation variational integrators introduced in [17] can be viewed as a special case of the shooting variational integrator. In particular, this involves the collocation method on the prolongation of the Euler–Lagrange vector field as the one-step method, and Euler–Maclaurin formula as the numerical quadrature method.

More general quadrature formulas. While we have primarily discussed numerical quadrature formulas that only depend on the integrand, one could consider more general quadrature formulas that depend on derivatives of the integrand, such as Gauss–Hermite quadrature. This would require information about the second- and higher-derivatives of q, which can be obtained by considering the prolongation of the Euler–Lagrange vector field. It is easy to show that prolongations of the vector field can be used to express all higher-derivatives of q in terms of q, $\dot{q}$, and derivatives of the Lagrangian L.

Therefore, the shooting method would provide the necessary information $(q, \dot{q})$ at the quadrature points to deduce the higher-derivatives $(\ddot{q}, q^{(3)}, \ldots)$, which in turn would allow the use of more general quadrature formulas.

Symmetric shooting discrete Lagrangians. We now show that self-adjoint one-step methods and symmetric quadrature formulas yield self-adjoint shooting discrete Lagrangians. This guarantees that the resulting variational integrator is symmetric, and hence has even order of accuracy.

Proposition 3 *Given a self-adjoint one-step method Ψ_h and a symmetric quadrature formula $(c_i + c_{n-i} = 1,\ b_i = b_{n-i})$, the associated shooting discrete Lagrangian is self-adjoint.*

Proof The discrete Lagrangian is given by

$$L_d(q_0, q_1; h) = h \sum_{i=0}^{n} b_i L(q^i, v^i),$$

where the sequence (q^i, v^i) satisfies

$$(q^{i+1}, v^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, v^i), \quad q^0 = q_0, \quad q^n = q_1.$$

The adjoint discrete Lagrangian is given by

$$L_d^*(q_0, q_1; h) = -L_d(q_1, q_0; -h) = -(-h) \sum_{i=0}^{n} b_i L(\widetilde{q}^i, \widetilde{v}^i),$$

where the sequence $(\widetilde{q}^i, \widetilde{v}^i)$ satisfies

$$(\widetilde{q}^{i+1}, \widetilde{v}^{i+1}) = \Psi_{(c_{i+1}-c_i)(-h)}(\widetilde{q}^i, \widetilde{v}^i), \quad \widetilde{q}^0 = q_1, \quad \widetilde{q}^n = q_0.$$

Since Ψ_h is self-adjoint, $\Psi_h = \Psi_h^* = \Psi_{-h}^{-1}$, and hence,

$$(\widetilde{q}^i, \widetilde{v}^i) = \Psi_{(c_{i+1}-c_i)h}(\widetilde{q}^{i+1}, \widetilde{v}^{i+1}).$$

Also, $c_i + c_{n-i} = 1$ implies that

$$c_{i+1} - c_i = c_{n-i} - c_{n-i-1}.$$

From these two properties, it is easy to see that the two sequences are related by $q^{n-i} = \widetilde{q}^i$, and $v^{n-i} = \widetilde{v}^i$. Indeed, direct substitution allows us to rewrite

$$(\widetilde{q}^i, \widetilde{v}^i) = \Psi_{(c_{i+1}-c_i)h}(\widetilde{q}^{i+1}, \widetilde{v}^{i+1})$$

as

$$(q^{n-i}, v^{n-i}) = \Psi_{(c_{n-i}-c_{n-i-1})h}(q^{n-i-1}, v^{n-i-1}),$$

which is equivalent to

$$(q^{i+1}, v^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, v^i).$$

Then

$$\begin{aligned} L_d^*(q_0, q_1; h) = & -(-h)\sum_{i=0}^{n} b_i L(\widetilde{q}^i, \widetilde{v}^i) \\ &= h\sum_{i=0}^{n} b_i L(q^{n-i}, v^{n-i}) \\ &= h\sum_{i=0}^{n} b_{n-i} L(q^i, v^i) \\ &= h\sum_{i=0}^{n} b_i L(q^i, v^i) \\ &= L_d(q_0, q_1; h), \end{aligned}$$

which implies that the discrete Lagrangian is self-adjoint, where we used the fact that $b_i = b_{n-i}$. □

Group-invariant shooting discrete Lagrangians. By the discrete Noether's theorem, a G-invariant discrete Lagrangian leads to a momentum-preserving variational integrator. In the Galerkin discrete Lagrangian construction, this can be achieved by choosing a G-equivariant interpolation space. We will see that G-equivariant one-step methods play a similar role in the construction of G-invariant shooting discrete Lagrangians.

Let $\Phi\colon G\times Q \to Q$ be a group action of G on Q, and let $T\Phi\colon G\times TQ \to TQ$ be the associated tangent-lifted action. A one-step method $\Psi_h\colon TQ \to TQ$ is G-equivariant if it commutes with $T\Phi$, i.e.,

$$\Psi_h \circ T\Phi_g = T\Phi_g \circ \Psi_h, \quad \forall\ g \in G.$$

Proposition 4 *Given a G-equivariant one-step method $\Psi_h\colon TQ \to TQ$ and a G-invariant Lagrangian $L\colon TQ \to \mathbb{R}$, the associated shooting discrete Lagrangian is G-invariant.*

Proof For notational simplicity, we will denote the group action by $\Phi_g(q) = gq$, and the tangent-lifted action by $T\Phi_g(q, v) = (gq, gv)$. By definition, the shooting discrete Lagrangian is given by

$$L_d(q_0, q_1) = h\sum_{i=0}^{n} b_i L(q^i, v^i),$$

where the sequence (q^i, v^i) satisfies

$$(q^{i+1}, v^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, v^i), \quad q^0 = q_0, \quad q^n = q_1.$$

Also,

$$L_d(gq_0, gq_1) = h\sum_{i=0}^{n} b_i L(\widetilde{q}^i, \widetilde{v}^i),$$

where the sequence $(\widetilde{q}^i, \widetilde{v}^i)$ satisfies

$$(\widetilde{q}^{i+1}, \widetilde{v}^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(\widetilde{q}^i, \widetilde{v}^i), \quad \widetilde{q}^0 = gq_0, \quad \widetilde{q}^n = gq_1.$$

By G-equivariance of the one-step method Ψ_h, we have

$$(gq^{i+1}, gv^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(gq^i, gv^i),$$

where

$$gq^0 = gq_0, \quad gq^n = gq_1.$$

From this, we conclude that the two sequences are related by

$$(\widetilde{q}^i, \widetilde{v}^i) = (gq^i, gv^i).$$

Hence,

$$L_d(gq_0, gq_1) = h\sum_{i=0}^{n} b_i L(gq^i, gv^i) = h\sum_{i=0}^{n} b_i L(q^i, v^i) = L_d(q_0, q_1),$$

where we used the G-invariance of the continuous Lagrangian. □

3.1 Implementation issues

While one can view the implicit definition of the discrete Lagrangian separately from the implicit discrete Euler–Lagrange equations

$$p_0 = -D_1 L_d(q_0, q_1; h), \quad p_1 = D_2 L_d(q_0, q_1; h),$$

in practice, one typically considers the two sets of equations together to implicitly define a one-step method:

$$L_d(q_0, q_1; h) = h\sum_{i=0}^{n} b_i L(q^i, v^i), \tag{12a}$$

$$(q^{i+1}, v^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, v^i), \quad i = 0, \dots n-1, \tag{12b}$$

$$q^0 = q_0, \tag{12c}$$

$$q^n = q_1, \tag{12d}$$

$$p_0 = -D_1 L_d(q_0, q_1; h), \tag{12e}$$

$$p_1 = D_2 L_d(q_0, q_1; h). \tag{12f}$$

Equation count. The unknowns in the equations are $(q^i, v^i)_{i=0}^n$, q_0, q_1, and p_0, p_1. If the dimension of the configuration space Q is m, then there are $2(n+3)m$ unknowns. Equations (12c)–(12f) yield $4m$ conditions, the propagation equations (12b) give $2nm$ conditions, and the initial conditions (q_0, p_0) give the last $2m$ conditions. While it is possible to solve this system directly using a nonlinear root finder, like Newton's method, another possibility is to adopt a shooting approach, which we will now describe.

Shooting implementation. Given initial conditions (q_0, p_0), we let $q^0 = q_0$, and guess an initial velocity v^0. Using the propagation equations

$$(q^{i+1}, v^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, v^i),$$

we obtain $(q^i, v^i)_{i=1}^n$. Then, we let $q_1 = q^n$, and compute

$$p_1 = D_2 L_d(q_0, q_1; h).$$

However, unless the initial velocity v^0 is chosen correctly, the equation

$$p_0 = -D_1 L_d(q_0, q_1; h)$$

will not be satisfied, and one needs to compute the sensitivity of $-D_1 L_d(q_0, q_1; h)$ on v^0, and iterate on v^0 so that $p_0 = -D_1 L_d(q_0, q_1; h)$ is satisfied. This gives a one-step method $(q_0, p_0) \mapsto (q_1, p_1)$. In practice, a good initial guess for v^0 can be obtained by inverting the continuous Legendre transformation $p = \partial L/\partial v$.

3.2 Generalizations

Hamiltonian Approach. It might be preferable to express the variational integrator in terms of the Hamiltonian, and an underlying one-step method for Hamilton's equations. We first note that

$$L(q, v) = pv - H(q, p)|_{v=\partial H/\partial p}\,.$$

This leads to the following discrete Lagrangian:

$$L_d(q_0, q_1; h) = h \sum_{i=0}^{n} b_i [p^i v^i - H(q^i, p^i)]_{v^i = \partial H/\partial p(q^i, p^i)},$$

where

$$(q^{i+1}, p^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, p^i), \quad q^0 = q_0, \quad q^n = q_1.$$

While it is possible, as before, to solve this together with the implicit discrete Euler–Lagrange equations as a nonlinear system, it can also be solved using a shooting method.

Shooting implementation. Given initial conditions (q_0, p_0), we let $q^0 = q_0$, guess an initial momentum p^0, and let

$$v^0 = \frac{\partial H}{\partial p}\,(q^0, p^0).$$

Using the propagation equations

$$(q^{i+1}, p^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, p^i)$$

and the continuous Legendre transformation

$$v^i = \frac{\partial H}{\partial p}\,(q^i, p^i),$$

we obtain $(q^i, v^i, p^i)_{i=1}^n$. Then, we let $q_1 = q^n$, and compute $p_1 = D_2 L_d(q_0, q_1; h)$. As before, unless p^0 is correctly chosen, $p_0 = -D_1 L_d(q_0, q_1; h)$ will not be satisfied, and one needs to compute the sensitivity of $-D_1 L_d(q_0, q_1; h)$ on p^0, and iterate on p^0 so that $p_0 = -D_1 L_d(q_0, q_1; h)$ is satisfied. This gives a one-step method $(q_0, p_0) \mapsto (q_1, p_1)$.

Type II *Hamiltonian Approach.* In some instances, for example, when the Hamiltonian is degenerate, the Type I generating function based approach to variational integrators using discrete Lagrangians is not applicable. A construction based on Type II generating functions, which corresponds to a discrete Hamiltonian [12,18], can be adopted. This is based on the exact discrete Hamiltonian

$$H_d^{+,E}(q_0, p_1; h) = p(h)q(h) - \int_0^h [p(t)v(t) - H(q(t), p(t))]_{v=\partial H/\partial p} \mathrm{d}t,$$

where $(q(t), p(t))$ is a solution of Hamilton's equations satisfying the boundary conditions $q(0) = q_0$, $p(h) = p_1$. A computable discrete Hamiltonian can be obtained by taking

$$H_d^{+}(q_0, p_1; h) = p^n q^n - h \sum_{i=0}^{n} b_i [p^i v^i - H(q^i, p^i)]_{v^i = \partial H/\partial p(q^i, p^i)},$$

where

$$(q^{i+1}, p^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, p^i), \quad q^0 = q_0, \quad p^n = p_1.$$

The discrete Hamiltonian then yields a symplectic map via the discrete Hamilton's equations

$$q_1 = D_2 H_d^+(q_0, p_1), \quad p_0 = D_1 H_d^+(q_0, p_1).$$

Shooting implementation. Given initial conditions (q_0, p_0), we let $q^0 = q_0$, guess an initial momentum p^0, and let $v^0 = \partial H/\partial p(q^0, p^0)$. Using the propagation equations

$$(q^{i+1}, p^{i+1}) = \Psi_{(c_{i+1}-c_i)h}(q^i, p^i)$$

and the continuous Legendre transformation $v^i = \partial H/\partial p(q^i, p^i)$, we obtain $(q^i, v^i, p^i)_{i=1}^n$. Then, we let $p_1 = p^n$, compute

$$q_1 = D_2 H_d^+(q_0, p_1),$$

and iterate on p^0 until

$$p_0 = D_1 H_d^+(q_0, p_1).$$

3.3 Examples

3.3.1 Second-order shooting variational integrator

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 which 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 equations for this Lagrangian system are

$$M\ddot{q} = -\nabla V(q), \tag{13}$$

which is Newton's second law. We rewrite it as a system of first-order equations:

$$\dot{q} = v, \quad \dot{v} = -M^{-1}\nabla V(q).$$

We introduce discrete variables that denote the following approximations:

$$q_0 \approx q(0), \quad q_1 \approx q(h), \quad v_0 \approx \dot{q}(0), \quad v_1 \approx \dot{q}(h).$$

For the corresponding one-step method, we choose the implicit midpoint rule:

$$q_1 = q_0 + h\,\frac{v_0+v_1}{2}, \tag{14a}$$

$$v_1 = v_0 - hM^{-1}\nabla V\Big(\frac{q_0+q_1}{2}\Big). \tag{14b}$$

Further, we apply the trapezoidal formula to construct the discrete Lagrangian. This will lead to a discrete Lagrangian which is second-order accurate:

$$L_d(q_0,q_1) = \frac{h}{2}\Big(\frac{1}{2}\,v_0^{\mathrm{T}} M v_0 - V(q_0) + \frac{1}{2}\,v_1^{\mathrm{T}} M v_1 - V(q_1)\Big),$$

where v_0, v_1 are functions depending on q_0, q_1, as determined by the one-step method (14). The discrete Hamilton's equations are

$$p_0 = -D_1 L_d(q_0,q_1) = -\frac{h}{2}\Big(\frac{\partial v_0}{\partial q_0}\cdot M v_0 - \nabla V(q_0) + \frac{\partial v_1}{\partial q_0}\cdot M v_1\Big),$$

$$p_1 = D_2 L_d(q_0,q_1) = \frac{h}{2}\Big(\frac{\partial v_0}{\partial q_1}\cdot M v_0 + \frac{\partial v_1}{\partial q_1}\cdot M v_1 - \nabla V(q_1)\Big),$$

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 (14) to find the expressions for $\frac{\partial v_i}{\partial q_j}$ and substitute them into the discrete Hamilton's equations. This, together with (14) leads to the system of four implicit equations with respect to the unknowns v_0, v_1, q_1 and p_1:

$$q_1 - q_0 - \frac{h}{2}\,(v_0+v_1) = 0, \tag{15a}$$

$$v_1 - v_0 + hM^{-1}\nabla V\Big(\frac{q_0+q_1}{2}\Big) = 0, \tag{15b}$$

$$p_0 - \frac{1}{2}\,M(v_0+v_1) - \frac{h}{2}\,\nabla V(q_0) - \frac{h^2}{8}\nabla\Big(\nabla V\Big(\frac{q_0+q_1}{2}\Big)\Big)(v_1-v_0) = 0, \tag{15c}$$

$$p_1 - \frac{1}{2}\,M(v_0+v_1) + \frac{h}{2}\,\nabla V(q_1) + \frac{h^2}{8}\nabla\Big(\nabla V\Big(\frac{q_0+q_1}{2}\Big)\Big)(v_1-v_0) = 0. \tag{15d}$$

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.

Remark 5 If, in the construction of the discrete Lagrangian, instead of the trapezoidal quadrature rule we apply the rectangle rule, we obtain

$$L_d(q_0, q_1) = h\Big(\frac{1}{2}\Big(\frac{v_0 + v_1}{2}\Big)^{\mathrm{T}} M\Big(\frac{v_0 + v_1}{2}\Big) - V\Big(\frac{q_0 + q_1}{2}\Big)\Big),$$

where v_0, v_1 are again functions depending on q_0, q_1 and determined by the one-step method (14), the corresponding discrete Hamiltonian map will be the standard midpoint rule:

$$p_0 = M\Big(\frac{q_1 - q_0}{h}\Big) + \frac{h}{2}\nabla V\Big(\frac{q_0 + q_1}{2}\Big), \tag{16a}$$

$$p_1 = M\Big(\frac{q_1 - q_0}{h}\Big) - \frac{h}{2}\nabla V\Big(\frac{q_0 + q_1}{2}\Big). \tag{16b}$$

3.3.2 Kepler problem

We test the numerical method (15) on the Kepler problem, a detailed discussion of which can be found in [8, § 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:

$$\dot{y}_1 = y_3, \quad \dot{y}_2 = y_4, \quad \dot{y}_3 = -\frac{y_1}{(y_1^2 + y_2^2)^{3/2}}, \quad \dot{y}_4 = -\frac{y_2}{(y_1^2 + y_2^2)^{3/2}}. \tag{17}$$

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 (15) to numerically solve (17) with initial conditions borrowed from [8, § 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. 2 that the shooting variational integrator (Fig. 2 (c)) produces an elliptic orbit almost identical to the one generated by the second-order implicit midpoint rule (Fig. 2 (a)). The Störmer–Verlet method (Fig. 2 (b)) appears to have a smaller numerical error in its orbit than SVIMID and MID. Fig. 2 (d) shows the reference solution generated by the 4th order symplectic Runge–Kutta method (Table 1).

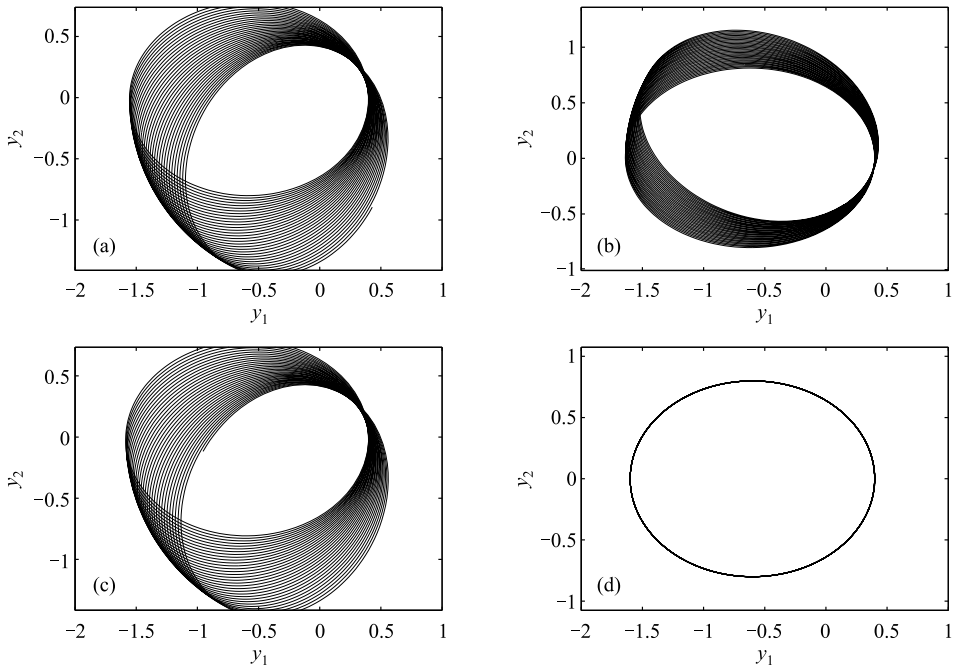

Fig. 2 Two-body Kepler problem. Step-size $h = 0.05$. First 4000 steps are shown. (a) Numerical solution using implicit midpoint rule (MID); (b) numerical solution using Störmer–Verlet method (SV); (c) numerical solution using 2nd order shooting variational integrator (SVIMID); (d) numerical solution using 4th order symplectic Runge–Kutta method.

Table 1 4th order symplectic Runge–Kutta method

$\frac{1}{2} + \frac{1}{2\sqrt{3}}$	$\frac{1}{4}$	$\frac{1}{4} + \frac{1}{2\sqrt{3}}$
$\frac{1}{2} - \frac{1}{2\sqrt{3}}$	$\frac{1}{4} - \frac{1}{2\sqrt{3}}$	$\frac{1}{4}$
	$\frac{1}{2}$	$\frac{1}{2}$

In Fig. 3, we plot the global error for the discrete Hamiltonian map versus the step-size h (Fig. 3 (a)) and CPU time (Fig. 3(b)), respectively. One can see that the shooting variational integrator SVIMID has the smallest global error per step-size. Concerning the cost comparison displayed in Fig. 3 (b), the Störmer–Verlet method is the most efficient, whereas the shooting variational integrator SVIMID, though requiring slightly more CPU time for every step-size h, is overall more accurate than the implicit midpoint rule MID.

3.3.3 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

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

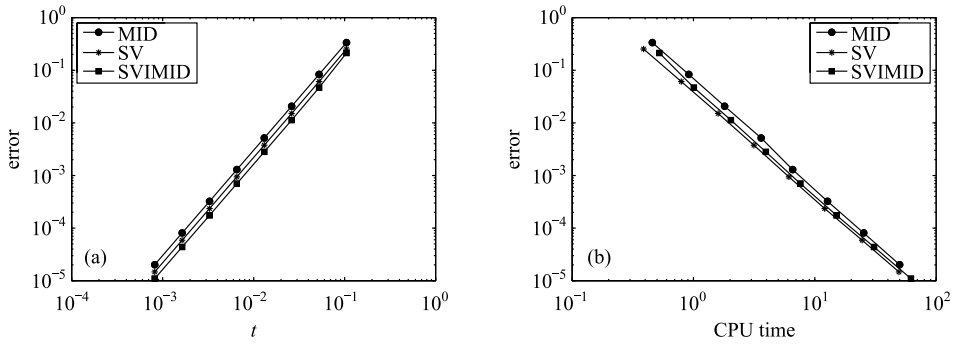

Fig. 3 Two-body Kepler problem.
(a) Global error vs. step-size; (b) global error vs. CPU time.

The total energy of the system is given by the Hamiltonian

$$H(q,p) = \frac{1}{2}\,p^2 - \cos(q).$$

The plots in Fig. 4 represent the comparison of the standard implicit midpoint integrator MID given by the formulas (16) and the shooting variational integrator SVIMID (15) when applied to the planar pendulum equations.

In Fig. 5, we compare the performance of the shooting variational integrator constructed from the explicit Runge–Kutta method (RK4) and the Simpson quadrature rule with the two-stage symplectic Runge–Kutta method (Table 1). Both methods have global error of order 4. The shooting method SVIRK4 also exhibits a slightly larger energy error as follows from Figure 5 (c).

4 Conclusions

We presented two general techniques for constructing discrete Lagrangians: (i) the Galerkin approach, which depends on a choice of a quadrature formula, and a finite-dimensional function space; (ii) the shooting approach, which depends on a choice of a quadrature formula, and a one-step method.

The order of approximation and momentum-conservation properties of a variational integrator are related to the order of approximation and the group-invariance of the discrete Lagrangian, respectively. This results in a substantial simplification in the analysis of variational integrators, since it is easier to verify the approximation and group-invariance properties of the discrete Lagrangian than it is to directly verify the order of accuracy and momentum-conservation properties of the associated variational integrator.

For Galerkin variational integrators, the group-invariance of the discrete Lagrangian can further be reduced to the group-equivariance of the finite-dimensional function space. For shooting variational integrators, the order of the discrete Lagrangian is related to the order of the quadrature formula and one-step method, and the group-invariance of the discrete Lagrangian is

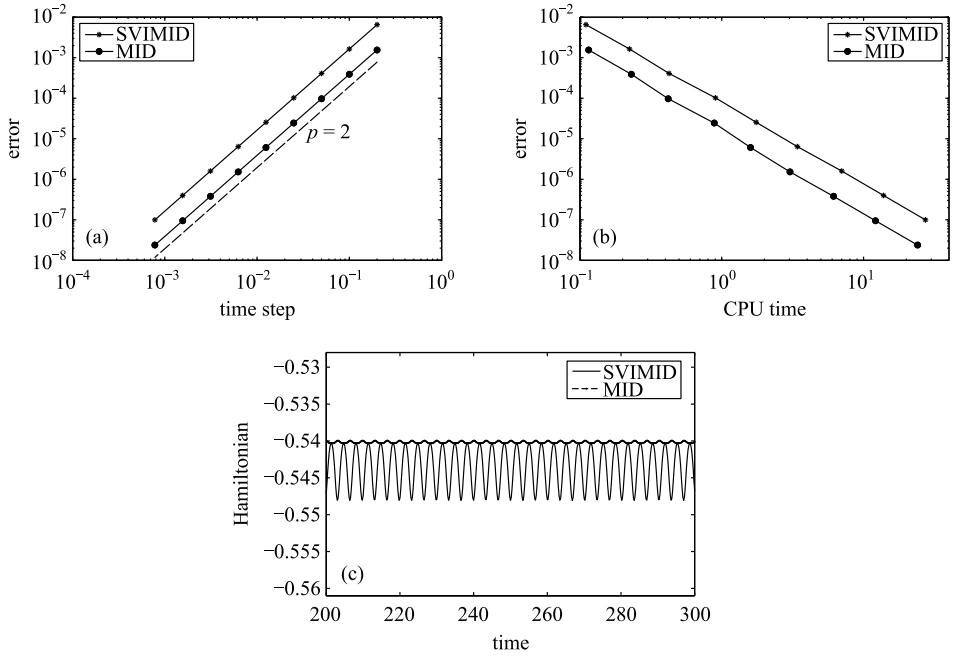

Fig. 4 Pendulum. Comparison of a shooting variational integrator (SVIMID) and implicit midpoint (MID). (a) Global error vs. time step. Dotted line is reference line for exact order. (b) Global error vs. CPU time. (c) Energy error, step-size $h = 0.2$.

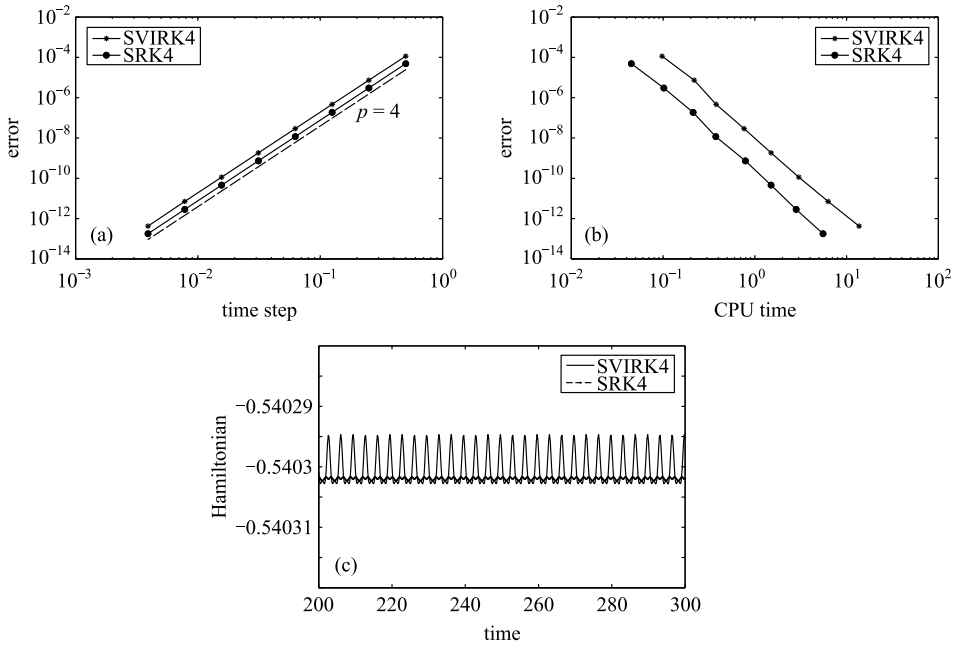

Fig. 5 Pendulum. Comparison of a shooting variational integrator (SVIRK4) and symplectic Runge–Kutta (SRK4). (a) Global error vs. time step. Dotted line is reference line for exact order. (b) Global error vs. CPU time. (c) Energy error, step size $h = 0.2$.

related to the group-equivariance of the one-step method. Furthermore, the shooting implementation allows the variational integrator to partially inherit the computational efficiencies of the underlying one-step method. In particular, a shooting variational integrator constructed from an explicit one-step method will be more computationally efficient than one based on an implicit one-step method.

These two approaches provide an explicit link between the construction of variational integrators, approximation theory, and one-step methods for ordinary differential equations. In particular, this allows one to leverage existing theoretical results and techniques in approximation theory and the numerical analysis of time-integration methods in the construction and analysis of variational integrators.

Acknowledgements This work was supported in part by the NSF Grant DMS-1001521 and the NSF CAREER Award DMS-1010687.

References

1. Benettin G, Giorgilli A. On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms. J Stat Phys, 1994, 74: 1117–1143
2. Bou-Rabee N, Owhadi H. Stochastic variational integrators. IMA J Numer Anal, 2009, 29(2): 421–443
3. Cortés J, Martínez S. Non-holonomic integrators. Nonlinearity, 2001, 14(5): 1365–1392
4. Cuell C, Patrick G. Geometric discrete analogues of tangent bundles and constrained Lagrangian systems. J Geom Phys, 2009, 59(7): 976–997
5. Fetecau R, Marsden J, Ortiz M, West M. Nonsmooth Lagrangian mechanics and variational collision integrators. SIAM Journal on Applied Dynamical Systems, 2003, 2(3): 381–416
6. Hairer E. Backward analysis of numerical integrators and symplectic methods. Scientific Computation and Differential Equations (Auckland, 1993). Ann Numer Math, 1994, 1(1-4): 107–132
7. Hairer E, Lubich C. The life-span of backward error analysis for numerical integrators. Numer Math, 1997, 76: 441–462
8. Hairer E, Lubich C, Wanner G. Geometric Numerical Integration. 2nd ed. Springer Series in Computational Mathematics, Vol 31. Berlin: Springer-Verlag, 2006
9. Iserles A, Munthe-Kaas H, Nørsett S, Zanna A. Lie-group methods. In: Acta Numerica, Vol 9. Cambridge: Cambridge University Press, 2000, 215–365
10. Kahan W. Further remarks on reducing truncation errors. Commun ACM, 1965, 8: 40
11. Keller H B. Numerical methods for two-point boundary value problems. New York: Dover Publications Inc, 1992
12. Lall S, West M. Discrete variational Hamiltonian mechanics. J Phys A, 2006, 39(19): 5509–5519
13. Lee T, Leok M, McClamroch N. Lie group variational integrators for the full body problem. Comput Methods Appl Mech Engrg, 2007, 196(29-30): 2907–2924
14. Lee T, Leok M, McClamroch N. Lie group variational integrators for the full body problem in orbital mechanics. Celestial Mech Dynam Astronom, 2007, 98(2): 121–144
15. Lee T, Leok M, McClamroch N. Lagrangian mechanics and variational integrators on two-spheres. Int J Numer Methods Eng, 2009, 79(9): 1147–1174

16. Leok M. Generalized Galerkin variational integrators: Lie group, multiscale, and pseudospectral methods. Preprint, 2004, arXiv: math.NA/0508360
17. Leok M, Shingel T. Prolongation-collocation variational integrators. IMA J Numer Anal (in press), arXiv: 1101.1995 [math.NA]
18. Leok M, Zhang J. Discrete Hamiltonian variational integrators. IMA J Numer Anal, 2011, 31(4): 1497–1532
19. Lew A, Marsden J E, Ortiz M, West M. Asynchronous variational integrators. Arch Ration Mech Anal, 2003, 167(2): 85–146
20. Leyendecker S, Marsden J, Ortiz M. Variational integrators for constrained mechanical systems. Z Angew Math Mech, 2008, 88: 677–708
21. Marsden J, Pekarsky S, Shkoller S. Discrete Euler–Poincaré and Lie–Poisson equations. Nonlinearity, 1999, 12(6): 1647–1662
22. Marsden J E, West M. Discrete mechanics and variational integrators. Acta Numer, 2001, 10: 357–514
23. Oliver M, West M, Wulff C. Approximate momentum conservation for spatial semi-discretizations of nonlinear wave equations. Numer Math, 2004, 97: 493–535
24. Patrick G, Spiteri R, Zhang W, Cuell C. On converting any one-step method to a variational integrator of the same order. In: 7th International Conference on Multi-body systems, Nonlinear Dynamics, and Control, Vol 4. 2009, 341–349
25. Reich S. Backward error analysis for numerical integrators. SIAM J Numer Anal, 1999, 36: 1549–1570
26. Stern A, Grinspun E. Implicit-explicit variational integration of highly oscillatory problems. Multiscale Model Simul, 2009, 7(4): 1779–1794