BIT Numer Math (2018) 58:457–488
https://doi.org/10.1007/s10543-017-0690-9

Lagrangian and Hamiltonian Taylor variational integrators

Jeremy Schmitt[1] · **Tatiana Shingel**[1] ·
Melvin Leok[1]

Received: 20 March 2017 / Accepted: 3 November 2017 / Published online: 24 November 2017
© Springer Science+Business Media B.V., part of Springer Nature 2017

Abstract In this paper, we present a variational integrator that is based on an approximation of the Euler–Lagrange boundary-value problem via Taylor's method. This can be viewed as a special case of the shooting-based variational integrator. The Taylor variational integrator exploits the structure of the Taylor method, which results in a shooting method that is one order higher compared to other shooting methods based on a one-step method of the same order. In addition, this method can generate quadrature nodal evaluations at the cost of a polynomial evaluation, which may increase its efficiency relative to other shooting-based variational integrators. A symmetric version of the method is proposed, and numerical experiments are conducted to exhibit the efficacy and efficiency of the method.

Keywords Geometric numerical integration · Variational integrators · Symplectic integrators · Hamiltonian mechanics

Mathematics Subject Classification 37M15 · 65P10 · 70H05

Communicated by Christian Lubich.

The authors were supported in part by NSF Grants DMS-1001521, CMMI-1029445, DMS-1065972, CMMI-1334759, DMS-1411792, and NSF CAREER award DMS-1010687.

✉ Melvin Leok
mleok@math.ucsd.edu

Jeremy Schmitt
j2schmit@ucsd.edu

Tatiana Shingel
t.shingel@gmail.com

[1] Department of Mathematics, University of California, San Diego, CA, USA

1 Introduction

This paper is concerned with the systematic construction and analysis of Lagrangian and Hamiltonian variational integrators of arbitrarily high-order derived from an underlying Taylor integrator. This can be viewed, on the Lagrangian side, as a special case of the shooting-based variational integrators introduced in [12], which provided a general framework for constructing a Lagrangian variational integrator from a given one-step method.

The main limitation of the shooting-based variational integrator approach is that in order to achieve higher-order accuracy, one requires multiple steps of the underlying one-step method in order to obtain approximations of the solution of the Euler–Lagrange boundary-value problem at the quadrature points. This is of course the best one can hope to achieve given a generic one-step method, but for one-step methods such as collocation methods or Taylor methods, one obtains a continuous approximation that can be evaluated at multiple points. As such, these methods only require a single step of the one-step method in order to obtain a continuous approximation of the Euler–Lagrange boundary-value problem that can be used to construct discrete Lagrangians and discrete Hamiltonians that generate symplectic integrators.

We focus on the use of Taylor integrators as the underlying one-step method, since they can be efficiently implemented to arbitrarily high-order for a broad range of problems by leveraging automatic differentiation techniques, and the resulting solution can be evaluated at additional quadrature points at the cost of a polynomial evaluation.

2 Discrete mechanics

Let Q be a differentiable manifold, often referred to as the ***configuration space***. Discrete Lagrangian mechanics [14] is based on a discrete analogue of Hamilton's principle, referred to as the ***discrete Hamilton's principle***,

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

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

$$\mathbb{S}_d(q_0, q_1, \ldots, q_n) = \sum_{i=0}^{n-1} L_d(q_k, q_{k+1}).$$

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

$$L_d^E(q_0, q_1; h) = \int_0^h L(q_{01}(t), \dot{q}_{01}(t))dt, \tag{2.1}$$

where $q_{01}(0) = q_0$, $q_{01}(h) = q_1$, and q_{01} satisfies the Euler–Lagrange equation in the time interval $(0, h)$.

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

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \tag{2.2}$$

which implicitly defines the ***discrete Lagrangian map*** $F_{L_d} : (q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$ 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,

$$p_k = -D_1 L_d(q_k, q_{k+1}), \qquad p_{k+1} = D_2 L_d(q_k, q_{k+1}), \tag{2.3}$$

which implicitly defines the ***discrete Hamiltonian map*** $\tilde{F}_{L_d} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$, where the discrete Lagrangian is the Type I generating function of the symplectic transformation. Furthermore, the discrete Hamiltonian map associated with the exact discrete Lagrangian $\tilde{F}_{L_d^E}$ is the time-h flow map of the Hamiltonian vector field. These observations serve as the basis by which the variational error analysis result of Sect. 2.1 is proven in [14]. In particular, variational error analysis relates the order to which a computable discrete Lagrangian approximates the exact discrete Lagrangian with the order of accuracy of the discrete Hamiltonian map when viewed as a one-step method for approximating the flow of Hamilton's equations.

2.1 Variational error analysis

The natural setting for analyzing the order of accuracy of a variational integrator is the variational error analysis framework introduced in [14]. In particular, Theorem 2.3.1 of [14] states that if a discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$, approximates the exact discrete Lagrangian, $L_d^E : Q \times Q \to \mathbb{R}$, given in (2.1) to order p, i.e.,

$$L_d(q_0, q_1; h) = L_d^E(q_0, q_1; h) + \mathscr{O}(h^{p+1}), \tag{2.4}$$

then the discrete Hamiltonian map, $\tilde{F}_{L_d} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$, viewed as a one-step method, is order p accurate.

3 Lagrangian Taylor variational integrator

The exact discrete Lagrangian (2.1) is given by the action integral evaluated along the solution of the Euler–Lagrange boundary-value problem. In turn, the boundary-value problem with boundary data (q_0, q_1) can be related to an initial-value problem with initial data (q_0, v_0), which satisfies the condition $q_1 = \pi_Q \circ \Phi_h(q_0, v_0)$, where $\pi_Q : TQ \to Q$ is the canonical projection onto Q and $\Phi_h : TQ \to TQ$ is the exact time-h flow map. This yields the following characterization of the exact discrete Lagrangian,

$$L_d^E(q_0, q_1; h) = \int_0^h L(\Phi_t(q_0, v_0))dt,$$

where $q_1 = \pi_Q \circ \Phi_h(q_0, v_0)$. The Taylor variational integrator is generated by a computable discrete Lagrangian obtained when the integral is approximated by a quadrature rule, and the Taylor method is used to approximate the flow map that relates the boundary data (q_0, q_1) with the initial-value data (q_0, v_0), and the trajectory associated with the initial data. The following summarizes the construction of the Taylor variational integrator.

(i) The approximation to $\dot{q}(0) = v_0$, denoted as $\tilde{v}_0$, is defined via the inverse problem,

$$q_1 = \pi_Q \circ \Psi_h^{(r+1)}(q_0, \tilde{v}_0), \tag{3.1}$$

where $\pi_Q : TQ \to Q$ is the canonical projection onto Q and $\Psi_h^{(r+1)} : TQ \to TQ$ denotes a $(r+1)$-order Taylor method.

(ii) Generate approximations to the quadrature nodal values, $q_{c_i} \approx q(c_i h)$ (excluding q_1 if needed, which is assumed to be given) and $v_{c_i} \approx \dot{q}(c_i h)$, via Taylor's method using $\tilde{v}_0$,

$$(q_{c_i}, v_{c_i}) = \Psi_{c_i h}^{(r)}(q_0, \tilde{v}_0). \tag{3.2}$$

(iii) Apply the quadrature rule to construct the associated discrete Lagrangian,

$$L_d(q_0, q_1; h) = h \sum_{i=1}^{m} b_i L(q_{c_i}, v_{c_i}). \tag{3.3}$$

(iv) Applying the discrete Legendre transforms implicitly defines the method,

$$\begin{aligned} p_0 &= -D_1 L_d(q_0, q_1; h), \\ p_1 &= D_2 L_d(q_0, q_1; h). \end{aligned}$$

Remark 3.1 It may seem like a waste to solve for $\tilde{v}_0$ using a $(r+1)$-order Taylor method, and then to use only a r-order method to solve the Euler–Lagrange boundary-value problem, but from an implementation perspective, no additional derivative evaluations are needed to solve (3.1), other than those already required in implementing the r-order Taylor method on TQ. In fact, it is an efficient use of the higher-derivative information we already needed to compute in order to construct the r-order Taylor method on TQ.

This apparent discrepancy can be resolved by thinking of Eq. (3.1) as being a $(r+1)$-order Taylor method for the second-order differential equation on Q, and (3.2) as a r-order Taylor method on the first-order differential equation on TQ. In particular, notice that because of the canonical projection π_Q in Eq. (3.1), we only need to compute up to $q^{(r+1)}(0)$ in order to solve for $\tilde{v}_0$, instead of the up to $q^{(r+2)}(0)$ that is necessary to define $\Psi_h^{(r+1)}$. But, we needed to compute up to $v^{(r)}(0) = q^{(r+1)}(0)$ in order to construct $\Psi_h^{(r)}$, the r-order Taylor method on TQ.

The approximation of $\tilde{v}_0$ by this procedure is particular to the Taylor method, and consequently, as the theorems below will show, leads to a variational integrator that is one order higher than would be obtained following the shooting construction of [12]. This is an immediate consequence of Lemma 3.1 below, which shows that the Taylor construction outlined above leads to a higher-order approximation of $\tilde{v}_0$ than would occur under the shooting framework of [12] (see Theorem 3.1 of [12]). The standard shooting method based on a r-order method for an IVP, results in a $(r-1)$-order method for the corresponding BVP (see Thm 2.2.2 [9]). However, by using a $(r+1)$-order Taylor method in (3.1) the resulting approximate solution of the Euler–Lagrange BVP is r-order, and this incurs no additional cost as explained above.

One might wonder what would happen if we replaced the Taylor method with the Runge–Kutta method in our approach. This would involve using a pair of Runge–Kutta methods, one of order p and another of order $p+1$, where the $(p+1)$-order method is used to obtain the approximation of $\tilde{v}_0$, and the p-order method is used to obtain approximations of (q_{c_i}, v_{c_i}) from $(q_0, \tilde{v}_0)$. The relative computational cost depend on the relative cost of one step of a Taylor method versus m steps of the Runge–Kutta method. The reason for the factor of m is that the cost of computing a Runge–Kutta step with the same initial data but different stepsize is almost as high as computing a Runge–Kutta step with different initial data, as all the internal stages, with the exception of the initial one, are different. As such, the cost of computing (q_{c_i}, v_{c_i}) with Runge–Kutta scales with the number of quadrature points, whereas each additional evaluation of the Taylor method with the same initial data but with different stepsizes is at the relatively negligible cost of a polynomial evaluation. Having said that, it is likely that the method obtained by replacing the Taylor method with the pair of Runge–Kutta methods, would be cheaper than the shooting-based method developed in [12], as the cost of approximating the initial velocity would be lower, since it would involve inverting one step of a $(p+1)$-order method, instead of inverting the composition of m steps of a p-order method.

The upshot is that the shooting method outlined here is an augmented form of the traditional shooting method, and the Taylor method is uniquely suited to take advantage of the overall approach, both because it is able to obtain a higher-order approximation of the initial velocity at no additional cost, and that the cost of computing the approximations at each additional quadrature point is at the cost of a polynomial evaluation as opposed to an evaluation of the underlying one-step method.

The following lemmas are needed for a theorem on the accuracy of the method. These lemmas can be proved using Lipschitz continuity and triangle inequalities (see "Appendix" for their proofs).

Lemma 3.1 *$\tilde{v}_0$ as defined by, (3.1), approximates v_0 to at least $\mathscr{O}(h^{r+1})$.*

Lemma 3.2 *A r-order Taylor method with initial conditions $(q_0, \tilde{v}_0)$, where $\tilde{v}_0$ is defined by (3.1), is accurate to at least $\mathscr{O}(h^{r+1})$ for the Euler–Lagrange boundary-value problem with boundary conditions (q_0, q_1).*

Theorem 3.1 *Assuming a Lagrangian L that is Lipschitz continuous in both variables, then for a r-order accurate Taylor method, $\Psi_h^{(r)}$, and a s-order accurate quadrature formula, the associated Taylor discrete Lagrangian (3.3) has order of accuracy at least* $\min(r+1, s)$.

Proof $(q_d(t), v_d(t))$, associated with the Taylor method Ψ_h of order r and initial data $(q_0, \tilde{v}_0)$, approximates the exact solution $(q(t), v(t))$ of the Euler–Lagrange boundary-value problem with the following error,

$$\begin{aligned} q(c_i h) &= q_{c_i} + \mathscr{O}(h^{r+1}), \\ v(c_i h) &= v_{c_i} + \mathscr{O}(h^{r+1}). \end{aligned}$$

If the numerical quadrature formula is order s accurate, then

$$\begin{aligned} L_d^E(q_0, q_1; h) &= \int_0^h L(q(t), v(t))dt \\ &= \left[h \sum_{i=1}^m b_i L(q(c_i h), v(c_i h)) \right] + \mathscr{O}(h^{s+1}) \\ &= \left[h \sum_{i=1}^m b_i L(q_{c_i} + \mathscr{O}(h^{r+1}), v_{c_i} + \mathscr{O}(h^{r+1})) \right] + \mathscr{O}(h^{s+1}) \\ &= \left[h \sum_{i=1}^m b_i L(q_{c_i}, v_{c_i}) \right] + \mathscr{O}(h^{r+2}) + \mathscr{O}(h^{s+1}) \\ &= L_d(q_0, q_1; h) + \mathscr{O}(h^{r+2}) + \mathscr{O}(h^{s+1}) \\ &= L_d(q_0, q_1; h) + \mathscr{O}(h^{\min(r+1,s)+1}), \end{aligned}$$

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

The choice of the Taylor method as the underlying one-step method has the advantage that it only requires one to precompute the prolongation of the Euler–Lagrange vector field once at the initial time, and the computational cost is not increased appreciably by having to compute the numerical solution at multiple quadrature nodes, since that only requires a polynomial evaluation. This efficiency in evaluation improves upon the methods outlined in [11] and [12], which utilized collocation and the shooting-method, respectively.

Example 3.1 Consider a first-order Taylor variational integrator that uses the rectangular quadrature rule about the initial point. We assume a Lagrangian of the form $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q)$. Then the integrator is constructed as follows:

(i) The inverse problem is,

$$q_1 = q_0 + h\tilde{v}_0.$$

This implies $\tilde{v}_0 = \frac{q_1 - q_0}{h}$, where q_0, q_1 are the given boundary conditions.

(ii) The quadrature nodal values are $q_{c_1} = q_0$ and $v_{c_1} = \tilde{v}_0 = \frac{q_1 - q_0}{h}$.

(iii) The corresponding discrete Lagrangian is given by,

$$\begin{aligned} L_d(q_0, q_1; h) &= hL\Big(q_0, \frac{q_1 - q_0}{h}\Big) \\ &= h\left[\frac{1}{2}\Big(\frac{q_1 - q_0}{h}\Big)^T M\Big(\frac{q_1 - q_0}{h}\Big) - V(q_0)\right]. \end{aligned}$$

(iv) The discrete Legendre transforms are given by,

$$\begin{aligned} p_0 &= M\Big(\frac{q_1 - q_0}{h}\Big) + h\nabla V(q_0), \\ p_1 &= M\Big(\frac{q_1 - q_0}{h}\Big). \end{aligned}$$

With some rearranging and substitution we see that this is symplectic Euler-A,

$$\begin{aligned} q_1 &= q_0 + hM^{-1}p_1, \\ p_1 &= p_0 - h\nabla V(q_0). \end{aligned}$$

If we use the rectangular quadrature rule about the end point, then the resulting method would be symplectic Euler-B. If instead we choose the trapezoid quadrature rule, then the resulting method will be Störmer–Verlet. All three of these classic symplectic integrators can be derived as Taylor variational integrators. However, there are also novel methods that come from the Taylor variational integrator framework, as the next example illustrates.

Example 3.2 Consider a second-order Taylor variational integrator, which utilizes a first-order Taylor method combined with the trapezoid rule to approximate the discrete Lagrangian. The approximate initial velocity is given by,

$$\tilde{v}_0 = \frac{q_1 - q_0}{h} - \frac{h}{2}M^{-1}\nabla V(q_0).$$

The resulting method is an explicit second-order method given by,

$$\begin{aligned} q_1 &= q_0 + hM^{-1}p_0 - \frac{h^2}{2}M^{-1}\nabla V(q_0) + \frac{h^4}{4}M^{-1}\nabla\nabla V(q_0)M^{-1}V(q_0), \\ p_1 &= M\tilde{v}_0 - \frac{h}{2}(\nabla V(q_0) + \nabla V(q_1)). \end{aligned}$$

Note that the method may not be explicit for nonseparable Lagrangians.

As demonstrated above, Lagrangian Taylor variational integrators provide a very general family of symplectic integrators that include not only some of the classic symplectic integrators, but also novel symplectic integrators. The Taylor variational integrator is amenable to the construction of higher-order symplectic integrators that can benefit from many of the numerical techniques that have enhanced the classical

Taylor method (see [8,21]). In particular, automatic differentiation allows for accurate and relatively cheap derivative evaluations (see [5,15,17]). In general, higher-order Taylor variational integrators will require solving a system of nonlinear equations, which can be dealt with using standard methods (see [6]). While it is clear that Taylor variational integrators will have a higher computational cost than the Taylor method, in many cases the Taylor variational integrator can preserve accuracy and structure for larger stepsizes, which may justify the higher cost per step. We will further examine these topics in Sect. 5. Next, we consider discrete Hamiltonian formulations and symmetric formulations of the Taylor variational integrator.

4 Hamiltonian and symmetric Taylor variational integrators

4.1 Hamiltonian Taylor variational integrators

Thus far, we have derived the Taylor variational integrator by approximating the discrete Lagrangian, which is a type I generating function of the symplectic map/integrator. However, we will also consider the discrete right and discrete left Hamiltonians (see [10,13]), which are type II and type III generating functions, respectively. The motivation being that for a degenerate Hamiltonian there may be no corresponding Lagrangian formulation, in which case the discrete Hamiltonian formulation may be the only way to construct a variational integrator. Also, it has recently been shown in [19] that even when the Legendre transform is a diffeomorphism, the discrete Lagrangian and discrete Hamiltonian formulation generated by a fixed approximation scheme can lead to different variational integrators.

The boundary-value formulation of the exact discrete right Hamiltonian is given by,

$$H_d^{+,E}(q_0, p_1; h) = \left(p_1^T q_1 - \int_0^T \left[p^T \dot{q} - H(q, p) \right] dt \right),$$

where $(q(t), p(t))$ satisfy Hamilton's equations with boundary conditions $q(0) = q_0$, $p(T) = p_1$. Now let us consider the construction of a Taylor discrete right Hamiltonian.

(i) Construct a r-order Taylor expansion on the cotangent bundle, T^*Q, and solve for $\tilde{p}_0$,

$$p_1 = \pi_{T^*Q} \circ \Psi_h^{(r)}(q_0, \tilde{p}_0),$$

where $\pi_{T^*Q} : (q, p) \mapsto p$.

(ii) Pick a quadrature rule of order s with quadrature weights and nodes given by (b_i, c_i) for $i = 1, \ldots, m$.

(iii) Use a r-order Taylor method to generate approximations of $(q(t), p(t))$ at the quadrature nodes,

$$(q_{c_i}, p_{c_i}) = \Psi_{c_i h}^{(r)}(q_0, \tilde{p}_0),$$

and use a $(r+1)$-order Taylor method on the configuration manifold to generate the approximation to the boundary term q_1,

$$\tilde{q}_1 = \pi_Q \circ \Psi_h^{(r+1)}(q_0, \tilde{p}_0).$$

(iv) Use the quadrature rule and approximate boundary term, $\tilde{q}_1$, to construct the discrete right Hamiltonian of order $\min(r+1, s)$,

$$H_d^+(q_0, p_1; h) = p_1^T \tilde{q}_1 - h \sum_{i=1}^{m} \Big[p_{c_i}^T \dot{q}_{c_i} - H\Big(\Psi_{c_i h}^{(r)}(q_0, \tilde{p}_0)\Big)\Big],$$

where $\dot{q}_{c_i}$ is obtained by using the continuous Legendre transform, $(q_{c_i}, \dot{q}_{c_i}) = \mathbb{F}H(q_{c_i}, p_{c_i})$.

(v) The method is implicitly defined by the implicit discrete right Hamilton's equations,

$$q_1 = D_2 H_d^+(q_0, p_1), \qquad p_0 = D_1 H_d^+(q_0, p_1). \tag{4.1}$$

The boundary-value formulation of the exact discrete left Hamiltonian is given by,

$$H_d^{-,E}(q_1, p_0; h) = -\left(p_0^T q_0 - \int_0^T \Big[p^T \dot{q} - H(q, p) \Big] dt \right),$$

where $(q(t), p(t))$ satisfy Hamilton's equations with boundary conditions $q(T) = q_1$, $p(0) = p_0$. Now let us consider the construction of a Taylor discrete left Hamiltonian.

(i) Construct a $(r+1)$-order Taylor expansion on the cotangent bundle, T^*Q, and solve for $\tilde{q}_0$,

$$q_1 = \pi_Q \circ \Psi_h^{(r+1)}(\tilde{q}_0, p_0).$$

(ii) Pick a quadrature rule of order s with quadrature weights and nodes given by (b_i, c_i) for $i = 1, \ldots, m$.

(iii) Use a r-order Taylor method to generate approximations of $(q(t), p(t))$ at the quadrature nodes,

$$(q_{c_i}, p_{c_i}) = \Psi_{c_i h}^{(r)}(\tilde{q}_0, p_0).$$

(iv) Use the quadrature rule and approximate boundary term, $\tilde{q}_0$, to construct the discrete left Hamiltonian of order $\min(r+1, s)$,

$$H_d^-(q_1, p_0; h) = -p_0^T \tilde{q}_0 - h \sum_{i=1}^{m} \Big[p_{c_i}^T \dot{q}_{c_i} - H\Big(\Psi_{c_i h}^{(r)}(\tilde{q}_0, p_0)\Big)\Big],$$

where $\dot{q}_{c_i}$ is obtained by using the continuous Legendre transform, $(q_{c_i}, \dot{q}_{c_i}) = \mathbb{F}H(q_{c_i}, p_{c_i})$.

(v) The method is implicitly defined by the implicit discrete left Hamilton's equations,

$$p_1 = -D_1 H_d^-(q_1, p_0; h), \qquad q_0 = -D_2 H_d^-(q_1, p_0; h). \tag{4.2}$$

The Störmer–Verlet method can be derived as a Lagrangian Taylor variational integrator by choosing $r = 0$ for the respective Taylor methods and using the trapezoid rule for the quadrature rule. This yields a discrete Lagrangian corresponding to the Störmer–Verlet method,

$$L_d(q_0, q_1; h) = \frac{h}{2}\left(\left(\frac{q_1 - q_0}{h}\right)^T M\left(\frac{q_1 - q_0}{h}\right) - V(q_0) - V(q_1)\right).$$

Choosing $r = 0$ and the trapezoid rule to construct a Hamiltonian Taylor variational integrator results in a discrete right Hamiltonian given by,

$$H_d^+(q_0, p_1; h) = p_1^T(q_0 + hM^{-1}p_1) - \frac{h}{2}\left(p_1^T M^{-1} p_1 - V(q_0 + hM^{-1}p_1)\right),$$

and a discrete left Hamiltonian given by,

$$H_d^-(q_1, p_0; h) = -p_0^T(q_1 - hM^{-1}p_0) - \frac{h}{2}\left(p_0^T M^{-1} p_0 - V(q_1 - hM^{-1}p_0)\right).$$

The corresponding methods are not Störmer–Verlet, in fact they are neither symmetric nor explicit. However, a simple calculation shows that these discrete Hamiltonians are adjoint to each other (see [19] for info on adjoint discrete Hamiltonians), i.e., $-H_d^+(q_1, p_0; -h) = H_d^-(q_1, p_0; h)$. Therefore, a symmetric method can be constructed by composing the two methods (see [4] II.4). We will denote the resulting symmetric method by SVHd, and we compare it to Störmer–Verlet in Sect. 5 (see Figs. 7, 12).

It should be noted that some approximations schemes do yield the same method when applied to a discrete Lagrangian and a discrete right/left Hamiltonian. For instance, choosing $r = 0$ and the rectangular rule about the end point will yield symplectic Euler-B for both the discrete Lagrangian and discrete right Hamiltonian approximation. When can we expect a fixed approximation scheme applied to a discrete Lagrangian and a discrete right Hamiltonian to yield the same method? The following theorem answers this question.

Theorem 4.1 *Assuming a regular Lagrangian, we consider a fixed approximation scheme used to construct a discrete Lagrangian, L_d, and a discrete right Hamiltonian, H_d^+. This results in two integrators, $\tilde{F}_{L_d} : (q_0, p_0) \mapsto (q_{1,L_d}, p_{1,L_d})$ and $\tilde{F}_{H_d^+} : (q_0, p_0) \mapsto (q_{1,H_d^+}, p_{1,H_d^+})$. If the discrete right Hamiltonian approximation satisfies $p_{1,H_d^+} = D_2 L_d(q_0, \hat{q}_1)$, where $\hat{q}_1$ is the approximated value of q_1, then the integrators represent the same map, i.e., $(q_{1,L_d}, p_{1,L_d}) = (q_{1,H_d^+}, p_{1,H_d^+})$.*

We have placed the proof of the above theorem in the appendix. It is important to note that even though the theorem guarantees the analytical equivalence of the integrators, this does not guarantee numerical equivalence (see [19]). It was shown in [13] that construction of a Galerkin variational integrator in terms of a discrete Lagrangian and also in terms of a discrete right Hamiltonian leads to the same symplectic map. An alternate proof of this result can be obtained from the theorem above.

4.2 Symmetric Lagrangian Taylor variational integrators

Consider the following variational derivation of the Störmer–Verlet method. Construct the discrete Lagrangian by using the trapezoid rule and approximating $\tilde{v}_0$ and $\tilde{v}_1$ with the inverse problems given by,

$$q_1 = \pi_Q \circ \Psi_h^{(1)}(q_0, \tilde{v}_0), \qquad q_0 = \pi_Q \circ \Psi_{-h}^{(1)}(q_1, \tilde{v}_1),$$

where $\Psi_h^{(r)}$ denotes the r-th order Taylor method (in this case, $r = 1$) with stepsize h with respect to the Euler–Lagrange vector field. Then, the velocity approximations are given by,

$$\tilde{v}_0 = \frac{q_1 - q_0}{h}, \qquad \tilde{v}_1 = \frac{q_1 - q_0}{h},$$

and the resulting discrete Lagrangian yields the Störmer–Verlet method,

$$L_d(q_0, q_1; h) = \frac{h}{2}\left(\left(\frac{q_1 - q_0}{h}\right)^T M\left(\frac{q_1 - q_0}{h}\right) - V(q_0) - V(q_1)\right).$$

It is well-known that Störmer–Verlet is a symmetric method, and that symmetric methods preserve important structure of time-reversible equations and are desirable for highly-oscillatory problems (see chapters V and XI of [4]). We can generalize the above approximation to yield a class of symmetric Taylor variational integrators. The approximation scheme uses a symmetric quadrature rule with weights and nodes $\{b_i, c_i\}_{i=1}^m$, and the Taylor method, and it is outlined as follows:

(i) Solve the inverse problems for $\tilde{v}_0$ and $\tilde{v}_1$,

$$q_1 = \pi_Q \circ \Psi_h^{(r)}(q_0, \tilde{v}_0), \qquad q_0 = \pi_Q \circ \Psi_{-h}^{(r)}(q_1, \tilde{v}_1). \tag{4.3}$$

(ii) Generate approximations to the quadrature nodes (q_{c_i}, v_{c_i}) via,

$$\begin{aligned} q_{c_i} &= c_i \pi_Q \circ \Psi_{c_i h}^{(r)}(q_0, \tilde{v}_0) + (1 - c_i)\pi_Q \circ \Psi_{-(1-c_i)h}^{(r)}(q_1, \tilde{v}_1) \\ v_{c_i} &= c_i \pi_{T_q Q} \circ \Psi_{c_i h}^{(r-1)}(q_0, \tilde{v}_0) + (1 - c_i)\pi_{T_q Q} \circ \Psi_{-(1-c_i)h}^{(r-1)}(q_1, \tilde{v}_1), \end{aligned}$$

where $\pi_{T_q Q}$ denotes the canonical projection onto the tangent space. Note that q_0, q_1, $\tilde{v}_0$, and $\tilde{v}_1$ are used as the approximations for their respective quadrature nodal values. Also, since the quadrature rule is assumed to be symmetric, $c_i = 1 - c_{m-i+1}$ and $b_i = b_{m-i+1}$.

(iii) Construct the discrete Lagrangian,

$$L_d(q_0, q_1; h) = h \sum_{i=1}^{m} b_i L(q_{c_i}, v_{c_i}).$$

(iv) Apply the discrete Legendre transforms to implicitly define the variational integrator,

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

Theorem 4.2 *The symmetric Taylor variational integrator is a symmetric method.*

Proof By Theorem 2.4.1 of [14], it is sufficient and necessary to show that the discrete Lagrangian of the symmetric Taylor variational integrator is self-adjoint, i.e., $L_d(q_0, q_1; h) = -L_d(q_1, q_0; -h)$. We will use (*) to denote the approximated values resulting from exchanging (q_0, q_1, h) for $(q_1, q_0, -h)$. Exchanging (q_0, q_1, h) for $(q_1, q_0, -h)$ transforms (4.3) into,

$$q_0 = \pi_Q \circ \Psi_{-h}^{(r)}(q_1, \tilde{v}_0^*), \qquad q_1 = \pi_Q \circ \Psi_h^{(r)}(q_0, \tilde{v}_1^*),$$

so that $\tilde{v}_0^* = \tilde{v}_1$ and $\tilde{v}_1^* = \tilde{v}_0$. Therefore,

$$\begin{aligned} q_{c_i}^* &= c_i \pi_Q \circ \Psi_{-c_i h}^{(r)}(q_1, \tilde{v}_0^*) + (1 - c_i)\pi_Q \circ \Psi_{(1-c_i)h}^{(r)}(q_0, \tilde{v}_1^*) \\ &= (1 - c_{m-i+1})\pi_Q \circ \Psi_{-(1-c_{m-i+1})h}^{(r)}(q_1, \tilde{v}_1) + c_{m-i+1}\pi_Q \circ \Psi_{c_{m-i+1}h}^{(r)}(q_0, \tilde{v}_0) \\ &= q_{c_{m-i+1}}. \end{aligned}$$

The second to last line follows from the fact that the quadrature rule is symmetric and therefore satisfies $1 - c_i = c_{m-i+1}$. The same steps show that $v_{c_i}^* = v_{c_{m-i+1}}$. The symmetric quadrature rule also implies that $b_i = b_{m-i+1}$, so that we have the following,

$$\begin{aligned} -L_d(q_1, q_0; -h) &= -(-h) \sum_{i=1}^{m} b_i L(q_{c_i}^*, v_{c_i}^*) \\ &= h \sum_{i=1}^{m} b_i L(q_{c_i}, v_{c_i}) = L_d(q_0, q_1; h). \end{aligned}$$

□

As the following theorem will show, the symmetric Taylor variational integrator is of order $r+1$, but only requires the derivatives of a r-order Taylor method, which makes it more efficient than the non-symmetric Taylor variational integrator, in addition to the qualitative benefits associated with its symmetry.

Theorem 4.3 *Given a regular Lagrangian, an odd r-order Taylor method, and a symmetric quadrature rule of order $r + 1$, then the resulting symmetric Taylor variational integrator is of order $r + 1$.*

Proof Lemmas 3.1 and 3.2 imply that the nodal value approximations, q_{c_i} and v_{c_i}, are of order $\mathscr{O}(h^r)$. Therefore,

$$\begin{aligned} L_d(q_0, q_1; h) &= h \sum_{i=1}^{m} b_i L(q_{c_i}, v_{c_i}) \\ &= h \sum_{i=1}^{m} b_i L(q(c_i h) + \mathscr{O}(h^r), \dot{q}(c_i h) + \mathscr{O}(h^r)) \\ &= h \sum_{i=1}^{m} b_i \left(L(q(c_i h), \dot{q}(c_i h)) + \mathscr{O}(h^r) \right) \\ &= h \sum_{i=1}^{m} b_i L(q(c_i h), \dot{q}(c_i h)) + \mathscr{O}(h^{r+1}) \\ &= \int_0^h L(q(t), \dot{q}(t)) dt + \mathscr{O}(h^{r+1}) \\ &= L_d^E(q_0, q_1; h) + \mathscr{O}(h^{r+1}). \end{aligned}$$

We have used the order of the nodal approximations, the error order of the quadrature rule, and the Lipschitz continuity of a regular Lagrangian. By Theorem 2.3.1 of [14], the resulting variational integrator, denoted by $\tilde{\Psi}_h$, is at least of order r, i.e.,

$$\begin{aligned} \tilde{\Psi}_h(q_0, p_0) &= \Phi_h(q_0, p_0) + \mathscr{O}(h^{r+1}) \\ &= \Phi_h(q_0, p_0) + C(q_0, p_0)h^{r+1} + \mathscr{O}(h^{r+2}), \end{aligned}$$

where Φ_h is the true flow of the Euler–Lagrange equations and C is a function of q_0, p_0 that is bounded on an open subset with compact closure of T^*Q over which the dynamics and integrator are defined (see [14]).

Finally, since the variational integrator is symmetric it must be even-order (see Section II.3 of [4]). The following calculation echoes the proof in [4].

$$\begin{aligned} \Phi_h(q_0, p_0) - C(q_0, p_0)h^{r+1} + \mathscr{O}(h^{r+2}) &= \Phi_h^*(q_0, p_0) - C(q_0, p_0)(-h)^{r+1} + \mathscr{O}(h^{r+2}) \\ &= \tilde{\Psi}_h^*(q_0, p_0) \\ &= \tilde{\Psi}_h(q_0, p_0) \\ &= \Phi_h(q_0, p_0) + C(q_0, p_0)h^{r+1} + \mathscr{O}(h^{r+2}), \end{aligned}$$

which implies $C(q_0, p_0)h^{r+1} = 0$, and the method is of order $r + 1$ as claimed. Note the first equality comes from the symmetry of the method and the assumption that $r + 1$ is even. The second equality follows from the definition of the adjoint as, $\Phi_h^*(q_0, p_0) = \Phi_{-h}^{-1}(q_0, p_0)$, and as a result their error expansions only may differ by a sign change. The next equality comes from the time-reversibility of the true flow, the last equality is equivalent to the method being at least of order r. □

The symmetric Taylor variational integrator is of order $r + 1$, but only requires the derivatives of a r-order Taylor method, which makes it more efficient than the non-symmetric Taylor variational integrator, in addition to the qualitative benefits associated with its symmetry. However, applying this approximation scheme to generate a discrete Hamiltonian will not directly lead to a symmetric method. Recall that the symmetric Taylor variational integrator was inspired by Störmer–Verlet, so it is likely that using this approximation scheme to generate a discrete right and left Hamiltonian will result in the discrete left and right Hamiltonian methods that are adjoint to each other. In that case, the composition of these methods should yield a symmetric method from the discrete Hamiltonian formulation. We conjecture that if an approximation scheme yields a symmetric discrete Lagrangian, then the corresponding discrete right and left Hamiltonians will be adjoint. We will explore this further in future work.

The primary method of indirectly constructing symmetric Hamiltonian Taylor variational integrators is to consider a $(2r + 1)$-order Hamiltonian Taylor variational integrator and compose it with its adjoint (see Definition 2.6 of [19]), which automatically yields a $(2r + 2)$-order symmetric method. In fact, this is precisely how Störmer–Verlet can be constructed, by first constructing the discrete right Hamiltonian with $r = 0$ and the rectangular quadrature about the end point. This yields symplectic Euler-B, which is a first-order method, then taking the adjoint yields symplectic Euler-A, and it is well-known that the composition of these two methods yields Störmer–Verlet, a symmetric second-order method. Symmetric Hamiltonian Taylor variational integrators can be constructed to arbitrary-order by following this approach.

5 Numerical implementation and experiments

We now discuss the numerical implementation of the methods introduced in this paper. Below, we present the algorithm for the Lagrangian Taylor variational integrator, and we discuss some of our observations about the implementation details. Additionally, we compare the methods to other kinds of variational integrators, and discuss their relative merits.

Algorithm Given (q_0, p_0), h, $L(q(t), \dot{q}(t))$, the Euler–Lagrange vector field, quadrature weights and nodes $\{(b_i, c_i)\}_{i=1:m}$, and the desired order of the method $r + 1$, then the Taylor variational integrator will output (q_1, p_1) and is implemented as follows:

1. Prolongate the Euler–Lagrange vector field to obtain derivatives $q^{(j)}(q(t), v(t))$ for $j = 1, \ldots, r + 1$.
2. Compute the partial derivatives $\frac{\partial q^{(j)}(q,v)}{\partial q}$ and $\frac{\partial q^{(j)}(q,v)}{\partial v}$.
3. Solve the following nonlinear system for q_1 and $\tilde{v}_0$:

$$\begin{cases} 0 = q_1 - q_0 - h\tilde{v}_0 - \sum_{j=2}^{r+1} q^{(j)}(q_0, \tilde{v}_0)\frac{h^j}{j!}, \\ 0 = p_0 + \frac{\partial L_d(q_0,q_1)}{\partial q_0}. \end{cases}$$

4. Finally, p_1 is given explicitly by,

$$p_1 = \frac{\partial L_d(q_0, q_1)}{\partial q_1}.$$

When solving the nonlinear system that arises above, the following points should be noted:

1. In general, the nonlinear system is not amenable to a fixed-point iteration, so a form of Newton's method is preferable.
2. Each iteration will require evaluation of

$$\begin{aligned} q_{c_i} &= q_0 + (c_i h)\tilde{v}_0 + \sum_{j=2}^{r} q^{(j)}(q_0, \tilde{v}_0)\frac{(c_i h)^j}{j!}, \\ v_{c_i} &= \tilde{v}_0 + \sum_{j=2}^{r+1} q^{(j)}(q_0, \tilde{v}_0)\frac{(c_i h)^{j-1}}{(j-1)!}. \end{aligned}$$

3. The following requires computing $\frac{\partial \tilde{v}_0}{\partial q_0}$,

$$\begin{aligned} -p_0 &= \frac{\partial L_d(q_0, q_1)}{\partial q_0} \\ &= h\sum_{i=1}^{m} b_i \left(\frac{\partial L(q_{c_i}, v_{c_i})}{\partial q_0} + \frac{\partial L(q_{c_i}, v_{c_i})}{\partial \tilde{v}_0}\frac{\partial \tilde{v}_0}{\partial q_0} \right)^T. \end{aligned}$$

Fortunately, this can be computed explicitly and need only be computed once at the beginning of the iteration,

$$\frac{\partial \tilde{v}_0}{\partial q_0} = \left(I + \sum_{j=2}^{r+1} \frac{\partial q^{(j)}(q_0, \tilde{v}_0)}{\partial \tilde{v}_0}\frac{(c_i h)^{j-1}}{j!} \right)^{-1} \left(\frac{-1}{h} I - \sum_{j=2}^{r+1} \frac{\partial q^{(j)}(q_0, \tilde{v}_0)}{\partial q_0}\frac{(c_i h)^{j-1}}{j!} \right).$$

4. Likewise, when solving $p_1 = \frac{\partial L_d(q_0, q_1)}{\partial q_1}$, it will be necessary to compute

$$\frac{\partial \tilde{v}_0}{\partial q_1} = \frac{1}{h}\left(I + \sum_{j=2}^{r+1} \frac{\partial q^{(j)}(q_0, \tilde{v}_0)}{\partial \tilde{v}_0}\frac{(c_i h)^{j-1}}{j!} \right)^{-1},$$

which is explicit and is composed of terms that have already been computed.

In summary, the Lagrangian Taylor variational integrator is constructed by solving the following system of equations:

$$q_1 = q_0 + h\tilde{v}_0 + \sum_{j=2}^{r+1} q^{(j)}(q_0, \tilde{v}_0)\frac{h^j}{j!}, \tag{5.1}$$

$$q_{c_i} = q_0 + (c_i h)\tilde{v}_0 + \sum_{j=2}^{r} q^{(j)}(q_0, \tilde{v}_0)\frac{(c_i h)^j}{j!}, \tag{5.2}$$

$$v_{c_i} = \tilde{v}_0 + \sum_{j=2}^{r+1} q^{(j)}(q_0, \tilde{v}_0)\frac{(c_i h)^{j-1}}{(j-1)!}, \tag{5.3}$$

$$-p_0 = h\sum_{i=1}^{m} b_i\left(\frac{\partial L(q_{c_i}, v_{c_i})}{\partial q_0} + \frac{\partial L(q_{c_i}, v_{c_i})}{\partial \tilde{v}_0}\frac{\partial \tilde{v}_0}{\partial q_0}\right)^T, \tag{5.4}$$

$$p_1 = h\sum_{i=1}^{m} b_i\left(\frac{\partial L(q_{c_i}, v_{c_i})}{\partial q_1} + \frac{\partial L(q_{c_i}, v_{c_i})}{\partial \tilde{v}_0}\frac{\partial \tilde{v}_0}{\partial q_1}\right)^T, \tag{5.5}$$

where h is the given stepsize, $r+1$ is the given order of the method, (b_i, c_i) are the given quadrature nodes and weights, and L is the given Lagrangian. Given the initial data (q_0, p_0), the system of equations is solved for the initial velocity $\tilde{v}_0$, the internal stages (q_{c_i}, v_{c_i}), and the output of the one-step method, (q_1, p_1).

Observe that good initial guesses for the nonlinear system are provided with little computational cost, by using a $(r+1)$-order Taylor method for q_1 and the Legendre transform of p_0 for $\tilde{v}_0$. Since this yields an approximate solution that is comparable in accuracy to the one obtained by the corresponding Taylor variational integrator, this yields a predictor–corrector implementation, where the Taylor variational integrator applies a symplectic correction that converges very rapidly. In general, when solving a nonlinear system as part of a symplectic method, the method becomes an almost symplectic method (see [24]) unless it is solved to within machine precision. This implies that the error tolerance of the nonlinear solver will dictate to what order the symplectic structure is preserved and consequently, how well near-energy conservation is preserved (see Fig. 1).

In practice, setting the nonlinear solver tolerance to h or h^2 times the global error of the integrator is sufficient to maintain symplecticity. For most Taylor variational integrators, the nonlinear solver with moderate tolerance converges in a few iterations, and often in one or no iterations. The symmetric Taylor variational integrator showed excellent nonlinear convergence, and only required one iteration of the nonlinear solver for the various experiments we ran.

5.1 Automatic differentiation

As with the Taylor method, an efficient general purpose implementation will require an efficient method of computing derivatives, such as automatic differentiation. For the following simulations, we used the AdiGator automatic differentiation package for MATLAB (see [16]). Implementation of a high-order Taylor variational integrator requires both the evaluation of higher time derivatives, $q^{(r+1)}(q_0, \tilde{v}_0)$, and the

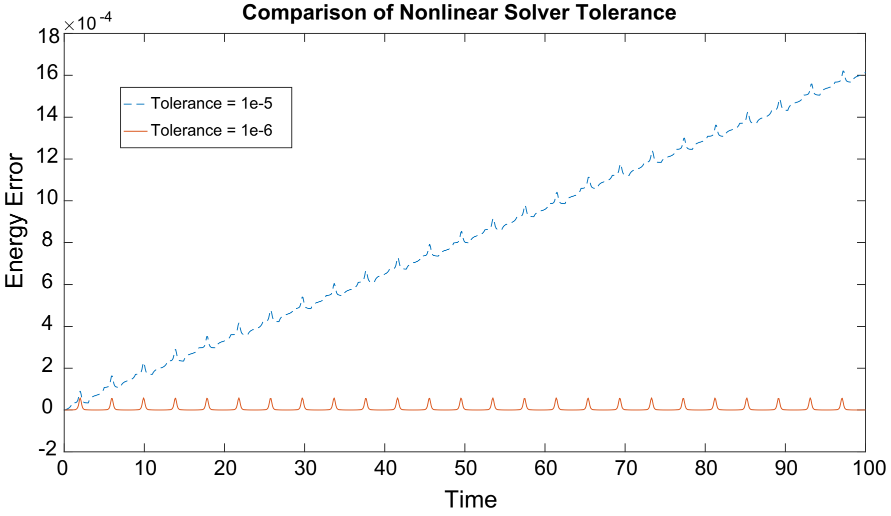

Fig. 1 The plot of the energy preservation of a 4th order Taylor variational integrator applied to the Kepler problem with two different tolerance levels for the nonlinear solver and a stepsize of 0.1. Energy drift is evident when the tolerance level is set at 10^{-5} or larger, but the drift disappears for smaller tolerance levels. The method had an average energy error around 6.5×10^{-5} for a tolerance of 10^{-6}, and an average energy error of 8.1×10^{-4} for a tolerance of 10^{-5}

evaluation of the Jacobians of the time derivatives w.r.t. q_0 and $\tilde{v}_0$. The Jacobian evaluations are the most expensive part of the method (see Fig. 2), especially for higher-dimensional systems, and for efficient high-order methods, the cost of Jacobian evaluations will need to be reduced to a level comparable to the time derivative.

Automatic differentiation greatly benefits from the way it is compiled, which means the more efficient implementations will be in languages such as Julia, Fortran or C++. In particular, Julia offers a language similar in syntax to both Python and MATLAB, combined with just-in-time compilation that has been shown to perform comparably to C++ [18]. Another aspect to consider is parallel implementation. Combining automatic differentiation and parallel computing techniques has been shown to significantly reduce computational time (see [1]).

5.2 Comparison of methods

The simulations compare the discrete Lagrangian form of the Taylor variational integrator (TVI), the discrete right Hamiltonian form of the Taylor variational integrator (HTVI), the symmetric Taylor variational integrator of 4th order (SV4), Taylor's method, and the Runge–Kutta shooting variational integrators (ShVI) (see [12]). Overall, high-order Taylor methods perform quite well in terms of computational time versus global error when high-order is desired. However, as the length of integration time becomes very large, the variational integrators begin to show their strength. Of the three variational integrators, the symmetric Taylor variational integrator is the most efficient. The first plot of Fig. 9 compares various orders of the Taylor variational inte-

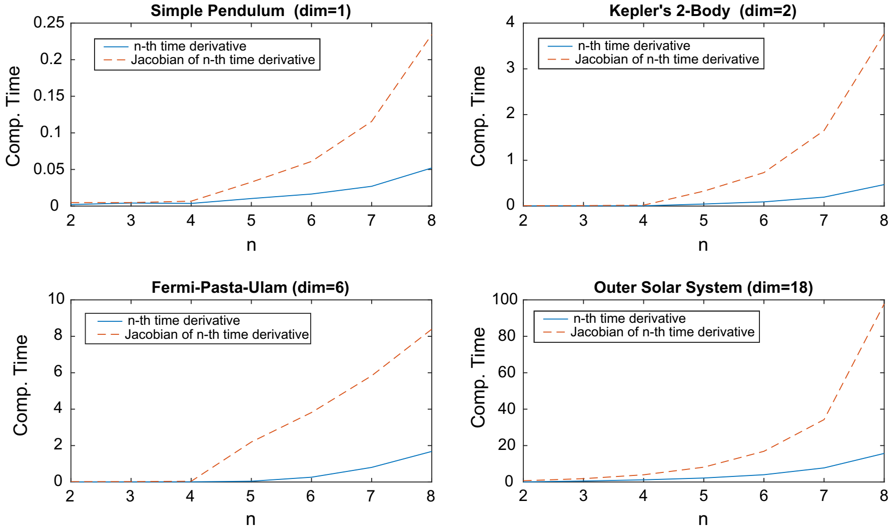

Fig. 2 The derivative order versus time plot of 100 evaluations of each derivative corresponding to 4 different models with increasing dimension. It is worth noting that the rate of growth in time needed for higher-order derivative evaluations appears to be independent of the dimension

grator applied to the Henon–Heiles model (see below for info on Henon–Heiles). The figure shows global error versus stepsize, and it is clear that the higher-order Taylor variational integrators are more efficient for a given global error.

In Figs. 5 and 9 we compare the various variational integrators. Comparison of the Lagrangian or Hamiltonian Taylor variational integrator to the Runge–Kutta shooting variational integrator does not result in a clear winner in terms of computational efficiency, although the Taylor variational integrator does appear to be more efficient on average. The computational cost for the RK shooting variational integrator must contend with increasing number of function evaluations, while the Taylor variational integrator will need higher-order derivative and Jacobian evaluations. It is well known that beyond 4th-order, Runge–Kutta (RK) methods require a higher number of stages/function evaluations, and the number of stages grows faster for vector differential equations as compared to scalar differential equations (see [2]). The number of order conditions grows quite quickly. For instance a 4th-order RK method has 8 order conditions, a 7th-order RK method has 85 order conditions, and a 25th-order method has 3,231,706,871 order conditions (see [22]). However, a 25th-order RK method only has 313 stages, so the function evaluations grow at a much slower rate. The Taylor method must contend with the increasing cost of evaluating higher-order derivatives, which for our implementation grows at a rate of 2^n, where n is the order of the derivative. For methods less than order 10 the difference in computational cost of the Taylor variational integrator and the Runge–Kutta based shooting variational integrator did not seem significant. However, the symmetric Taylor variational integrator did exhibit lower evaluation costs than the other methods. It should be noted that the most efficient implementations of the Taylor method involve variable stepsizes, and symplectic integrators are not predisposed to variable stepsizes.

Figure 5 shows error plots for both Lagrangian and Hamiltonian variational integrators. With respect to Theorem 4.1, it can be shown that the Lagrangian and Hamiltonian variational integrators presented in the plot correspond to different symplectic maps. Only the first-order Taylor variational integrators appear to generate the same map from either the Lagrangian or Hamiltonian construction. In terms of efficiency, both appear to be comparable to each other.

In Fig. 10 the global error versus stepsize is plotted for orders 4, 6, and 8 of the Lagrangian Taylor variational integrator and the Runge–Kutta based shooting variational integrator applied the the Henon–Heiles model with $H = \frac{1}{6}$. For a given stepsize, the Taylor variational integrator exhibits a smaller global error as compared to the Runge–Kutta based shooting variational integrator. It should also be noted that a $(r+1)$-order Runge–Kutta based shooting variational integrator is constructed used a $(r+1)$-order Runge–Kutta method, while a $(r+1)$-order Taylor variational integrator is constructed using a r-order Taylor method, and this plot numerically confirms the theoretical orders of these methods.

The following simulations were implemented in MATLAB.

5.3 Simple pendulum

Consider the simple pendulum with unit mass and length in a gravitational field with $g = -9.8\,\mathrm{m/s}^2$, where q is parametrized by the angle between the y-axis and the pendulum. The corresponding Lagrangian is,

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^2 - g(1 - \cos(q)).$$

The Euler–Lagrange equation yields,

$$\ddot{q} = -g\sin(q).$$

In Fig. 3, the level sets of the corresponding Hamiltonian are compared to the trajectories generated by a 2nd-order Taylor variational integrator (TVI2) (see Example 2). The numerical solutions appear nearly identical to the level sets of the Hamiltonian, which indicates that the variational integrator exhibited good energy behavior for a variety of initial conditions.

The simulation in Fig. 4 used initial conditions $(q_0, p_0) = (\frac{\pi}{2}, 0)$. The 6th-order Taylor variational integrator performed well at a stepsize of $h = 0.5$, while the 6th-order Taylor method failed to generate a reasonable approximation for this stepsize. The ability of the Taylor variational integrator to perform well at larger stepsizes may give it an advantage over traditional Taylor methods.

In Fig. 5, we compare various types of Taylor variational integrators against the shooting-based variational integrator (ShVI). The plots compare the energy error versus computational time for methods of various order. It is clear the symmetric Taylor variational integrator (SV4) is the most efficient in this respect, but it is not so clear whether the non-symmetric Taylor variational integrators (TVI and HTVI) are more efficient than ShVI.

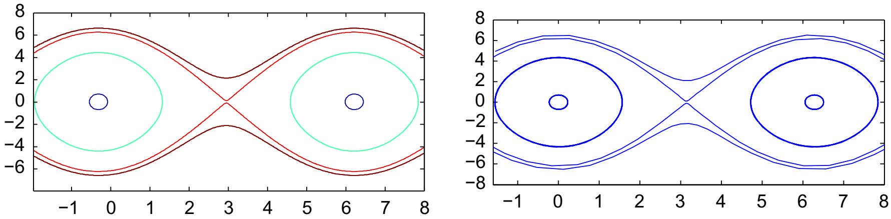

Fig. 3 **a** The level sets of the Hamiltonian of the simple pendulum corresponding to a variety of initial conditions. **b** The trajectories generated by TVI2 using the same initial conditions with a stepsize $h = 0.1$ for the time interval [0, 20]

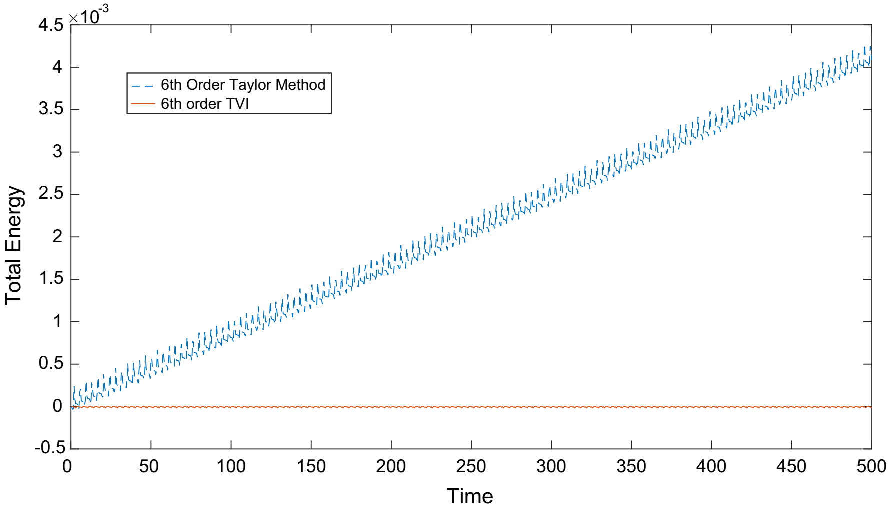

Fig. 4 A plot of the Simple Pendulum total energy versus time of the sixth-order integrators TVI6 and Taylor's method for a stepsize of $h = 0.5$. At this stepsize and time interval, Taylor's method has significant energy drift, and as a result its accuracy suffers

5.4 Kepler's planar 2-body problem

Consider two bodies interacting under mutual gravity and set one body as the center of the coordinate system (see [4]). Thus, constraining them to lie in a plane, we have Kepler's planar 2-body problem with corresponding Lagrangian,

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}(\dot{q}_1^2 + \dot{q}_2^2) + (q_1^2 + q_2^2)^{-1/2}.$$

Note here we are using q_1 and q_2 as the first and second components of $\mathbf{q}$. This in turn gives us the Euler–Lagrange equations,

$$\ddot{\mathbf{q}} = \begin{bmatrix} \frac{-q_1}{(q_1^2+q_2^2)^{3/2}} \\ \frac{-q_2}{(q_1^2+q_2^2)^{3/2}} \end{bmatrix}.$$

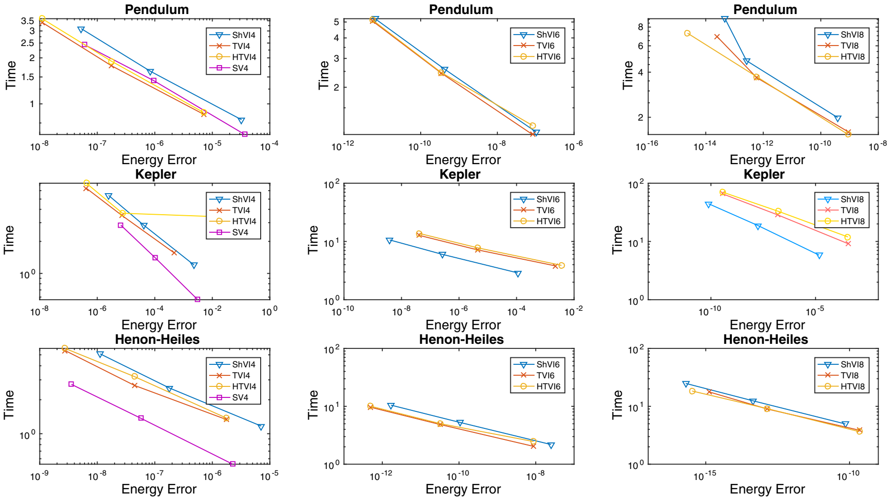

Fig. 5 Plots of the average energy error versus computational time for the various variational integrators. The 4th-order symmetric Taylor variational integrator (SV4) is the clear winner in terms of efficiency, while comparisons of TVI, HTVI, and ShVI are mixed

Our simulations used initial conditions $\mathbf{q}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{p}_0 = \begin{bmatrix} 0 \\ 0.8 \end{bmatrix}$. Figure 6 compares various Taylor variational integrators to Taylor methods of the same order using a stepsize of $h = 0.25$. The trajectories of the Taylor methods for this stepsize behave poorly, while variational integrators show good qualitative performance.

Figure 7 compares the Störmer–Verlet method (SV) to the discrete Hamiltonian composition method (SVHd) discussed in Sect. 4.1. Given that the Störmer–Verlet method is explicit, while SVHd is implicit, it is no surprise that the Störmer–Verlet method has lower computational cost. SVHd does exhibit lower energy error and performs slightly better qualitatively, so we will examine the two methods further on a nonseparable problem.

5.5 Nonseparable Hamiltonian

Consider the Hamiltonian, $H(q, p) = \left(1 + \frac{p^2}{2}\right)^2 \left(1 + q^2\right)$, which is nonseparable Hamiltonian and near the origin exhibits behavior similar to the harmonic oscillator. Using two different sets of initial conditions, $(q_0, p_0) = (\frac{1}{4}, 0)$ and $(q_0, p_0) = (\frac{1}{4}, 2)$ we compare the performance of Störmer–Verlet and SVHd in Fig. 8. Störmer–Verlet shows better performance, in terms of global error versus computational time, for both initial conditions.

SVHd can be interpreted as first applying the symmetric Taylor variational integrator construction to build a discrete right Hamiltonian and a discrete left Hamiltonian, and then SVHd comes from composing the two resulting variational integrators. Störmer–Verlet can be interpreted as first constructing Hamiltonian variational integrator with $r = 0$ and the trapezoid rule, then composing the resulting variational

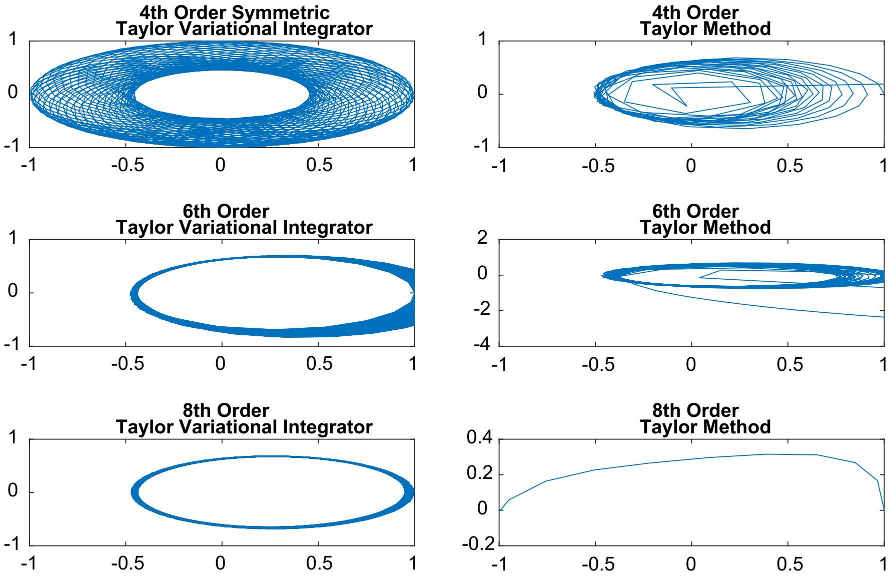

Fig. 6 Position plots of Kepler's planar 2-body problem as generated by various integrators with a stepsize of $h = 0.25$ over a time interval of [0, 250]. The Taylor variational integrators exhibit close to the correct behavior, while the various Taylor methods all fail to capture the behavior of the system

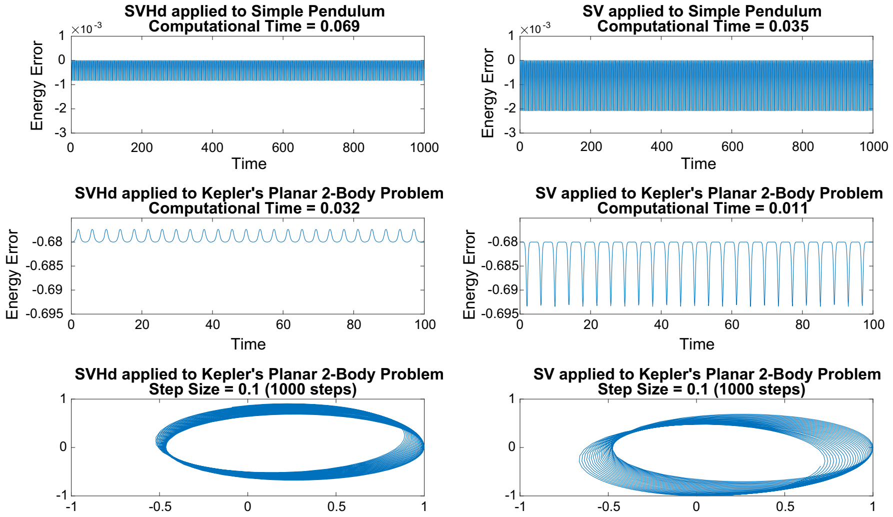

Fig. 7 This plot compares the performance of Störmer–Verlet (SV) and the discrete Hamiltonian composition method (SVHd) from Sect. 4.1. SVHd exhibits a much smaller amplitude in the energy error, as compared to SV, but the implicit nature of SVHd is reflected in the increased computational cost

integrator with the adjoint. The results of this experiment lend support in favor of the second approach for constructing symmetric variational integrators on the Hamiltonian side.

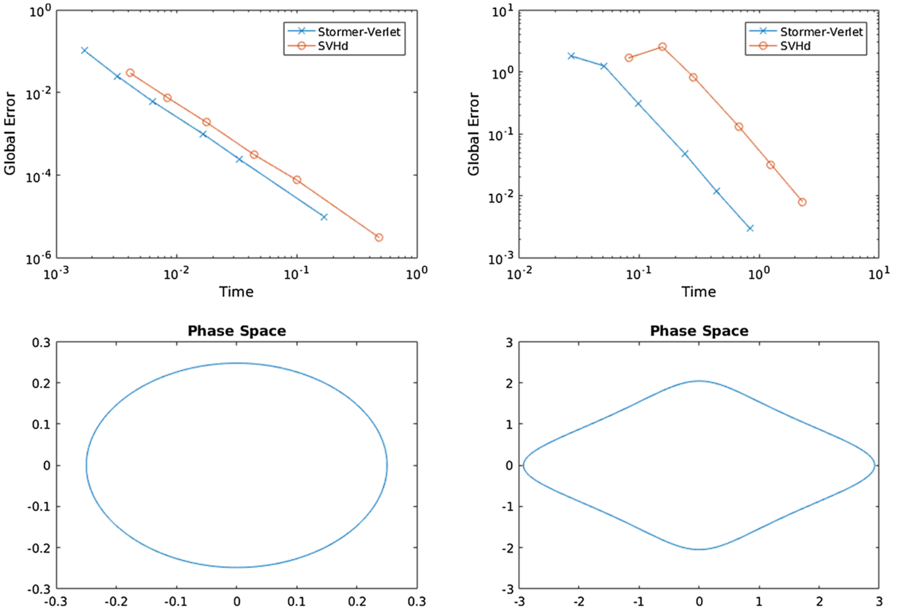

Fig. 8 This plot compares the performance of Störmer–Verlet (SV) and the discrete Hamiltonian composition method SVHd applied to a nonseparable Hamiltonian with two different initial conditions. It is clear that Störmer–Verlet is the more efficient method in both cases

5.6 Henon–Heiles model

The Henon–Heiles model attempts to capture the dynamics of a galaxy with cylindrical symmetry (see [4] for more info). The Hamiltonian is given by, $H(p,q) = \frac{1}{2}(p_1^2 + p_2^2) + U(q)$, where $U(q) = \frac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3$. The corresponding Euler–Lagrange equation is,

$$\ddot{\mathbf{q}} = \begin{bmatrix} -q_1 - 2q_1q_2 \\ -q_2 - q_1^2 + q_2^2 \end{bmatrix}.$$

We compare the global error versus computational time for the Taylor variational integrators and Runge–Kutta shooting-based variational integrators in Fig. 9, and the global error versus stepsize in Fig. 10. It is known that the dynamics become chaotic at higher energy levels. The following simulations were conducted with an initial energy level of $H_0 = \frac{1}{12}$ (see Fig. 11), which corresponds to a chaotic system.

In Fig. 11, we compare the 6th-order Taylor variational integrator (TVI6), the 6th-order Runge–Kutta shooting-based variational integrator (ShVI6), and the 4th-order symmetric Taylor variational integrator (SV4) applied to the Henon–Heiles model with $H_0 = \frac{1}{12}$. For global errors between 10^{-1} and 10^{-5}, SV4 is the more efficient method. As the global error decreases below 10^{-5}, TVI6 and ShVI6 appear to be the

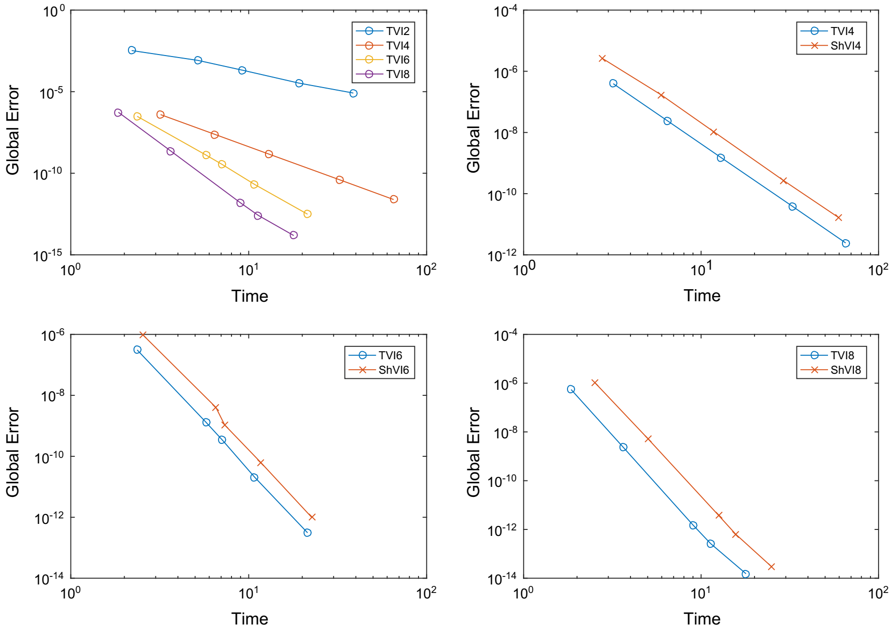

Fig. 9 The first plot compares the global error versus computational time for the second, fourth, sixth, and eighth order Taylor variational integrator (TVI) applied to the Henon–Heiles model with $H_0 = \frac{1}{6}$. The other plots compare the global error versus computational time for the TVI and the Runge–Kutta shooting-based variational integrator (ShVI). For this model the TVI has the edge in terms of computational efficiency

more efficient methods, but a 6th-order symmetric Taylor variational integrator would most likely offer more efficiency gains.

5.7 Fermi–Pasta–Ulam model

The Fermi–Pasta–Ulam (FPU) model has a particularly distinguished place in the history of numerical simulations and nonlinear dynamics (see [3]). We apply the modified model as outlined in [4], consisting of a sequence of 6 mass points, fixed at both ends connected on opposite sides by a series of soft nonlinear springs and stiff linear springs. Letting $\{q_i, p_i\}_{i=1}^{6}$ denote the displacements and velocities of the mass points, the corresponding Hamiltonian is given by,

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{3}(p_{2i-1}^2 + p_{2i}^2) + \frac{\omega^2}{4}\sum_{i=1}^{6}(q_{2i} - q_{2i-1})^2 + \sum_{i=0}^{6}(q_{2i+1} - q_{2i})^4,$$

where $\omega = 50$. By using the change of variables,

$$x_{0,i} = (q_{2i} + q_{2i-1})/\sqrt{2}, \qquad x_{1,i} = (q_{2i} - q_{2i-1})/\sqrt{2},$$
$$y_{0,i} = (p_{2i} + p_{2i-1})/\sqrt{2}, \qquad y_{1,i} = (p_{2i} - p_{2i-1})/\sqrt{2},$$

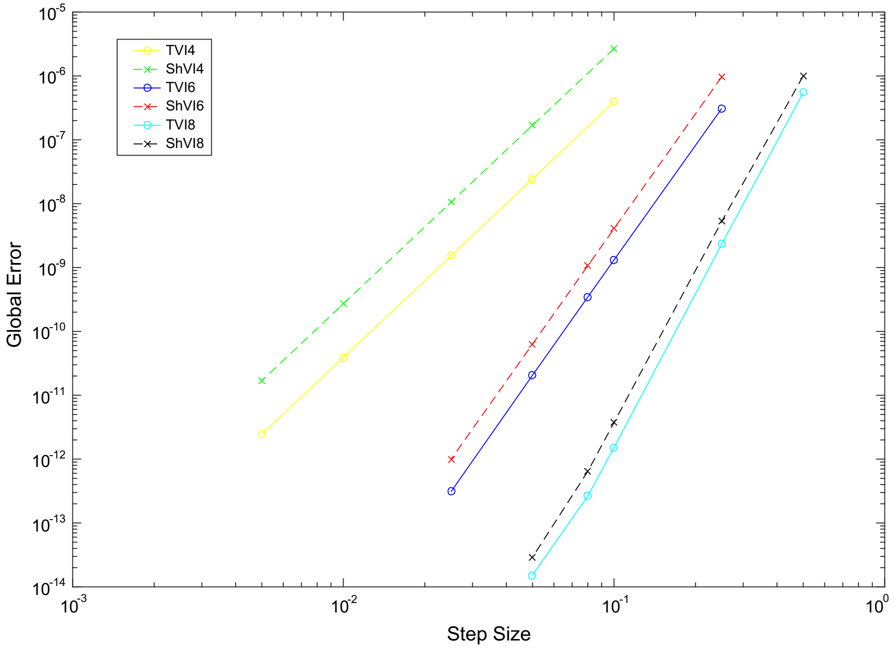

Fig. 10 A plot of the global error versus stepsize for orders 4, 6 and 8 of both the Lagrangian Taylor variational integrator and the Runge–Kutta shooting variational integrator applied to the Henon–Heiles model with $H_0 = \frac{1}{6}$. Note that the Taylor variational integrator consistently showed a lower error for a given stepsize and order

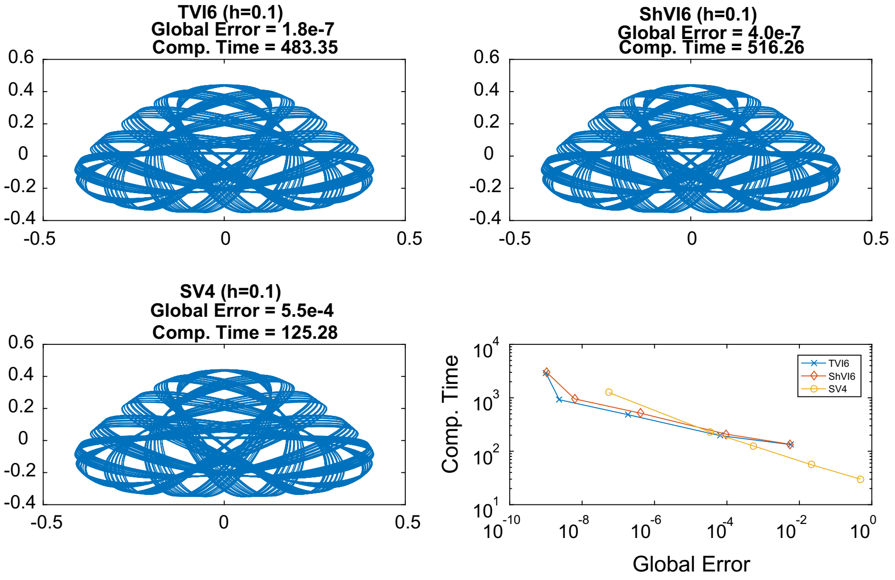

Fig. 11 The Henon–Heiles model simulated over the time interval [0, 1000]. The bottom right plot compares the global error versus computational time of the 6th-order Taylor variational integrator (TVI6), the 6th-order Runge–Kutta based shooting variational integrator (ShVI6), and the 4th-order symmetric Taylor method (SV4)

the resulting Hamiltonian system has a nearly conserved quantity $I = I_1 + \cdots + I_3$, where

$$I_j(x_{1,j}, y_{1,j}) = \frac{1}{2}(y_{1,j}^2 + \omega^2 x_{1,j}^2)$$

is the energy of the jth stiff spring. Despite the significant energy exchange between individual springs, the total oscillatory energy, I, remains near constant. Our simulations used initial values of,

$$\begin{bmatrix} x_{0,1} \\ x_{0,2} \\ x_{0,3} \\ x_{1,1} \\ x_{1,2} \\ x_{1,3} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1/\omega \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} y_{0,1} \\ y_{0,2} \\ y_{0,3} \\ y_{1,1} \\ y_{1,2} \\ y_{1,3} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Figure 12 compares the Störmer–Verlet method to SVHd. The first couple of plots use a stepsize of $h = 0.03$, which is in the region of the linear stability for Störmer–Verlet (i.e. $h\omega = 1.5 < 2$), but with the additional nonlinearity this stepsize is not small enough for either method. SVHd does appear to be qualitatively more accurate, but neither method does well at this stepsize. For $h = 0.01$, both methods give a much better qualitative representation of the system, but their global errors are still too large to be considered accurate. None of the methods in this paper are appropriate for a highly-oscillatory model such as the FPU model. For an accurate solution, one should consider either the IMEX method (see [23]) or Filon-type methods (see [7]). The

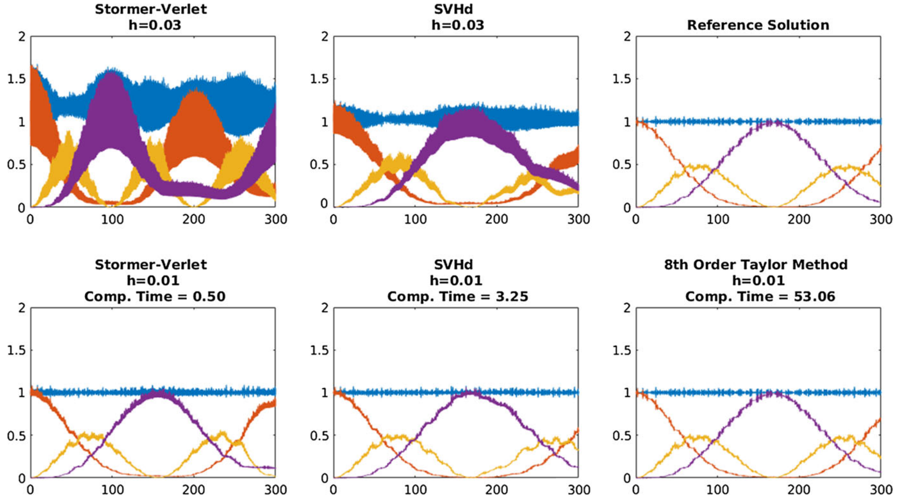

Fig. 12 A comparison of Störmer–Verlet, SVHd and the 8th-order Taylor method for the Fermi–Pasta–Ulam model. For $h = 0.03$, the Störmer–Verlet method is on the cusp of being linearly unstable. For $h = 0.01$, the methods all present a similar picture to the reference solution, but their global errors are quite large and none of them exhibit good accuracy

combination of exponential type integrators with symplectic and energy-preserving integrators was also recently considered in [20].

5.8 Outer solar system

Consider the motion of the five outer planets (including Pluto) relative to the sun. The corresponding Hamiltonian for this N-body problem is given by,

$$H(p, q) = \frac{1}{2}\sum_{i=0}^{5}\frac{1}{m_i}p_i^T p_i - G\sum_{i=1}^{5}\sum_{j=0}^{i-1}\frac{m_i m_j}{\|q_i - q_j\|},$$

where $G = 2.95912208286 \cdot 10^{-4}$. The initial data and masses is taken from Section 1.2.4 of [4], and corresponds to September 5, 1994 at 0h00. In Fig. 13, we compare the 4th and 6th-order Taylor variational integrators to the 4th and 6th-order Taylor methods. The simulations was over the time period [0, 200000], and the stepsize was $h = 400$ (days). The 4th-order methods did not produce a useful simulation at this stepsize, but both 6th-order integrators give a good representation of the system.

6 Conclusions and future directions

The Taylor variational integrators provide a way to build high-order symplectic integrators and include some of the classic symplectic integrators as special cases, i.e., symplectic Euler and Störmer–Verlet. This provides a framework for importing the large body of literature on the efficient construction of high-order Taylor integrators in order to construct similarly high-order symplectic integrators.

In particular, these methods can be viewed as a symplectic correction to higher-order Taylor methods that typically converges in a small number of iterations. By viewing these as predictor–corrector methods, one can interpolate between Taylor methods and Taylor variational integrators, and it would be interesting to see the extent to

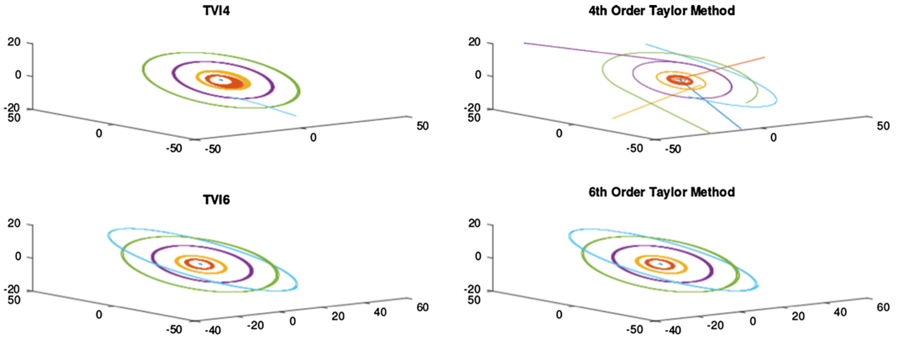

Fig. 13 The sun and 5 outer planets simulated over the time interval [0, 200000] with a stepsize of $h = 400$ (days). The stepsize is too large for the 4th-order methods to give a qualitatively accurate representation, but both 6th-order methods performed well qualitatively

which a fixed number of iterations of the symplectic corrector can improve upon the performance of Taylor integrators for realistic problems.

The numerical simulations demonstrate that the geometric structure-preserving properties of symplectic integrators can be important for achieving numerical stability of long time simulations, so it should be of great interest to the computational astrophysics community to combine the high-order accuracy of high-order Taylor integrators with the geometric structure-preserving properties of variational integrators.

The most efficient implementations of the Taylor method utilize a variable stepsize, and extending variable stepsizes to the variational integrator framework is an area that deserves continued research. We are currently considering an approach based on the combination of Hamiltonian variational integrators and the Poincaré transformation that is quite promising. In particular, we note that the use of Hamiltonian as opposed to Lagrangian variational integrators is critical, as the Poincaré transformed Hamiltonian is degenerate, and there is no corresponding Lagrangian formulation.

A Detailed proofs

Given an Euler–Lagrange equation of the form,

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

we denote the exact solution of the Euler–Lagrange boundary-value problem with boundary conditions (q_0, q_1) by $(q(t), v(t))$. We seek an estimate of the true initial velocity, v_0, for the corresponding Euler–Lagrange initial-value problem, with order of accuracy r. Let us denote this estimate by $\tilde{v}_0$. Given a one-step method, $\Psi_h^{(r+1)} : TQ \to TQ$, with order of accuracy $r + 1$, we solve for the initial velocity $\tilde{v}_0$, such that,

$$\pi_Q \circ \Psi_h^{(r+1)}(q_0, \tilde{v}_0) = q_1, \tag{A.1}$$

where $\pi_Q : TQ \to Q$ is the canonical projection. Let $\Phi_h : TQ \to TQ$ be the exact time-h flow map of the Euler–Lagrange initial-value problem. By definition, the exact Euler–Lagrange flow applied to the initial condition (q_0, v_0) is a solution of the Euler–Lagrange boundary-value problem with boundary conditions (q_0, q_1), where

$$\pi_Q \circ \Phi_h(q_0, v_0) = q_1. \tag{A.2}$$

Consider a Taylor method with order of accuracy r and $r + 1$,

$$\Psi_h^{(r)}(q_0, \tilde{v}_0) = \left(\sum_{k=0}^{r} \frac{h^k}{k!} q^{(k)}(0), \sum_{k=1}^{r+1} \frac{h^{k-1}}{(k-1)!} q^{(k)}(0) \right) \tag{A.3}$$

and

$$\Psi_h^{(r+1)}(q_0, \tilde{v}_0) = \left(\sum_{k=0}^{r+1} \frac{h^k}{k!} q^{(k)}(0), \sum_{k=1}^{r+2} \frac{h^{k-1}}{(k-1)!} q^{(k)}(0) \right), \tag{A.4}$$

where $q^{(k)}(0)$ is calculated by considering the prolongations of the Euler–Lagrange vector field, and evaluating it at $(q_0, \tilde{v}_0)$. An analogous approach, involving the prolongation of the Euler–Lagrange vector field at both the initial and final time, which can be viewed as a two-point Taylor method, was used to develop a prolongation–collocation variational integrator in [11].

Lemma A.1 *$\tilde{v}_0$ as defined by,* (A.1) *and* (A.4)*, approximates v_0 to at least $\mathscr{O}(h^{r+1})$.*

Proof Solving $\pi_Q \circ \Psi_h^{(r+1)}(q_0, \tilde{v}_0) = q_1$ for $\tilde{v}_0$ yields,

$$\tilde{v}_0 = \frac{q_1 - q_0}{h} - \sum_{k=1}^{r} \frac{h^k}{(k+1)!}(f^{(k-1)}(q_0, \tilde{v}_0, 0)).$$

Since the exact solution $q(t) \in C^{r+2}([0,h])$, using Taylor's theorem, we have,

$$q_1 = q_0 + v_0 h + \sum_{k=2}^{r+1} \frac{h^k}{k!} f^{(k-2)}(q_0, v_0, 0) + R_{r+1}(h).$$

Solving for v_0 yields,

$$v_0 = \frac{q_1 - q_0}{h} - \sum_{k=1}^{r} \frac{h^k}{(k+1)!} f^{(k-1)}(q_0, v_0, 0) - \frac{R_{r+1}(h)}{h}.$$

Now evaluating the norm of the difference we have,

$$\|\tilde{v}_0 - v_0\| = \left\| - \sum_{k=1}^{r} \frac{h^k}{(k+1)!}(f^{(k-1)}(q_0, \tilde{v}_0, 0) - f^{(k-1)}(q_0, v_0, 0)) + \frac{R_{r+1}(h)}{h} \right\|.$$

Since $q(t) \in C^{r+2}([0,h])$ each of $f^{(i-1)}$ is Lipschitz continuous in its arguments for $i = 1, 2, \ldots, r$. Let M_i be the Lipschitz constant for $f^{(i-1)}$ over the compact interval $[0, C]$ with respect to velocity, and $C > 0$ can be chosen so that M_i, $i = 1, 2, \ldots, r$, is bounded. Using the triangle inequality, we have,

$$\|\tilde{v}_0 - v_0\| \leq \sum_{k=1}^{r} \frac{h^k}{(k+1)!} M_k \|\tilde{v}_0 - v_0\| + \left\| \frac{R_{r+1}(h)}{h} \right\|.$$

Rearranging, we have,

$$\|\tilde{v}_0 - v_0\| \left(1 - \sum_{k=1}^{r} \frac{h^k}{(k+1)!} M_k \right) \leq \left\| \frac{R_{r+1}(h)}{h} \right\| \leq \frac{\mathscr{O}(h^{r+2})}{h} = \mathscr{O}(h^{r+1}).$$

By continuity, there exists $\tilde{C}$ satisfying $0 < \tilde{C} < C$, such that for all h satisfying $0 < h < \tilde{C}$, the term inside the parenthesis on the leftmost expression is positive and bounded away from zero. That concludes the proof. □

Remark A.1 It is worth noting that a similar proof may be given for any $(r+1)$-order one-step method. This is due to the fact that any $(r+1)$-order one-step method agrees with the $(r+1)$-order Taylor's method up to a local truncation error of order $\mathscr{O}(h^{r+2})$. Thus, the only change in the error term in the proof would be to replace $R_{r+1}(h)$ by the sum of the local truncation error of the one-step method and $R_{r+1}(h)$, which are both $\mathscr{O}(h^{r+2})$. Thus, this result can be generalized to any one-step method of the desired order.

Using this result, we can show that starting our r-order Taylor method at $\tilde{v_0}$, rather than at v_0, will not affect the order of accuracy of the method.

Lemma A.2 *A r-order Taylor method, defined by (A.3), with initial conditions $(q_0, \tilde{v}_0)$, where $\tilde{v}_0$ is defined by (A.1), is accurate to at least $\mathscr{O}(h^{r+1})$ for the Euler–Lagrange boundary-value problem with boundary conditions (q_0, q_1).*

Proof As before, we denote the solution to the Euler–Lagrange boundary-value problem with boundary condition (q_0, q_1) by $(q(t), v(t))$ for $t \in [0, h]$. This solution also satisfies the Euler–Lagrange initial-value problem with initial conditions (q_0, v_0), where v_0 satisfies (A.2). We denote the solution of the Euler–Lagrange initial-value problem with initial conditions $(q_0, \tilde{v}_0)$ by $(\tilde{q}(t), \tilde{v}(t))$. Let $(q_d(t), v_d(t))$ denote the values generated by r-order Taylor method with initial conditions $(q_0, \tilde{v}_0)$. Noting that the Euler–Lagrange initial-value problem is well-posed, we denote the Lipschitz constant with respect to initial velocity by M.

$$\|(q(t), v(t)) - (\tilde{q}(t), \tilde{v}(t))\| \leq M\|v_0 - \tilde{v}_0\| \leq \mathscr{O}(h^{r+1}).$$

Combining this inequality with our r-order method yields,

$$\begin{aligned}\|(q(t), v(t)) - (q_d(t), v_d(t))\| &= \|(q(t), v(t)) - (\tilde{q}(t), \tilde{v}(t)) \\ &\quad + (\tilde{q}(t), \tilde{v}(t)) - (q_d(t), v_d(t))\| \\ &\leq \|(q(t), v(t)) - (\tilde{q}(t), \tilde{v}(t))\| \\ &\quad + \|(\tilde{q}(t), \tilde{v}(t)) - (q_d(t), v_d(t))\| \\ &\leq \mathscr{O}(h^{r+1}),\end{aligned}$$

where we used the triangle inequality, and the fact that the local truncation error of a r-order Taylor method is $\mathscr{O}(h^{r+1})$ to bound the second term in line two, since $(\tilde{q}(t), \tilde{v}(t))$ and $(q_d(t), v_d(t))$ correspond to the exact solution and rth order Taylor approximation, respectively, of the Euler–Lagrange initial-value problem with initial data $(q_0, \tilde{v}_0)$. □

Theorem A.1 *Assuming a regular Lagrangian, we consider a fixed approximation scheme used to construct a corresponding discrete Lagrangian, L_d, and a discrete right Hamiltonian, H_d^+. This results in two integrators, $\tilde{F}_{L_d} : (q_0, p_0) \mapsto (q_{1,L_d}, p_{1,L_d})$ and $\tilde{F}_{H_d^+} : (q_0, p_0) \mapsto (q_{1,H_d^+}, p_{1,H_d^+})$. If the discrete right Hamiltonian approximation satisfies $p_{1,H_d^+} = D_2 L_d(q_0, \hat{q}_1)$, where $\hat{q}_1$ is the approximated value of q_1, then the integrators represent the same map, i.e., $(q_{1,L_d}, p_{1,L_d}) = (q_{1,H_d^+}, p_{1,H_d^+})$.*

Proof Let $\hat{p}_0$ be defined by $-\hat{p}_0 = D_1 L_d(q_0, \hat{q}_1)$, where we consider $\hat{q}_1$ as an independent variable. The discrete right Hamiltonian is given by,

$$H_d^+(q_0, p_{1,H_d^+}) = p_{1,H_d^+}^T \hat{q}_1 - L_d(q_0, \hat{q}_1).$$

Note that here $\hat{q}_1$ is being considered as a function of q_0 and p_{1,H_d^+}, as defined implicitly by the assumption $p_{1,H_d^+} = D_2 L_d(q_0, \hat{q}_1)$. Then

$$\begin{aligned} p_0 &= D_1 H_d^+(q_0, p_{1,H_d^+}) \\ &= \frac{\partial \hat{q}_1}{\partial q_0}^T p_{1,H_d^+} - \left(D_1 L_d(q_0, \hat{q}_1) + \frac{\partial \hat{q}_1}{\partial q_0}^T D_2 L_d(q_0, \hat{q}_1) \right) \\ &= \frac{\partial \hat{q}_1}{\partial q_0}^T \left(p_{1,H_d^+} - D_2 L_d(q_0, \hat{q}_1) \right) + \hat{p}_0 \\ &= \hat{p}_0, \end{aligned}$$

where the last line follows by the assumption $p_{1,H_d^+} = D_2 L_d(q_0, \hat{q}_1)$. Therefore, $p_0 = \hat{p}_0$, which then implies $-p_0 = D_1 L_d(q_0, \hat{q}_1)$ and consequently $\hat{q}_1 = q_{1,L_d}$. Applying the next discrete Legendre transform yields,

$$\begin{aligned} q_{1,H_d^+} &= D_2 H_d^+(q_0, p_{1,H_d^+}) \\ &= \frac{\partial \hat{q}_1}{\partial p_{1,H_d^+}}^T p_{1,H_d^+} + \hat{q}_1 - \left(\frac{\partial \hat{q}_1}{\partial p_{1,H_d^+}}^T D_2 L_d(q_0, \hat{q}_1) \right) \\ &= \frac{\partial \hat{q}_1}{\partial p_{1,H_d^+}}^T \left(p_{1,H_d^+} - D_2 L_d(q_0, \hat{q}_1) \right) + \hat{q}_1 \\ &= \hat{q}_1. \end{aligned}$$

Therefore, $\hat{q}_1 = q_{1,H_d^+}$, which implies $q_{1,L_d} = q_{1,H_d^+}$. Now we have,

$$\begin{aligned} p_{1,H_d^+} &= D_2 L_d(q_0, \hat{q}_1) \\ &= D_2 L_d(q_0, q_{1,L_d}) \\ &= p_{1,L_d}. \end{aligned}$$

□

References

1. Bucker, H.M., Lang, B., Mey, D. Bischof, C.: Bringing together automatic differentiation and OpenMP. In: ICS '01 Proceedings of the 15th International conference on Supercomputing, pp. 246–251 (2001)
2. Butcher, J.C.: On fifth and sixth order Runge–Kutta methods: order conditions and order barriers. Can. Appl. Math. Q. **17**(3), 433–445 (2009)

3. Ford, J.: The Fermi–Pasta–Ulam problem: paradox turns discovery. Phys. Rev. **213**(5), 271–310 (1992)
4. Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration, Volume 31 of Springer Series in Computational Mathematics, 2nd edn. Springer, Berlin (2006)
5. Haro, A.: Automatic differentiation methods in computational dynamical systems. In: New Directions Short Course: Invariant Objects in Dynamical Systems and their Application, Minneapolis, June 2011. Institute for Mathematics and its Applications (IMA). http://www.maia.ub.edu/~alex/ima/ima1.pdf
6. Iserles, A.: A First Course in the Numerical Analysis of Differential Equations. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge (2009)
7. Iserles, A., Norsett, S.P.: On quadrature methods for highly oscillatory integrals and their implementation. BIT Numer. Math. **44**(4), 755–772 (2004)
8. Jorba, A., Zou, M.: A software package for the numerical integration of ODEs by means of high-order Taylor methods. Exp. Math. **14**, 99–117 (2005)
9. Keller, H.B.: Numerical Methods for Two-Point Boundary Value Problems. Dover Publications Inc., New York (1992)
10. Lall, S., West, M.: Discrete variational Hamiltonian mechanics. J. Phys. A **39**(19), 5509–5519 (2006)
11. Leok, M., Shingel, T.: Prolongation–collocation variational integrators. IMA J. Numer. Anal. **32**(3), 1194–1216 (2012)
12. Leok, M., Shingel, T.: General techniques for constructing variational integrators. Front. Math. China **7**(2), 273–303 (2012)
13. Leok, M., Zhang, J.: Discrete Hamiltonian variational integrators. IMA J. Numer. Anal. **31**(4), 1497–1532 (2011)
14. Marsden, J.E., West, M.: Discrete mechanics and variational integrators. Acta Numer. **10**, 357–514 (2001)
15. Neidinger, R.D.: Introduction to automatic differentiation and MATLAB object-oriented programming. SIAM Rev. **52**(3), 545–563 (2010)
16. Patterson, M., Weinstein, M., Rao, A.: An efficient overloaded method for computing derivatives of mathematical functions in MATLAB. ACM Trans. Math. Softw. **39**(3), 17 (2013)
17. Rall, L.B.: Automatic Differentiation: Techniques and Applications. Lecture Notes in Computer Science, vol. 120. Springer, Berlin (1981)
18. Revels, J., Lubin, M., Papamarkou, T.: Forward-mode automatic differentiation in Julia. In: AD2016—7th International Conference on Algorithmic Differentiation (2016)
19. Schmitt, J.M., Leok, M.: Properties of Hamiltonian variational integrators. IMA J. Numer. Anal. **36**, drx010 (2017). https://doi.org/10.1093/imanum/drx010
20. Shen, X., Leok, M.: Geometric exponential integrators. arXiv:1703.00929 [math.NA] (2017)
21. Simo, C.: Principles of Taylor methods for analytic, non-stiff, ODE. In: Advanced Course on Long Time Integration. University of Barcelona (2007)
22. Sommeijer, B.P.: An explicit Runge–Kutta method of order twenty-five. CWI Q. **11**(1), 75–82 (1998)
23. Stern, A., Grinspun, E.: Implicit-explicit variational integration of highly oscillatory problems. Multiscale Model. Simul. **7**(4), 1779–1794 (2009)
24. Tan, X.: Almost symplectic Runge–Kutta schemes for Hamiltonian systems. J. Comput. Phys. **203**, 250–273 (2005)