UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Properties of Hamiltonian Variational Integrators**

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Jeremy M. Schmitt

Committee in charge:

Professor Melvin Leok, Chair
Professor Henry Abarbanel
Professor Randolph Bank
Professor Michael Holst
Professor Petr Krysl

2017

The dissertation of Jeremy M. Schmitt is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2017

DEDICATION

To my wife Birthe, my mother, my father, my brothers and sister, and Zoomie.

# EPIGRAPH

*Nothing in life is to be feared; it is only to be understood.*
—Marie Curie

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My advisor, Professor Melvin Leok, has always been generous with his time, and his intuition and mathematical rigor helped guide my research. The mathematics department at the University of California, San Diego is run by an amazing group of people from staff to faculty, which made this process both possible and enjoyable. The University of California, San Diego continues to provide an excellent educational experience accessible to those from many walks of life, and I am grateful for the opportunities I've been given. My fellow math grad students and officemates were always supportive and full of helpful advice. Finally, I must thank my wife for her patience and support during this process.

Chapter 2, in full, is a reprint of the material that has been accepted for publication by IMA Journal of Numerical Analysis, 2017. Schmitt, Jeremy; Leok, Melvin, Oxford University Press, 2017. The dissertation author was the primary investigator and author of this material.

Chapter 3, in full, is a reprint of the material that has been submitted for publication to BIT Numerical Mathematics, 2017. Schmitt, Jeremy; Shingel, Tatianna; Leok, Melvin, Springer, 2017. The dissertation author was the primary investigator and author of this material.

Chapter 4, in full, is currently being prepared for submission for publication of the material. Schmitt, Jeremy; Leok, Melvin. The dissertation author was the primary investigator and author of this material.

## VITA

| | |
|---|---|
| 2010 | B. A. in Joint Mathematics/Economics, University of California, San Diego |
| 2014 | M. A. in Applied Mathematics, University of California, San Diego |
| 2015 | C. Phil. in Mathematics, University of California, San Diego |
| 2017 | Ph. D. in Mathematics, University of California, San Diego |

## PUBLICATIONS

J.M. Schmitt, M. Leok, *Properties of Hamiltonian Variational Integrators,* to appear in IMA Journal of Numerical Analysis, 2017. Early Access

J.M. Schmitt, T. Shingel, M. Leok, *Lagrangian and Hamiltonian Variational Integrators,* submitted to BIT Numerical Mathematics, 2017. arXiv

J.M. Schmitt, M. Leok, *Adaptive Variational Integrators,* in preparation for submission, 2017.

ABSTRACT OF THE DISSERTATION

**Properties of Hamiltonian Variational Integrators**

by

Jeremy M. Schmitt

Doctor of Philosophy in Mathematics

University of California, San Diego, 2017

Professor Melvin Leok, Chair

This dissertation, *Properties of Hamiltonian Variational Integrators*, by Jeremy Schmitt, explores Hamiltonian varational integrators. Variational integrators are a common class of symplectic integrators, which have primarily been analyzed and constructed by discretizing Hamilton's principle. Hamiltonian variational integrators are derived by discretizing Hamilton's phase space principle and have not been studied as thoroughly. In this dissertation, new error analysis theorems and other related results extend the theory of Hamiltonian variational integrators. It is shown that these two formulations of variational integrators are not always numerically equivalent, even when they analytically represent the same map. Numerical simulations show there can be important differences between these

two formulations of variational integrators, particularly for averaging methods.

Next, a new class of variational integrators is developed based on the Taylor method combined with an augmented shooting method. A symmetric and more computationally efficient version is also developed, as well as a comparison of Lagrangian and Hamiltonian formulations of the integrator. Error analysis results are presented, and in addition, a proof is given of a sufficient condition for the equivalence of a Hamiltonian and Lagrangian variational integrator. Numerical simulations are presented, as well as a discussion on the role of automatic differentiation in the implementation of Taylor variational integrators.

The last topic focuses on an adaptive framework for symplectic integrators. The Poincaré transformation is used to construct an extended Hamiltonian system, which allows for variable step sizes. However, it is shown that the resulting Hamiltonian is often degenerate, and the only plausible framework for variational integrators is to use Hamiltonian variational integrators. Furthermore, the degeneracy of the Hamiltonian is discussed with regards to error analysis and the invertibility of the discrete Legendre transforms. A few monitor functions are considered, and numerical simulations demonstrate the significant gains in efficiency when using adaptive variational integrators.

# Chapter 1

# Introduction and Background

In the early 1800's Abel showed that even for the well understood polynomial equation we may need to settle for an approximate solution, rather than the exact solution [43]. Perhaps no mathematical field has benefited more from approximation methods than the field of differential equations. Sir Isaac Newton raised differential equations and dynamical systems to the fore of mathematics and the sciences when he developed the differential calculus and the laws of mechanics. Finding the solution of a differential equation was not merely a mathematical curiosity, but a powerful way of understanding the world around us. Unfortunately, for differential equations, and so many things in mathematics, the nonlinear case often requires methods of approximation. Linear approximations are often the first choice, but their value is generally constrained to a local region. How can we achieve a global approximation? A good place to start answering this question is to decide what are the relevant global properties of the dynamical system. The key idea is to view an approximation method as a discrete dynamical system, then it seems reasonable to seek a discrete dynamical system that has global properties similar to the exact dynamical system (see [4] for a related discussion). These statements will be made precise in the following pages, but this is the big idea motivating many of the topics in this dissertation, and I believe it is an idea that many areas of numerical approximation have yet to fully realize.

The global properties we consider will be mainly geometric properties, but

topological properties also merit consideration. Euclid's development of geometry is perhaps the most well-known, yet it is the notion of geometry proposed by Felix Klein and his Erlangen program that is most relevant to this discussion (see [23]). Geometry can be defined by the mathematical objects that remain unchanged or invariant under certain transformations, and Euclidean geometry centers around the study of rigid transformations and the invariants preserved under such transformations. Conservation laws in the physical sciences are intimately related to the notion of invariance under transformation. Emmy Noether showed that for many physical systems symmetry implies conservation (see [1], Chapter 4). We will discuss such systems, known as Hamiltonian systems, and the geometry that arises, called symplectic geometry. We first present the prerequisite material needed to make these notions precise, then an overview of this dissertation is provided.

## 1.1 Geometric Numerical Integration

Geometric numerical integration concerns algorithms that aim to approximate or preserve structure and qualitative properties of a dynamical system (see [19]). While there are many different notions of a geometric integrator, one particularly relevant characterization involves smooth manifolds and Lie groups.

The set of all diffeomorphisms on a smooth manifold $M$ forms a Lie group, $\mathfrak{G}$, and the associated set of all smooth vector fields (along with the commutator bracket) forms the corresponding Lie algebra, $\mathfrak{g}$. Let $\mathfrak{b} \subset \mathfrak{g}$ be a linear subspace, corresponding to $\mathfrak{B} \subset \mathfrak{G}$ via

$$\mathfrak{b} = \{F \in \mathfrak{g} | \Phi_{t,F} \in \mathfrak{B}\}.$$

The discrete time-$h$ map $\Psi_h : M \to M$ is a *geometric integrator* for $\mathfrak{b}$ if $\Psi_{h,F} \in \mathfrak{B}$ for all $F \in \mathfrak{b}$. Symplectic integrators correspond to setting $\mathfrak{b}$ equal to the set of all Hamiltonian vector fields on $M$, equipped with the commutator bracket, and setting $\mathfrak{B}$ equal to the set of all symplectic diffeomorphisms from $M$ to $M$. The discovery of symplectic transformations began in classical mechanics.

## 1.2 Continuous and Discrete Mechanics

Let $Q$ be some smooth manifold, often called the configuration manifold, and let $q \in C^2([0,T],Q)$ be a curve/trajectory on the manifold. Then, for a given function $L : TQ \to \mathbb{R}$, known as the *Lagrangian*, we can define the *action* functional

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

Curves which satisfy Hamilton's principle of stationary action,

$$\delta S = 0,$$

where $\delta$ represents variations with respect to $q$, will satisfy the Euler–Lagrange equations,

$$\frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0.$$

These concepts form the basis of Lagrangian mechanics, and the field of variational integrators originated from the discretization of Lagrangian mechanics. Consider a discrete curve $q = \{q_k\}_{i=0}^N$, then given a *discrete Lagrangian*, $L_d : Q \times Q \to \mathbb{R}$, define the discrete action as

$$S_d(q_0, q_1, \ldots, q_N; h) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}).$$

A discrete curve that is a stationary point of the discrete action, must satisfy the discrete Euler–Lagrange equations

$$D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1}) = 0.$$

This implies $D_2L_d(q_{k-1}, q_k) = -D_1L_d(q_k, q_{k+1})$, and by defining the conjugate momenta as $p_k = -D_1L_d(q_k, q_{k+1})$, then requiring the momenta to match for any given pair $(q_k, q_{k+1})$ yields an equivalent form of the discrete Euler–Lagrange equations,

$$p_k = -D_1L_d(q_k, q_{k+1}), \qquad p_{k+1} = D_2L_d(q_k, q_{k+1}).$$

These equations implicitly define the map (and one-step method) $\tilde{F}_{L_d} : (q_k, p_k) \to (q_{k+1}, p_{k+1})$, which is known as a variational integrator. The goal is to approximate

the continuous flow of the Euler–Lagrange equations, and we need a way to relate the discrete map $\tilde{F}_{L_d}$ to the flow of the Euler–Lagrange equations. This relation is made by an object known classically as a type I generating function, but in the language of variational integrators it is called the *exact discrete Lagrangian*

$$L_d^E(q_0, q_1; h) = \underset{\substack{q \in C^2([0,h],Q) \\ q(0)=q_0 \\ q(h)=q_1}}{ext} \int_0^h L(q(t), \dot{q}(t))dt.$$

Continuous mechanics is connected to discrete mechanics via the following theorem from [35].

**Theorem 1.** *If a discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$, approximates the exact discrete Lagrangian, $L_d^E : Q \times Q \to \mathbb{R}$ to order $r$, i.e.,*

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

*then the discrete Hamiltonian map, $\tilde{F}_{L_d} : (q_0, p_0) \mapsto (q_1, p_1)$, viewed as a one-step method, is order $r$ accurate.*

Variational integrators are symplectic integrators, and in order to see where symplecticity fits in we must examine Hamiltonian mechanics.

Hamiltonian mechanics is related to Lagrangian mechanics by the *Legendre transform*, which can be interpreted as a map from the tangent bundle $TQ$, to the cotangent bundle $T^*Q$. The *conjugate momentum*, $p$, can be defined by the Legendre transform $p = \frac{\partial L}{\partial \dot{q}}$, assuming a nondegenerate Lagrangian. The Hamiltonian $H : T^Q \to T^*Q$ is defined as

$$H(q, p) = < p, \dot{q} > -L(q, \dot{q}),$$

where $\dot{q}$ is defined via the Legendre transform and $< \cdot, \cdot >$ is the usual pairing between the tangent and cotangent bundle. The Euler–Lagrange equations are equivalent to Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q},$$

and reveal many intrinsic properties of the dynamical system. An (autonomous) Hamiltonian is constant along the flow of Hamilton's equations as the following calculation shows

$$\begin{aligned}\frac{dH(q,p)}{dt} &= \frac{\partial H}{\partial q}\frac{dq}{dt} + \frac{\partial H}{\partial p}\frac{dp}{dt} \\ &= -\dot{p}\dot{q} + \dot{q}\dot{p} \\ &= 0.\end{aligned}$$

The Hamiltonian often coincides with the sum of the kinetic energy and potential energy of a physical system, and the previous result is then more commonly known as the conservation of energy. The mathematical structure we are interested in is called a *symplectic* structure. A symplectic manifold is a smooth manifold, $M$, with an associated closed nondegenerate differential 2-form, $\omega$. In particular the cotangent bundle has a natural symplectic structure $(T^*Q, \omega)$, where $\omega$ takes the local coordinate form $dp \wedge dq$. A symplectic transformation $F : M_1 \to M_2$, is a map between two symplectic manifolds, $(M_1, \omega_1)$ and $(M_2, \omega_2)$, such that $F^*\omega_2 = \omega_1$, where $F^*$ is the pullback with respect to $F$. The flow of Hamilton's equations $\Phi_{H,t}$, is a symplectic transformation from $T^*Q$ to $T^*Q$.

$$\begin{aligned}\frac{d}{dt}\omega &= d\dot{q} \wedge dp + dq \wedge d\dot{p} \\ &= dH_p \wedge dp - dq \wedge dH_q \quad (by\ Hamilton's\ Eqts) \\ &= d(H_p dp + H_q dq) \qquad (Property\ of\ Exterior\ Derivative) \\ &= d(dH) \ = 0. \qquad\qquad (Exact \ \Rightarrow Closed)\end{aligned}$$

The following calculation shows that, assuming sufficiently independent coordinates, every symplectic transformation is locally associated with a generating function.

$$\begin{aligned}&\Phi_{H,t}(q_0, p_0) = (q_1, p_1) \\ &\Rightarrow dq_0 \wedge dp_0 = dq_1 \wedge dp_1 \\ &\Rightarrow dq_0 \wedge dp_0 - dq_1 \wedge dp_1 = 0 \\ &\Rightarrow d(-p_0 dq_0 + p_1 dq_1) = 0\end{aligned}$$

By Poincare's lemma, there exists (locally) a function $S$ of $q_0$ and $q_1$ such that $dS = -p_0 dq_0 + p_1 dq_1$. This is known as a type I generating function, since it is a function of $q_0$ and $q_1$, and the exact discrete Lagrangian is a type I generating function of the time-$h$ flow of Hamiltion's equations with boundary conditions $q(0) = q_0$ and $q(h) = q_1$. The fact that the flow of Hamilton's equations can be generated by a generating function is an alternate proof of the symplecticity of the flow on $T^*Q$. In a similar manner, this can be used to show that the mapping generated by a discrete Lagrangian, $\tilde{F}^h_{L_d}$, is a symplectic mapping, which implies variational integrators are also symplectic integrators. Variational integrators discretize the generating function of the Hamiltonian flow, and if the generating function approximates the exact discrete Lagrangian, then it is called a Lagrangian varaitional integrator. For a detailed overview of discrete mechanics and variational integrators, I recommend [35], and some excellent sources on Lagrangian and Hamiltonian mechanics include [1], [34], [14], and [27].

My research has focused on examining a discretization of type II and type III generating functions, known as the exact discrete right or exact discrete left Hamiltonian, and the resulting variational integrators, known as Hamiltonian variational integrators. Hamiltonian variational integrators were first introduced in [26], and further developed in [31]. [26] derived the discrete Hamiltonian from a discrete Lagrangian, and [31] showed that the discrete Hamiltonian could be derived independently of a discrete Lagrangian.

## 1.3 Backwards Error Analysis

Backwards error analysis was first used in [52] for numerical linear algebra. The development of backwards analysis for symplectic integrators (see [8], [15]) rigorously justified the energy performance of symplectic methods, and it also explained the poor behavior of symplectic integrators when combined with variable step sizes. Given a problem $\mathcal{P}$ and its solution $\mathcal{S}$, backwards error analysis attempts to show that a given approximation scheme $\tilde{\mathcal{S}}$ solves a problem $\tilde{\mathcal{P}}$, where $\tilde{\mathcal{P}}$ is in

some sense close to $\mathcal{P}$. Let $\Psi_h$ be a symplectic integrator with step size $h$ applied to a Hamiltonian vector field, $\dot{z} = F(z)$, then let $\Phi_{t,F}$ be the associated time-$t$ flow map. By comparing Taylor expansions, one can construct a modified vector field $\tilde{F}$ such that

$$\Psi_h^n(z_0) = \Phi_{nh,\tilde{F}}(z_0), \qquad \text{for } n = 1, 2, 3, \ldots$$

and furthermore, it can be shown that $\tilde{F}$ is also a Hamiltonian vector field. The following theorem shows not only the existence of a modified Hamiltonian, but also that the symplectic integrator nearly preserves the Hamiltonian over exponentially long time intervals.

**Theorem 2.** *Let $H$ be an analytic Hamiltonian associated to the Hamiltonian vector field $F$, then for a symplectic integrator $\Psi_h$ that stays within a compact set there exists a constant $h_0$ and a modified Hamiltonian, $\tilde{H}$ such that,*

$$\tilde{H}(\Psi_h^n(y_0)) = \tilde{H}(y_0) + \mathcal{O}(e^{-h_0/2h}),$$
$$H(\Psi_h^n(y_0)) = H(y_0) + \mathcal{O}(h^p),$$

*over exponentially long time intervals $nh \leq e^{h_0/2h}$.*

This explains the excellent long-time near energy preservation of symplectic integrators, but these results require a fixed step size $h$. Otherwise, each time $h$ varies the symplectic integrator becomes associated with a different modified Hamiltonian, which can lead to a drift away from the orignal Hamiltonian. Although this may seem a death sentence for adaptive symplectic methods, there has been progress made in this direction, which we will discuss in the fourth chapter, and we will show that discrete Hamiltonians are part of the solution.

## 1.4 Overview

The second chapter establishes error analysis theorems and results on the adjoint of a discrete Hamiltonian and Hamiltonian variational integrators. Next the question of when different generating functions lead to the same numerical method is addressed. This question is phrased as a question of commutativity for

the composition of discretization and a form of the Legendre transform, and lack of commutativity is shown for the general case. It is shown that even when the generating functions generate the same map analytically there can be a difference numerically. In particular, an averaging method is examined and shown to have significantly different behavior, which depends on the type of generating function involved. For this case, the different behavior can be attributed to the boundary values associated with each type of generating function. This lends significant support to the notion that in general, when constructing a variational integrator, one should consider not only the scheme for approximating the generating function, but also which type of generating function should be approximated.

The third chapter presents a new type of variational integrator that is constructed using the Taylor method, and it is called a Taylor variational integrator. In particular, a shooting-like scheme is developed that takes advantage of the Taylor method to give a shooting method that is in general one order higher than the usual shooting method. After deriving these error analysis results, a symmetric form of the integrator is developed, which has computational efficiency advantages compared to the unsymmetric form. Next, the Lagrangian and Hamiltonian variational integrators are compared, and a result is presented that gives a sufficient condition for the equivalence of a Lagrangian and Hamiltonian variational integrator. The chapter concludes with a series of numerical results, and a discussion of the role of automatic differentiation in developing more efficient implementions of Taylor variational integrators.

The fourth chapter discusses efforts to combine symplectic and variational integrators with adaptivity. In particular, the most common approach for symplectic integrators is to use the Poincaré transformation to generate a new Hamiltonian system with respect to a fictive time and subsequently fictive constant step sizes. This allows established symplectic integrators to be applied to the transformed Hamiltonian system with some noteable caveats. However, this approach has not been successfully derived at the level of the generating function, and it is shown

that this is due to a degeneracy of the transformed Hamiltonian. As a result, the way forward requires the construction of discrete Hamiltonians and Hamiltonian variational integrators, which until now have not been considered. A modification of the error analysis theorem is discussed for Hamiltonian variational integrators based on the Poincaré transformed Hamiltonian. Finally, numerical results are presented to show the efficiency advantages of adaptive variational integrators over non-adaptive variational integrators.

# Chapter 2

# Properties of Hamiltonian Variational Integrators

## 2.1 Introduction

Geometric numerical integration is a field of numerical analysis that develops numerical methods with the goal of preserving geometric properties of dynamical systems (see [19]). Variational integrators are geometric numerical integrators derived from discretizing Hamilton's principle from classical mechanics (see [35]). They have many desirable properties such as symplecticity, momentum-preservation, and near-energy preservation, which results in excellent long-term stability. While the Lagrangian formulation of variational integrators has been thoroughly investigated (see [9; 30; 32; 33; 35; 36]), only recently has the Hamiltonian formulation of variational integrators been established (see [26; 31]).

In this paper we will continue the investigation of Hamiltonian variational integrators, and establish theorems on error analysis, symmetry of the method, and provide numerical experiments to elucidate the relative numerical advantages and disadvantages of the Lagrangian and Hamiltonian formulations. In particular, evidence is presented to show that for oscillatory problems the discrete Lagrangian and discrete Hamiltonian variational integrators have differing resonance and conditioning properties. In addition, it is shown that some approximation methods

will only yield a symmetric method when derived from a specific type of generating function. The upshot is that the numerical properties of a variational integrator are determined both by the approximation scheme used to construct it and by the type of the generating function being approximated.

### 2.1.1 Discrete Mechanics

Lagrangian variational integrators are based on a discrete analogue of Hamilton's principle, and Hamiltonian variational integrators are based on a discrete analogue of Hamilton's phase space variational principle. The fundamental objects in the discretization are generating functions of symplectic maps, and in the Hamiltonian case, they are obtained by approximating the exact Type II generating function associated with a Hamiltonian flow, which we refer to as the exact discrete right Hamiltonian,

$$H_d^{+,E}(q_0,p_1) = \operatorname*{ext}_{\substack{(q,p)\in C^2([0,T],T^*Q)\\ q(0)=q_0,p(T)=p_1}} \left( p_1 q_1 - \int_0^T \left[ p\dot{q} - H(q,p) \right] dt \right). \tag{2.1}$$

This can be viewed as the solution at time $T$ of the Type II Hamilton–Jacobi equation,

$$\frac{\partial S_2(q_0,p,t)}{\partial t} = H\left( \frac{\partial S_2}{\partial p}, p \right), \tag{2.2}$$

which more generally describes the Type II generating function which generates the time-$t$ Hamiltonian flow map,

$$S_2(q_0,p,t) = \operatorname*{ext}_{\substack{(q,p)\in C^2([0,t],T^*Q)\\ q(0)=q_0,p(t)=p}} \left( p(t)q(t) - \int_0^t \left[ p(s)\dot{q}(s) - H(q(s),p(s)) \right] ds \right). \tag{2.3}$$

Similarly, the exact discrete left Hamiltonian is given by,

$$H_d^{-,E}(p_0,q_1;h) = \operatorname*{ext}_{\substack{(q,p)\in C^2([0,T],T^*Q)\\ q(0)=q_0,p(T)=p_1}} \left( -p_0 q_0 - \int_0^T \left[ p\dot{q} - H(q,p) \right] dt \right). \tag{2.4}$$

and it can be viewed as a solution at time $T$ of the Type III Hamilton–Jacobi equation,

$$\frac{\partial S_3(p_0,q,t)}{\partial t} = H\left( q, -\frac{\partial S_3}{\partial q} \right). \tag{2.5}$$

Given discrete Hamiltonians, $H_d^+(q_k, p_{k+1}; h)$ and $H_d^-(p_k, q_{k+1}; h)$, the discrete Hamilton's equations are given by,

$$q_{k+1} = D_2 H_d^+(q_k, p_{k+1}; h), \tag{2.6}$$

$$p_k = D_1 H_d^+(q_k, p_{k+1}; h), \tag{2.7}$$

and,

$$q_k = -D_1 H_d^-(p_k, q_{k+1}; h), \tag{2.8}$$

$$p_{k+1} = -D_2 H_d^-(p_k, q_{k+1}; h). \tag{2.9}$$

These can also be expressed in terms of the discrete Legendre transformations, $\mathbb{F}^{\pm} H_d^+ : (q_k, p_{k+1}) \to T^*Q$,

$$\mathbb{F}^+ H_d^+(q_k, p_{k+1}; h) = (D_2 H_d^+(q_k, p_{k+1}; h), p_{k+1}), \tag{2.10}$$

$$\mathbb{F}^- H_d^+(q_k, p_{k+1}; h) = (q_k, D_1 H_d^+(q_k, p_{k+1}; h)), \tag{2.11}$$

and $\mathbb{F}^{\pm} H_d^- : (p_k, q_{k+1}) \to T^*Q$,

$$\mathbb{F}^+ H_d^-(p_k, q_{k+1}; h) = (q_{k+1}, -D_2 H_d^-(p_k, q_{k+1}; h)), \tag{2.12}$$

$$\mathbb{F}^- H_d^-(p_k, q_{k+1}; h) = (-D_1 H_d^-(p_k, q_{k+1}; h), p_k). \tag{2.13}$$

We observe that the Hamiltonian maps $\tilde{F}_{H_d^{\pm}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ can be expressed as

$$\tilde{F}_{H_d^{\pm}} = \mathbb{F}^+ H_d^{\pm} \circ (\mathbb{F}^- H_d^{\pm})^{-1}. \tag{2.14}$$

## 2.2 Error Analysis and Symmetric Methods

### 2.2.1 Error Analysis

Variational integrators are able to benefit from and adopt many traditional techniques and methods of numerical analysis (see [30]). This can be largely attributed to the following theorem from [35].

**Theorem 3** (Theorem 2.3.1, Marsden and West [35])**.** *If a discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$, approximates the exact discrete Lagrangian, $L_d^E : Q \times Q \to \mathbb{R}$ to order $r$, i.e.,*

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

*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 $r$ accurate.*

Thus, in order to generate a variational integrator of a particular order, one can leverage techniques from numerical analysis with the goal of approximating the exact discrete Lagrangian, then the associated discrete Hamiltonian map yields the variational integrator. We first present the corresponding theorem for discrete Hamiltonian variational integrators, which draws much of its inspiration from the theorem and proof of the above result as detailed in [35].

**Theorem 4.** *If a discrete right Hamiltonian, $H_d^+ : T^*Q \to \mathbb{R}$, approximates the exact discrete right Hamiltonian, $H_d^{+,E} : T^*Q \to \mathbb{R}$ to order $r$, i.e.,*

$$H_d^+(q_0, p_1; h) = H_d^{+,E}(q_0, p_1; h) + \mathcal{O}(h^{r+1}),$$

*and the Hamiltonian is continuously differentiable, then the discrete map, $\tilde{F}_{H_d^+}^h : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$, viewed as a one-step method, is order $r$ accurate.*

Note that the following proof can be easily adjusted to prove an equivalent theorem for the discrete left Hamiltonian case. First, we will need the following lemma.

**Lemma 1.** *Let $f_1, g_1, e_1, f_2, g_2, e_2 \in C^r$ be such that*

$$f_1(x, h) = g_1(x, h) + h^{r+1} e_1(x, h),$$

$$f_2(x, h) = g_2(x, h) + h^{r+1} e_2(x, h).$$

*Then, there exists functions $e_{12}$ and $\bar{e}_1$ bounded on compact sets such that*

$$f_2(f_1(x, h), h) = g_2(g_1(x, h), h) + h^{r+1} e_{12}(g_1(x, h), h),$$

$$f_1^{-1}(y) = g_1^{-1}(y) + h^{r+1} \bar{e}_1(y).$$

*Proof.*

$$\begin{aligned} f_2(f_1(x,h),h) &= f_2(g_1(x,h)+h^{r+1}e_1(x,h),h) \\ &= g_2(g_1(x,h)+h^{r+1}e_1(x,h),h)+h^{r+1}e_2(g_1(x,h)+h^{r+1}e_1(x,h),h) \\ &= g_2(g_1(x,h),h)+h^{r+1}\tilde{e}_1(g_1(x,h),h)+h^{r+1}e_2(g_1(x,h) \\ &\quad +h^{r+1}e_1(x,h),h), \end{aligned}$$

where $\tilde{e}_1$ is bounded on compact set. This last line comes from combining compactness of the set with the smoothness of the functions to obtain a Lipschitz property of the form,

$$\|g_2(g_1(x,h)+h^{r+1}e_1(x,h),h)-g_2(g_1(x,h),h)\| \leq Ch^{r+1}.$$

For each choice of $(x,h)$, equality holds for a particular choice of constant, which defines $\tilde{e}_1$ and establishes its smoothness as a function. Adding $e_2$ to $\tilde{e}_1$ we obtain a function $e_{12}$, which is also bounded on compact sets such that,

$$f_2(f_1(x,h),h) = g_2(g_1(x,h),h)+h^{r+1}e_{12}(g_1(x,h),h).$$

Let $y = f_1(x,h)$, and note that by definition,

$$f_1^{-1}(f_1(x,h)) = g_1^{-1}(g_1(x,h)).$$

Since $g_1^{-1}(y) = g_1^{-1}(g_1(x,h)+h^{r+1}e_1(x,h))$, then

$$\|g_1^{-1}(y)-f_1^{-1}(y)\| = \|g_1^{-1}(y)-g_1^{-1}(g_1(x,h))\| \leq \bar{C}h^{r+1}.$$

From this, it follows that there exists a function $\bar{e}_1$ bounded on compact sets such that,

$$f_1^{-1}(y) = g_1^{-1}(y)+h^{r+1}\bar{e}_1(y).$$

□

Now we are ready for the proof of the theorem.

*Proof.* By assumption there is some bounded continuously differentiable function $e$ such that,

$$H_d^+(q(0),p(h);h) = H_d^{+,E}(q(0),p(h),h)+h^{r+1}e(q(0),p(h),h).$$

Differentiating yields,

$$D_1 H_d^+(q(0), p(h); h) = D_1 H_d^{+,E}(q(0), p(h); h) + h^{r+1} D_1 e(q(0), p(h), h),$$

where $\|D_1 e(q(0), p(h), h)\| \leq \tilde{C}$. This implies,

$$\|\mathbb{F}^- H_d^+(q(0), p(h); h) - \mathbb{F}^- H_d^{+,E}(q(0), p(h); h)\| \leq \tilde{C} h^{r+1}.$$

Now combining this with the fact that $\tilde{F}_{H_d^+} = \mathbb{F}^+ H_d^+ \circ (\mathbb{F}^- H_d^+)^{-1}$ and applying Lemma 1, we have,

$$\tilde{F}^h_{H_d^+} = \tilde{F}^h_{H_d^{+,E}} + \mathcal{O}(h^{r+1}).$$

□

Determining the order of a variational integrator is greatly simplified via the above theorems, which relate the order of the integrator to the order to which the associated discrete Lagrangian or discrete right Hamiltonian approximates the corresponding exact generating function. Similarly, it was shown in [35] that one can determine whether or not the variational integrator is a symmetric method by examining the corresponding discrete Lagrangian. We would like to extend this result to the case of discrete Hamiltonians.

### 2.2.2 Symmetric Methods

**Definition 1** (see Chapters II.3 and V of [19]). *A numerical one-step method $\Phi_h$ is called **symmetric** or **time-reversible**, if it satisfies*

$$\Phi_h \circ \Phi_{-h} = id$$

*or equivalently*

$$\Phi_h = \Phi_{-h}^{-1}.$$

*The **adjoint** of a numerical one-step method, denoted $\Phi_h^*$, is defined as*

$$\Phi_h^* = \Phi_{-h}^{-1}.$$

A numerical one-step method is a symmetric method if it is self-adjoint, i.e., $\Phi_h = \Phi_h^*$. The adjoint of a discrete Lagrangian, $L_d^*$, is defined as

$$L_d^*(q_0, q_1, h) = -L_d(q_1, q_0, -h).$$

The discrete Lagrangian is called self-adjoint if $L_d^*(q_0, q_1, h) = L_d(q_0, q_1, h)$. The following theorem from [35] relates the self-adjointness of the discrete Lagrangian with the self-adjointness of the corresponding variational integrator.

**Theorem 5** (Theorem 2.4.1 of [35])**.** *The discrete Lagrangian (or an equivalent discrete Lagrangian), $L_d$, is self-adjoint if and only if the method associated to the corresponding discrete Hamiltonian map is self-adjoint, i.e., symmetric.*

In many cases it is easier to check if the discrete Lagrangian is self-adjoint, rather than checking the variational integrator itself. We seek a definition for the adjoint of a discrete right Hamiltonian.

The adjoint of a one-step method $(q_1, p_1) = \Phi_h(q_0, p_0)$ can be obtained by reversing the direction of time, and reversing the roles of the initial data and terminal solution, i.e., $(q_0, p_0) = \Phi_{-h}^*(q_1, p_1)$. This corresponds to swapping out $(q_0, p_0, q_1, p_1, h)$ for $(q_1, p_1, q_0, p_0, -h)$. This motivates the definition of the adjoint of a Type II generating function as a Type III generating function and vice versa. In particular, given a Type II discrete Hamiltonian $H_d^+$, we seek a definition for the Type III adjoint $(H_d^+)^*$ that will satisfy $\tilde{F}^h_{(H_d^+)^*} = (\tilde{F}^h_{H_d^+})^*$. Let $\tilde{F}^h_{(H_d^+)^*}(q_0, p_0) = (q_1, p_1)$. Then, we want

$$\begin{aligned}(q_1, p_1) &= \tilde{F}^h_{(H_d^+)^*}(q_0, p_0) \\ &= (\tilde{F}^h_{H_d^+})^*(q_0, p_0) \\ &= (\tilde{F}^{-h}_{H_d^+})^{-1}(q_0, p_0).\end{aligned}$$

This implies $\tilde{F}^{-h}_{H_d^+}(q_1, p_1) = (q_0, p_0)$, which together with $\tilde{F}^h_{(H_d^+)^*}(q_0, p_0) = (q_1, p_1)$ yield the respective sets of equations,

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

and

$$p_1 = -D_2(H_d^+)^*(p_0, q_1; h),$$
$$q_0 = -D_1(H_d^+)^*(p_0, q_1; h).$$

Comparing these equations we see that setting $(H_d^+)^*(p_0, q_1; h) = -H_d^+(q_1, p_0; -h)$ satisfies $\tilde{F}^h_{(H_d^+)^*} = (\tilde{F}^h_{H_d^+})^*$. A similar calculation yields an analogous expression for the adjoint of a Type III generating function $H_d^-$.

**Definition 2.** *Given a Type II/III generating function, $H_d^\pm$, define the* ***adjoint*** *as the Type III/II generating function, $(H_d^\pm)^*$, where $\tilde{F}^h_{(H_d^\pm)^*}(q_0, p_0) = (q_1, p_1)$, as*

$$(H_d^+)^*(p_0, q_1; h) = -H_d^+(q_1, p_0; -h), \tag{2.15}$$

$$(H_d^-)^*(q_0, p_1; h) = -H_d^-(p_1, q_0; -h). \tag{2.16}$$

**Example 1.** *The symplectic Euler-A method for a Lagrangian of the form $L(q, \dot{q}) = \frac{1}{2}\dot{q}M\dot{q} - V(q)$ is given by,*

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

*The corresponding discrete right Hamiltonian is given by*

$$\begin{aligned} H_d^+(q_0, p_1, h) &= p_1(q_0 + hM^{-1}p_1) - h[p_1M^{-1}p_1 - H(q_0, p_1)], \\ &= p_1q_0 + hH(q_0, p_1). \end{aligned}$$

*The adjoint of this method is given by symplectic Euler-B,*

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

*We now derive the corresponding adjoint of the discrete right Hamiltonian for symplectic Euler-A.*

$$\begin{aligned} (H_d^+)^*(p_0, q_1; h) &= -H_d^+(q_1, p_0; -h) \\ &= -p_0(q_1 - hM^{-1}p_0) - h[p_0M^{-1}p_0 - H(q_1, p_0)] \end{aligned}$$

$$= -p_0 q_1 + hH(q_1, p_0).$$

*We can verify that this generates symplectic Euler-B by applying the discrete left Hamilton's equations,*

$$\begin{aligned} q_0 &= -D_1(H_d^+)^*(p_0, q_1; h) \\ &= D_2 H_d^+(q_1, p_0; -h) \\ &= q_1 - hM^{-1}p_0, \\ p_1 &= -D_2(H_d^+)^*(p_0, q_1; h) \\ &= D_1 H_d^+(q_1, p_0; -h) \\ &= p_0 - h\nabla V(q_1). \end{aligned}$$

*Solving the first equation for $q_1$ gives symplectic Euler-B, as expected.*

**Theorem 6.** $(H_d^\pm)^{**} = H_d^\pm$.

*Proof.* We consider the case of the Type II generating function $H_d^+$. Let $\tilde{F}^h_{(H_d^+)^{**}}(q_0, p_0) = (q_1, p_1)$. Since $(H_d^+)^*$ is a Type III generating function, applying the definition of the adjoint twice gives

$$\begin{aligned} (H_d^+)^{**}(q_0, p_1; h) &= -(H_d^+)^*(p_1, q_0; -h) \\ &= H_d^+(q_0, p_1; h), \end{aligned}$$

and a similar calculation shows that this holds for the Type III generating function $H_d^-$ as well. □

Since the notion of the adjoint that we introduced converts a Type II to a Type III generating function, for a discrete Hamiltonian to be self-adjoint, we need to compare the adjoint to the Legendre transformation of the discrete Hamiltonian, which is given by,

$$H_d^-(p_k, q_{k+1}; h) = -p_k q_k - p_{k+1} q_{k+1} + H_d^+(q_k, p_{k+1}; h),$$

where we view $p_{k+1}$ and $q_k$ as functions of $p_k$ and $q_{k+1}$. Then, the following calculation shows that these two generating functions generate the same symplectic

map, i.e., $\tilde{F}_{H_d^-} = \tilde{F}_{H_d^+}$,

$$\begin{aligned}
-D_1 H_d^-(p_k, q_{k+1}; h) &= q_k + p_k \frac{\partial q_k}{\partial p_k} + \frac{\partial p_{k+1}}{\partial p_k} q_{k+1} - D_1 H_d^+(q_k, p_{k+1}; h) \frac{\partial q_k}{\partial p_k} \\
&\quad - D_2 H_d^+(q_k, p_{k+1}; h) \frac{\partial p_{k+1}}{\partial p_k} \\
&= q_k + \left(p_k - D_1 H_d^+(q_k, p_{k+1}; h)\right) \frac{\partial q_k}{\partial p_k} \\
&\quad + \left(q_{k+1} - D_2 H_d^+(q_k, p_{k+1}; h\right) \frac{\partial p_{k+1}}{\partial p_k}, \\
-D_2 H_d^-(p_k, q_{k+1}; h) &= p_k \frac{\partial q_k}{\partial q_{k+1}} + \frac{\partial p_{k+1}}{\partial q_{k+1}} q_{k+1} + p_{k+1} \\
&\quad - D_1 H_d^+(q_k, p_{k+1}; h) \frac{\partial q_k}{\partial q_{k+1}} - D_2 H_d^+(q_k, p_{k+1}; h) \frac{\partial p_{k+1}}{\partial q_{k+1}} \\
&= p_{k+1} + \left(p_k - D_1 H_d^+(q_k, p_{k+1}; h)\right) \frac{\partial q_k}{\partial q_{k+1}} \\
&\quad + \left(q_{k+1} - D_2 H_d^+(q_k, p_{k+1}; h)\right) \frac{\partial p_{k+1}}{\partial q_{k+1}}.
\end{aligned}$$

**Definition 3.** *A Type II/III generating function is* ***self-adjoint****, if it is equal (up to equivalency) to the Legendre transform of its adjoint.*

Note that this definition implies that a discrete right Hamiltonian is self-adjoint if its adjoint is equal (up to equivalency) to the associated discrete left Hamiltonian, i.e., $(H_d^+)^* = H_d^-$.

**Theorem 7.** *Given a self-adjoint discrete right Hamiltonian, i.e., $H_d^- = (H_d^+)^*$, the method associated to the discrete right Hamiltonian map is self-adjoint. Likewise, if a method coming from a discrete right Hamiltonian map is self-adjoint, then the associated discrete right Hamiltonian is self-adjoint.*

*Proof.* Assume $H_d^- = (H_d^+)^*$. Then,

$$(\tilde{F}_{H_d^+})^* = \tilde{F}_{(H_d^+)^*} = \tilde{F}_{H_d^-} = \tilde{F}_{H_d^+},$$

and so, by definition, the map is self-adjoint. Now assume $\tilde{F}_{H_d^+} = (\tilde{F}_{H_d^+})^*$. Then,

$$\tilde{F}_{H_d^-} = \tilde{F}_{H_d^+} = (\tilde{F}_{H_d^+})^* = \tilde{F}_{(H_d^+)^*},$$

which implies $(H_d^+)^* = H_d^-$ (up to equivalency) and, by definition, the discrete right Hamiltonian is self-adjoint. □

A similar proof can be used to prove an identical theorem for the discrete left Hamiltonian. The previous theorem provides an easy way to check if a variational integrator is self-adjoint. Assuming the Hamiltonian flow is time-reversible, it follows that the exact discrete right Hamiltonian is self-adjoint. This can also be shown using the definition of a self-adjoint exact discrete right Hamiltonian, and the same result can be shown for the discrete left Hamiltonian with only minor adjustments.

**Corollary 1.** *The exact discrete right Hamiltonian, $H_d^{+,E}$, is self-adjoint.*

*Proof.* A direct calculation shows that

$$\begin{aligned}
(H_d^{+,E})^*(p_0, q_1; h) &= -H_d^{+,E}(q_1, p_0; -h) \\
&= -(\tilde{p}(-h)\tilde{q}(-h) - \int_0^{-h} [\tilde{p}(\tau)\tilde{q}(\tau) - H(\tilde{q}(\tau), \tilde{p}(\tau))]d\tau) \\
&= -p(-h+h)q(-h+h) - \int_{-h}^{0} [p(\tau+h)q(\tau+h) \\
&\quad - H(q(\tau+h), p(\tau+h))]d\tau \\
&= -p(0)q(0) - \int_0^h [p(t)q(t) - H(q(t), p(t))]dt \\
&= H_d^{-,E}(p_0, q_1; h),
\end{aligned}$$

where we used the fact that the time-reversed solution $(\tilde{q}(\tau), \tilde{p}(\tau))$ over the time domain $[-h, 0]$ with $(q_1, p_0)$ boundary data is related to the solution curve $(q(t), p(t))$ over the time domain $[0, h]$ with $(q_0, p_1)$ boundary data by $(\tilde{q}(\tau), \tilde{p}(\tau)) = (q(\tau + h), p(\tau + h))$. □

The definition of the adjoint also provides a simple way to construct symmetric methods. Given any method defined by $H_d$, we can construct a symmetric method using composition, for example, $\tilde{F}_{H_d}^{\frac{h}{2}} \circ \tilde{F}_{H_d^*}^{\frac{h}{2}}$, which is nothing more than composing a half-step of the adjoint method with a half-step of the method. It is well-known that this leads to a symmetric method, as the following calculation demonstrates,

$$(\tilde{F}_{H_d}^{\frac{h}{2}} \circ \tilde{F}_{H_d^*}^{\frac{h}{2}})^* = (\tilde{F}_{H_d^*}^{\frac{h}{2}})^* \circ (\tilde{F}_{H_d}^{\frac{h}{2}})^*$$

$$= \tilde{F}^{\frac{h}{2}}_{H_d^{**}} \circ \tilde{F}^{\frac{h}{2}}_{H_d^{*}}$$
$$= \tilde{F}^{\frac{h}{2}}_{H_d} \circ \tilde{F}^{\frac{h}{2}}_{H_d^{*}},$$

where the last line used Theorem 2.7. More generally, a composition method of the form,

$$\tilde{F}^{\alpha_s h}_{H_d} \circ \tilde{F}^{\beta_s h}_{H_d^*} \circ \cdots \circ \tilde{F}^{\beta_2 h}_{H_d^*} \circ \tilde{F}^{\alpha_1 h}_{H_d} \circ \tilde{F}^{\beta_1 h}_{H_d^*},$$

where $\alpha_{s+1-i} = \beta_i$ for $i = 1, \ldots, s$, will be symmetric. For a more in depth discussion of symmetric composition methods, see Chapter V.3 of [19].

## 2.3 Discrete Lagrangians versus Discrete Hamiltonians

### 2.3.1 Composition of Discretization and the Legendre Transform

A transformation of one type of a generating function into another type of generating function, which preserves the associated sympletic map is given by the following Legendre transforms.

**Definition 4.** *(i) The Legendre transform of a Type I generating into a Type II generating function is given by the equation,*

$$H_d^+(q_k, p_{k+1}; h) = p_{k+1} q_{k+1} - L_d(q_k, q_{k+1}; h),$$

*where $q_{k+1}$ is implicitly defined by $p_{k+1} = D_2 L_d(q_k, q_{k+1}; h)$. The Legendre transform of a Type II generating function into a Type I is given by the same equation, where $p_{k+1}$ is implicitly defined by $q_{k+1} = D_2 H_d^+(q_k, p_{k+1}; h)$.*

*(ii) The Legendre transform of a Type I generating function into a Type III generating function is given by the equation,*

$$H_d^-(p_k, q_{k+1}; h) = -p_k q_k - L_d(q_k, q_{k+1}; h),$$

*where $q_k$ is implicitly defined by $p_k = -D_1 L_d(q_k, q_{k+1}; h)$. The Legendre transform of a Type III generating function into a Type I is given by the same equation, where $p_k$ is implicitly defined by $q_k = -D_1 H_d^-(p_k, q_{k+1}; h)$.*

*(iii) The Legendre transform of a Type II generating function into a Type III generating function is given by the equation,*

$$H_d^-(p_k, q_{k+1}; h) = -p_k q_k - p_{k+1} q_{k+1} + H_d^+(q_k, p_{k+1}; h),$$

*where $q_k$ and $p_{k+1}$ are implicitly defined by the set of equations $p_k = D_1 H_d^+(q_k, p_{k+1}; h)$ and $q_{k+1} = D_2 H_d^+(q_k, p_{k+1}; h)$. The Legendre transform of a Type III generating into a Type II is given by the same equation, where $q_{k+1}$ and $p_k$ are implicitly defined by the set of equations $q_k = -H_d^-(p_k, q_{k+1}; h)$ and $p_{k+1} = -H_d^-(p_k, q_{k+1})$.*

Variational integrators are derived by first discretizing the exact generating function, which results in a new generating function of the same type. Using the Legendre transforms defined above, the discretized generating function can be transformed into an equivalent generating function of a different type. This can be viewed as composing the Legendre transform and the discretization. Likewise, one could first take the Legendre transform of the given exact generating function and then apply the discretization. The question we address next is whether or not changing the order of this composition results in the same symplectic map, and for one particular type of discretization this question has already been answered.

It was shown in [31] that the Galerkin variational integrator construction leads to equivalent discrete Lagrangian and discrete Hamiltonian methods for the same choice of quadrature rule and finite-dimensional function space, and the result is given in the following theorem.

**Theorem 8** (Proposition 4.1 of [31])**.** *If the continuous Hamiltonian $H(q, p)$ is hyperregular and we construct a Lagrangian $L(q, \dot{q})$ by the Legendre transformation, then the generalized Galerkin Hamiltonian variational integrator (see [31]) and the generalized Galerkin Lagrangian variational integrator, associated with the same choice of basis functions and numerical quadrature formula, are equivalent.*

Will this hold for other types of variational integrators? To begin to address this question, we propose the following examples.

**Example 2.** *Consider a Lagrangian of the form* $L(q,\dot{q}) = \frac{1}{2}\dot{q}M\dot{q} - V(q)$*, where* $M$ *is symmetric positive-definite and* $V$ *is sufficiently smooth.*

*The exact discrete Lagrangian, which is defined as,*

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

*where* $q_{01}(0) = q_0$, $q_{01}(h) = q_1$, *and* $q_{01}$ *satisfies the Euler–Lagrange equation in the time interval* $(0,h)$*. Letting* $q_0$ *and* $q_1$ *be fixed, then a first-order finite difference approximation of the velocities yields,*

$$\dot{q}_0 \approx \frac{q_1 - q_0}{h},$$
$$\dot{q}_1 \approx \frac{q_1 - q_0}{h}.$$

*Using the rectangular quadrature rule about the initial point results in the following discrete Lagrangian (i.e. Type I generating function),*

$$L_d(q_0,q_1;h) = \frac{1}{2}\left(\frac{q_1 - q_0}{h}\right) M \left(\frac{q_1 - q_0}{h}\right) - V(q_0) \tag{2.17}$$

*Applying the implicit discrete Euler–Lagrange equations yields,*

$$p_0 = M\frac{q_1 - q_0}{h} + h\nabla V(q_0), \qquad p_1 = M\frac{q_1 - q_0}{h}.$$

*Finally re-arranging these equations results in the variational integrator, also known as symplectic Euler-A,*

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

*Now we apply this same approximation scheme to the associated exact Type II generating function. The boundary-value formulation of the exact discrete right Hamiltonian is given by,*

$$H_d^{+,E}(q_0,p_1) = \left(p_1q_1 - \int_0^h \left[p\dot{q} - H(q,p)\right] dt\right),$$

*where $(q(t), p(t))$ satisfy Hamilton's equations with boundary conditions $q(0) = q_0$, $p(h) = p_1$. The first-order finite difference approximations of the velocities yields the following equations for the momentum,*

$$p_0 \approx M\frac{q_1 - q_0}{h},$$
$$p_1 \approx M\frac{q_1 - q_0}{h}.$$

*Solving for $q_1$ in terms of $q_0$ and $p_1$ yields the approximation $q_1 \approx q_0 + hM^{-1}p_1$ and this simplifies the approximation to $p_0$ as $p_0 \approx p_1$. Now using these approximations in combination with the rectangular rule about the initial point yields the discrete right Hamiltonian,*

$$H_d^+(q_0, p_1; h) = p_1(q_0 + hM^{-1}p_1) - h\left(\frac{1}{2}p_1M^{-1}p_1 - V(q_0)\right).$$

*After applying the implicit discrete Hamilton's equations and re-arranging terms the resulting method is again symplectic Euler-A,*

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

The composition of the Legendre transform and the discretization of a generating function will be called commutative, if regardless of the order of this composition, either resulting generating function leads to the same symplectic map. In this context, the previous example shows that this particular discretization scheme, which includes both the trajectory approximation and quadrature rule, commutes with the Legendre transform between Type I and Type II generating functions.

Now we look at the same velocity approximation, but using the rectangular rule about the end point rather than the initial point.

**Example 3.** *As before we first build the discrete Lagrangian with $q_0$, $q_1$ fixed and velocity approximations,*

$$\dot{q}_0 \approx \frac{q_1 - q_0}{h},$$
$$\dot{q}_1 \approx \frac{q_1 - q_0}{h}.$$

*Applying the rectangular rule about the endpoint yields the discrete Lagrangian,*

$$L_d(q_0, q_1; h) = \frac{1}{2}\left(\frac{q_1 - q_0}{h}\right)M\left(\frac{q_1 - q_0}{h}\right) - V(q_1), \tag{2.18}$$

*and the associated variational integrator is symplectic Euler-B,*

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

*Now applying the same approximation scheme to construct the discrete right Hamiltonian results in the following Type II generating function,*

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

*Applying the implicit discrete Hamilton's equations and re-arranging terms results in the method,*

$$q_1 = q_0 + hM^{-1}p_0, \qquad p_1 = p_0 - h\nabla V(q_0 + hM^{-1}p_1),$$

*which is not symplectic Euler-B.*

We have just proven the following theorem.

**Theorem 9.** *In general, the composition of the Legendre transform of a generating function and the discretization of a generating function do not commute.*

Therefore, the answer to our original question is that in general, a fixed approximation scheme used to construct a discrete Lagrangian will not generate the same method when it is used to construct a discrete Hamiltonian. In general, how might the two resulting methods differ? A complete characterization of this issue is subtle, and beyond the scope of this paper, but it will be a topic of future work. For now, we will consider how the two approaches differ when combined with the method of averaging, which will also serve to illustrate how the type of generating function and the associated boundary data can affect the numerical properties of the method.

### 2.3.2 Averaged Lagrangians and Hamiltonians

Averaging methods have played a role in solving differential equations since at least as far back as the time of Lagrange (see [51]), and they continue to play a key role particularly in the field of numerical differential equations applied to nearly

integrable systems or problems with multiple timescales. We consider perturbed Hamiltonian systems with Hamiltonians of the form,

$$H = H^{(A)} + \epsilon H^{(B)}, \tag{2.19}$$

where $\epsilon \ll 1$ and the dynamics of the Hamiltonian system corresponding to $H^{(A)}$ is exactly solvable or at the very least cheap to approximate. We call this an almost-integrable system. The motivation being that the dynamics of the system are largely influenced by an integrable Hamiltonian with simpler dynamics, but smaller influences also play a role in the overall dynamics. An example is the classic $n$-body problem of the solar system, where a particular planet's trajectory is largely influenced by the sun, but other planets and nearby objects also play a role. Averaging methods can be constructed to exploit the larger influence of $H^{(A)}$ on the dynamics of the system by averaging out the smaller influences. Ideally, averaging techniques will allow for larger time steps to be used while still yielding a reasonable approximation to the solution.

A variational integrator for such a system was proposed in [10] using a discrete Lagrangian formulation, which drew inspiration from the kick-drift-kick leapfrog method (see [53]). We will discuss the original Lagrangian formulation (hereafter referred to as the averaged Lagrangian) and in addition construct an analogous method in terms of a discrete right Hamiltonian (referred to as the averaged Hamiltonian). The Lagrangian corresponding to (2.19) is given by,

$$L = L^{(A)} + \epsilon L^{(B)}, \tag{2.20}$$

and we will make the assumption that $L^{(B)}(q(t), \dot{q}(t)) = -V^{(B)}(q(t))$.

The averaging method of interest has a local truncation error of $\mathcal{O}(\epsilon^2 h^3)$, and is defined in terms of a discrete Lagrangian, $L_d$. This method, proposed in [10], uses a discrete Lagrangian of the form,

$$\begin{aligned} L_d(q_0, q_1, h) &= L_d^{(A),E}(q_0, q_1; h) + \epsilon \int_0^h L^{(B)}(q_A(q_0, q_1, t), \dot{q}_A(q_0, q_1, t))dt \\ &= L_d^{(A),E}(q_0, q_1; h) - \epsilon \int_0^h V^{(B)}(q_A(q_0, q_1, t))dt, \end{aligned}$$

where we denote the trajectory corresponding to $L^{(A)}$ with boundary conditions $(q_0, q_1)$ by $(q_A(t), \dot{q}_A(t))$. The idea is to use the dynamics of $L^{(A)}$, which is either solved for exactly or efficiently and accurately approximated, to average the contribution of $L^{(B)}$ to the dynamics. The corresponding discrete Hamiltonian map is given implicitly by

$$-p_0 = D_1 L_d^{(A),E}(q_0, q_1; h) - \epsilon \int_0^h D_1 V^{(B)}(q_A(q_0, q_1, t))dt, \tag{2.21a}$$

$$p_1 = D_2 L_d^{(A),E}(q_0, q_1; h) - \epsilon \int_0^h D_2 V^{(B)}(q_A(q_0, q_1, t))dt. \tag{2.21b}$$

As shown in [10], this method has local truncation error of size $\mathcal{O}(\epsilon^2 h^3)$. Using the notation $p_0^A(q_0, q_1) = -D_1 L_d^{(A),E}(q_0, q_1; h)$ and $p_1^A(q_0, q_1) = D_2 L_d^{(A),E}(q_0, q_1; h)$, we rearrange the above equations to get

$$p_0 - \epsilon \int_0^h D_1 V^{(B)}(q_A(q_0, q_1, t))dt = p_0^A(q_0, q_1), \tag{2.22a}$$

$$p_1 = p_1^A(q_0, q_1) - \epsilon \int_0^h D_2 V^{(B)}(q_A(q_0, q_1, t))dt. \tag{2.22b}$$

In [10], it is noted that $-\epsilon \int_0^h D_1 V^{(B)}(q_A(q_0, q_1, t))dt$ is an average along the trajectory generated by $L^{(A)}$ which, in general, gives more weight to the initial periods of the trajectory, while $-\epsilon \int_0^h D_2 V^{(B)}(q_A(q_0, q_1, t))dt$ is an average along the trajectory generated by $L^{(A)}$ that, in general, favors the latter periods of the trajectory.

Now let us consider the discrete right Hamiltonian given by the same form of approximation,

$$H_d^+(q_0, p_1; h) = H_d^{(A),+,E}(q_0, p_1; h) + \epsilon \int_0^h V^{(B)}(q_A(q_0, p_1, t))dt.$$

The discrete right Hamiltonian map is given implicitly by

$$p_0 = D_1 H_d^{(A),+,E}(q_0, p_1; h) + \epsilon \int_0^h D_1 V^{(B)}(q_A(q_0, p_1, t))dt,$$

$$q_1 = D_2 H_d^{(A),+,E}(q_0, p_1; h) + \epsilon \int_0^h D_2 V^{(B)}(q_A(q_0, p_1, t))dt.$$

Using the notation $p_0^A(q_0, p_1) = D_1 H_d^{(A),+,E}(q_0, p_1; h)$ and $q_1^A(q_0, p_1) = D_2 H_d^{(A),+,E}(q_0, p_1; h)$, we rearrange the equations to yield

$$p_0 - \epsilon \int_0^h D_1 V^{(B)}(q_A(q_0, p_1, t))dt = p_0^A(q_0, p_1), \tag{2.23a}$$

$$q_1 = q_1^A(q_0, p_1) + \epsilon \int_0^h D_2 V^{(B)}(q_A(q_0, p_1, t))dt. \tag{2.23b}$$

**Theorem 10.** *The method defined implicitly by* (2.23) *has local truncation error* $\mathcal{O}(\epsilon^2 h^3)$.

*Proof.* Using variational error analysis, we need to show

$$\begin{aligned}\mathcal{O}(\epsilon^2 h^3) &= H_d^{E,+} - H_d^+ \\ &= \Delta_A + \epsilon \Delta_B,\end{aligned}$$

where $\Delta_A$ is given by

$$\begin{aligned}p(h)q(h) - \int_0^h [p(t)\dot{q}(t) - H^{(A)}(q(t), p(t))]dt \\ - \left( p_A(h)q_A(h) - \int_0^h [p_A(t)\dot{q}_A(t) - H^{(A)}(q_A(t), p_A(t))]dt \right),\end{aligned}$$

and $\epsilon\Delta_B$ is given by

$$\epsilon \int_0^h \left[ V^{(B)}(q(t)) - V^{(B)}(q_A(t)) \right] dt.$$

Using a functional Taylor expansion, $\Delta_A$ becomes

$$\begin{aligned}\Delta_A = \frac{\delta}{\delta q_A} \left( \int_0^h [p_A(t)\dot{q}_A(t) - H^{(A)}(q_A(t), p_A(t))]dt \right) \delta q_A \\ + \frac{\delta^2}{\delta q_A^2} \left( \int_0^h [p_A(t)\dot{q}_A(t) - H^{(A)}(q_A(t), p_A(t))]dt \right) \delta q_A^2 + \mathcal{O}(\delta q_A^3),\end{aligned}$$

where $\delta q_A$ is the difference between $q$ and $q_A$. Noting that $q$ and $q_A$ differ in forces of order $\epsilon h^2$ and $p$ differs from $p_A$ to first order in $\epsilon h$, implies that $\delta q_A$ is on the order of $\mathcal{O}(\epsilon h)$. This can be seen explicitly by comparing Taylor expansions about time zero. Since $q_A$ satisfies Hamilton's equations for $H^{(A)}$, the first variation vanishes (see Lemma 2.1 of [31]) leaving a term on the order of $h\delta q_A^2$. Therefore, we have

$$\Delta_A = \mathcal{O}(\epsilon^2 h^3).$$

Likewise, a functional Taylor expansion for $\Delta_B$ yields,

$$\Delta_B = \frac{\delta}{\delta q_A} \left[ \int_0^h V^{(B)}(q_A(t))dt \right] \delta q_A + \mathcal{O}(\delta q_A^2).$$

Noting that $V^{(B)}$ is only a function of $q_A$ and that $q$ differs from $q_A$ on the order of $\epsilon h^2$, implies $\epsilon\Delta_B = \mathcal{O}(\epsilon^2 h^3)$. □

**Theorem 11.** *Assuming the flow associated with $L^{(A)}$ is time-reversible, then both methods, defined respectively by* (2.22) *and* (2.23)*, are symmetric methods.*

*Proof.* The discrete Lagrangian associated with (2.22) is given by,

$$L_d(q_0, q_1; h) = L_d^{(A),E}(q_0, q_1; h) - \epsilon \int_0^h V^{(B)}(q^A(q_0, q_1, t))dt.$$

The adjoint of the discrete Lagrangian is given by,

$$\begin{aligned}
(L_d(q_0, q_1; h))^* &= -L_d(q_1, q_0; -h) \\
&= -L_d^{(A),E}(q_1, q_0; -h) + \epsilon \int_0^{-h} V^{(B)}(q^A(q_1, q_0, t))dt \\
&= -L_d^{(A),E}(q_1, q_0; -h) - \epsilon \int_0^{h} V^{(B)}(q^A(q_1, q_0, t))dt \\
&= L_d^{(A),E}(q_0, q_1; h) - \epsilon \int_0^{h} V^{(B)}(q^A(q_0, q_1, t))dt \\
&= L_d(q_0, q_1; h).
\end{aligned}$$

The third equality comes from the time-reversibility of the flow associated with $L^{(A)}$, and the fourth equality uses this property together with the fact that the exact discrete Lagrangian is self-adjoint.

The discrete right Hamiltonian associated with (2.23) is given by,

$$H_d^+(q_0, p_1; h) = H_d^{(A),+,E}(q_0, p_1; h) + \epsilon \int_0^h V^{(B)}(q_A(q_0, p_1, t))dt.$$

The adjoint of the discrete right Hamiltonian is given by,

$$\begin{aligned}
(H_d^+)^*(p_0, q_1; h) &= -H_d^+(q_1, p_0; -h) \\
&= -H_d^{(A),+,E}(q_1, p_0; -h) - \epsilon \int_0^{-h} V^{(B)}(q_A(q_1, p_0, t))dt \\
&= -H_d^{(A),+,E}(q_1, p_0; -h) + \epsilon \int_0^{h} V^{(B)}(q_A(q_1, p_0, t))dt \\
&= H_d^{(A),-,E}(p_0, q_1; h) + \epsilon \int_0^{h} V^{(B)}(q_A(p_0, q_1, t))dt
\end{aligned}$$

$$= H_d^-(p_0, q_1; h),$$

where the third equality comes from the time-reversibility of the flow associated with $H^{(A)}$, and the fourth equality uses this property together with the fact that the exact discrete Hamiltonian is self-adjoint. By Definition 2.8 and Theorem 2.9, the method is symmetric. □

It can be shown that methods (2.22) and (2.23) do not result in the same symplectic map. How do these respective maps differ? To gain insight into this question we now turn to numerical experimentation.

## 2.4 Numerical Results

### 2.4.1 Exact Generating Functions

First, we consider the unperturbed harmonic oscillator boundary-value problem,

$$\ddot{q}(t) + q(t) = 0, \qquad q(0) = q_0,\ q(h) = q_1. \tag{2.24}$$

Analytically, the boundary-value problem is not well-posed when $h$ is an integer multiple of $\pi$, and in particular there are infinitely many solutions. Recall that the exact discrete Lagrangian is given by,

$$L_d^E(q_0, q_1; h) = \int_0^h L(q_{01}(t), \dot{q}_{01}(t))dt, \tag{2.25}$$

where $q_{01}(0) = q_0$, $q_{01}(h) = q_1$, and $q_{01}(t)$ satisfies the Euler–Lagrange equation in the time interval $(0, h)$. Thus, the exact Type I generating function is ultimately defined in terms of such a boundary-value problem. The integrator obtained from the exact discrete Lagrangian is given by,

$$\begin{aligned} q_1 &= q_0 \cos(h) + p_0 \sin(h), \\ p_1 &= q_1 \cot(h) - q_0 \csc(h). \end{aligned}$$

This integrator is analytically the true solution of the harmonic oscillator initial-value problem, where $(q(0), p(0)) = (q_0, p_0)$, and consequently the local truncation

error is zero. However, noting that $\cot(h)$ and $\csc(h)$ both involve dividing by $\sin(h)$, we expect increased round-off error for values of $h$ that are near integer multiples of $\pi$.

Similarly, the exact discrete right Hamiltonian is given by,

$$H_d^{+,E}(q_0, p_1; h) = p_1 q_1 - \int_0^h [p_{01}(t)\dot{q}_{01}(t) - H(q_{01}(t), p_{01}(t))]dt, \tag{2.26}$$

where $q_{01}(0) = q_0$, $p_{01}(h) = p_1$, and $(q_{01}(t), p_{01}(t))$ satisfies Hamilton's equations in the time interval $(0, h)$. This is related to the unperturbed harmonic oscillator boundary-value problem given by,

$$\dot{q}(t) = p(t), \qquad \dot{p}(t) = -q(t), \qquad q(0) = q_0,\ p(h) = p_1. \tag{2.27}$$

This boundary-value problem is not well-posed for values of $h$ that are odd multiples of $\frac{\pi}{2}$ and there are infinitely many solutions for such values of $h$. The integrator obtained from the exact discrete right Hamiltonian for the harmonic oscillator is given by,

$$\begin{aligned} p_1 &= p_0 \cos(h) - q_0 \sin(h), \\ q_1 &= p_1 \tan(h) + q_0 \sec(h). \end{aligned}$$

This integrator is analytically the true solution to the harmonic oscillator initial-value problem, where $(q(0), p(0)) = (q_0, p_0)$ and the local truncation error will be zero. Noting that the method involves $\tan(h)$ and $\sec(h)$, we expect increased round-off error around odd multiples of $\frac{\pi}{2}$.

Both of the integrators given by the exact discrete Lagrangian and the exact discrete right Hamiltonian have been implemented for the harmonic oscillator with initial conditions $(q_0, p_0) = (1, 0)$ over the time interval $[0, 10000]$, and the energy error is shown in Figure 2.1. Note the jump in round-off error corresponding to values of $h$ that are odd multiples of $\pi$ (for the discrete Lagrangian) and odd multiples of $\frac{\pi}{2}$ (for the discrete right Hamiltonian). The bottom plot takes the minimum error of the two methods, and this indicates that a step-size causing noticeable round-off error for one method will work just fine for the other method.

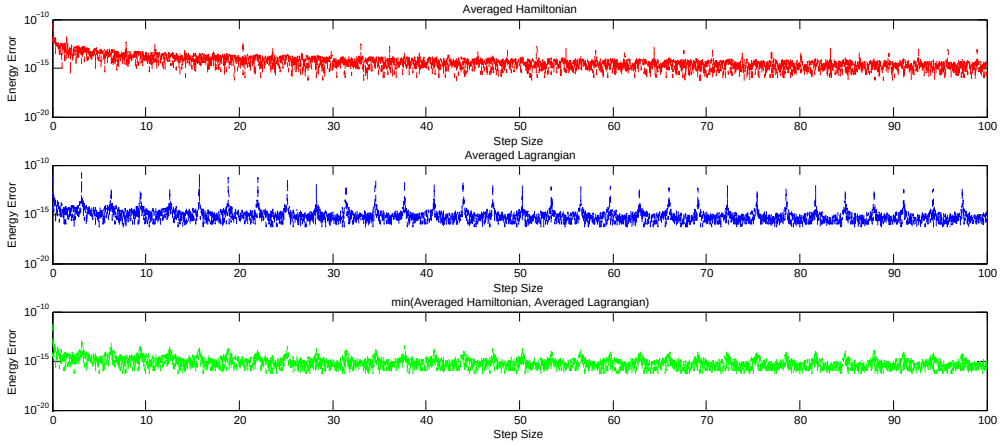



**Figure 2.1**: Energy error versus step size for exact generating functions. The first plot is the energy error versus step size for the exact discrete right Hamiltonian applied to the harmonic oscillator. The second plot shows the energy error versus step size for the exact discrete Lagrangian, while the third plot takes the minimum of the energy error from either method.

In this particular case, we can conclude that the numerical difference between the symplectic maps generated by the respective exact discrete Lagrangian and exact discrete right Hamiltonian is a matter of numerical conditioning, which is inherited from the underlying ill-posedness of the associated boundary-value problem. Despite the fact that the methods are applied to an initial-value problem, numerical properties can be attributed to a boundary-value problem that is no longer visible in the methods themselves. Considering many symplectic integrators are derived independently of the variational integrator formulation, perhaps some of their numerical properties can be better understood by reinterpreting them in the framework of variational integrators.

### 2.4.2 Averaged variational integrators for nonlinearly perturbed harmonic oscillator

Now we consider the previous averaging methods applied to a Hamiltonian of the form,

$$H(q,p) = \frac{1}{2}(p^2+q^2) + \frac{\epsilon}{3}q^3, \tag{2.28}$$

which is the Hamiltonian for a nonlinearly perturbed harmonic oscillator. The corresponding averaged Lagrangian is given by

$$L_d(q_0, q_1, h) = \int_0^h \frac{1}{2}(\dot{q}_A(t)^2 - q_A(t)^2)dt - \int_0^h \frac{\epsilon}{3} q_A(t)^3 dt, \tag{2.29}$$

where $(q_A(t), \dot{q}_A(t))$ is the solution corresponding to the Lagrangian $L^{(A)}(q, \dot{q}) = \frac{1}{2}(\dot{q}^2 - q^2)$ with boundary conditions $(q_0, q_1)$. Analogously, the averaged Hamiltonian is given by

$$H_d^+(q_0, p_1, h) = p_1 q_A(h) - \int_0^h \frac{1}{2}(p_A(t)^2 - q_A(t)^2)dt + \frac{\epsilon}{3}\int_0^h q_A(t)^3 dt, \tag{2.30}$$

where $(q_A(t), p_A(t))$ is the solution corresponding to the Hamiltonian $H^{(A)}(q, p)$ with boundary conditions $(q_0, p_1)$. Applying the discrete right and left Legendre transforms implicitly defines the discrete Hamiltonian map for $L_d(q_0, q_1, h)$ and the discrete right Hamiltonian map for $H_d^+(q_0, p_1, h)$, which yields the respective one-step methods. Numerical simulations were run over a time-span from 0 to 10000 or the nearest integer value to 10000 for the respective time-step. The initial conditions are given by $(q_0, p_0) = (1, 0)$.

Figures 2.2 and 2.3 show plots of the energy error versus step size for two different values of $\epsilon$. The third plot in each of the figures hints that the discrete Lagrangian and discrete right Hamiltonian have numerical resonance that is nearly dual, in some sense, with respect to step size. The discrete Lagrangian exhibits excessive numerical resonance for step sizes near odd multiples of $\pi$, while the discrete right Hamiltonian exhibits excessive numerical resonance for step sizes near odd multiples of $\frac{\pi}{2}$. It should be noted that the arbitrary value of $10^6$ was substituted for output that was either near infinite or NaN. What is particularly striking is that the occurence of the numerical resonance is intimately connected to the corresponding boundary-values for each generating function.

Now this by no means provides a rigorous analysis of the numerical resonances, nor does it fully explain all of the resonance effects, but it does provide motivation and insight into the numerical differences between the discrete Lagrangian and discrete right Hamiltonian. A more in-depth analysis might be provided by applying something similar to modulated Fourier expansions (see [16; 18], and

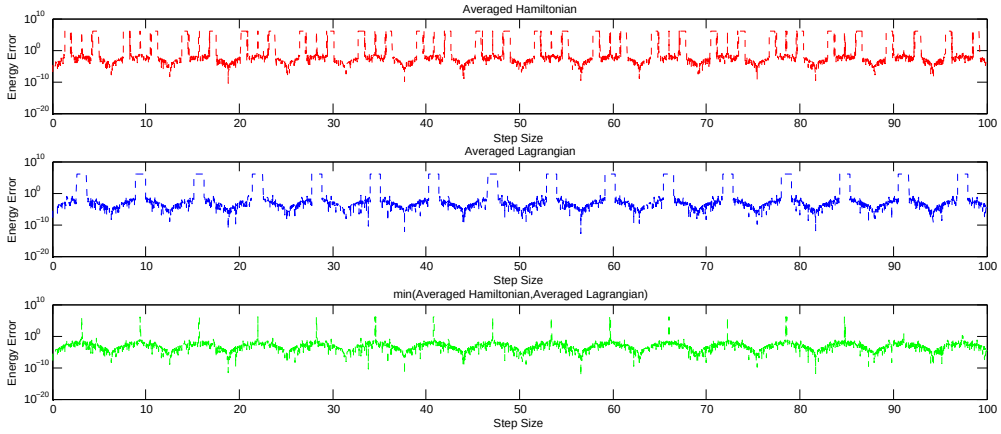



**Figure 2.2**: Energy error versus stepsize for $\epsilon = 0.1$. Three plots of step size versus energy error with fixed $\epsilon = 0.1$. The first plot corresponds to the averaged Hamiltonian, and it suffers from numerical resonance around odd integer multiples of $\frac{\pi}{2}$ and exactly at odd multiples $\pi$. The second plot corresponds to the averaged Lagrangian which suffers from numerical resonance around odd multiples of $\pi$. The last plot takes the minimum error of the respective methods.

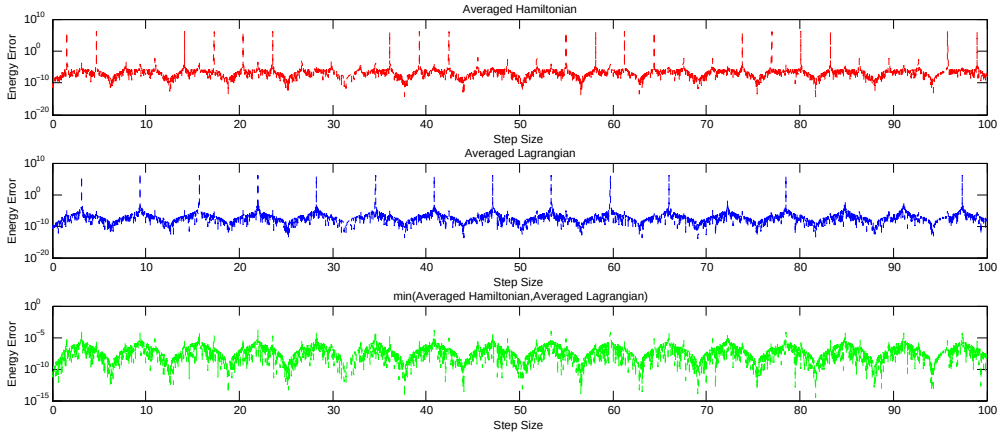



**Figure 2.3**: Energy error versus stepsize for $\epsilon = 0.001$. Three plots of step size versus energy error with fixed $\epsilon = 0.001$. The first plot corresponds to the averaged Hamiltonian, and it suffers from numerical resonance at some odd integer multiples of $\frac{\pi}{2}$. The second plot corresponds to the averaged Lagrangian which suffers from numerical resonance around odd multiples of $\pi$. The last plot takes the minimum error of the respective methods.

Chapter XIII of [19]). Modulated Fourier expansions are particularly well-suited for oscillatory problems when large step sizes are sought. The standard backward error analysis relies on $h\omega \to 0$, which is not the case for high oscillatory problems when seeking large step sizes. Modulated Fourier expansions can provide a tool for deriving many of the same results as backward error analysis, such as long-term energy preservation. Furthermore, it can be quite useful for examining the step sizes that lead to excessive numerical resonance. However, it should be noted that while modulated Fourier expansions have been used quite successfully to analyze explicit trigonometric integrators, it is not quite as clear how easily it can deal with implicit integrators such as those obtained from the discrete averaged Lagrangian and discrete averaged Hamiltonian.

## 2.5 Conclusion

Error analysis and symmetry results have now been extended to cover discrete Hamiltonian variational integrators. Furthermore, examples have been presented indicating that the underlying well-posedness in terms of the boundary conditions of the exact generating function can be directly related to numerical resonance. In conclusion, it is clear that the numerical properties of variational integrators are dependent on both the approximation scheme used in constructing the generating function and the type of generating function being approximated.

This paper indicates that the class of variational integrators generated using the Hamiltonian formulation are not necessarily equivalent to the ones obtained from the Lagrangian formulation, and it would therefore be of interest to continue developing methods based on the discrete Hamiltonian variational integrator formulation. In particular, the results presented suggest that further work remains to be done to better understand the circumstances under which it is preferable to favor one approach over the other.

Chapter 2, in full, is a reprint of the material that has been accepted for publication by IMA Journal of Numerical Analysis, 2017. Schmitt, Jeremy; Leok, Melvin, Oxford University Press, 2017. The dissertation author was the primary

investigator and author of this material.

# Chapter 3

# Lagrangian and Hamiltonian Taylor Variational Integrators

## 3.1 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 [30], 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.

## 3.2 Discrete Mechanics

Discrete Lagrangian mechanics [35] 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

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

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***,

$$L_d^E(q_0, q_1; h) = \int_0^h L(q_{01}(t), \dot{q}_{01}(t))dt, \tag{3.1}$$

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,

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \tag{3.2}$$

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,

$$p_k = -D_1 L_d(q_k, q_{k+1}), \qquad p_{k+1} = D_2 L_d(q_k, q_{k+1}), \tag{3.3}$$

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 §3.2.1 is proven in [35]. 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.

### 3.2.1 Variational error analysis

The natural setting for analyzing the order of accuracy of a variational integrator is the variational error analysis framework introduced in [35]. In particular, Theorem 2.3.1 of [35] 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 (3.1) to order $p$, i.e.,

$$L_d(q_0, q_1; h) = L_d^E(q_0, q_1; h) + \mathcal{O}(h^{p+1}), \tag{3.4}$$

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.

## 3.3 Lagrangian Taylor Variational Integrator

The exact discrete Lagrangian (3.1) 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 $q_1 = \pi_Q \Phi_h(q_0, v_0)$, 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 $q_1 = \pi_Q \Phi_h(q_0, v_0)$. 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,

$$q_1 = \pi_Q \circ \Psi_h^{(r+1)}(q_0, \tilde{v}_0), \tag{3.5}$$

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 $\tilde{v}_0$,

$$(q_{c_i}, v_{c_i}) = \Psi_{c_i h}^{(r)}(q_0, \tilde{v}_0). \tag{3.6}$$

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

$$L_d(q_0, q_1; h) = h \sum_{i=1}^{m} b_i L(q_{c_i}, v_{c_i}). \tag{3.7}$$

(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.** *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* (3.5)*, 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 equation* (3.5) *as being a* $(r+1)$*-order Taylor method for the second-order differential equation on* $Q$*, and* (3.6) *as a* $r$*-order Taylor method on the first-order differential equation on* $TQ$*. In particular, notice that because of the canonical projection* $\pi_Q$ *in equation* (3.5)*, 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 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 Appendices for their proofs).

**Lemma 2.** $\tilde{v}_0$ *as defined by,* (3.5)*, approximates* $v_0$ *to at least* $\mathcal{O}(h^{r+1})$*.*

**Lemma 3.** *A* $r$*-order Taylor method with initial conditions* $(q_0, \tilde{v}_0)$*, where* $\tilde{v}_0$ *is defined by* (3.5)*, is accurate to at least* $\mathcal{O}(h^{r+1})$ *for the Euler–Lagrange boundary-value problem with boundary conditions* $(q_0, q_1)$*.*

**Theorem 12.** *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* (3.7) *has order of accuracy at least* $\min(r+1, s)$*.*

*Proof.* $(q_d(t), v_d(t))$, associated with the Taylor method $\Psi_h$ of order $r$ and initial data $(q_0, \tilde{v}_0)$, approximates the exact solution $(q_{01}(t), v_{01}(t))$ of the Euler–Lagrange boundary-value problem with the following error,

$$\begin{aligned} q_{01}(c_i h) &= q_d(c_i h) + \mathcal{O}(h^{r+1}), \\ v_{01}(c_i h) &= v_d(c_i h) + \mathcal{O}(h^{r+1}). \end{aligned}$$

If the numerical quadrature formula is order $s$ accurate, then

$$\begin{aligned} L_d^E(q_0, q_1; h) &= \int_0^h L(q_{01}(t), v_{01}(t))dt \\ &= \left[h \sum\nolimits_{i=1}^m b_i L(q_{01}(c_i h), v_{01}(c_i h))\right] + \mathcal{O}(h^{s+1}) \end{aligned}$$

$$
\begin{aligned}
&= \Big[h \sum_{i=1}^{m} b_i L(q_d(c_i h) + \mathcal{O}(h^{r+1}), v_d(c_i h) + \mathcal{O}(h^{r+1}))\Big] + \mathcal{O}(h^{s+1}) \\
&= \Big[h \sum_{i=1}^{m} b_i L(q_d(c_i h), v_d(c_i h))\Big] + \mathcal{O}(h^{r+2}) + \mathcal{O}(h^{s+1}) \\
&= L_d(q_0, q_1; h) + \mathcal{O}(h^{r+2}) + \mathcal{O}(h^{s+1}) \\
&= L_d(q_0, q_1; h) + \mathcal{O}(h^{\min(r+1,s)+1}),
\end{aligned}
$$

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 [29] and [30], which utilized collocation and the shooting-method, respectively.

**Example 4.** *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,*

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

$$p_1 = M\Big(\frac{q_1 - q_0}{h}\Big).$$

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

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

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 5.** *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,*

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

As demonstrated above, Lagrangian Taylor variational integrators provide a very general family of symplectic integrators that include not only 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 [24], [46]). In particular, automatic differentiation allows for accurate and relatively cheap derivative evaluations (see [20], [41], [39]). In general, higher-order Taylor variational integrators will require solving a system of nonlinear equations, which can be dealt with using standard methods (see [21]).

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 step sizes, which may justify the higher cost per step. We will further examine these topics in section 3.5. Next, we consider discrete Hamiltonian formulations and symmetric formulations of the Taylor variational integrator.

## 3.4 Hamiltonian and Symmetric Taylor Variational Integrators

### 3.4.1 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 [26], [31]), 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 [44] 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 $\min(r+1, s)$,

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

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

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

$$q_1 = D_2 H_d^+(q_0, p_1), \qquad p_0 = D_1 H_d^+(q_0, p_1). \tag{3.8}$$

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

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

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 $\min(r+1, s)$,

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

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

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

$$p_1 = -D_2 H_d^-(q_1, p_0; h), \qquad q_0 = -D_1 H_d^-(q_1, p_0; h). \tag{3.9}$$

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,

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

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

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

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 [44] 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. We will denote the resulting symmetric method by SVHd, and we compare it to Störmer–Verlet in section 3.5 (see Figure 3.7 and Figure 3.9).

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 13.** *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 [44]).

### 3.4.2 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,

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

where $\Psi_h^r$ denotes the $r$-th order Taylor method with step size $h$. 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 [19]). 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 $\tilde{v}_1$,

$$q_1 = \Psi_h^{(r)}(q_0, \tilde{v}_0), \qquad q_0 = \Psi_{-h}^{(r)}(q_1 \tilde{v}_1). \tag{3.10}$$

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

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

Note $q_0$, $q_1$, $\tilde{v}_0$, 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 14.** *The symmetric Taylor variational integrator is a symmetric method.*

*Proof.* By theorem 2.4.1 of [35], 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 (3.10) into,

$$q_0 = \Psi^{(1)}_{-h}(q_1, \tilde{v}_0^*), \qquad q_1 = \Psi^{(1)}_{h}(q_0, \tilde{v}_1^*),$$

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

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

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,

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

□

**Theorem 15.** *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 2 and 3 imply that the nodal value approximations, $q_{c_i}$ and $v_{c_i}$, are of order $r$ and $r-1$ respectively. Therefore,

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

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 [35], the resulting variational integrator, denoted by $\tilde{\Psi}_h$, is at least of order $r$, i.e.,

$$\begin{aligned}
\tilde{\Psi}_h(q_0, v_0) &= \Phi_h(q_0, v_0) + \mathcal{O}(h^{r+1}) \\
&= \Phi_h(q_0, v_0) + C(q_0, v_0) h^{r+1} + \mathcal{O}(h^{r+2}),
\end{aligned}$$

where $\Phi_h$ is the true flow of the Euler–Lagrange equations, and the last equality is a consequence of the implicit function theorem.

Finally, since the variational integrator is symmetric and $r+1$ is even, the method will be of order $r+1$ as the following implies.

$$\begin{aligned}
\Phi_h(q_0, v_0) - C(q_0, v_0) h^{r+1} + \mathcal{O}(h^{r+2}) &= \Phi_h^*(q_0, v_0) - C(q_0, v_0)(-h)^{r+1} + \mathcal{O}(h^{r+2}) \\
&= \tilde{\Psi}_h^*(q_0, v_0) \\
&= \tilde{\Psi}_h(q_0, v_0) \\
&= \Phi_h(q_0, v_0) + C(q_0, v_0) h^{r+1} + \mathcal{O}(h^{r+2}),
\end{aligned}$$

which implies $C(q_0, v_0) h^{r+1} = 0$, and the method is of order $r+1$ as claimed. □

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.

## 3.5 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 $\tilde{v}_0$:

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

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

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

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 found 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.

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 [50]) 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 Figure 3.1).

In practice, setting the nonlinear solver tolerance one or two orders above the order 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.

### 3.5.1 Automatic Differentiation

As with the Taylor method, an efficient general purpose implementation will require an efficient means of computing derivatives, such as automatic differentiation. For the following simulations, we used the AdiGator automatic differentiation package for MATLAB (see [40]). Implementation of a high-order Taylor variational integrator requires both the evaluation of higher time derivatives, $q^{(p+1)}(q_0, \tilde{v}_0)$, and the 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 Figure 3.2),

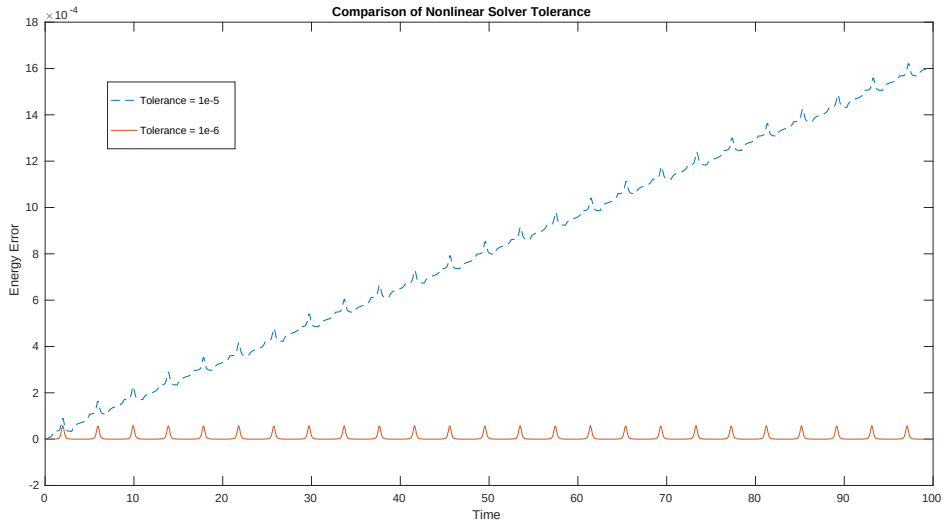



**Figure 3.1**: Pseudo-symplectic behavior. The plot of the energy preservation of a 4th order Taylor variational integrator applied to the simple pendulum with two different tolerance levels for the nonlinear solver and a step size of 0.1. Energy drift is evident when the tolerance level is set at $10^{-5}$ or larger, but the drift disappears for smaller tolerance levels. The method had an average energy error around $6.5 \cdot 10^{-5}$ for a tolerance of $10^{-6}$, and an average energy error of $8.1 \cdot 10^{-4}$ for a tolerance of $10^{-5}$.

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. There appears to be some relationships between the Jacobians and the time derivatives that could potentially be exploited to decrease the evaluation costs. For instance,

$$\frac{\partial q^{(3)}(q_0, \tilde{v}_0)}{\partial \tilde{v}_0} = \left[ q^{(3)}\left(q_0, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \;\; q^{(3)}\left(q_0, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \right],$$

which allows us to replace expensive Jacobian evaluations with cheaper time derivative evaluations. Additionally, Jacobians of higher-order time derivatives appear to have some relations to Jacobians of lower-order time derivatives, such as,

$$\frac{\partial q^{(4)}(q_0, \tilde{v}_0)}{\partial \tilde{v}_0} = -2\frac{\partial q^{(3)}(q_0, \tilde{v}_0)}{\partial q_0}.$$

Hopefully, a good implementation of automatic differentiation will already take advantage of such relationships.

Automatic differentiation greatly benefits from the way it is compiled, which means the more efficient implementations will be in languages such as Fortran or C++. Another aspect to consider is parallel implementation. Combining automatic differentiation and parallel computing techniques has been shown to significantly reduce computational time (see [5]).

One possible implementation for the algorithm would be to construct the Taylor discrete Lagrangian, then apply automatic differentiation to the discrete Lagrangian in combination with a nonlinear solver to recover the discrete Legendre transforms and consequently $(q_1, p_1)$. In fact, this could provide a more general framework for the derivation of all implicit variational integrators.

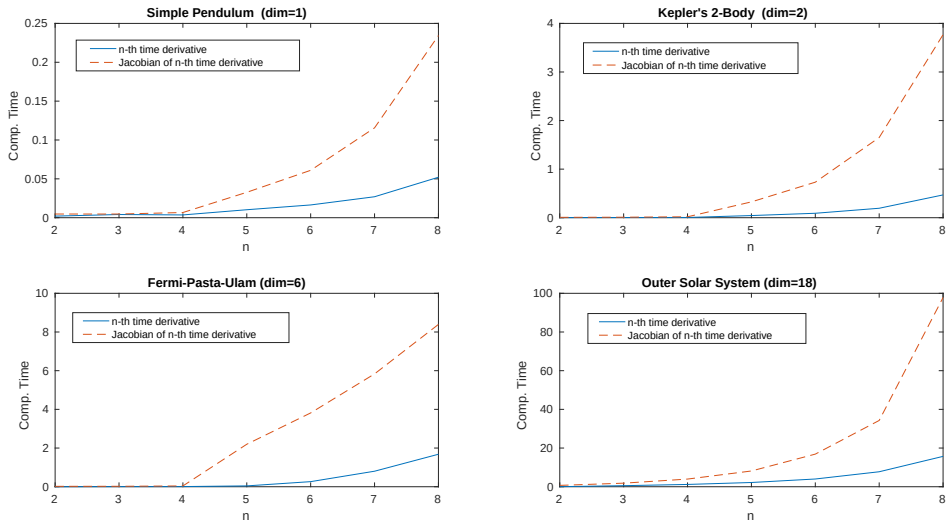



**Figure 3.2**: Computational cost of full derivative and partial Jacobian evaluations. The derivative order versus time plot of 100 evaluations of each derivative corresponding to 4 different models with increasing dimension. It is worth noting that the rate of growth in time needed for higher-order derivative evaluations appears to be independent of the dimension.

### 3.5.2 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 [30]). Overall, high-order Taylor methods perform quite well in terms of computational time versus global error. 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.

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. 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 [6]). 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 [47]). 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.

The following simulations were implemented in MATLAB.

### 3.5.3 Simple Pendulum

Consider the simple pendulum with unit mass and length in a gravitational field with $g = -9.8\text{m/s}^2$, 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 Figure 3.3, 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.

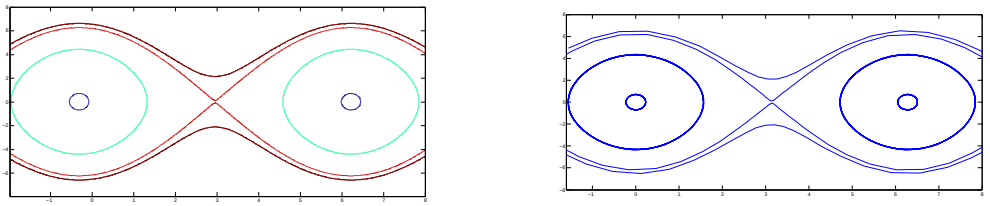
The simulation in Figure 3.4 used initial conditions $(q_0, p_0) = (\frac{\pi}{2}, 0)$. The 6th-order Taylor variational integrator performed well at a stepsize of $h = 0.5$, while the 6th-order Taylor method failed to generate a reasonable approximation for this stepsize. The ability of the Taylor variational integrator to perform well at larger stepsizes may gives it an advantage over traditional Taylor methods.

In Figure 3.5, 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 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.

### 3.5.4 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 [19]). Thus, constraining them to lie in a plane, we have Kepler's planar 2-body problem with corresponding Lagrangian,

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



(a) (b)

**Figure 3.3**: (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 step size $h = 0.1$ for the time interval $[0, 20]$.

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,

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

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 3.6 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 3.7 compares the Störmer–Verlet method (SV) to the discrete Hamiltonian composition method (SVHd) discussed in section 3.4.1. 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. However, SVHd does exhibit lower energy error and performs slightly better qualitatively, so when the problem is non-separable (and SV is implicit), SVHd may be a better alternative.

### 3.5.5 Henon-Heiles Model

The Henon–Heiles model attempts to capture the dynamics of a galaxy with cylindrical symmetry (see [19] for more info). The Hamiltonian is given by,

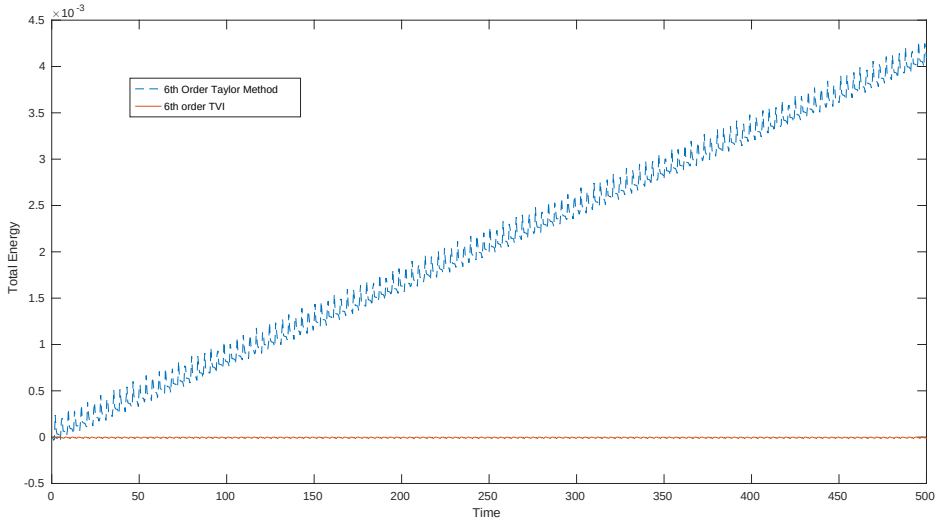



**Figure 3.4**: Simple Pendulum energy versus time. A plot of the Simple Pendulum total energy vs. time of the sixth-order integrators TVI6 and Taylor's method for a step size of $h = 0.5$. At this step size and time interval, Taylor's method has significant energy drift, and as a result its accuracy suffers.

$H(p,q) = \frac{1}{2}(p_1^2+p_2^2)+U(q)$, where $U(q) = \frac{1}{2}(q_1^2+q_2^2)+q_1^2q_2-\frac{1}{3}q_2^3$. The corresponding Euler–Lagrange equation is,

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

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 Figure 3.8) and $H_0 = \frac{1}{8}$ (see Figure 3.5). The second energy value corresponds to a chaotic system.

In Figure 3.8, 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. Amongst the higher-order methods, TVI6 and ShVI6 appear to be the more efficient methods. A 6th-order symmetric Taylor variational integrator would be even more efficient for higher-order accuracy.

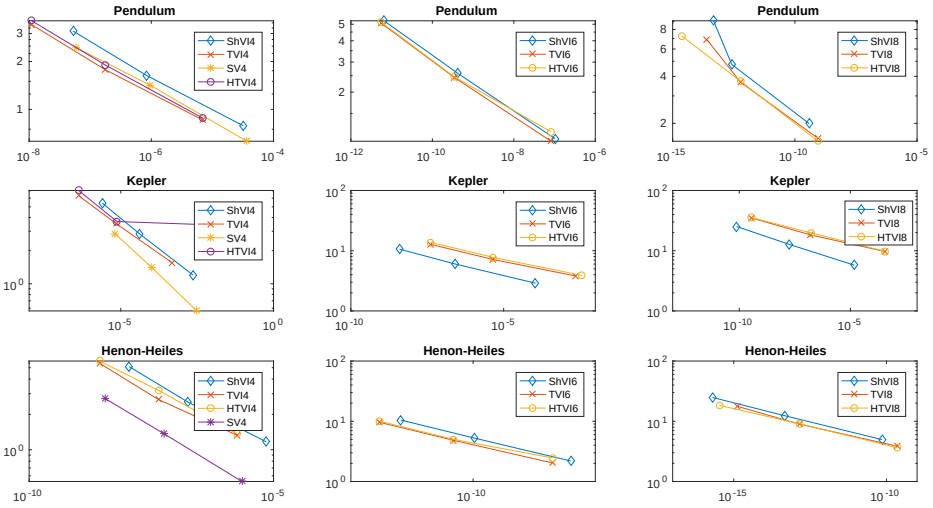



**Figure 3.5**: Plots of the average energy error versus computational time for the various variational integrators. The 4th-order symmetric Taylor variational integrator (SV4) is the clear winner in terms of efficiency, while comparisons of TVI, HTVI, and ShVI are mixed.

### 3.5.6 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 [11]). We apply the modified model as outlined in [19], 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,

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

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

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

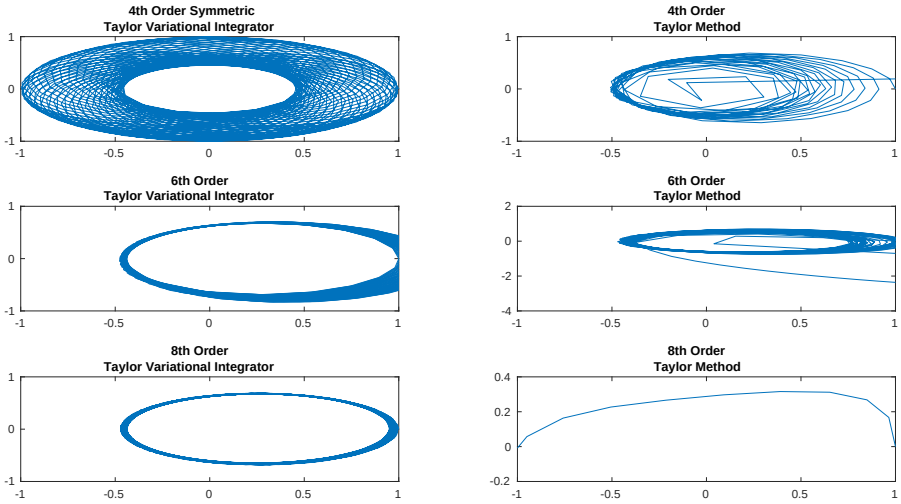



**Figure 3.6**: Kepler's planar 2-body problem. Position plots of Kepler's planar 2-body problem as generated by various integrators with a time step of $h = 0.25$ over a time interval of $[0, 250]$. The Taylor variational integrators exhibit close to the correct behavior, while the various Taylor methods all fail to capture the behavior of the system.

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

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

Figure 3.9 compares the Störmer–Verlet method to SVHd. The first couple of plots use a stepsize of $h = 0.03$, which is on the boundary of the linear stability of Störmer–Verlet (i.e. $h\omega = 1.5$). 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

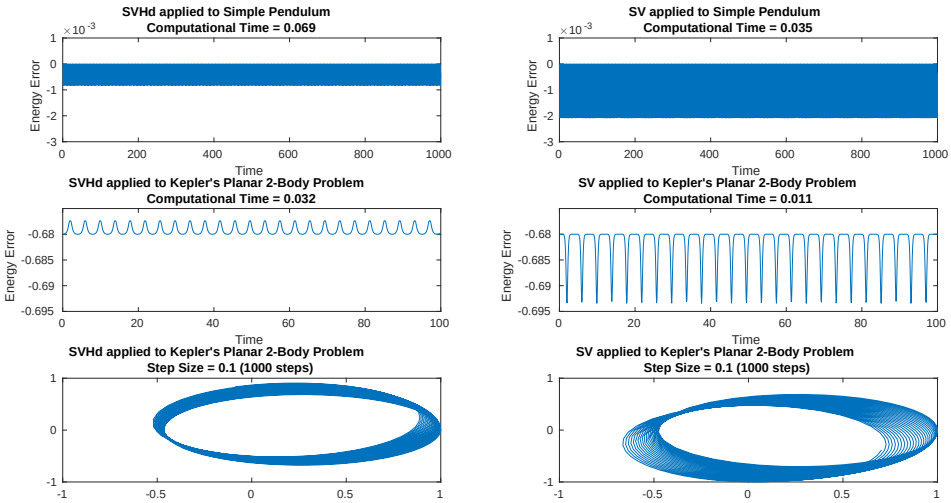



**Figure 3.7**: Comparison of Störmer–Verlet and SVHd. This plot compares the performance of Störmer–Verlet (SV) and the discrete Hamiltonian composition method (SVHd) from section 3.4.1. SVHd exhibits a much smaller amplitude in the energy error, as compared to SV, but the implicit nature of SVHd is reflected in the increased computational cost. Clearly, SV is preferable for separable problems, but for non-separable problems SVHd may be the better choice.

accurate solution, one should consider either the IMEX method (see [48]) or Filon-type methods (see [22]). The combination of exponential type integrators with symplectic and energy-preserving integrators was also recently considered in [**?** ].

### 3.5.7 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 [19], and corresponds to September 5, 1994 at 0h00. In Figure 3.10, 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

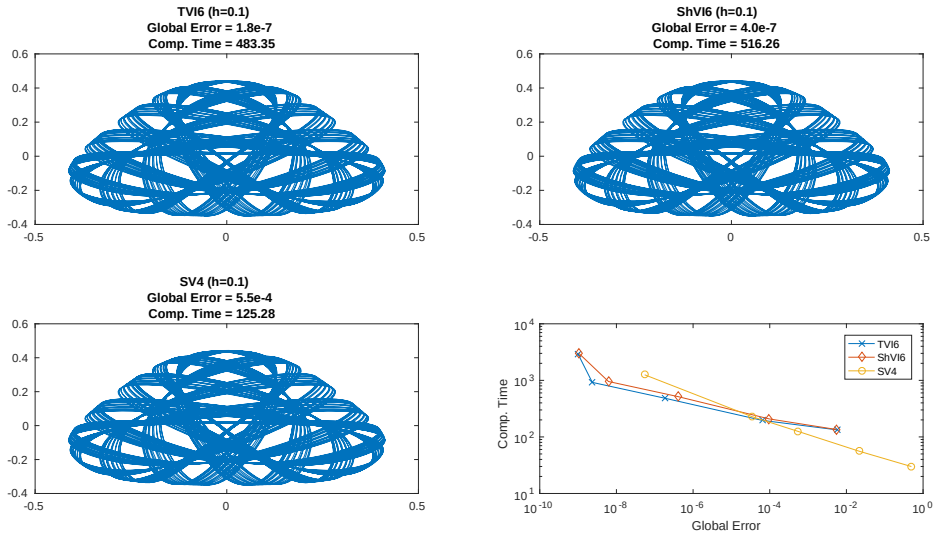



**Figure 3.8**: The Henon-Heiles model simulated over the time interval $[0, 1000]$. The bottom right plot compares the global error versus computational time of the 6th-order Taylor variational integrator (TVI6), the 6th-order Runge–Kutta based shooting variational integrator (ShVI6), and the 4th-order symmetric Taylor method (SV4).

a useful simulation at this stepsize, but both 6th-order integrators give a good representation of the system.

## 3.6 Conclusions and Future Directions

The Taylor variational integrators provide a way to build high-order symplectic integrators and include many 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 which a fixed number of iterations of the symplectic corrector can improve upon the performance of Taylor integrators for realistic

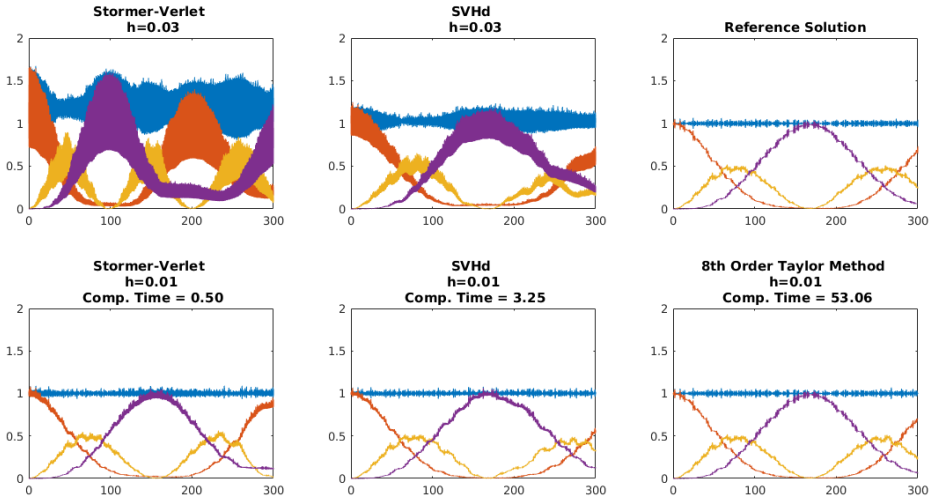



**Figure 3.9**: A comparison of Störmer–Verlet, SVHd, and the 8th-order Taylor method for the Fermi-Pasta-Ulam model. For $h = 0.03$, the Störmer–Verlet method is on the cusp of being linearly unstable. For $h = 0.01$, the methods all present a similar picture to the reference solution, but their global errors are quite large and none of them exhibit good accuracy.

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.

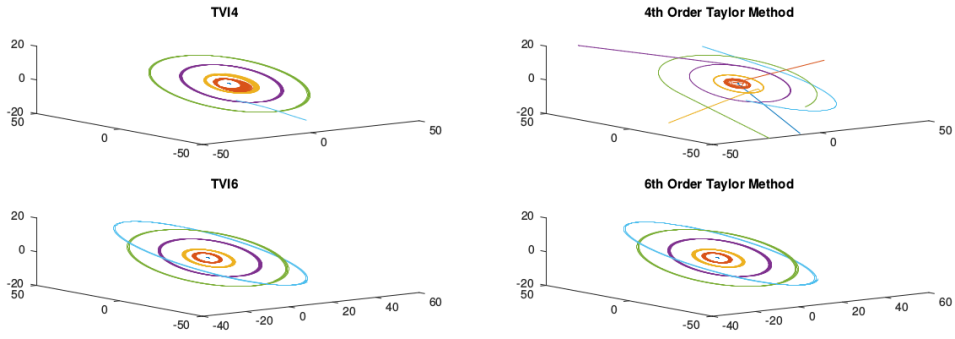



**Figure 3.10**: The sun and 5 outer planets simulated over the time interval [0,200000] with a step size of $h = 400$ (days). The stepsize is too large for the 4th-order methods to give a qualitatively accurate representation, but both 6th-order methods performed well qualitatively.

## 3.7 Appendix: 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, $\hat{\Psi}_h : TQ \to TQ$, with order of accuracy $r + 1$, we solve for the initial velocity $\tilde{v}_0$, such that,

$$\pi_Q \circ \hat{\Psi}_h(q_0, \tilde{v}_0) = q_1, \tag{3.11}$$

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

$$\pi_Q \circ \Phi_h(q_0, v_0) = q_1. \tag{3.12}$$

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

$$\Psi_h(q_0, \tilde{v}_0) = \left( \sum\nolimits_{k=0}^{r} \frac{h^k}{k!} q^{(k)}(0), \sum\nolimits_{k=1}^{r+1} \frac{h^{k-1}}{(k-1)!} q^{(k)}(0) \right) \tag{3.13}$$

and

$$\hat{\Psi}_h(q_0, \tilde{v}_0) = \left( \sum\nolimits_{k=0}^{r+1} \frac{h^k}{k!} q^{(k)}(0), \sum\nolimits_{k=1}^{r+2} \frac{h^{k-1}}{(k-1)!} q^{(k)}(0) \right), \tag{3.14}$$

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 [29].

**Lemma 4.** *$\tilde{v}_0$ as defined by,* (3.11) *and* (3.14)*, approximates $v_0$ to at least $\mathcal{O}(h^{r+1})$.*

*Proof.* Solving $\pi_Q \circ \hat{\Psi}_h(q_0, \tilde{v}_0) = q_1$ for $\tilde{v}_0$ yields,

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

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

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

Solving for $v_0$ yields,

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

Now evaluating the norm of the difference we have,

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

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 $M_i$, $i = 1, 2, \ldots, p$, is bounded. Using the triangle inequality, we have,

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

Rearranging, we have,

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

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.** *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 $\mathcal{O}(h^{r+2})$. 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 $\mathcal{O}(h^{r+2})$. 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 5.** *A $r$-order Taylor method, defined by* (3.13)*, with initial conditions $(q_0, \tilde{v}_0)$, where $\tilde{v}_0$ is defined by* (3.11)*, is accurate to at least $\mathcal{O}(h^{r+1})$ 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 (3.12). 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 $M$.

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

Combining this inequality with our $r$-order method yields,

$$\begin{aligned}\|(q(t), v(t)) - (q_d(t), v_d(t))\| &= \|(q(t), v(t)) - (\tilde{q}(t), \tilde{v}(t)) \\ &\quad + (\tilde{q}(t), \tilde{v}(t)) - (q_d(t), v_d(t))\| \\ &\leq \|(q(t), v(t)) - (\tilde{q}(t), \tilde{v}(t))\|\end{aligned}$$

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

where we used the triangle inequality, and the fact that the local truncation error of a $r$-order Taylor method is $\mathcal{O}(h^{r+1})$ 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 $r$-th order Taylor approximation, respectively, of the Euler–Lagrange initial-value problem with initial data $(q_0, \tilde{v}_0)$. □

**Theorem 16.** *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,

$$
q_{1,H_d^+} = D_2 H_d^+(q_0, p_{1,H_d^+})
$$

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

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

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

□

Chapter 3, in full, is a reprint of the material that has been submitted for publication to BIT Numerical Mathematics, 2017. Schmitt, Jeremy; Shingel, Tatianna; Leok, Melvin, Springer, 2017. The dissertation author was the primary investigator and author of this material.

# Chapter 4

# Adaptive Hamiltonian Variational Integrators

## 4.1 Introduction

Symplectic integrators are a class of geometric integrators that when applied to a Hamiltonian system yield a discrete approximation of the flow that preserves the symplectic 2-form (see [19]). The preservation of symplecticity results in the preservation of many qualitative aspects of the underlying dynamical system. In particular, when applied to conservative Hamiltonian systems, symplectic integrators show excellent long-time near-energy preservation. However, when symplectic integrators were first used in combination with variable time-steps, the near-energy preservation was lost and the integrators performed poorly (see [7], [13]). Backwards error analysis provided justification both for the excellent long-time near-energy preservation of symplectic integrators and for the poor performance experienced when using variable time-steps (see Chapter IX of [19]). Backward error analysis shows that symplectic integrators can be associated with a modified Hamiltonian in the form of a powers series in terms of the time-step. Changing the time-step results in a different modified Hamiltonian each time the time-step is varied. This is the source of the poor energy conservation. There has been a great effort to circumvent this problem, and there have been many suc-

cesses. However, there has yet to be a unified general framework for constructing adaptive symplectic integrators. In this paper, we attempt to add to this effort by extending variable time-steps into the domain of variational integrators. After a brief introduction to variational integrators, we present a framework for variable time-step variational integrators, and contrast our method with existing work on the matter.

## 4.2 Variational Integrators

Variational integrators are symplectic integrators derived by discretizing Hamilton's principle, versus discretizing Hamilton's equations directly. As a result, variational integrators are symplectic, preserve many invariants and momentum maps, as well as having excellent long-time near-energy preservation (see [35]). Traditionally, variational integrators have focused on the type I generating function known as the discrete Lagrangian, $L_d : Q \times Q \mapsto \mathbb{R}$. The exact discrete Lagrangian of the true flow of Hamilton's equations can be represented in both a variational form and in a boundary-value form. The latter is given by

$$L_d^E(q_0, q_1; h) = \int_0^h L(q_{01}(t), \dot{q}_{01}(t))dt, \tag{4.1}$$

where $q_{01}(0) = q_0$, $q_{01}(h) = q_1$, and $q_{01}$ satisfies the Euler–Lagrange equations over the time interval $[0, h]$. A variational integrator is defined by constructing an approximation to (4.1), $L_d : Q \times Q \mapsto \mathbb{R}$, and then applying the discrete Euler–Lagrange equations,

$$p_k = -D_1 L_d(q_k, q_{k+1}), \qquad p_{k+1} = D_2 L_d(q_k, q_{k+1}), \tag{4.2}$$

which implicitly define the integrator, $\tilde{F}_{L_d} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$. The error analysis is greatly simplified via Theorem 2.3.1 of [35], which states that if a discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$, approximates the exact discrete Lagrangian, $L_d^E : Q \times Q \to \mathbb{R}$, to order $r$, i.e.,

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

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 $r$ accurate.

Many other properties of the integrator, such as symmetry of the method, can be determined by analyzing the associated discrete Lagrangian, as opposed to analyzing the integrator directly. More recently, variational integrators have been extended to the framework of type II and type III generating functions, commonly referred to as discrete Hamiltonians (see [26], [31]). Hamiltonian variational integrators are derived by discretizing Hamilton's phase space principle. The boundary-value formulation of the exact type II generating function of the time-$h$ flow of Hamilton's equations is given by the exact discrete right Hamiltonian,

$$H_d^{+,E}(q_0, p_1; h) = p_1^T q_1 - \int_0^h \left[ p(t)^T \dot{q}(t) - H(q(t), p(t)) \right] dt, \tag{4.3}$$

where $(q(t), p(t))$ satisfy Hamilton's equations with boundary conditions $q(0) = q_0$ and $p(h) = p_1$. A type II Hamiltonian variational integrator is constructed by using an approximate discrete Hamiltonian, $H_d^+$, and applying the discrete right Hamilton's equations,

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

which implicitly defines the integrator, $\tilde{F}_{H_d^+} : (q_0, p_0) \mapsto (q_1, p_1)$.

Various methods for constructing and analyzing Hamiltonian variational integrators can be found in [31], [44], and [45]. In particular, there is an analogous error analysis theorem as in the case of Lagrangian variational integrators. If a discrete right Hamiltonian, $H_d^+$, approximates the exact discrete right Hamiltonian, $H_d^{+,E}$, to order $r$, i.e.,

$$H_d^+(q_0, p_1; h) = H_d^{+,E}(q_0, p_1; h) + \mathcal{O}(h^{r+1}),$$

then the discrete right Hamilton's map, $\tilde{F}_{H_d^+} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$, viewed as a one-step method, is order $r$ accurate.

Hamiltonian and Lagrangian variational integrators are not always equivalent. In particular, it was shown in [44] that in some cases even when the Hamiltonian and Lagrangian integrators are analytically equivalent they can have different

numerical properties. Even more to the point, Lagrangian variational integrators cannot always be constructed when the underlying Hamiltonian is degenerate, and in that situation, Hamiltonian variational integrators are the more natural choice. In the next section we examine a transformation commonly used to construct variable time-step symplectic integrators, which in most cases of interest results in a degenerate Hamiltonian. Our approach is to apply Hamiltonian variational integrators to the resulting transformed Hamiltonian system.

## 4.3 The Poincaré Transformation and Discrete Hamiltonians

Given a Hamiltonian, $H(q,p)$, and a desired transformation of time, $t \mapsto \tau$, given by $\frac{dt}{d\tau} = g(q,p)$, a new Hamiltonian system is given by the Poincaré transformation,

$$\bar{H}(\bar{q},\bar{p}) = g(q,p)\left(H(q,p) + p^t\right), \tag{4.4}$$

where $(\bar{q},\bar{p}) = \left(\begin{bmatrix} q \\ q^t \end{bmatrix}, \begin{bmatrix} p \\ p^t \end{bmatrix}\right)$. We will follow the common choice of setting $q^t = t$ and $p^t = -H(q(0),p(0))$, so that $\bar{H}(\bar{q},\bar{p}) = 0$ along all integral curves through $(q(0),p(0))$. The time $t$ shall be referred to as the physical time, and $\tau$ as the fictive time. The corresponding Hamilton's equations are given by,

$$\dot{\bar{q}} = \begin{bmatrix} \nabla_p g(q,p) \\ 0 \end{bmatrix} (H(q,p) + p^t) + \begin{bmatrix} \frac{\partial H}{\partial p} \\ 1 \end{bmatrix} g(q,p) \tag{4.5}$$

$$\dot{\bar{p}} = -\begin{bmatrix} \nabla_q g(q,p) \\ 0 \end{bmatrix} (H(q,p) + p^t) - \begin{bmatrix} \frac{\partial H}{\partial q} \\ 0 \end{bmatrix} g(q,p). \tag{4.6}$$

When the initial conditions are $(q(0),p(0))$, then $H(q,p) + p^t = 0$ and

$$\dot{\bar{q}} = \begin{bmatrix} g(q,p)\frac{\partial H}{\partial p} \\ g(q,p) \end{bmatrix}, \qquad \dot{\bar{p}} = \begin{bmatrix} -g(q,p)\frac{\partial H}{\partial q} \\ 0 \end{bmatrix}. \tag{4.7}$$

In general,

$$\frac{\partial^2 \bar{H}}{\partial \bar{p}^2} = \begin{bmatrix} \frac{\partial H}{\partial p}\nabla_p g(q,p)^T + g(q,p)\frac{\partial^2 H}{\partial p^2} + \nabla_p g(q,p)\frac{\partial H}{\partial p}^T & \nabla_p g(q,p) \\ \nabla_p g(q,p)^T & 0 \end{bmatrix},$$

which can be singular in many cases. Most of the papers cited here on variable time-step symplectic integrators focus exclusively on using a monitor function, $g$, that is only a function of position, in which case the resulting transformed Hamiltonian is degenerate and there is no corresponding Lagrangian formulation. Therefore, Hamiltonian variational integrators are the most general and natural way to derive variable time-step variational integrators.

The exact type II generating function for the transformed Hamiltonian is given by,

$$\bar{H}_d^{+,E}(\bar{q}_0, \bar{p}_1; h) = \bar{p}_1^T \bar{q}_1 - \int_0^h \left( \bar{p}(\tau)^T \dot{\bar{q}}(\tau) - \bar{H}(\bar{q}(\tau), \bar{p}(\tau)) \right) d\tau, \tag{4.8}$$

where $(\bar{q}(\tau), \bar{p}(\tau))$ satisfy Hamilton's equations (4.7), with boundary conditions $\bar{q}(0) = \bar{q}_0$, $\bar{p}(h) = \bar{p}_1$.

The above exact discrete right Hamiltonian implicitly defines a symplectic map with respect to the symplectic form $\bar{\omega}(\bar{p}_k, \bar{q}_k)$ on $T^*\bar{Q}$ via the discrete Legendre transforms given by,

$$\bar{p}_0 = \frac{\partial \bar{H}_d^{+,E}}{\partial \bar{q}_0}, \qquad \bar{q}_1 = \frac{\partial \bar{H}_d^{+,E}}{\partial \bar{p}_1}.$$

Our approach is to construct Hamiltonian variational integrators by using a discrete right Hamiltonian, $\bar{H}_d^+$, that approximates (4.8) to order $r$, then the resulting integrator will be a variable time-step symplectic integrator. It is important to note that this method will be symplectic in two different ways. It will be symplectic both with respect to the symplectic form $d\bar{p} \wedge d\bar{q}$ and with respect to the symplectic form $dp \wedge dq$. Since $p_t$ is constant(i.e. $p_0^t = \frac{\partial \bar{H}_d^+}{\partial q_0^t} = p_1^t$), the symplectic form in generalized coordinates is given by

$$\begin{aligned}
\bar{\omega}(\bar{p}_k, \bar{q}_k) &= d\bar{p}_k \wedge d\bar{q}_k \\
&= \sum_{i=1}^{n+1} d\bar{p}_{k,i} \wedge d\bar{q}_{k,i} \\
&= \sum_{i=1}^{n} dp_{k,i} \wedge dq_{k,i} + dp_k^t \wedge dq_k^t \\
&= \sum_{i=1}^{n} dp_{k,i} \wedge dq_{k,i}
\end{aligned}$$

$$= \omega(p_k, q_k).$$

A symplectic variable time-step method was proposed independently in [17] and [42], which applied a symplectic integrator to the Hamiltonian system resulting from the Poincaré transformation. In [17], it is noted that one of the first applications of the Poincaré transformation was by Levi-Civita, who applied it to the three-body problem. A more in-depth discussion of such time transformations can be found in [49]. There has been further work using this transformation in papers such as [2] and [3], which focus on developing symplectic, explicit, splitting methods with variable time-steps.

Our approach is to discretize the type II generating function for the flow of Hamilton's equations, where the Hamiltonian is given by the Poincaré transformation. Therefore, we are constructing variational integrators, and in particular Hamiltonian variational integrators (see [26], [31]). This approach works seamlessly with existing methods and theorems of Hamiltonian variational integrators, but now the system under consideration is the transformed Hamiltonian system resulting from the Poincaré transformation. It should be noted that the methods of [17] and [42] include the possibility of applying a variational integrator to the Poincaré transformed Hamilton's equations. Our approach gives a framework for constructing variational integrators by using the Poincaré transformed discrete right Hamiltonian. In most cases, these two approaches will produce equivalent integrators, but our new approach allows for the method to analyzed at the level of the generating function, and indicates that most such symplectic methods are best interpreted as coming from a type II or III generating function, as opposed to a type I generating function.

**Remark.** *Other approaches to variable time-step variational integrators can be found in [25], [37] and [38]. In particular, [25] is inspired by the result of Ge and Marsden ([12]), which states that constant time-step symplectic integrators of autonomous Hamiltonian systems cannot exactly conserve the energy unless it agrees with the exact flow map up to a time reparametrization. Therefore, they sought a variable time-step energy-conserving symplectic integrator. However, symplecticity is with respect to the space-time symplectic form* $dp \wedge dq + dH \wedge dt$. *The*

*time-step is determined by enforcing discrete energy conservation, which arises as a consequence of the fact that energy is the Noether quantity associated with time translational symmetry. An extended Hamiltonian is used that is similar in spirit to the Poincaré transformation. A nearly identical approach and definition of symplecticity was used in [38].*

*In [37], the approach involves a transformation of the Lagrangian, which is motivated by the Poincaré transformation, but it is not equivalent. The lack of equivalence is not surprising, since the Hamiltonian given by the Poincaré transformation is degenerate for their choice of monitor functions. As a consequence, the phase space path is not preserved, but the state space path is preserved up to a rescaling of the velocity.*

## 4.4 Variational Error Analysis

The standard error analysis theorem for Hamiltonian variational integrators assumes a non-degenerate Hamiltonian, i.e., $\det(\frac{\partial^2 \bar{H}}{\partial \bar{p}^2}) \neq 0$ (see [44]). The non-degeneracy implies that the usual implicit function theorem can be applied to the discrete right Hamilton's equations. In particular, the proof of the error analysis theorem relies upon the following lemma, which follows from the implicit function theorem.

**Lemma 6.** *Let $f_1, g_1, e_1, f_2, g_2, e_2 \in C^r$ be such that*

$$f_1(x,h) = g_1(x,h) + h^{r+1}e_1(x,h),$$

$$f_2(x,h) = g_2(x,h) + h^{r+1}e_2(x,h).$$

*Then, there exists functions $e_{12}$ and $\bar{e}_1$ bounded on compact sets such that*

$$f_2(f_1(x,h),h) = g_2(g_1(x,h),h) + h^{r+1}e_{12}(g_1(x,h),h),$$

$$f_1^{-1}(y) = g_1^{-1}(y) + h^{r+1}\bar{e}_1(y).$$

Combining this lemma with the discrete right Hamiltonian map,

$$\tilde{F}_{H_d^+}(q_0,p_0) = \mathbb{F}^+H_d^+ \circ (\mathbb{F}^-H_d^+)^{-1}(q_0,p_0) = (q_1,p_1),$$

ensures the order of accuracy of the integrator is at least of the order to which $H_d^+$ approximates $H_d^{+,E}$. Since the usual implicit function theorem does not apply, we need to justify the invertibility of $\mathbb{F}^- H_d^+$, which comes down to whether $\bar{p}_0 = D_1 \bar{H}_d^+(\bar{q}_0, \bar{p}_1; h)$ can be solved for $\bar{p}_1$.

We assume the original Hamiltonian, $H(q,p)$, is nondegenerate. Then, we will show that the exact discrete right Hamiltonian can be reduced to a particular form and the extended variables $p_1^t$ and $q_1^t$ can be solved for explicitly. As a result, the implicit function theorem is not needed with respect to these variables. Hamilton's equations of the transformed Hamiltonian, $\bar{H}(\bar{q}, \bar{p}) = g(q,p)\,(H(q,p) + p^t)$, are

$$\dot{\bar{q}} = \begin{bmatrix} \nabla_p g(q,p) \\ 0 \end{bmatrix} (H(q,p) + p^t) + \begin{bmatrix} \frac{\partial H}{\partial p} \\ 1 \end{bmatrix} g(q,p)$$

$$\dot{\bar{p}} = -\begin{bmatrix} \nabla_q g(q,p) \\ 0 \end{bmatrix} (H(q,p) + p^t) - \begin{bmatrix} \frac{\partial H}{\partial q} \\ 0 \end{bmatrix} g(q,p).$$

The corresponding exact discrete right Hamiltonian is of the form

$$\begin{aligned} \bar{H}_d^{+,E}(\bar{q}_0, \bar{p}_1; h) &= \bar{p}_1^T \bar{q}_1 - \int_0^h \left( \bar{p}(\tau)^T \dot{\bar{q}}(\tau) - \bar{H}(\bar{q}(\tau), \bar{p}(\tau)) \right) d\tau \\ &= p_1^T q_1 + p_1^t q_1^t - \int_0^h (p(\tau)^T \dot{q}(\tau) + p^t(\tau) g(q(\tau), p(\tau)) \\ &\quad - g(q(\tau), p(\tau)) p^t(\tau) - g(q(\tau), p(\tau)) H(q(\tau), p(\tau))) d\tau \\ &= p_1^T q_1 + p_1^t q_1^t - \int_0^h \left( p(\tau)^T \dot{q}(\tau) - g(q(\tau), p(\tau)) H(q(\tau), p(\tau)) \right) d\tau. \end{aligned}$$

As a result, only one part of this exact discrete right Hamiltonian requires approximations of the extended variable $q^t$ and $p^t$. Furthermore, since $\dot{p}^t = 0$ this implies $p_1^t = p_0^t$. Now, let $\bar{H}_d^+(\bar{q}_0, \bar{p}_1; h)$ be an approximation to the exact discrete right Hamiltonian of the form

$$\bar{H}_d^+(\bar{q}_0, \bar{p}_1; h) = p_1^T \hat{q}_1(q_0, p_1; h) + p_1^t \hat{q}_1^t(q_0^t, q_0, p_1; h) - I(q_0, p_1; h),$$

where $\hat{\cdot}$ denotes an approximated value and $I(q_0, p_1; h)$ is an approximation of

$$\int_0^h \left( p(\tau)^T \dot{q}(\tau) - g(q(\tau), p(\tau)) H(q(\tau), p(\tau)) \right) d\tau.$$

Then, the discrete right Legendre transforms, $\bar{p}_0 = D_1\bar{H}_d^+(\bar{q}_0, \bar{p}_1; h)$ and $\bar{q}_1 = D_2\bar{H}_d^+(\bar{q}_0, \bar{p}_1; h)$, give the following explicit relations for $p_1^t$ and $q_1^t$,

$$\begin{bmatrix} p_0 \\ p_0^t \end{bmatrix} = \begin{bmatrix} \frac{\partial \hat{q}_1}{\partial q_0}^T p_1 + p_1^t \frac{\partial \hat{q}_1^t}{\partial q_0} - \frac{\partial I}{\partial q_0} \\ \frac{\partial \hat{q}_1^t}{\partial q_0^t} p_1^t \end{bmatrix},$$
$$\begin{bmatrix} q_1 \\ q_1^t \end{bmatrix} = \begin{bmatrix} \hat{q}_1 + \frac{\partial \hat{q}_1}{\partial p_1}^T p_1 + \frac{\partial \hat{q}_1}{\partial p_1}^T p_1^t - \frac{\partial I}{\partial p_1} \\ \hat{q}_1^t \end{bmatrix}.$$

Now, since the analytic solution satisfies $p_1^t = p_0^t$, there is no need to approximate $p_1^t$. Therefore, $\frac{\partial \hat{q}_1^t}{\partial q_0^t} = 1$, and $p_1^t$ is given independently of the other values. The upshot is a system that can be solved by first setting $p_1^t = p_0^t$, then implicity solving for $p_1$ in terms of $(q_0^t, q_0, p_1^t, p_1)$, explicitly solving for $q_1$ and finally explictly solving for $q_1^t$. Since $p_1$ is not determined by $q_1^t$, the implicit function theorem is simply needed for finding $p_1$. Therefore, we need $\det(\frac{\partial^2 \bar{H}}{\partial p^2}) \neq 0$, which is the same as $\det(\frac{\partial H}{\partial p}\nabla_p g(q,p)^T + g(q,p)\frac{\partial^2 H}{\partial p^2} + \nabla_p g(q,p)\frac{\partial H}{\partial p}^T) \neq 0$. Note this holds for nondegenerate Hamiltonians $H$ and $p$-independent monitor functions.

**Theorem 17.** *Given a nondegenerate Hamiltonian $H$, and a monitor function $g \in C^1([0,h])$, such that $\det(\frac{\partial H}{\partial p}\nabla_p g(q,p)^T + g(q,p)\frac{\partial^2 H}{\partial p^2} + \nabla_p g(q,p)\frac{\partial H}{\partial p}^T) \neq 0$. Then, if the discrete right Hamiltonian $\bar{H}_d^+$, approximates the exact discrete right Hamiltonian $\bar{H}_d^{+,E}$, to order $r$, i.e.,*

$$\bar{H}_d^+(\bar{q}_0, \bar{p}_1; h) = \bar{H}_d^{+,E}(\bar{q}_0, \bar{p}_1; h) + \mathcal{O}(h^{r+1}),$$

*then the discrete right Hamilton's map $\tilde{F}_{\bar{H}_d^+} : (\bar{q}_k, \bar{p}_k) \mapsto (\bar{q}_{k+1}, \bar{p}_{k+1})$, viewed as a one-step method, is order $r$ accurate.*

## 4.5 Adaptive Hamiltonian Taylor Variational Integrators

We will demonstrate the approach using Hamiltonian Taylor variational integrators (see [45]), which are constructed as follows:

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

$$\bar{p}_1 = \pi_{T^*\bar{Q}} \circ \Psi_h^{(r)}(\bar{q}_0, \tilde{\bar{p}}_0),$$

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

(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 $(\bar{q}(t), \bar{p}(t))$ at the quadrature nodes,

$$(\bar{q}_{c_i}, \bar{p}_{c_i}) = \Psi_{c_i h}^{(r)}(\bar{q}_0, \tilde{\bar{p}}_0),$$

and use a $(r+1)$-order Taylor method on the configuration manifold to generate the approximation to the boundary term $\bar{q}_1$,

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

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

$$\bar{H}_d^+(\bar{q}_0, \bar{p}_1; h) = \bar{p}_1^T \tilde{\bar{q}}_1 - h \sum_{i=1}^{m} \left[ \bar{p}_{c_i}^T \dot{\bar{q}}_{c_i} - \bar{H}\left( \Psi_{c_i h}^{(r)}(\bar{q}_0, \tilde{\bar{p}}_0) \right) \right].$$

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

$$\bar{q}_1 = D_2 \bar{H}_d^+(\bar{q}_0, \bar{p}_1), \qquad \bar{p}_0 = D_1 \bar{H}_d^+(\bar{q}_0, \bar{p}_1). \tag{4.9}$$

For a lucid exposition, we will at first assume $g(q,p) = g(q)$ and $H(q,p) = \frac{1}{2} p^T M^{-1} p + V(q)$. Consider the discrete right Hamiltonian given by approximating $\bar{q}_1$ with a first-order Taylor method about $\bar{q}_0$, approximating $\bar{p}_0$ with a zeroth-order Taylor expansion about $\bar{p}_0$, and using the rectangular quadrature rule about the initial point, which yields

$$\bar{H}_d^+ = p_1^T (q_0 + h g(q_0) M^{-1} p_1) + p_1^t (q_0^t + h g(q_0)) - h g(q_0) \left[ \frac{1}{2} p_1^T M^{-1} p_1 - V(q_0) \right]. \tag{4.10}$$

The corresponding variational integrator is given by,

$$\bar{p}_1 = \begin{bmatrix} p_0 - hg(q_0)\nabla V(q_0) - h\nabla g(q_0)\left(\frac{1}{2}p_1^T M^{-1} p_1 + V(q_0) + p_0^t\right) \\ p_0^t \end{bmatrix}, \tag{4.11}$$

$$\bar{q}_1 = \begin{bmatrix} q_0 + hg(q_0)M^{-1}p_1 \\ q_0^t + hg(q_0) \end{bmatrix}. \tag{4.12}$$

The resulting integrator is merely symplectic Euler-B applied to the transformed Hamiltonian system,

$$\bar{q}_1 = \bar{q}_0 + h\frac{\partial \bar{H}(\bar{q}_0, \bar{p}_1)}{\partial \bar{p}},$$
$$\bar{p}_1 = \bar{p}_0 - h\frac{\partial \bar{H}(\bar{q}_0, \bar{p}_1)}{\partial \bar{q}}.$$

In fact, this is precisely the adaptive symplectic integrator first proposed in [17] and also presented on page 254 of [28]. Most existing symplectic integrators can be interpreted as variational integrators, but there are also new methods that are most naturally derived as variational integrators. We will also consider a fourth-order Hamiltonian variational integrator recently developed in [45], which is distinct from any existing symplectic method.

One of the most important aspects of implementing a variable time-step symplectic integrator of this form is a well chosen monitor function, $g(q)$. We need $g$ to be positive-definite, so that we never stall or march backward in time. Noting that the above integrator is first-order, a natural choice is to use the second-order truncation error given by $-\frac{(q_1^t - q_0^t)^2}{2}M^{-1}\nabla V(q_0)$. Let $tol$ be some desired level of accuracy, then one choice for $g$ would be,

$$g(q_0) = \frac{tol}{\|\frac{(q_1^t - q_0^t)^2}{2}g(q_0)M^{-1}\nabla V(q_0)\|}. \tag{4.13}$$

Noting that $q_1^t - q_0^t = hg(q_0)$, we have,

$$g(q_0) = \frac{tol}{\|\frac{h^2 g(q_0)^3}{2}M^{-1}\nabla V(q_0)\|}, \tag{4.14}$$

which yields,

$$g(q_0) = \left(\frac{tol}{\|\frac{h^2}{2}M^{-1}\nabla V(q_0)\|}\right)^{\frac{1}{4}}. \tag{4.15}$$

This justifies our choice for $g$ as,

$$g(q_0) = \frac{tol}{\|\frac{h^2}{2} M^{-1} \nabla V(q_0)\|}, \tag{4.16}$$

which achieves an error which is comparable to the chosen value of $tol$.

Alternative choices of $g$, proposed in [17], include the $p$-independent arclength parameterization given by,

$$g(q) = (2(H_0 - V(q)) + \nabla V(q)^T M^{-1} \nabla V(q))^{-\frac{1}{2}}, \tag{4.17}$$

and a choice particular to Kepler's two-body problem,

$$g(q) = q^T q, \tag{4.18}$$

which is motivated by Kepler's second law, which states that a line segment joining the two bodies sweeps out equal areas during equal intervals of time.

We have tested the algorithm given by (4.12) on Kepler's planar two-body problem, with an eccentricity of 0.9, using the three choices of $g$ given by (4.16), (4.17), and (4.18). Of these three choices, (4.18) is particular to Kepler's two-body problem, while (4.16) and (4.17) are more general choices. However, since (4.16) is based on the truncation error, the cost of computing this function will increase as the order of the method increases. In contrast, (4.17) is independent of the order. Simulations using Kepler's two-body problem with an eccentricity of 0.9 over a time interval of $[0, 1000]$ were run using the three different choices of $g$ and the usual symplectic Euler-B. Results indicate that symplectic Euler-B takes the most steps and computational time to achieve a level of accuracy around $10^{-5}$. To achieve a level of accuracy around $10^{-5}$, the choice of the truncation error monitor function, (4.16), resulted in the least number of steps, and the second lowest computational time. The lowest computational time belonged to (4.18), but it used significantly more steps than (4.16). The lower computational cost can be attributed to the cheaper evaluation cost of the monitor function and its derivative. Finally, the monitor function (4.17) required the most steps and computational time of the adaptive algorithms, but it is still a good choice in general given its broad applicability. See Figures 4.1, 4.2, and 4.3.

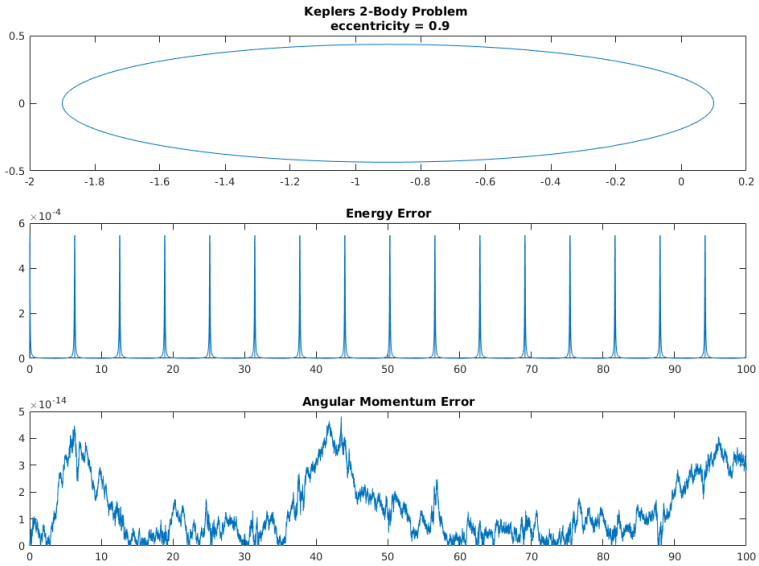



**Figure 4.1**: A time-step of $h = 0.00001$ was used, and it took 10,000,000 steps. Global error $= 5.5 \cdot 10^{-4}$.

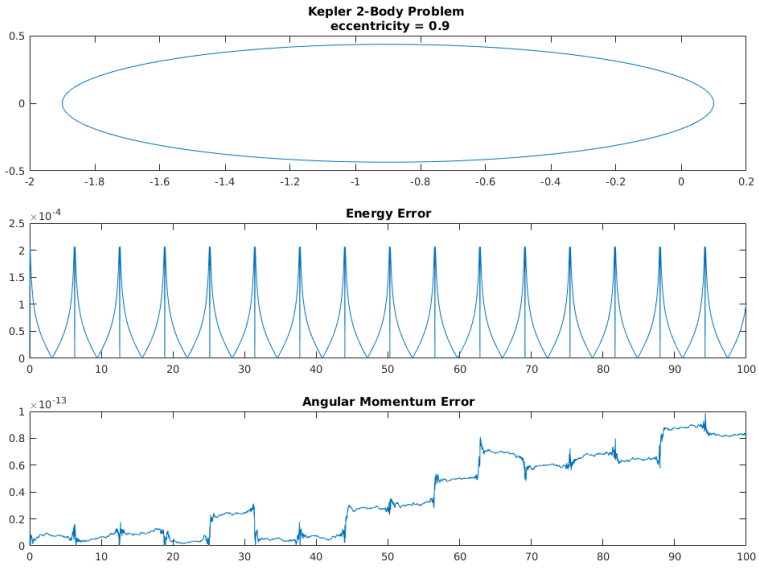



**Figure 4.2**: The tolerance was set to $10^{-5}$ and it took 1,123,116 steps. Global error $= 4.2 \cdot 10^{-5}$.

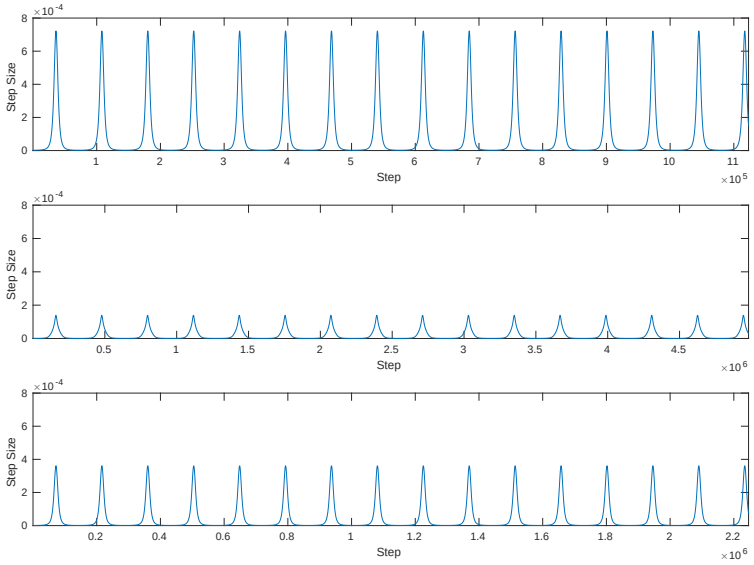



**Figure 4.3**: The top plot corresponds to (4.16), the middle plot corresponds to (4.17), and the bottom plot corresponds to (4.18). All of the monitor functions appear to increase and decrease the step size at the same points along the trajectory, but clearly (4.16) allowed for the larger steps to be taken.

Next, we consider the fourth-order Hamiltonian Taylor variational integrator constructed using Taylor methods up to order 3 and Simpson's quadrature rule. We will now drop the assumption of $p$-independent monitor functions and consider $g(q,p)$. The following monitor functions were considered,

$$g(q) = \left(q^T q\right)^{\gamma} \text{ for } \gamma = \frac{1}{2}, 1 \tag{4.19}$$

$$g(q) = \left(2(H_0 - V(q)) + \nabla V(q)^T M^{-1} \nabla V(q)\right)^{-\frac{1}{2}} \tag{4.20}$$

$$g(q,p) = \|p^t - L(q, M^{-1}p)\|_2^{-1} \tag{4.21}$$

The monitor function (4.21) was originally intended to be $\|p^t + H(q,p)\|_2^{-1}$, but an accidental error led to the conclusion that (4.21) is the better choice. We will discuss the shortcomings of using the inverse energy error in the next paragraph. Note that $\|L(q, M^{-1}p)\|_2^{-1}$ also performs decently, but the addition of $p^t = -H(q_0, p_0)$ showed noticeable improvement. It was noted in [17] that the inverse Lagrangian has been considered as a possible choice for $g$ in the Poincaré transformation, but not in the framework of symplectic integration. While the

choice of (4.19) was generally the most efficient, (4.21) was very close in terms of efficiency and offers a more general monitor function. This also implies that efficiency is not limited to only $q$ or $p$-independent monitor functions. However, various attempts to construct seperable transformed Hamiltonians (see [2], [3]) required the use of $q$ or $p$-independent monitor functions, so this is where such monitor functions are most useful.

The truncation error monitor function, (4.16), performed quite well for first-order methods, and this motivated the choice of using Taylor variational integrators, since derivatives would be readily available. However, its success cannot as easily be applied to higher-order methods. This is due to the fact that for higher-order truncation errors, one obtains an implicit differential-algebraic definition of the monitor function. This deviates from the first-order case, where the monitor function can be solved for explicitly. Another seemingly natural choice for the monitor function is the inverse of the energy error. However, Taylor variational integrators are constructed using Taylor expansions about the initial point, and consequently the monitor function is largely evaluated about the initial point. If the initial point is at a particularly tricky part of the dynamics and requires a small first step, then the energy error at the first step will not reflect this, since initially the energy error is zero. In contrast, the inverse Lagrangian will be small at an initial point that requires a small first step. The inverse energy error may work well for methods that primarily evaluate the energy error at the end point rather than the initial point.

Additionally, it is often advantageous to bound the time-step below or above. As noted on page 248 of [28], this can be done by setting $a = \frac{\Delta t_{\min}}{\Delta \tau}$ and $b = \frac{\Delta t_{\max}}{\Delta \tau}$, then defining the new monitor function as,

$$\hat{g} = b\frac{g+a}{g+b}. \tag{4.22}$$

Note that for methods such as the Taylor variational integrator, bounding $g(q,p)$ does bound the step-size, but not directly (see the tables below for a comparison of bounds, computationals time, steps, and error).

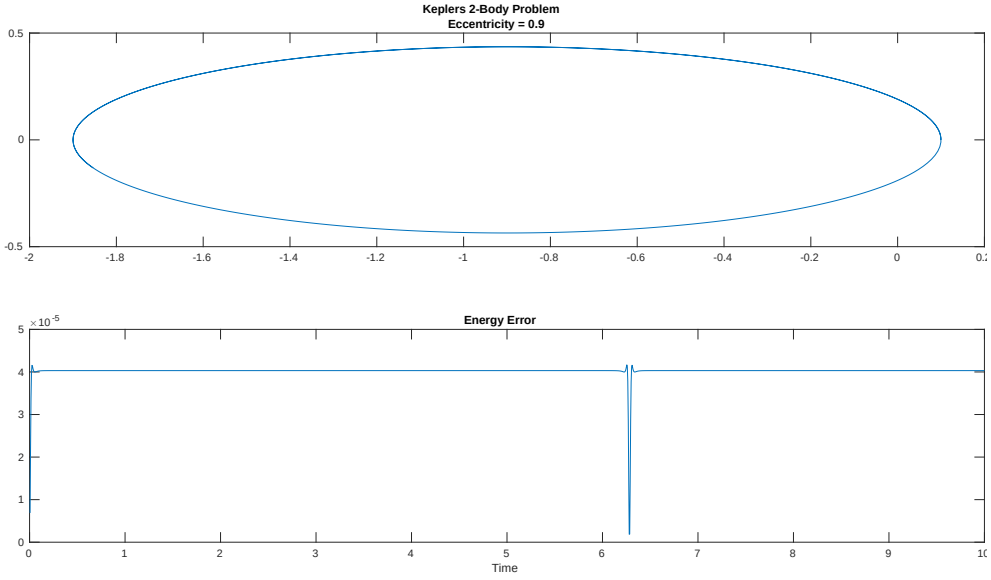



**Figure 4.4**: It was applied to Kepler's planar two-body problem over a time interval of $[0, 10]$ with an eccentricity of 0.9, and the method required 2000 steps to achieve a global error of around $6.2 \cdot 10^{-5}$.

Compared to non-adaptive variational integrators, the adaptive methods showed a significant gain in efficiency for Kepler's 2-body planar problem with high eccentricity, while low eccentricity models do not need nor do they benefit from adaptivity. A Hamiltonian dynamical system with regions of high curvature in the vector field and its norm will in general benefit from an adaptive scheme such as the one outlined here.

**Table 4.1**: A comparison of different choices of monitor functions for Kepler's 2-body problem with an eccentricity of 0.9

| Kepler Planar two-Body Problem, Eccentricity = 0.9 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Method | Monitor | $h$ | min Step | max Step | min $g$ | max $g$ | Energy Error | Global Error | Steps | Time |
| HTVI4 | Gamma | 0.1 | 0.0020 | 0.2493 | 0.01 | 8 | 1.43E-05 | 7.09E-06 | 181 | 26.9 |
| HTVI4 | Energy | 0.1 | 0.0051 | 0.1809 | 0.0001 | 2 | 1.93E-06 | 4.76E-06 | 146 | 28.3 |
| HTVI4 | Arclength | 0.1 | 0.0040 | 0.1458 | 0.003 | 0.3 | 1.10E-04 | 3.69E-05 | 185 | 70.2 |
| HTVI4 | - | 0.0025 | 0.0025 | 0.0025 | - | - | 2.50E-06 | 2.89E-05 | 4000 | 120 |

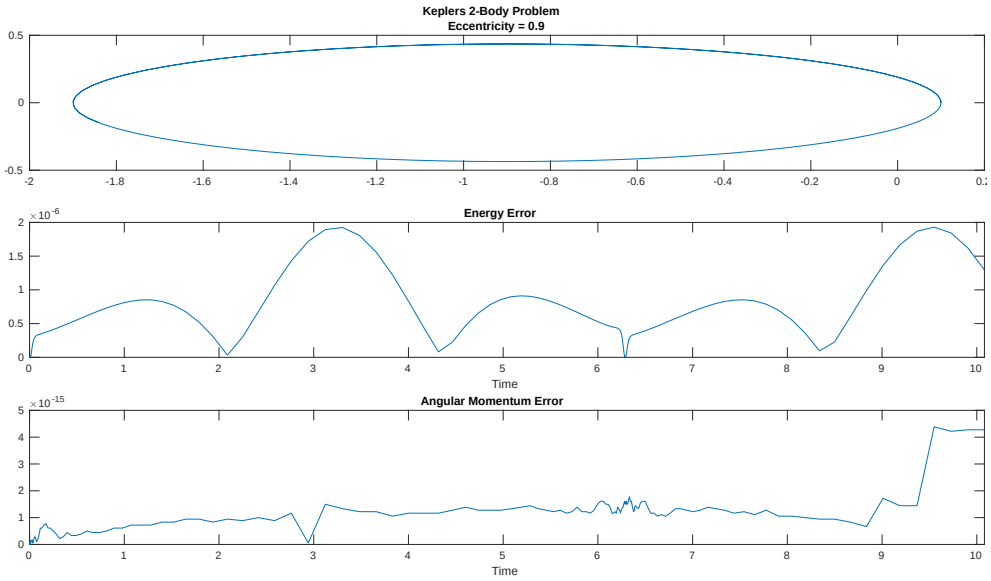



**Figure 4.5**: It was applied to Kepler's planar two-body problem over a time interval of $[0, 10]$ with an eccentricity of 0.9, and it took 146 steps and had a global error $= 4.76 \cdot 10^{-6}$.

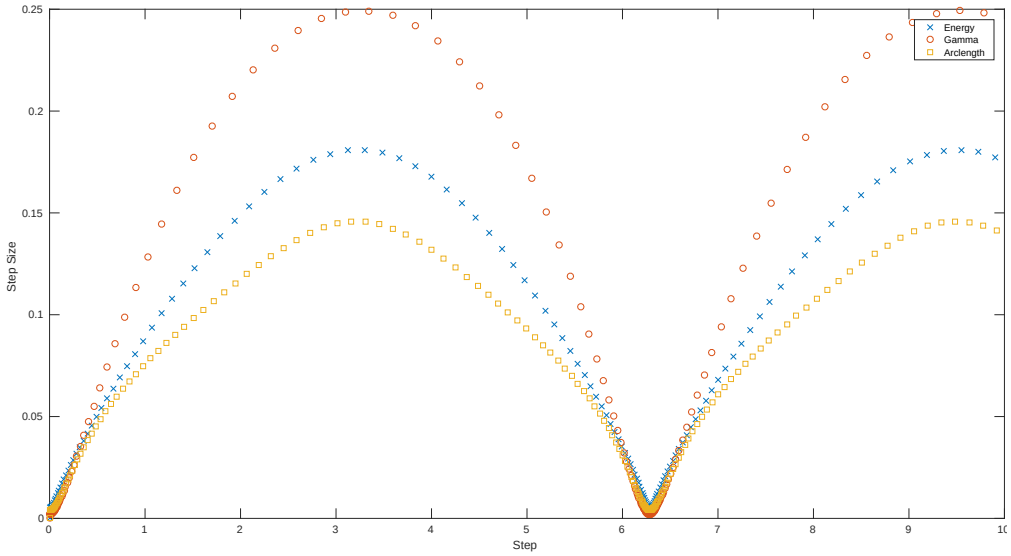



**Figure 4.6**: Energy is the monitor function (4.21), gamma is the monitor function (4.19), and arc length is the monitor function (4.20). The energy monitor and gamma monitor function performed the best in terms of fewest steps, lowest computational cost and lowest global error. Notice that (4.21) did not take the largest steps nor the smallest steps.

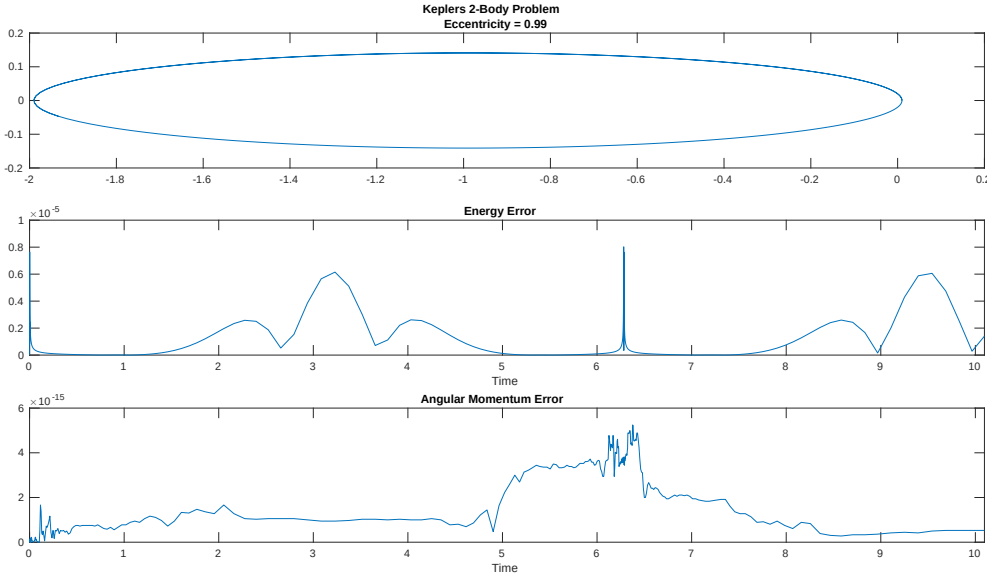



**Figure 4.7**: This choice of monitor function resulted in the fewest steps for an accuracy of $10^{-5}$ or better.

**Table 4.2**: A comparison of different choices of monitor functions for Kepler's 2-body problem with an eccentricity of 0.99

| Kepler Planar two-Body Problem, Eccentricity = 0.99 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Method | Monitor $g(q,p)$ | $h$ | min Step | max Step | min $g$ | max $g$ | Energy Error | Global Error | Steps | Time |
| HTVI4 | Gamma | 0.1 | 0.00006 | 0.2648 | 0.0005 | 8 | 4.88E-05 | 5.60E-06 | 372 | 49.3 |
| HTVI4 | Energy | 0.03 | 0.00015 | 0.1462 | 1E-6 | 5 | 9.13E-06 | 4.63E-06 | 383 | 58.4 |
| HTVI4 | Arclength | 0.1 | 0.00005 | 0.1379 | 0.0008 | 10 | 1.31E-05 | 1.49E-05 | 691 | 146.0 |
| HTVI4 | - | 0.0005 | 0.0005 | 0.0005 | - | - | 1.38E-01 | 7.83E-01 | 20000 | 525.2 |
| SV | - | 5E-7 | 5E-7 | 5E-7 | - | - | 3.34E-06 | 2.68E-05 | 2E7 | 189.2 |

## 4.6 Conclusion

Due to the degeneracy of the Hamiltonian, adaptive variational integrators based on the Poincaré transformation should be constructed using discrete Hamiltonians, which are type II or III generating functions. This has potential implications for the numerical properties of such integrators, and might explain why there has only been a limited amount of work on the construction of adaptive variational integrators based on the traditional Lagrangian perspective. The standard variational error analysis has been extended to include this particular form of a degenerate Hamiltonian. The efficiency of the resulting integrator is

largely based upon a proper choice of the monitor function $g$, and more research is needed to find a general choice of $g$ that maintains a decent level of efficiency. Galerkin variational integrators are likely to be a more promising choice than Taylor variational integrators, since the cost of evaluating the monitor function and its derivatives should be lower. In addition, the Galerkin approximation scheme may help inform a better choice of monitor function, due to the extensive literature on efficient *a posteriori* error estimation.

Chapter 4, in full, is currently being prepared for submission for publication of the material. Schmitt, Jeremy; Leok, Melvin. The dissertation author was the primary investigator and author of this material.

# Chapter 5

# Conclusions and Future Directions

This dissertation has extended the theory and algorithmic framework for Hamiltonian variational integrators and their associated type II and type III generating functions. It has been shown that the type of generating function used can affect the numerical properties of the resulting variational integrator. Averaging methods are particularly affected by the choice of using a Lagrangian variational integrator versus a Hamiltonian variational integrator. Furthermore, it was shown that discretization does not always commute with the Legendre transforms for generating functions, and a sufficient condition was provided for when this composition is commutative. A new class of variational integrators was developed that exploits the structure of the Taylor method to gain a higher order of accuracy for the particular shooting problem that arises in the construction of variational integrators. The framework for adaptive symplectic integrators, based on the Poincaré transformation, has been extended to variational integrators, and due to degeneracy issues it requires discrete Hamiltonians as opposed to discrete Lagrangians. The standard variational error analysis theorem has been extended to this particular degenerate case.

The computational efficiency of Taylor varaitional integrators ultimately depends upon bringing down the cost of the Jacobian evaluations. Alternative

automatic differentiation packages may help here, but the more promising route is to exploit the potential scalability of automatic differentiation for shared or distributed computing. Variational integrators require the use of discrete Legendre transforms, which generally involve partial differentiation, and higher order variational integrators are generally implicit. A general computational framework that applies automatic differentiation to a discrete Hamiltonian or discrete Lagrangian in combination with a compatible nonlinear solver could greatly simplify the implementation of variational integrators. This would greatly increase the accessbility of variational integrators to the general scientific community.

Further research on the differing numerical properties of Lagrangian and Hamiltonian variational integrators for particular classes of variational integrators could yield more interesting results. However, it is intriguing that Galerkin variational integrators are equivalent for either formulation, and this may indicate that furthering their computational development is the best way forward. In particular, Galerkin variational integrators might be the best candidates to implement in the adaptive framework. Additionally, it has been brought to my attention that type IV generating functions are of interest for some areas in statistical mechanics, and this type of generating functionhas yet to be established in a variational setting for deriving integrators. Also, more research is needed for choosing a monitor function in adaptive implementation. This is another area where Galerkin variational integrators woud be interesting to consider, as the monitor function might benefit from being based on the Galerkin approximation error. The monitor function is an *a priori* error estimator, but variational integrators in general could benefit from the development of *a posteriori* error indicators. The final area of further research would be the development of error analysis theorems for more general degenerate Hamiltonians and degenerate Lagrangians.

At the very least the work in this thesis indicates that Hamiltonian variational integrators may deserve more attention than they have recieved thus far.

# Bibliography


[1] V. I. Arnol′d. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1989. Translated from the Russian by K. Vogtmann and A. Weinstein.

[2] Sergio Blanes and Chris J. Budd. Explicit adaptive symplectic (easy) integrators: A scaling invariant generalisation of the Levi-Civita and KS regularisations. *Celestial Mechanics and Dynamical Astronomy*, 89(4):383–405, 2004.

[3] Sergio Blanes and Arieh Iserles. Explicit adaptive symplectic integrators for solving Hamiltonian systems. *Celestial Mechanics and Dynamical Astronomy*, 114(3):297–317, 2012.

[4] D.S. Broomhead and A. Iserles. *The Dynamics of Numerics and the Numerics of Dynamics*. Oxford University Press, New York, 1992.

[5] H. Martin Bucker, Bruno Lang, Dieter an Mey, and Christian Bischof. Bringing together automatic differentiation and OpenMP. In *ICS '01 Proceedings of the 15th International conference on Supercomputing*, pages 246–251, 2001.

[6] J.C. Butcher. On fifth and sixth order Runge-Kutta methods: Order conditions and order barriers. *Canadian Applied MAthematics Quarterly*, 17(3): 433–445, 2009.

[7] J.P. Calvo and J.M. Sanz-Serna. The development of variable-step symplectic integrators, with application to the two-body problem. *SIAM J. Sci. Comp.*, 14(4):936–952, 1993.

[8] M.P. Calvo, A. Murua, and J.M. Sanz-Serna. Modified equations for ODEs. *Contemporary Mathematics*, 172:63–74, 1994.

[9] Manuel de León, David Martín de Diego, and Aitor Santamaría-Merino. Discrete variational integrators and optimal control theory. *Adv. Comput. Math.*, 26(1-3):251–268, 2007.

[10] Will M. Farr. Variational integrators for almost-integrable systems. *Celestial*

*Mechanics and Dynamical Astronomy*, 102(2):105–118, 2009.

[11] Joseph Ford. The Fermi-Pasta-Ulam problem: Paradox turns discovery. *Physics Review*, 213(5):271–310, 1992.

[12] Zhong Ge and Jerrold Marsden. Lie–Poisson Hamilton–Jacobi theory and Lie-Poisson integrators. *Physics Letters A*, 133(3):134–139, 1988.

[13] B. Gladman, M. Duncan, and J. Candy. Symplectic integrators for long-time integrations in celestial mechanics. *Celestial Mech. Dynamical Astronomy*, 52: 221–240, 1991.

[14] H. Goldstein, C.P. Poole, and J.L. Safko. *Classical Mechanics*. Addison Wesley, 2002. ISBN 9780201657029.

[15] E. Hairer and C. Lubich. The life-span of backward error analysis for numerical integrators. *Numerische Mathematik*, 76(4):441–462, 1997.

[16] Ernest Hairer and Christian Lubich. Long-time energy conservation of numerical methods for oscillatory differential equations. *SIAM Journal of Numerical Analysis*, 38:414–441, 2001.

[17] Ernst Hairer. Variable time step integration with symplectic methods. *Applied Numerical Mathematics*, 25(2-3):219–227, 1997.

[18] Ernst Hairer and Christian Lubich. Modulated fourier expansions for continuous and discrete oscillatory systems. In *Foundations of Computational Mathematics, Budapest 2011*, London Mathematical Society Lecture Note Series, pages 113–128. Cambridge University Press, Cambridge, 2012.

[19] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[20] Alex Haro. 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).

[21] Arieh Iserles. *A First Course in the Numerical Analysis of Differential Equations*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2009.

[22] Arieh Iserles and S.P. Norsett. On quadrature methods for highly oscillatory integrals and their implementation. *BIT Numerical Mathematics*, 40(4), 2000.

[23] L. Ji and A. Papadopoulos. *Sophus Lie and Felix Klein: The Erlangen Pro-*

*gram and Its Impact in Mathematics and Physics*. IRMA lectures in mathematics and theoretical physics. 2015. ISBN 9783037196489.

[24] Angel Jorba and Maorong Zou. A software package for the numerical integration of ODEs by means of high-order Taylor methods. *Experimental Mathematics*, 14:99–117, 2005.

[25] C. Kane, J.E. Marsden, and M. Ortiz. Symplectic-energy-momentum preserving variational integrators. *Journal of Mathematical Physics*, 40(7), 1999.

[26] S. Lall and M. West. Discrete variational Hamiltonian mechanics. *J. Phys. A*, 39(19):5509–5519, 2006.

[27] C. Lanczos. *The Variational Principles of Mechanics*. Dover Books on Physics. Dover Publications, 2012. ISBN 9780486134703.

[28] Benedict Leimkuhler and Sebastian Reich. *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2004.

[29] Melvin Leok and Tatiana Shingel. Prolongation-collocation variational integrators. *IMA J. Numer. Anal.*, 32(3):1194–1216, 2012.

[30] Melvin Leok and Tatiana Shingel. General techniques for constructing variational integrators. *Front. Math. China*, 7(2):273–303, 2012.

[31] Melvin Leok and Jingjing Zhang. Discrete Hamiltonian variational integrators. *IMA Journal of Numerical Analysis*, 31(4):1497–1532, 2011.

[32] A. Lew, J. E. Marsden, M. Ortiz, and M. West. Asynchronous variational integrators. *Arch. Ration. Mech. Anal.*, 167(2):85–146, 2003.

[33] S. Leyendecker, J. E. Marsden, and M. Ortiz. Variational integrators for constrained mechanical systems. *Z. Angew. Math. Mech.*, 88:677–708, 2008.

[34] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to mechanics and symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1999.

[35] Jerrold E. Marsden and Matthew West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001.

[36] Jerrold E. Marsden, George W. Patrick, and Steve Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Comm. Math. Phys.*, 199(2):351–395, 1998.

[37] Klas Modin and Claus Führer. Time-step adaptivity in variational integrators

with application to contact problems. *Z. Angew. Math. Mech.*, 86(10):785–794, 2006.

[38] Sujit Nair. Time adaptive variational integrators: A space-time geodesic approach. *Physica D: Nonlinear Phenomena*, 241(4):315–325, 2012.

[39] Richard D. Neidinger. Introduction to Automatic Differentiation and MATLAB Object-Oriented Programming. *SIAM Review*, 52(3):545–563, 2010.

[40] M. Patterson, M. Weinstein, and A. Rao. An efficient overloaded method for computing derivatives of mathematical functions in MATLAB. *ACM Transactions on Mathematical Software*, 39(3), 2013.

[41] Louis B. Rall. *Automatic Differentiation: Techniques and Applications*, volume 120 of *Lecture Notes in Computer Science*. Springer Verlag, Berlin, 1981.

[42] S. Reich. Backward error analysis for numerical integrators. *SIAM J. Numer. Anal.*, 36:1549–1570, 1999.

[43] M. Rosen. Niels Hendrick Abel and equations of the fith degree. *Amer. Math. Monthly*, 102:495–505, 1995.

[44] J.M. Schmitt and M. Leok. Properties of Hamiltonian variational integrators. *IMA Journal of Numerical Analysis*, 36(2), 2017.

[45] J.M. Schmitt, T. Shingel, and M. Leok. Lagrangian and Hamiltonian Taylor variational integrators. *submitted to BIT Numerical Mathematics*, 2017.

[46] Carles Simo. Principles of Taylor methods for analytic, non-stiff, ODE. In *Advanced Course on Long Time Integration*, University of Barcelona, 2007.

[47] B.P. Sommeijer. An explicit Runge-Kutta method of order twenty-five. *CWI Quarterly*, 11(1):75–82, 1998.

[48] A. Stern and E. Grinspun. Implicit-explicit variational integration of highly oscillatory problems. *Multiscale Model. Simul.*, 7(4):1779–1794, 2009.

[49] Jürgen Struckmeier. Hamiltonian dynamics on the symplectic extended phase space for autonomous and non- autonomous systems. *Journal of Physics A*, 38:1257–1278, 2005.

[50] Xiaobo Tan. Almost symplectic Runge–Kutta schemes for Hamiltonian systems. *Journal of Computational Physics*, 203:250–273, 2005.

[51] Ferdinand Verhulst. *Non-linear Differential Equations and Dynamical Systems*. Springer Verlag, Berlin, 2000.

[52] J.H. Wilkinson. Error analysis of floating-point computation. *Numer. Math.*, 2:319–340, 1960.

[53] J. Wisdom and M. Holman. Symplectic maps for the n-body problem. *Astronomical Journal*, 102(2):1528–1538, 1991.