UNIVERSITY OF CALIFORNIA, SAN DIEGO

Convergence of Galerkin Variational Integrators for Vector Spaces and Lie Groups

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

in

Mathematics with a specialization in Computational Science

by

James Hall

Committee in charge:

Professor Melvin Leok, Chair
Professor Randolph Bank
Professor Jorge Cortes
Professor Michael Holst
Professor Petr Krysl

2013

The dissertation of James Hall is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

DEDICATION

To my family, who have never wavered in their love and support.

EPIGRAPH

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.
—T.S. Elliot

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Epigraph v

Table of Contents vi

List of Figures viii

Acknowledgements x

Vita xi

Abstract of the Dissertation xii

Chapter 1 Introduction 1
- 1.1 Motivation 3
  - 1.1.1 Kepler N-Body Problems and Stability of the Solar System 3
  - 1.1.2 Lie Group Problems and the Rigid Body 4
- 1.2 Preliminaries 8
  - 1.2.1 Geometric Mechanics and Discrete Geometric Mechanics 8
  - 1.2.2 Some Results for Functional Analysis 15
  - 1.2.3 Taylor Expansion of Functionals 16
- 1.3 Galerkin Variational Integrators 17

Chapter 2 Spectral Variational Integrators 20
- 2.1 Introduction 22
  - 2.1.1 Discrete Mechanics 24
- 2.2 Construction 26
  - 2.2.1 Generalized Galerkin Variational Integrators 26
  - 2.2.2 Spectral Variational Integrators 28
- 2.3 Existence, Uniqueness and Convergence 30
  - 2.3.1 Existence and Uniqueness 31
  - 2.3.2 Order Optimal and Geometric Convergence 37
  - 2.3.3 Minimization of the Action 44
  - 2.3.4 Convergence of Galerkin Curves and Noether Quantities 47
- 2.4 Numerical Experiments 63
  - 2.4.1 Harmonic Oscillator 64

2.4.2 N-body Problems 64
2.5 Conclusions and Future Work 73
2.5.1 Future Work 74
2.6 Appendix 75
2.6.1 Proofs of Geometric Convergence of Spectral Variational Integrators 75

Chapter 3 Lie Group Galerkin Variational Integrators 85
3.1 Introduction 87
3.1.1 Discrete Mechanics 89
3.2 Construction 91
3.2.1 General Galerkin Variational Integrators 91
3.2.2 Lie Group Galerkin Variational Integrators 93
3.3 Convergence 102
3.3.1 Geometric and Optimal Convergence 102
3.3.2 Stationary Points are Minimizers 109
3.3.3 Convergence of Galerkin Curves 115
3.4 Cayley Transform Based Method on the Special Orthogonal Group 121
3.4.1 Free Rigid Body 121
3.4.2 Construction 122
3.4.3 Discrete Euler Poincaré Equations 130
3.5 Numerical Experiments 134
3.5.1 Cayley Transform Method for the Rigid Body 135
3.5.2 Cayley Transform Method for the 3D Pendulum . . 136
3.6 Conclusions and Future Work 140
3.6.1 Future Work 144
3.7 Appendix 145

Chapter 4 Conclusions 153
4.1 Summary of Work 154
4.2 Future Applications 155

LIST OF FIGURES

Figure 1.1: A comparison of numerical integrations of the Solar System using different numerical methods. 5
Figure 1.2: Angular momenta of numerical integrations of the rigid body. . . . 7

Figure 2.1: A visual schematic of the curve $\tilde{q}_n(t) \in \mathbb{M}^n([0,h],Q)$. 26
Figure 2.2: Conserved and approximately conserved Noether quantities and the resulting constrained solution space. 53
Figure 2.3: Geometric convergence of the spectral variational integration of the harmonic oscillator. 65
Figure 2.4: Geometric convergence of the energy error of the spectral variational integration of the harmonic oscillator. 65
Figure 2.5: Energy stability of the spectral variational integration of the harmonic oscillator. 66
Figure 2.6: Geometric convergence of the Kepler 2-body problem. 66
Figure 2.7: Geometric convergence of the energy error of Kepler 2-body problem. 67
Figure 2.8: Geometric convergence of the angular momentum of the Kepler 2-body problem. 67
Figure 2.9: Order optimal convergence of the Kepler 2-body problem. 68
Figure 2.10: Convergence of the Kepler 2-body problem energy with h refinement. 68
Figure 2.11: Convergence of the angular momentum Kepler 2-body problem with h refinement. 69
Figure 2.12: Stability of energy for Kepler 2-body problem. 69
Figure 2.13: Stability of angular momentum for the Kepler 2-body problem. . . 71
Figure 2.14: Orbital diagram for the inner Solar System. 72
Figure 2.15: Orbital diagram for the outer Solar System. 72

Figure 3.1: Geometric convergence of the Lie group spectral variation integrator for the rigid body. 137
Figure 3.2: Order optimal convergence of the Lie group Galerkin variational integrator for the rigid body. 137
Figure 3.3: Energy behavior of the Lie group Galerkin variational integrator for the rigid body. 138
Figure 3.4: Conserved quantities for the Lie group Galerkin variational integrator for the rigid body. 139
Figure 3.5: Geometric convergence of the Lie group spectral variational integrator for the 3D pendulum. 140
Figure 3.6: Order optimal convergence of the Lie group Galerkin variational integrator for the 3D pendulum. 141
Figure 3.7: Energy behavior of the Lie group Galerkin variational integrator for the 3D pendulum. 141
Figure 3.8: Dynamics of the numeric integration of the 3D pendulum. 142

Figure 3.9: Energy error of the Lie group Galerkin variational integrator for the 3D pendulum. 143

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Melvin Leok for his support as my advisor and chair of my thesis committee. I could not have asked for a better mentor during my time as a graduate student at UCSD. Without his guidance, insight, and encouragement, this work would not have been possible, and I would not have had the great opportunity to work in such a fascinating and beautiful field.

During my time at UCSD, I have been blessed with excpetional friends. I cannot imagine having gone through grad school without them. I would like to offer my deep gratitude to them for sharing this experience with me.

I would also like to thank the other members of my thesis committee– Randolph Bank, Jorge Cortes, Michael Holst, and Petr Krysl, for their time and input. Additionally, I would like to thank the professors of Center for Computational Mathematics, for deeply enriching my time at UCSD.

Chapter 2, in full, has been submitted for publication of the material as it may appear in Numerische Mathematik, 2012. Hall, James; Leok, Melvin, Springer, 2012. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Hall, James; Leok, Melvin. The dissertation author was the primary investigator and author of this material.

VITA

2006	Bachelor of Science, *magna cum laude*, Georgetown University
2006	Technical Staff, Jet Propulsion Laboratory
2007	Technical Staff, The Aerospace Corporation
2007-2008	Master of Science, Oxford University
2008-2012	Teaching Assistant, University of California, San Diego
2009-2013	Senior MATLAB Teaching Assistant, University of California, San Diego
2011	Master of Arts, Applied Mathematics, University of California, San Diego
2012	Candidate of Philosophy, University of California, San Diego
2013	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

James Hall, Melvin Leok, "Spectral Variational Integrators", *submitted for publication*, 2012. arXiv:math.NA/1211.4534

James Hall, "Wavelets for Road Surface Classification", *Master of Science Thesis*, University of Oxford, 2008

James Hall, James Cutts, Tibor Balint, Craig Peterson, "Rapid Cost Assessment of Space Mission Concepts through Application of Complexity-Based Cost Indices", *Proceeding of 2008 IEEE Aerospace Conference*, 2008

ABSTRACT OF THE DISSERTATION

Convergence of Galerkin Variational Integrators for Vector Spaces and Lie Groups

by

James Hall

Doctor of Philosophy in Mathematics with a specialization in Computational Science

University of California, San Diego, 2013

Professor Melvin Leok, Chair

In this dissertation we discuss the construction and convergence of high order structure preserving numerical methods for problems in Lagrangian mechanics. Specifically, we make use of the discrete mechanics framework to construct symplectic integrators which have convergence which is optimal, in a certain sense. We then demonstrate how this optimal convergence can be leveraged to construct integrators of arbitrarily high order or with geometric convergence. We further show how these methods can be used to construct continuous approximations to the dynamics of a Lagrangian system, and show that these approximations are also very high order and have excellent structure preserving properties.

We discuss the formulation of symplectic integrators on both vector spaces and

Lie groups, using a Galerkin construction to induce a variational integrator. We begin with the formulation for vector spaces, and then extend this construction to Lie groups through the use of a convenient coordinate chart. We provide the necessary conditions for optimal convergence for both the vector space and Lie group constructions, and demonstrate that many canonical systems automatically satisfy these conditions. We close with several numerical experiments demonstrating the predicted convergence, and discuss further extensions of this work.

Chapter 1

Introduction

Research into geometric mechanics has proven to be an important development in the study of the behavior of mechanical systems. Broadly speaking, the study of geometric mechanics offers insight into the behavior of mechanical systems by examining the underlying geometric properties of the system, particularly conserved quantities that arise from symmetries of the system. These conserved quantities constrain the evolution of the system in important ways; by understanding these constraints, we both study the behavior of many important systems which arise across diverse fields including engineering, physics, chemistry, etc., even when analytic solutions are not available to us.

One of the important practical consequences of the research into geometric mechanics has been the development of numerical methods which conserve or approximately conserve the geometric invariants revealed by the continuous theory. The development of these methods is important because they generally construct more faithful approximations to the behavior of systems with structure than methods which do not conserve the geometric structure. Numerical methods which respect the underlying geometric structure of the systems they approximate are referred to as *geometric methods* or *structure-preserving methods*, and there is a vast collection of them. Because the important underlying structures change from system to system, different methods are appropriate for different types of problems.

This dissertation focuses on the development of a new class of numerical methods for problems in Lagrangian mechanics. Working from the framework of discrete mechanics, a general construction for symplectic integrators is presented. By leveraging results from approximation theory, functional analysis, and geometric mechanics, it is shown that the collection of methods which result from this construction have exceptional properties. Specifically, these methods are high-order, symplectic, momentum preserving, highly stable, and robust over long term simulations. Furthermore, they can be adapted to address the specialized needs of the problems, and as such represent a significant development in the field of structure preserving numerical methods.

The remainder of this introduction presents two example problems with significant geometric structure, and some of the preliminary mathematical theory, which will be briefly reviewed in Chapters 2 and 3. A brief summary of the approach and major results of the dissertation will also be presented. Chapter 2, which is a reprint of a paper which has been submitted for publication, discusses the construction for problems where the configuration space is a vector space, and the convergence of the numerical method in this case. Chapter 3, which is a preprint of a paper being prepared for publication, discusses the extension of the vector space method from Chapter 2 to Lie groups. The final chapter, the conclusion, summarizes the major results and suggest further directions for this research.

1.1 Motivation

It is natural to wonder whether there is much to gain from utilizing numerical methods which preserve the geometric structure of a mechanical system. Structure preserving numerical methods can be more difficult to construct and analyze than classical methods, and there already exists a well developed theory of numerical methods for general numerical ordinary differential equations. In this section, two example problems are presented which provide examples of cases where the geometry of the mechanical system offers significant insight into the behavior of the system, and where geometric numerical methods obviously outperform their classical counterparts. Furthermore, these examples are neither exotic nor pathological; they both arise from simple and common mechanical systems, and understanding their behavior remains an important field which is still under active research.

1.1.1 Kepler N-Body Problems and Stability of the Solar System

The first example problem is the Kepler N-body problem and the dynamics of the Solar System. The Kepler N-body problem describes the dynamics of N celestial bodies under mutual gravitational attraction. Kepler's law of gravitation states that the acceleration of two bodies under the mutual influence of gravity is proportional to the masses of the two bodies and inversely proportional to the square of the distance between them,

$$\ddot{q}_i = \frac{G m_i m_j}{\left\| q_i - q_j \right\|_2^3} \left(q_j - q_i \right),$$

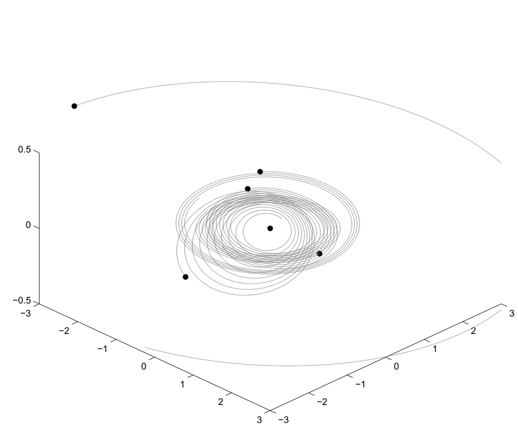
where we have used q_i to index the position of body i. This problem, while very well studied, still lacks a general closed form solution for $N > 2$. Perhaps more surprisingly, the long term stability of the configuration of our Solar System under these laws of gravitation is an open question. Numerical integration has been an important tool for investigating stability. However, if standard numerical techniques are used to numerically integrate the Solar System, the resulting approximations are deeply flawed, as can be seen in Figures 1.1a and 1.1b.

While experience tells us that the simulations in Figures 1.1a and 1.1b are hopelessly incorrect in their qualitative behavior, geometric mechanics can offer a deeper insight into the flaws of these numerics. Plots of the evolution of energy of the system over time demonstrates that the numerical solutions have incorrect energy behavior, as energy is an invariant for the Kepler N-body system, and these numerical simulations have dramatic energy drift. Geometric mechanics places our intuition on firm ground; while we cannot say for certain that all of the planets will maintain stable orbits over long periods of time, we know that the numerical simulations of decaying orbits are certainly incorrect, as they wildly violate the conservation laws associated with this system.

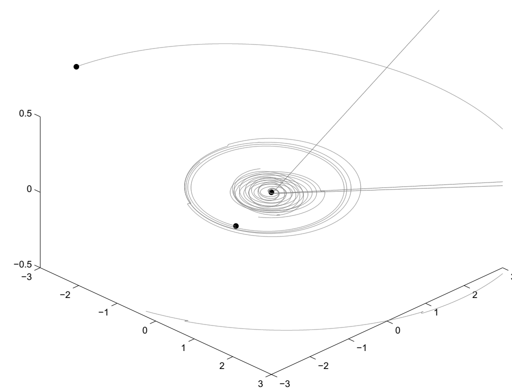
However, when a symplectic integrator is applied to this problem, as in Figure 1.1c, the numerical solution has much better qualitative behavior. While the symplectic integrator has the same order of accuracy as the classical methods, it provides a solution which approximately conserves energy. Thus, because the numerical solution in Figure 1.1c has similar geometric structure to the true solution, the symplectic integrator provides an approximation which much more closely resembles the qualitative behavior of the system, and hence offers a better understanding of the dynamics of the system. In fact, the invention of specially constructed symplectic integrators was a key development for much of the numerical investigation of long term Solar System behavior, as described in Sussman and Wisdom [7].

1.1.2 Lie Group Problems and the Rigid Body

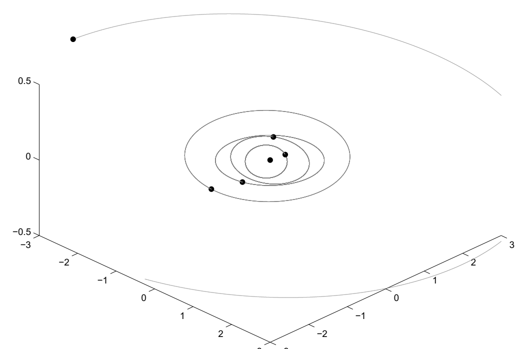
The second example problem is the evolution of the rigid body in the absence of external forces. In addition to having conserved quantities associated with symmetries of the system, this problem has the added structure that it evolves on the group



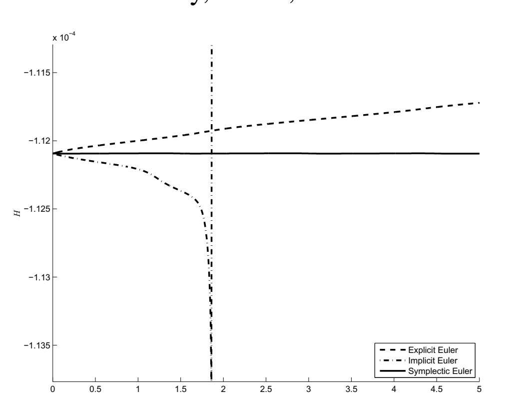
(a) Numerical integration of the Solar System using the explicit Euler method with time step $h = 0.005$. This figure only shows the inner planets.



(b) Numerical integration of the Solar System using the implicit Euler method with time step $h = 0.005$. This figure only shows the inner planets, including escape trajectories for Mercury, Venus, and Earth.



(c) Numerical integration of the Solar System using the symplectic Euler method with time step $h = 0.005$. This figure only shows the inner planets.



(d) Comparison of the energy behavior of the different numerical simulations for the Solar System.

Figure 1.1: A comparison of numerical integrations of the Solar System using different numerical methods. All three methods used are $\mathcal{O}(h)$ where h is the step size for the one-step map.

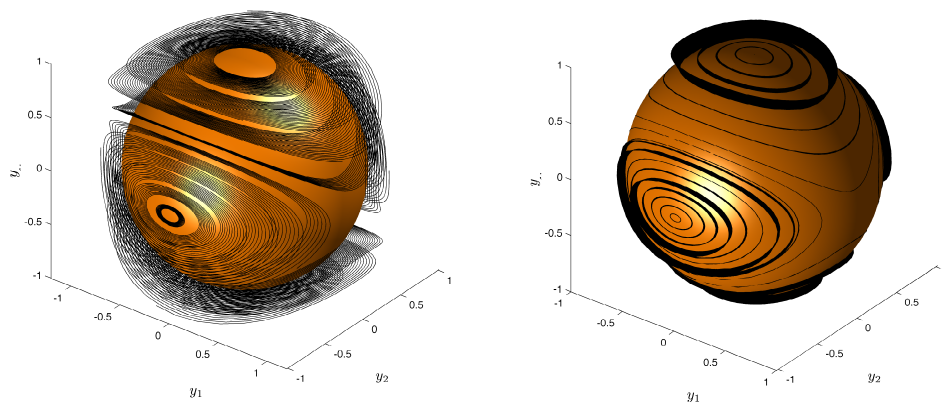
of rotations, $SO(3)$. As the system evolves, the dynamics can be described through a matrix valued function $Q(t)$ with the properties $Q(t)^T Q(t) = I$ and $\det Q(t) = 1$ for all t. When formulated as a differential equation, it is most convenient to describe the evolution of this system with the system of equations

$$\begin{aligned}\dot{y} &= y \times w \\ y &= \Theta w \\ \dot{Q} &= QW\end{aligned}$$

where y is a vector in $\mathbb{R}^3$ that describes the angular momentum of the rigid body, $\Theta = \operatorname{diag}(I_1, I_2, I_3)$ is the diagonal matrix of the moments of inertia of the rigid body, w is a vector in $\mathbb{R}^3$ that describes the angular velocity of the rigid body, and W is the hat map of w, given by

$$W = \hat{w} = \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix}.$$

Naively applying standard discretizations to these methods is disastrous. Not only will the method fail to conserve two important invariants, the energy $E = \frac{1}{2}\sum_{i=1}^{3} I_i^{-1} y_i^2$ and the angular momentum $\Omega = \frac{1}{2}\sum_{i=1}^{3} y_i^2$, but the approximate solution will quickly lose orthogonality, causing the solution to drift out of the Lie group. Such behaviors are qualitatively undesirable, and render the approximate solution essentially meaningless, as the numerical solution will no longer be in the configuration space $SO(3)$. Again, geometric methods produce much better results. By designing methods that preserve conserved quantities and respect the group structure of the analytic solution, numerical methods can be constructed for problems that are inherently out of reach of methods that do not respect the group structure.



(a) Numerical integration of the rigid body using the explicit Euler method.

(b) Numerical integration of the rigid body using the MATLAB routine ode45.

Figure 1.2: Behavior of the angular momenta of the numerical integration of the rigid body. The sphere represents the invariant $\frac{1}{2}\sum_{i=1}^{3} y_i^2 = \Omega$ which is conserved in the analytic solution of the rigid body. Furthermore, the analytic solution of the rigid body also has the invariant $H = \frac{1}{2}\sum_{i=1}^{3} I_i^{-1} y_i^2$. The solution is constrained to intersections of the surfaces described by these two invariants, and hence should lay on closed curves on the surface of the sphere depicted. Clearly, neither numerical solutions exhibit this behavior.

1.2 Preliminaries

Both Chapter 1 and Chapter 2 are self contained papers, however, the following section will briefly review some of the results that are used throughout both works. Some of the relevant material, particularly that pertaining to discrete geometric mechanics, will be reviewed in the chapters.

1.2.1 Geometric Mechanics and Discrete Geometric Mechanics

Geometric Mechanics

Geometric mechanics is the study of mechanical systems that makes use of the geometric structure of the systems. There are two basic perspectives of geometric mechanics: Hamiltonian and Lagrangian mechanics. The work done in later chapters is almost exclusively from the Lagrangian perspective, but there are several important structures that arise naturally from Hamiltonian mechanics, so for completeness, a brief discussion of both perspectives is included below.

Hamiltonian Mechanics

Even though the primary perspective in this dissertation is the Lagrangian variational formulation, Hamiltonian mechanics will be introduced first. This is because some of the important geometric structure, particularly the symplectic form, arise very naturally from the Hamiltonian formulation of mechanics. This is only a short introduction; for a more extensive discussion, the interested reader is referred to Marsden and Ratiu [6].

Consider a dynamical system where at each point in time, the state of the system can be described by a point on a n-dimensional manifold Q. Q is referred to as the *configuration manifold* or *configuration space*, and the cotangent bundle associated with a configuration space, T^*Q, is referred to as *phase space*. On the configuration space, local coordinates are denoted by q and on phase space local coordinates are given by (q, p).

A dynamical system is said to be *Hamiltonian* if there exists a Hamiltonian function $H : T^*Q \to \mathbb{R}$ such that the dynamics of the system over the time interval $[0, T]$

satisfy

$$\delta \int_0^T \langle p, \dot{q} \rangle - H(q,p)\,\mathrm{d}t = 0$$

for fixed $q(0)$ and $q(T)$, where δ denotes the variational derivative. This variational formulation is equivalent to the first-order system, referred to as *Hamilton's equations*:

$$\begin{aligned} \dot{q} &= \frac{\partial H}{\partial p}(q,p) \\ \dot{p} &= -\frac{\partial H}{\partial q}(q,p). \end{aligned} \tag{1.1}$$

Define the canonical two-form Ω, which will be referred to as the *symplectic form*, associated with T^*Q as

$$\Omega(p,q) = dq^i \wedge dp^i.$$

If $Q = \mathbb{R}^n$, then T^*Q can easily be identified with $\mathbb{R}^{2n}$, and the canonical two-form is represented as a quadratic form on $\mathbb{R}^{2n}$, $\Omega(v,w) = v^T \Omega w$, with the symplectic matrix

$$\Omega = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

where I is the $n \times n$ identity matrix. Given a Hamiltonian function H, the Hamiltonian X_H vector field (1.1) associated with H is given by

$$i_{X_H}\Omega = dH$$

where i_{X_H} is the interior product or contraction by X_H and d is the exterior derivative. In $\mathbb{R}^n$, the vector field X has the coordinate expression

$$X(q,p) = \Omega^{-1}\nabla H(q,p),$$

which is simply (1.1) written in terms of the symplectic matrix. For any flow of a Hamiltonian vector field, ϕ_t, the symplectic form is conserved along the flow, that is

$\phi_t^*\Omega = \Omega$ for any $t \in [0,T]$. Flow maps that have this property are called *symplectic*.

Lagrangian Mechanics

While in the Hamiltonian formulation the fundamental object was the Hamiltonian $H : T^*Q \to \mathbb{R}$, the fundamental object in Lagrangian mechanics is the *Lagrangian*, $L : TQ \to \mathbb{R}$, which is a function on the tangent bundle of Q. The Lagrangian induces a functional over the space of second differentiable curves on Q, known as the *action*, $\mathfrak{S} : C^2([0,T],Q) \to \mathbb{R}$, given by

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

For mechanical systems, *Hamilton's principle* states the dynamics of the system are the stationary points of the action for a certain Lagrangian with prescribed boundary conditions $q(0) = q_0$, $q(T) = q_T$, that is, the solutions satisfy

$$\delta\mathfrak{S}(q) = \delta\int_0^T L(q(t),\dot{q}(t))\,\mathrm{d}t = 0. \tag{1.2}$$

An alternative formulation is to view solutions as extremizers of the action, that is if $\bar{q}(t)$ describes the dynamics of the system, then $\bar{q}(t)$ satisfies

$$\bar{q}(t) = \underset{\substack{q(t)\in C^2([0,T],Q)\\ q(0)=q_0,q(T)=q_T}}{\operatorname{argext}} \int_0^T L(q,\dot{q})\,\mathrm{d}t. \tag{1.3}$$

Using the calculus of variations it can be shown that (1.2) and (1.3) are equivalent to the solutions satisfying the *Euler-Lagrange equations*:

$$\frac{\partial L}{\partial q}(q,\dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}(q,\dot{q}) = 0. \tag{1.4}$$

Lagrangian functions typically take the form of kinetic energy minus potential energy, and for Lagrangians of the canonical form $L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$, where M is a mass matrix and $V : Q \to \mathbb{R}$ is the potential energy, the Euler-Lagrange equations reduce to Newton's second law of motion: $M\ddot{q} = -\nabla V(q)$.

For sufficiently smooth Lagrangians, there exists a correspondence between the Lagrangian formulation for the dynamics on the tangent bundle, TQ, and a Hamiltonian formulation for the dynamics in phase space, T^*Q. Given a Lagrangian L, define the *Legendre transform* $\mathbb{F}L : TQ \to T^*Q$,

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

or

$$p = \frac{\partial L}{\partial \dot{q}}(q,\dot{q}). \tag{1.5}$$

A Lagrangian is said to be *hyperregular* if the Legendre transform is a global isomorphism. For any hyperregular Lagrangian it is possible to define a Hamiltonian map $H : T^*Q \to \mathbb{R}$

$$H(q,p) = \langle p, \dot{q} \rangle - L(q,\dot{q}).$$

where $\dot{q}$ is considered a function of (q,p) through (1.5). Using this definition, an elementary calculation shows that the Euler-Lagrange equations (1.4) are equivalent to Hamilton's equations (1.1).

Formulating mechanics from the Lagrangian viewpoint offers many insights into the geometric structure of the dynamics. For example for hyperregular Lagrangians, the flow of the Euler-Lagrange equations ϕ_t preserves the symplectic form on the phase space of Q,

$$\Omega_L = \phi_t^* \Omega_L,$$

where Ω_L is the symplectic form Ω pulled back to TQ by the Legendre transform.

Additionally, Noether's Theorem reveals that every symmetry of a Lagrangian gives rise to a conserved quantity.

Theorem 1.2.1. (Noether's Theorem) *Consider a system with Hamiltonian $H(q,p)$ and Lagrangian $L(q,\dot{q})$. Suppose $\{g_s : s \in \mathbb{R}\}$ is a one-parameter group of transformations*

$(g_s \circ g_r = g_{s+r})$ *which leaves the Lagrangian invariant:*

$$L\left(g_s(q), g_s'(q)\dot{q}\right) = L(q, \dot{q})$$

for all s and all $(q, \dot{q})$. *Let*

$$a(q) = \left.\frac{d}{ds}\right|_{s=0} g_s(q)$$

be defined as the vector field with flow $g_s(q)$. *Then*

$$I(p, q) = \langle p, a(q) \rangle,$$

where $\langle \cdot, \cdot \rangle$ *denotes the canonical pairing between a vector and a one-form, is a conserved quantity along the flow of the Hamiltonian system.*

Of course, Noether's Theorem also induces a conserved quantity along the flow of the Lagrangian system through the Legendre transform.

An important special case of Noether's Theorem is when the Lagrangian is invariant under the action of a Lie group, G, where the invariance of the Lagrangian leads to the theory of equivariant *momentum maps*, $J : T^*Q \to \mathfrak{g}^*$ and $J_L : TQ \to \mathfrak{g}^*$, where $\mathfrak{g}$ is the Lie Algebra associated with the Lie Group G and J_L is J pulled back to TQ by the Legendre transform. A classical example of an equivariant momentum map induced by a Lie group is the invariance of the Lagrangian under the action of $\mathrm{SO}(n)$, which leads to the principle of conservation of angular momentum.

Discrete Geometric Mechanics

The developments in this dissertation are deeply rooted in the theory of discrete mechanics. Discrete mechanics provides a powerful theoretical framework for constructing numerical methods by developing a discrete theory of mechanics which closely parallels the continuous geometric theory. By using a variational approach, discrete mechanical systems are endowed with many structures which are analogous to structures in continuous mechanics. Integrators constructed from this framework hence have conservation properties similar to the conservation properties of the exact flow. An extensive overview can be found in Marsden and West [5].

The fundamental object in discrete mechanics is the discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$. The discrete Lagrangian can be viewed as an approximation to the action of the Lagrangian over a short interval,

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

The discrete Lagrangian is used to form the discrete action sum, which approximates the action of the entire time interval $[0, T]$,

$$\mathbb{S}\left(\{q_i\}_{i=1}^N\right) = \sum_{i=1}^{N-1} L_d(q_i, q_{i+1}) \approx \int_0^T L(q, \dot{q})\, \mathrm{d}t, \tag{1.6}$$

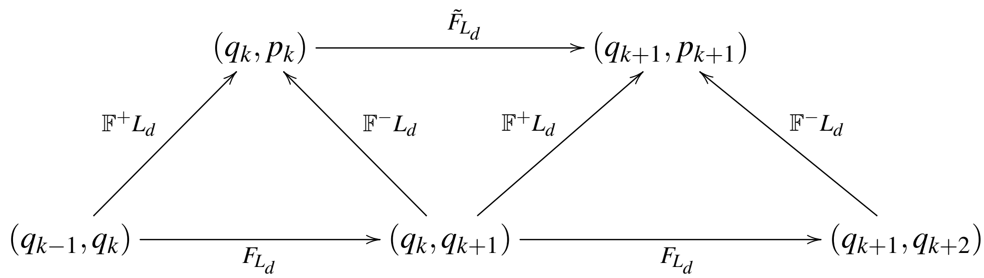
where $T = Nh$. Given a fixed q_1 and q_N, requiring the stationarity of the action sum yields the discrete Euler-Lagrange equations,

$$D_2 L_d(q_{i-1}, q_i) + D_1 L_d(q_i, q_{i+1}) = 0. \tag{1.7}$$

The discrete Euler-Lagrange equations implicitly define an update map, $F_{L_d} : Q \times Q \to Q \times Q$, where $F_{L_d}(q_{i-1}, q_i) = (q_i, q_{i+1})$ if q_{i-1}, q_i, and q_{i+1} satisfy the discrete Euler Lagrange equations. This update map is known as the *discrete Lagrangian flow map*. They additionally define a pair of discrete Legendre transforms $\mathbb{F}L_d^{\pm} : Q \times Q \to T^*Q$, given by

$$\begin{aligned}
\mathbb{F}L^{+}L_d(q_0, q_1) &\to (q_1, p_1) = (q_1, D_2 L_d(q_0, q_1)), \\
\mathbb{F}L^{-}L_d(q_0, q_1) &\to (q_0, p_0) = (q_0, -D_1 L_d(q_0, q_1)).
\end{aligned}$$

Combining the discrete Legendre transforms with (1.7) gives an update map, $\tilde{F}_{L_d} : T^*Q \to T^*Q$ through phase space known as the *discrete Hamiltonian flow map*. The relationship between the discrete Lagrangian flow map, the discrete Legendre transforms, and the discrete Hamiltonian flow map is summarized by the following commutative diagram,



.

The discrete Lagrangian flow map generates a sequence of points which can be considered a discrete analogue to the flow of the Euler-Lagrange equations. Hence, the discrete Lagrangian flow map can be viewed as a numerical method for approximating the flow of the Euler-Lagrange equations. Numerical methods constructed this way are known as *variational integrators*, since they are induced from a discrete variational principle (1.6). Like the flow of the Euler-Lagrange equations, the discrete flow has several important geometric properties.

1. The discrete flow map is symplectic, that is

$$F_{L_d}^* \Omega_{L_d} = \Omega_{L_d},$$

   where $\Omega_{L_d} = \left(\mathbb{F}L_d^{\pm}\right)^* \Omega$.

2. There exists conserved momentum maps through a discrete Noether's theorem (as stated in Hairer et al. [1]),

   Theorem 1.2.2. *(Discrete Noether's Theorem) Suppose a discrete Lagrangian $L_d(q_k, q_{k+1})$ is invariant under the action of a one-parameter group of transformations $\{g_s : s \in \mathbb{R}\}$,*

$$L_d(g_s(q_k), g_s(q_{k+1})) = L_d(q_k, q_{k+1}) \textit{ for all } (q_k, q_{k+1}) \textit{ and } s.$$

   Let $a(q) = \frac{d}{ds}\big|_{s=0} g_s(q)$, and define

$$I(p,q) = \langle p, q \rangle .$$

Then $I(p_k, q_k) = I(p_{k+1}, q_{k+1})$ *for all k.*

3. The energy of the system remains bounded for exponentially long periods of time.

It is the existence of these geometric properties that make variational integrators extremely powerful numerical methods. The discrete mechanics framework gives a systematic method for constructing integrators which are automatically symplectic, momentum preserving, and which have stable energy behavior. Because of these favorable properties, there has been significant recent interest into constructing variational integrators.

1.2.2 Some Results for Functional Analysis

Many of the main results of this paper are a result of studying the behavior of stationary points of functionals that approximate a limiting functional. As such, several results and techniques from functional analysis will be used repeatedly. They are briefly summarized below, and more thorough discussions can be found in Kurdila and Zabarankin [2] and Larsson and Thomée [3].

Poincaré's Inequality

Theorem 1.2.3. *(Poincaré's Inequality) If* Ω *is a bounded domain in* $\mathbb{R}^n$*, then there exists a constant* $C = C(\Omega)$ *such that*

$$\|v\|_{L^2(\Omega)} \leq C \|\nabla v\|_{L^2(\Omega)}, \forall v \in H_0^1(\Omega).$$

When this theorem is used, $\Omega = [0, h]$, an interval of time; as a result we there exists a bound for the Poincaré constant,

$$C \leq \frac{h^2}{\pi^2},$$

and hence the Poincaré inequality gives a specific bound

$$\|v\|_{L^2([0,h])} \leq \frac{h^2}{\pi^2} \|\dot{v}\|_{L^2([0,h])}, \forall v \in H_0^1([0, h]).$$

1.2.3 Taylor Expansion of Functionals

The second major result from functional analysis that will be used repeatedly is the Taylor expansion of functionals. Considering a vector space X and a functional $F : X \to \mathbb{R}$, define its Gateaux derivative to be

$$DF(x_0)[\delta x] = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left(F(x_0 + \varepsilon \delta x) - F(x_0)\right).$$

and its Frechet derivative to be

$$DF(x_0)[\delta x] = \lim_{\|\delta x\| \to 0} \frac{1}{\|\delta x\|}\left(F(x_0 + \delta x) - F(x_0)\right).$$

It should be noted that when the Frechet derivative exists, so does the Gateaux derivative and they are equivalent. Furthermore, it should be noted that the Gateaux derivative corresponds exactly to taking a variation in the sense of variational calculus, and hence in normed vector spaces, the stationarity of the actions corresponds with requiring the first Gateaux derivative of the action to be zero. Even though some of the work presented in this dissertation relates to methods which are not for vector spaces, all functionals will act on vector spaces, and these definitions will always apply.

Now consider the expression for the second Frechet derivative

$$D^2F(x_0)[\delta x_1][\delta x_2] = \lim_{\delta x_2 \to 0} \frac{1}{\|\delta x_2\|}\left(DF(x_0 + \delta x_2)[\delta x_1] - DF(x_0)[\delta x_1]\right).$$

This expression allows the computation of a Taylor expansion of the functional F,

$$F(x_0 + \delta x) = F(x_0) + DF(x_0)[\delta x] + \frac{1}{2}D^2F(z)[\delta x][\delta x]$$

where z lays on the line that connects x_0 and δx. These types of Taylor expansions will be critical for the error analysis which will be presented later.

Finally, most functionals in this dissertation will act on $C^2([0,h],Q)$, and will be of the form

$$F(q) = \int_0^h f(q)\,\mathrm{d}t.$$

Considering the Gateaux derivatives of these curves, which are also the Frechet derivatives, it can be see that

$$\begin{aligned} DF(q)[\delta q] &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^h \left(f(q+\varepsilon q) - f(q)\right) \mathrm{d}t \\ &= \int_0^h \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(f(q+\varepsilon \delta q) - f(q)\right) \mathrm{d}t \\ &= \int_0^h \nabla f(q)[\delta q] \, \mathrm{d}t. \end{aligned}$$

and a similar computation yields

$$D^2 F(q)[\delta q_1][\delta q_2] = \int_0^h \nabla^2 f(q)[\delta q_1][\delta q_2].$$

These expressions establish a clear relationship between the Frechet derivatives of the functional induced by integrating f and the derivatives of f itself, which will be used repeatedly throughout this dissertation.

1.3 Galerkin Variational Integrators

In this dissertation, variational integrators are constructed by taking a Galerkin approach to approximating the discrete Lagrangian. Essentially, a finite dimensional approximation space over the interval $[0,h]$, $\mathbb{M}^n([0,h],Q)$ is chosen and the discrete Lagrangian is constructed by replacing $C^2([0,h],Q)$ with this approximation space and the integral with a quadrature rule,

$$\begin{aligned} L_d^G(q_0,q_1,h) &= \operatorname*{ext}_{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_1}} h \sum_{j=1}^m b_j L\left(q_n\left(c_j h\right), \dot{q}_n\left(c_j h\right)\right) \\ &\approx \operatorname*{ext}_{\substack{q(t)\in C^2([0,h],Q) \\ q(0)=q_0, q(h)=q_1}} \int_0^h L(q(t), \dot{q}(t)) \, \mathrm{d}t. \end{aligned}$$

This type of construction, proposed in Leok [4], naturally arises from considering the variational formulation of the discrete Lagrangian. However, determining the convergence of such types of methods is surprisingly involved. This is because the convergence of Galerkin variational integrators is related to how the stationary points of a sequence of functionals converges to the stationary point of a limiting functional. This type of convergence, known as Γ-convergence, can be difficult to analyze.

This dissertation presents several major results relating to Galerkin variational integrators. In Chapter 2, it is established that, given a certain time step limitation, Galerkin variational integrators on Lagrangians of the canonical form on vector spaces:

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$$

will converge at the same rate as the optimal approximation in the approximation space $\mathbb{M}^n([0,h],Q)$. Additionally, it is established that under similar assumptions, Galerkin integrators will converge geometrically, that is with $\mathscr{O}(K^n)$ for some $K < 1$ when the dimension n of $\mathbb{M}^n([0,h],Q)$ is increased and the time step h is held constant. Finally, it is established that from a Galerkin variational integrator, it is possible to recover a continuous approximation to the exact solution of the Euler-Lagrange equations which converges to the exact solution of the Euler-Lagrange equations, but at a lower rate than the variational integrator.

In Chapter 3, the results from Chapter 2 are extended to Lie groups. Specifically, the Galerkin integrator construction is extended to Lie groups by using a chart function to locally reduce the dynamics of the system to the Lie algebra associated with the Lie group, and the apply to the vector space construction to this local reduction. It is established that as long as the chart function is sufficiently well behaved, the results from Chapter 2 extend easily, and Lie group variational integrators exhibit the same excellent convergence properties.

The optimality results of the error analysis performed in this dissertation are significant because they reduce the convergence of Galerkin variational integrators to a question of approximation theory. Since Galerkin variational integrators can be shown to converge at the same rate as the best possible approximation in the finite-dimensional approximation space $\mathbb{M}^n([0,h],Q)$, one can understand the behavior of Galerkin variational integrators by understanding the properties of the approximation spaces used to construct them. These optimality results provide a systematic method for constructing highly accurate methods.

Finally, through several numerical examples, a variety of features of Galerkin variational integrators that are appealing are demonstrated. Perhaps the most dramatic is that Galerkin variational integrators are stable even with extremely large time steps, and by enriching the function space $\mathbb{M}^n([0,h],Q)$, convergence can be achieved even with large time steps. This opens the possibility for constructing highly accurate long term numerical simulations using these types of methods, as the number of time steps required for a certain length of integration does not increase as the solution is refined. Additionally, in practice, the reconstructed continuous approximations to the solutions of the Euler-Lagrange equations converge at the same rate as the one-step map, even though they have theoretically lower rates of convergence.

References

[1] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[2] Andrew J Kurdila and Michael Zabarankin. *Convex functional analysis*. Springer, 2005.

[3] S. Larsson and V. Thomée. *Partial differential equations with numerical methods*, volume 45 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 2003.

[4] M. Leok. Generalized Galerkin variational integrators: Lie group, multiscale, and pseudospectral methods. (preprint, arXiv:math.NA/0508360), 2004.

[5] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001.

[6] J.E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. Texts in Applied Mathematics. Springer, 1999. ISBN 9780387986432. URL http://books.google.com/books?id=I2gH9ZIs-3AC.

[7] Gerald J Sussman and Jack Wisdom. Chaotic evolution of the solar system. Technical report, DTIC Document, 1992.

Chapter 2

Spectral Variational Integrators

SPECTRAL VARIATIONAL INTEGRATORS

JAMES HALL AND MELVIN LEOK

ABSTRACT. In this paper, we present a new variational integrator for problems in Lagrangian mechanics. Using techniques from Galerkin variational integrators, we construct a scheme for numerical integration that converges geometrically, and is symplectic and momentum preserving. Furthermore, we prove that under appropriate assumptions, variational integrators constructed using Galerkin techniques will yield numerical methods that are in a certain sense optimal, converging at the same rate as the best possible approximation in a certain function space. We further prove that certain geometric invariants also converge at an optimal rate, and that the error associated with these geometric invariants is independent of the number of steps taken. We close with several numerical examples that demonstrate the predicted rates of convergence.

2.1 Introduction

There has been significant recent interest in the development of structure preserving numerical methods for variational problems. One of the key points of interest is developing high-order symplectic integrators for Lagrangian systems. The generalized Galerkin framework has proven to be a powerful theoretical and practical tool for developing such methods. This paper presents a high-order Galerkin variational integrator that exhibits geometric convergence to the true flow of a Lagrangian system. In addition, this method is symplectic, momentum-preserving, and stable even for very large time steps.

Galerkin variational integrators fall into the general framework of discrete mechanics. For a general and comprehensive introduction to the subject, the reader is referred to Marsden and West [23]. Discrete mechanics develops mechanics from discrete variational principles, and, as Marsden and West demonstrated, gives rise to many discrete structures which are analogous to structures found in classical mechanics. By taking these structures into account, discrete mechanics suggests numerical methods which often exhibit excellent long term stability and qualitative behavior. Because of these qualities, much recent work has been done on developing numerical methods from the discrete mechanics viewpoint. See, for example, Hairer et al. [11] for a broad overview of the field of geometric numerical integration, and Marsden and West [23], Müller and Ortiz [26], Patrick and Cuell [28] discuss the error analysis of variational integrators. Various extensions have also been considered, including, Lall and West [14], Leok and Zhang [21] for Hamiltonian systems; Fetecau et al. [10] for nonsmooth problems with collisions; Lew et al. [22], Marsden et al. [24] for Lagrangian PDEs; Cortés and Martínez [5], Fedorov and Zenkov [9], McLachlan and Perlmutter [25] for nonholonomic systems; Bou-Rabee and Owhadi [2, 3] for stochastic Hamiltonian systems; Bou-Rabee and Marsden [1], Lee et al. [16, 17] for problems on Lie groups and homogeneous spaces.

The fundamental object in discrete mechanics is the discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$, where Q is a configuration manifold. The discrete Lagrangian is chosen to be an approximation to the action of a Lagrangian over the time step $[0,h]$,

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

or simply $L_d(q_0,q_1)$ when h is assumed to be constant. Discrete mechanics is formulated by finding stationary points of a discrete action sum based on the sum of discrete Lagrangians,

$$\mathbb{S}(\{q_k\}_{k=1}^n) = \sum_{k=1}^{n-1} L_d(q_k,q_{k+1}) \approx \int_{t_1}^{t_2} L(q,\dot{q})\,\mathrm{d}t.$$

For Galerkin variational integrators specifically, the discrete Lagrangian is induced by constructing a discrete approximation of the action integral over the interval $[0,h]$ based on a finite-dimensional function space and quadrature rule. Once this discrete action is constructed, the discrete Lagrangian can be recovered by solving for stationary points of the discrete action subject to fixed endpoints, and then evaluating the discrete action at these stationary points,

$$L_d(q_0,q_1,h) = \operatorname*{ext}_{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q)\\ q_n(0)=q_0,q_n(h)=q_1}} h\sum_{j=1}^{m} b_j L\left(q\left(c_j h\right),\dot{q}\left(c_j h\right)\right). \tag{2.1}$$

Because the rate of convergence of the approximate flow to the true flow is related to how well the discrete Lagrangian approximates the true action, this type of construction gives a method for constructing and analyzing high-order methods. The hope is that the discrete Lagrangian inherits the accuracy of the function space used to construct it, much in the same way as standard finite-element methods. We will show that for certain Lagrangians, Galerkin constructions based on high-order approximation spaces do in fact result in correspondingly high order methods.

Significant work has already been done constructing and analyzing these types of Galerkin variational integrators. In Leok [18], a number of different possible constructions based on the Galerkin framework are presented. In Leok and Shingel [19], Hermite polynomials are used to construct globally smooth high-order methods. What separates this work from the work that precedes it is the use of a spectral approximation paradigm, which induces methods that exhibit geometric convergence. This type of convergence is established theoretically and demonstrated through numerical examples.

2.1.1 Discrete Mechanics

Before discussing the construction and convergence of spectral variational integrators, it is useful to review some of the fundamental results from discrete mechanics that are used in our analysis. We have already introduced the *discrete Lagrangian* $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$,

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

and the *discrete action sum*,

$$\mathbb{S}\left(\{q_k\}_{k=1}^n\right) = \sum_{k=1}^{n-1} L_d(q_k, q_{k+1}) \approx \int_{t_1}^{t_2} L(q, \dot{q})\, \mathrm{d}t.$$

Taking variations of the discrete action sum and using discrete integration by parts leads to the discrete Euler-Lagrange equations,

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

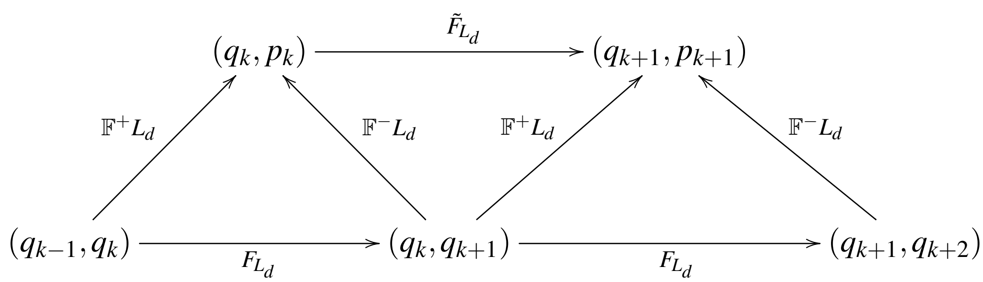
where D_1 denotes differentiation with respect to the first argument and D_2 denotes differentiation with respect to the second argument. Given (q_{k-1}, q_k), these equations implicitly define an update map, known as the *discrete Lagrangian flow map*, $F_{L_d} : Q \times Q \to Q \times Q$, given by $F_{L_d}(q_{k-1}, q_k) = (q_k, q_{k+1})$, where $(q_{k-1}, q_k), (q_k, q_{k+1})$ satisfy (2.2). Furthermore, the discrete Lagrangian defines the *discrete Legendre transforms*, $\mathbb{F}^{\pm} L_d : Q \times Q \to T^*Q$:

$$\begin{aligned}
\mathbb{F}^{+} L_d &: (q_0, q_1) \to (q_1, p_1) = (q_1, D_2 L_d(q_0, q_1)), \\
\mathbb{F}^{-} L_d &: (q_0, q_1) \to (q_0, p_0) = (q_0, -D_1 L_d(q_0, q_1)).
\end{aligned}$$

Using the discrete Legendre transforms, we define the *discrete Hamiltonian flow map*, $\tilde{F}_{L_d} : T^*Q \to T^*Q$,

$$\tilde{F}_{L_d} : (q_0, p_0) \to (q_1, p_1) = \mathbb{F}^+ L_d \left(\left(\mathbb{F}^- L_d \right)^{-1} (q_0, p_0) \right).$$

The following commutative diagram illustrates the relationship between the discrete Hamiltonian flow map, discrete Lagrangian flow map, and the discrete Legendre transforms,



We now introduce the *exact discrete Lagrangian* L_d^E,

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

An important theoretical result for the error analysis of variational integrators is that the discrete Hamiltonian and Lagrangian flow maps associated with the exact discrete Lagrangian produces an exact sampling of the true flow, as was shown in Marsden and West [23]. Using this result, Marsden and West [23] shows that there is a fundamental relationship between how well a discrete Lagrangian L_d approximates the exact discrete Lagrangian L_d^E and how well the corresponding discrete Hamiltonian flow maps, discrete Lagrangian flow maps and discrete Legendre transforms approximate each other. Since the exact discrete Lagrangian produces an exact sampling of the true flow, this in turn leads to the following theorem regarding the error analysis of variational integrators, also found in Marsden and West [23]:

Theorem 2.1.1. (Variational Error Analysis) *Given a regular Lagrangian L and corresponding Hamiltonian H, the following are equivalent for a discrete Lagrangian* L_d*:*

1. *the discrete Hamiltonian flow map for L_d has error $\mathscr{O}\left(h^{p+1}\right)$,*

2. *the discrete Legendre transforms of L_d have error $\mathscr{O}\left(h^{p+1}\right)$,*

3. *L_d approximates the exact discrete Lagrangian with error $\mathscr{O}\left(h^{p+1}\right)$.*

We will make extensive use of this theorem later when we analyze the convergence of spectral variational integrators.

In addition, in Marsden and West [23], it is shown that integrators constructed in this way, which are referred to as *variational integrators*, have significant geometric structure. Most importantly, variational integrators always conserve the canonical symplectic form, and a discrete Noether's Theorem guarantees that a discrete momentum map is conserved for any continuous symmetry of the discrete Lagrangian. The preservation of these discrete geometric structures underlie the excellent long term behavior of variational integrators.

2.2 Construction

2.2.1 Generalized Galerkin Variational Integrators

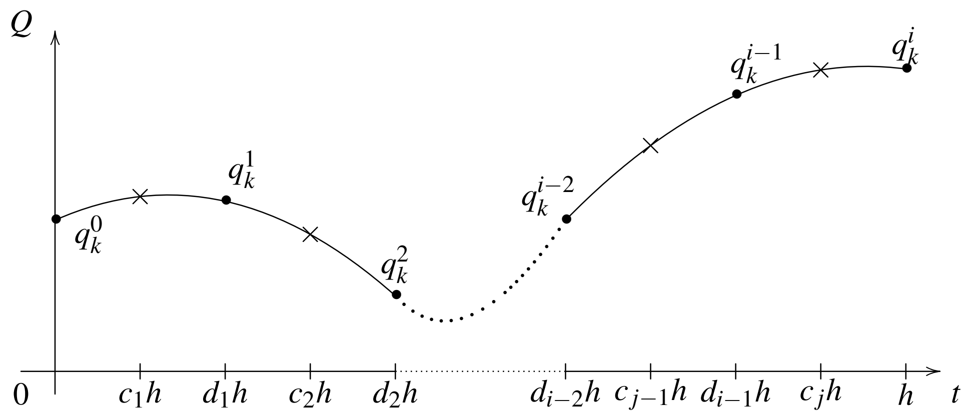


Figure 2.1: A visual schematic of the curve $\tilde{q}_n(t) \in \mathbb{M}^n([0,h],Q)$. The points marked with $(\times)$ represent the quadrature points, which may or may not be the same as interpolation points d_ih. In this figure we have chosen to depict a curve constructed from interpolating basis functions, but this is not necessary in general.

The construction of spectral variational integrators falls within the framework of generalized Galerkin variational integrators, discussed in Leok [18] and Marsden and West [23]. The motivating idea is that we replace the generally non-computable exact discrete Lagrangian $L_d^E(q_k, q_{k+1})$ with a computable discrete analogue, $L_d^G(q_k, q_{k+1})$. Galerkin variational integrators are constructed by using a finite-dimensional function space to discretize the action of a Lagrangian. Specifically, given a Lagrangian $L : TQ \to \mathbb{R}$, to construct a Galerkin variational integrator:

1. choose an n-dimensional function space $\mathbb{M}^n([0,h], Q) \subset C^2([0,h], Q)$, with a finite set of basis functions $\{\phi_i(t)\}_{i=1}^n$,
2. choose a quadrature rule $\mathcal{G}(\cdot) : F([0,h], \mathbb{R}) \to \mathbb{R}$, so that

$$\mathcal{G}(f) = h \sum_{j=1}^{m} b_j f(c_j h) \approx \int_0^h f(t)\,\mathrm{d}t,$$

   where F is some appropriate function space,

and then construct the discrete action $\mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right) : \prod_{i=1}^n Q_i \to \mathbb{R}$, (not to be confused with the discrete action sum $\mathbb{S}(\{q_k\}_{k=1}^\infty)$),

$$\begin{aligned}\mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right) &= \mathcal{G}\left(L\left(\sum_{i=1}^n q_k^i \phi_i(t), \sum_{i=1}^n q_k^i \dot{\phi}_i(t)\right)\right) \\ &= h\sum_{j=1}^m b_j L\left(\sum_{i=1}^n q_k^i \phi_i(c_j h), \sum_{i=1}^n q_k^i \dot{\phi}_i(c_j h)\right),\end{aligned}$$

where we use superscripts to index the weights associated with each basis function, as in Marsden and West [23].

Once the discrete action has been constructed, a discrete Lagrangian can be induced by finding stationary points $\tilde{q}_n(t) = \sum_{i=1}^n q_k^i \phi_i(t)$ of the action under the conditions $\tilde{q}_n(0) = \sum_{i=1}^n q_k^i \phi_i(0) = q_k$ and $\tilde{q}_n(h) = \sum_{i=1}^n q_k^i \phi_i(h) = q_{k+1}$ for some given q_k and q_{k+1},

$$L_d(q_k, q_{k+1}, h) = \operatorname*{ext}_{\substack{\tilde{q}_n(0)=q_k \\ \tilde{q}_n(h)=q_{k+1}}} \mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right) = \operatorname*{ext}_{\substack{\tilde{q}_n(0)=q_k \\ \tilde{q}_n(h)=q_{k+1}}} h\sum_{j=1}^m b_j L\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right).$$

A discrete Lagrangian flow map that result from this type of discrete Lagrangian is referred to as a Galerkin variational integrator.

2.2.2 Spectral Variational Integrators

There are two defining features of spectral variational integrators. The first is the choice of function space $\mathbb{M}^n([0,h],Q)$, and the second is that convergence is achieved not by shortening the time step h, but by increasing the dimension n of the function space.

Choice of Function Space

Restricting our attention to the case where Q is a linear space, spectral variational integrators are constructed using the basis functions $\phi_i(t) = l_i(t)$, where $l_i(t)$ are Lagrange interpolating polynomials based on the points $h_i = \frac{h}{2}\cos\left(\frac{i\pi}{n}\right) + \frac{h}{2}$ which are the Chebyshev points $t_i = \cos\left(\frac{i\pi}{n}\right)$, rescaled and shifted from $[-1,1]$ to $[0,h]$. The resulting finite dimensional function space $\mathbb{M}^n([0,h],Q)$ is simply the polynomials of degree at most n on Q. However, the choice of this particular set of basis functions offer several advantages over other possible bases for the polynomials:

1. the restriction on variations $\sum_{i=1}^{n}\delta q_k^i\phi_i(0) = \sum_{i=1}^{n}\delta q_k^i\phi_i(h) = 0$ reduces to $\delta q_k^1 = \delta q_k^n = 0$,
2. the condition $\tilde{q}_n(0) = q_k$ reduces to $q_k^1 = q_k$,
3. the induced numerical methods have generally better stability properties because of the excellent approximation properties of the interpolation polynomials at the Chebyshev points.

Using this choice of basis functions, for any chosen quadrature rule, the discrete Lagrangian becomes,

$$L_d(q_k,q_{k+1},h) = \underset{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q)\\ q_k^1=q_k,\,q_k^n=q_{k+1}}}{\operatorname{ext}} h\sum_{j=1}^{m} b_j L\left(\tilde{q}_n\left(c_j h\right),\dot{\tilde{q}}_n\left(c_j h\right)\right).$$

Requiring the curve $\tilde{q}_n(t)$ to be a stationary point of the discretized action provides $n-2$ internal stage conditions:

$$h\sum_{j=1}^{m} b_j\left(\frac{\partial L}{\partial q}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\phi_i\left(c_jh\right)+\frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\dot{\phi}_i\left(c_jh\right)\right)=0 \tag{2.3}$$

for $i=2,...,n-1$. Combining these internal stage conditions with the discrete Euler-Lagrange equations,

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

and the continuity condition $q_k^1=q_k$ yields the following set of n nonlinear equations,

$$q_k^1=q_k,$$

$$h\sum_{j=1}^{m} b_j\left(\frac{\partial L}{\partial q}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\phi_i\left(c_jh\right)+\frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\dot{\phi}_i\left(c_jh\right)\right)=0$$

for $i=2,..,n-1$,

$$h\sum_{j=1}^{m} b_j\left(\frac{\partial L}{\partial q}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\phi_1\left(c_jh\right)+\frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\dot{\phi}_1\left(c_jh\right)\right)=p_{k-1}, \tag{2.4}$$

which must be solved at each time step k, and the momentum condition:

$$h\sum_{j=1}^{m} b_j\left(\frac{\partial L}{\partial q}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\phi_n\left(c_jh\right)+\frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\dot{\phi}_n\left(c_jh\right)\right)=p_k,$$

which defines (2.4) for the next time step. Evaluating $q_{k+1}=\tilde{q}_n(h)$ defines the next step for the discrete Lagrangian flow map:

$$F_{L_d}\left(q_{k-1},q_k\right)=\left(q_k,q_{k+1}\right),$$

and because of the choice of basis functions, this is simply $q_{k+1}=q_k^n$.

n-Refinement

As is typical for spectral numerical methods (see, for example, Boyd [4], Trefethen [29]), convergence for spectral variational integrators is achieved by increasing the dimension of the function space, $\mathbb{M}^n([0,h],Q)$. Furthermore, because the order of the discrete Lagrangian also depends on the order of the quadrature rule $\mathcal{G}$, we must also refine the quadrature rule as we refine n. Hence, for examining convergence, we must also consider the quadrature rule as a function of n, $\mathcal{G}_n$. Because of the dependence on n instead of h, we will often examine the discrete Lagrangian L_d as a function of $Q \times Q \times \mathbb{N}$,

$$\begin{aligned}
L_d(q_k, q_{k+1}, n) &= \operatorname*{ext}_{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q)\\ q_k^1=q_k,q_k^n=q_{k+1}}} \mathcal{G}_n\left(L\left(\tilde{q}_n(t), \dot{\tilde{q}}_n(t)\right)\right) \\
&= \operatorname*{ext}_{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q)\\ q_k^1=q_k,q_k^n=q_{k+1}}} h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n\left(c_{n_j}h\right), \dot{\tilde{q}}_n\left(c_{n_j}h\right)\right),
\end{aligned}$$

as opposed to the more conventional

$$\begin{aligned}
L_d(q_k, q_{k+1}, h) &= \operatorname*{ext}_{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q)\\ q_k^1=q_k,q_k^n=q_{k+1}}} \mathcal{G}\left(L\left(\tilde{q}_n(t), \dot{\tilde{q}}_n(t)\right)\right) \\
&= \operatorname*{ext}_{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q)\\ q_k^1=q_k,q_k^n=q_{k+1}}} h\sum_{j=1}^{m} b_j L\left(\tilde{q}_n\left(c_j h\right), \dot{\tilde{q}}_n\left(c_j h\right)\right).
\end{aligned}$$

This type of refinement is the foundation for the exceptional convergence properties of spectral variational integrators.

2.3 Existence, Uniqueness and Convergence

In this section, we will discuss the major important properties of Galerkin variational integrators and spectral variational integrators. The first will be the existence of unique solutions to the internal stage equations (2.3) for certain types of Lagrangians. The second is the convergence of the one-step map that results from the Galerkin and spectral variational constructions, which will be shown to be optimal in a certain sense. The third is the convergence of continuous approximations to the Euler-Lagrange flow which can easily be constructed from Galerkin and spectral variational integrators, and the behavior of geometric invariants associated with the approximate continuous flow. We will show a number of different convergence results associated with these quantities, which demonstrate that Galerkin and spectral variational integrators can be used to compute continuous approximations to the exact solutions of the Euler-Lagrange equations which have excellent convergence and geometric behavior.

2.3.1 Existence and Uniqueness

In general, demonstrating that there exists a unique solution to the internal stage equations for a spectral variational integrator is difficult, and depends on the properties of the Lagrangian. However, assuming a Lagrangian of the form

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q),$$

it is possible to show the existence and uniqueness of the solutions to the implicit equations for the one-step method under appropriate assumptions.

Theorem 2.3.1. (Existence and Uniqueness of Solutions to the Internal Stage Equations) *Given a Lagrangian* $L : TQ \to \mathbb{R}$ *of the form*

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q),$$

if ∇V *is Lipschitz continuous,* $b_j > 0$ *for every* j *and* $\sum_{i=1}^{m} b_j = 1$*, and* M *is symmetric positive-definite, then there exists an interval* $[0,h]$ *where there exists a unique solution to the internal stage equations for a spectral variational integrator.*

Proof. We will consider only the case where $q(t) \in \mathbb{R}$, but the argument generalizes easily to higher dimensions. To begin, we note that for a Lagrangian of the form,

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$$

the equations

$$
\begin{aligned}
q_k^1 =& q_k \\
h\sum_{j=1}^{m} b_j\left(\frac{\partial L}{\partial q}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\phi_i\left(c_jh\right)+\frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\dot{\phi}_i\left(c_jh\right)\right) =& 0,
\end{aligned}
$$

for $i = 2,...,n-1$, and

$$
h\sum_{j=1}^{m} b_j\left(\frac{\partial L}{\partial q}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\phi_1\left(c_jh\right)+\frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\dot{\phi}_1\left(c_jh\right)\right) = p_{k-1},
$$

take the form

$$
Aq^i - f\left(q^i\right) = 0, \tag{2.5}
$$

where q^i is the vector of internal weights, $q^i = \left(q_k^1, q_k^2, ..., q_k^n\right)^T$, A is a matrix with entries defined by

$$
\begin{aligned}
A_{1,1} =& 1, && \quad (2.6)\\
A_{1,l} =& 0, & l = 2,...,n, & \quad (2.7)\\
A_{i,l} =& h\sum_{j=1}^{m} b_j M\dot{\phi}_l\left(c_jh\right)\dot{\phi}_i\left(c_jh\right), & i = 2,...,n; l = 1,...,n, & \quad (2.8)
\end{aligned}
$$

and f is a vector valued function defined by

$$
f\left(q^i\right) = \begin{pmatrix} q_k \\ h\sum_{j=1}^{m} b_j \nabla V\left(\sum_{i=1}^{n} q_k^i \phi_i\left(c_jh\right)\right)\phi_2 \\ \vdots \\ h\sum_{j=1}^{m} b_j \nabla V\left(\sum_{i=1}^{n} q_k^i \phi_i\left(c_jh\right)\right)\phi_{n-1} \\ p_{k-1} \end{pmatrix}.
$$

It is important to note that the entries of A depend on h. For now we will assume A is invertible, and that $\left\|A^{-1}\right\| < \left\|A_1^{-1}\right\|$, for where A_1 is the matrix A generated on the interval $[0,1]$. Of course, the properties of A depend on the choice of basis functions $\{\phi_i\}_{i=1}^{n}$, but we will establish these properties for the polynomial basis later. Defining the map:

$$\Phi\left(q^i\right) = A^{-1} f\left(q^i\right),$$

it is easily seen that (2.5) is satisfied if and only if $q^i = \Phi\left(q^i\right)$, that is, q^i is a fixed point of $\Phi(\cdot)$. If we establish that $\Phi(\cdot)$ is a contraction mapping,

$$\left\|\Phi\left(w^i\right) - \Phi\left(v^i\right)\right\|_\infty \leq k \left\|w^i - v^i\right\|_\infty,$$

for some $k < 1$, we can establish the existence of a unique fixed point, and thus show that the steps of the one step method are well-defined. Here, and throughout this section, we use $\|\cdot\|_p$ to denote the vector or matrix p-norm, as appropriate.

To show that $\Phi(\cdot)$ is a contraction mapping, we consider arbitrary w^i and v^i:

$$\begin{aligned}
\left\|\Phi\left(w^i\right) - \Phi\left(v^i\right)\right\|_\infty &= \left\|A^{-1} f\left(w^i\right) - A^{-1} f\left(v^i\right)\right\|_\infty \\
&= \left\|A^{-1}\left(f\left(w^i\right) - f\left(v^i\right)\right)\right\|_\infty \\
&\leq \left\|A^{-1}\right\|_\infty \left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty .
\end{aligned}$$

Considering $\left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty$, we see that

$$\begin{aligned}
&\left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty \\
&= \left|\sum_{j=1}^{m} b_j \left[\nabla V\left(\sum_{i=1}^{n} w_k^i \phi_i\left(c_j h\right)\right) - \nabla V\left(\sum_{i=1}^{n} v_k^i \phi_i\left(c_j h\right)\right)\right] \phi_{l^*}\left(c_j h\right)\right|,
\end{aligned} \tag{2.9}$$

for some appropriate index l^*. Note that the first and last terms of $\left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty$ will vanish, so the maximum element must take the form of (2.9). Let

$$\phi^i(t) = \left(\phi_1(t), \phi_2(t), \ldots, \phi_n(t)\right).$$

Let C_L be the Lipschitz constant for $\nabla V(q)$. Now

$$\left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty$$

$$
\begin{aligned}
&= \left| h \sum_{j=1}^{m} b_j \left[\nabla V \left(\sum_{i=1}^{n} w_k^i \phi_i \left(c_j h \right) \right) - \nabla V \left(\sum_{i=1}^{n} v_k^i \phi_i \left(c_j h \right) \right) \right] \phi_{p^*} \left(c_j h \right) \right| \\
&\leq h \sum_{j=1}^{m} \left| b_j \right| \left| \left[\nabla V \left(\sum_{i=1}^{n} w_k^i \phi_i \left(c_j h \right) \right) - \nabla V \left(\sum_{i=1}^{n} v_k^i \phi_i \left(c_j h \right) \right) \right] \right| \left| \phi_{p^*} \left(c_j h \right) \right| \\
&\leq h \sum_{j=1}^{m} b_j C_L \left| \sum_{i=1}^{n} w_k^i \phi_i \left(c_j h \right) - \sum_{i=1}^{n} v_k^i \phi_i \left(c_j h \right) \right| \left| \phi_{p^*} \left(c_j h \right) \right| \\
&= h \sum_{j=1}^{m} b_j C_L \left| \sum_{i=1}^{n} \left(w_k^i - v_k^i \right) \phi_i \left(c_j h \right) \right| \left| \phi_{p^*} \left(c_j h \right) \right| \\
&\leq h \sum_{j=1}^{m} b_j C_L \left\| w^i - v^i \right\|_\infty \left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \\
&\leq h \sum_{j=1}^{m} b_j C_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \left\| w^i - v^i \right\|_\infty \\
&= h C_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \left\| w^i - v^i \right\|_\infty .
\end{aligned}
$$

Hence, we derive the inequality

$$
\left\| \Phi \left(w^i \right) - \Phi \left(v^i \right) \right\|_\infty \leq h \left\| A^{-1} \right\|_\infty C_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \left\| w^i - v^i \right\|_\infty ,
$$

and since by assumption $\left\| A^{-1} \right\|_\infty \leq \left\| A_1^{-1} \right\|_\infty$,

$$
\left\| \Phi \left(w^i \right) - \Phi \left(v^i \right) \right\|_\infty \leq h \left\| A_1^{-1} \right\|_\infty C_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \left\| w^i - v^i \right\|_\infty .
$$

Thus if:

$$
h < \left(\left\| A_1^{-1} \right\|_\infty C_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \right)^{-1} ,
$$

then

$$
\left\| \Phi \left(w^i \right) - \Phi \left(v^i \right) \right\|_\infty \leq k \left\| w^i - v^i \right\|_\infty ,
$$

where $k < 1$, which establishes that $\Phi(\cdot)$ is a contraction mapping, and establishes the existence of a unique fixed point, and thus the existence of unique steps of the one step method. □

A critical assumption made during the proof of existence and uniqueness is that the matrix A is nonsingular. This property depends on the choice of basis functions ϕ_i. However, using a polynomial basis, like Lagrange interpolation polynomials, it can be shown that A is invertible.

Lemma 2.3.1. (A is invertible) *If* $\{\phi_i\}_{i=1}^{n}$ *is a polynomial basis of* P_n*, the space of polynomials of degree at most n, M is symmetric positive-definite, and the quadrature rule is order at least* $2n+1$*, then A defined by* (2.6) – (2.8) *is invertible.*

Proof. We begin by considering the equation:

$$Aq^i = 0.$$

Let $\tilde{q}_n(t) = \sum_{i=1}^{n} q_k^i \phi_i(t)$. Considering the definition of A, $Aq^i = 0$ holds if and only if the following equations hold:

$$\begin{aligned} \tilde{q}_n(0) &= 0, \\ h \sum_{j=1}^{m} b_j M \dot{\tilde{q}}_n(c_j h)\, \dot{\phi}_i(c_j h) &= 0, \qquad\qquad i = 1, \ldots, (n-1). \end{aligned} \tag{2.10}$$

It can easily be seen that $\{\dot{\phi}_i\}_{i=1}^{n-1}$ is a basis of P_{n-1}. Using the assumption that the quadrature rule is of order at least $2n-1$ and that M is symmetric positive-definite, we can see that (2.10) implies:

$$\int_0^h M \dot{\tilde{q}}_n(t)\, \dot{\phi}_i(t)\, \mathrm{d}t = 0, \qquad\qquad i = 1, \ldots, (n-1),$$

but,

$$\int_0^h M \dot{\tilde{q}}_n(t)\, \dot{\phi}_i(t)\, \mathrm{d}t = 0$$

implies

$$\left\langle \dot{\tilde{q}}_n, \dot{\phi}_i \right\rangle = 0,$$

where $\langle \cdot, \cdot \rangle$ is the standard L^2 inner product on $[0,h]$. Since $\{\dot{\phi}_i\}_{i=1}^{n-1}$ forms a basis for

P_{n-1}, $\dot{\tilde{q}}_n \in P_{n-1}$, and $\langle \cdot, \cdot \rangle$ is non-degenerate, this implies that $\dot{\tilde{q}}_n(t) = 0$. Thus,

$$\tilde{q}_n(0) = 0$$
$$\dot{\tilde{q}}_n(t) = 0$$

which implies that $\tilde{q}_n(t) = 0$ and hence $q^i = 0$. Thus, if $Aq^i = 0$ then $q^i = 0$, from which it follows that A is non-singular. □

Another subtle difficulty is that the matrix A is a function of h. Since we assumed that $\left\|A^{-1}\right\|_\infty$ is bounded to prove Theorem 2.3.1, we must show that for any choice of h, the quantity $\left\|A^{-1}\right\|_\infty$ is bounded. We will do this by establishing $\left\|A^{-1}\right\|_\infty \leq \left\|A_1^{-1}\right\|_\infty$, where A_1 is A defined with $h = 1$. By Lemma 2.3.1, we know that $\left\|A_1^{-1}\right\|_\infty < \infty$, which establishes the upper bound for $\left\|A^{-1}\right\|_\infty$. This argument is easily generalized for a higher upper bound on h.

Lemma 2.3.2. ($\left\|A^{-1}\right\|_\infty \leq \left\|A_1^{-1}\right\|_\infty$) *For the matrix A defined by* (2.6) – (2.8)*, if $h < 1$, $\left\|A^{-1}\right\|_\infty < \left\|A_1^{-1}\right\|_\infty$ where A_1 is A defined on the interval* $[0,1]$.

Proof. We begin the proof by examining how A changes as a function of h. First, let $\{\phi_i\}_{i=1}^n$ be the basis for the interval $[0,1]$. Then for the interval $[0,h]$, the basis functions are

$$\phi_i^h(t) = \phi_i\left(\frac{t}{h}\right)$$

and hence the derivatives are:

$$\dot{\phi}_i^h(t) = \frac{1}{h}\dot{\phi}_i\left(\frac{t}{h}\right).$$

Thus, if A_1 is the matrix defined by (2.6) – (2.8) on the interval $[0,1]$, then for the interval $[0,h]$,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{h} I_{(n-1)\times(n-1)} \end{pmatrix} A_1,$$

where $I_{n\times n}$ is the $n\times n$ identity matrix. This gives

$$A^{-1} = A_1^{-1}\begin{pmatrix} 1 & 0 \\ 0 & hI_{(n-1)\times(n-1)} \end{pmatrix}$$

which gives

$$\begin{aligned} \left\|A^{-1}\right\|_\infty &= \left\|A_1^{-1}\begin{pmatrix} 1 & 0 \\ 0 & hI_{(n-1)\times(n-1)} \end{pmatrix}\right\|_\infty \\ &\leq \left\|A_1^{-1}\right\|_\infty \left\|\begin{pmatrix} 1 & 0 \\ 0 & hI_{(n-1)\times(n-1)} \end{pmatrix}\right\|_\infty \\ &= \left\|A_1^{-1}\right\|_\infty, \end{aligned}$$

which proves the statement. □

2.3.2 Order Optimal and Geometric Convergence

To determine the rate of convergence for spectral variational integrators, we will utilize Theorem 2.1.1 and a simple extension of Theorem 2.1.1:

Theorem 2.3.2. (Extension of Theorem 2.1.1 to Geometric Convergence) *Given a regular Lagrangian L and corresponding Hamiltonian H, the following are equivalent for a discrete Lagrangian $L_d(q_0,q_1,n)$:*

1. *there exists a positive constant K, where $K<1$, such that the discrete Hamiltonian map for L_d has error $\mathcal{O}(K^n)$,*

2. *there exists a positive constant K, where $K<1$, such that the discrete Legendre transforms of L_d have error $\mathcal{O}(K^n)$,*

3. *there exists a positive constant K, where $K<1$, such that L_d is equivalent to a discrete Lagrangian with error $\mathcal{O}(K^n)$.*

This theorem provides a fundamental tool for the analysis of Galerkin variational methods. Its proof is almost identical to that of Theorem 2.1.1, and can be found in the appendix. The critical result is that the order of the error of the discrete Hamiltonian flow map, from which we construct the discrete flow, has the same order as the discrete Lagrangian from which it is constructed. Thus, in order to determine the order of the error of the flow generated by spectral variational integrators, we need only determine how well the discrete Lagrangian approximates the exact discrete Lagrangian. This is a key result which greatly reduces the difficulty of the error analysis of Galerkin variational integrators.

Naturally, the goal of constructing spectral variational integrators is constructing a variational method that has geometric convergence. To this end, it is essential to establish that Galerkin type integrators inherit the convergence properties of the spaces which are used to construct them. The order optimality result is related to the problem of Γ-convergence (see, for example, Dal Maso [6]), as the Galerkin discrete Lagrangians are given by extremizers of an approximating sequence of variational problems, and the exact discrete Lagrangian is the extremizer of the limiting variational problem. The Γ-convergence of variational integrators was studied in Müller and Ortiz [26], and our approach involves a refinement of their analysis. We now state our results, which establish not only the geometric convergence of spectral variational integrators, but also order optimality of all Galerkin variational integrators under appropriate smoothness assumptions.

Theorem 2.3.3. (Order Optimality of Galerkin Variational Integrators) *Given an interval $[0,h]$ and a Lagrangian $L : TQ \to \mathbb{R}$, let $\bar{q}$ be the exact solution to the Euler-Lagrange equations subject to the conditions $\bar{q}(0) = q_0$ and $\bar{q}(h) = q_h$, and let $\tilde{q}_n$ be the stationary point of a Galerkin variational discrete action, i.e. if $L_d^G : Q \times Q \times \mathbb{R} \to \mathbb{R}$,*

$$\begin{aligned} L_d^G(q_0, q_h, h) &= \underset{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0,\, q_n(h)=q_h}}{\operatorname{ext}} \mathbb{S}_d\left(\{q_i\}_{i=1}^n\right) \\ &= \underset{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0,\, q_n(h)=q_h}}{\operatorname{ext}} h \sum_{j=1}^m b_j L\left(q_n\left(c_j h\right), \dot{q}_n\left(c_j h\right)\right), \end{aligned}$$

then

$$\tilde{q}_n = \underset{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}}{\operatorname{argmin}} h\sum_{j=1}^{m} b_j L\left(q_n\left(c_j h\right), \dot{q}_n\left(c_j h\right)\right).$$

If:

1. *there exists a constant* C_A *independent of* h*, such that, for each* h*, there exists a curve* $\hat{q}_n \in \mathbb{M}^n\left([0,h],Q\right)$*, such that,*

$$\left|\left(\hat{q}_n(t), \dot{\hat{q}}_n(t)\right) - \left(\bar{q}(t), \dot{\bar{q}}(t)\right)\right| \leq C_A h^n,$$

2. *there exists a closed and bounded neighborhood* $U \subset TQ$*, such that* $\left(\bar{q}(t), \dot{\bar{q}}(t)\right) \in U$*,* $\left(\hat{q}_n(t), \dot{\hat{q}}_n(t)\right) \in U$ *for all* t*, and all partial derivatives of* L *are continuous on* U*,*

3. *for the quadrature rule* $\mathscr{G}(f) = h\sum_{j=1}^{m} b_j f\left(c_j h\right) \approx \int_0^h f(t)\,dt$*, there exists a constant* C_g*, such that,*

$$\left|\int_0^h L\left(q_n(t), \dot{q}_n(t)\right) dt - h\sum_{j=1}^{m} b_j L\left(q_n\left(c_j h\right), \dot{q}_n\left(c_j h\right)\right)\right| \leq C_g h^{n+1},$$

   for any $q_n \in \mathbb{M}^n\left([0,h],Q\right)$*,*

4. *and the stationary points* $\bar{q}$*,* $\tilde{q}_n$ *minimize their respective actions,*

then

$$\left|L_d^E\left(q_0, q_h, h\right) - L_d^G\left(q_0, q_h, h\right)\right| \leq C_{op} h^{n+1},$$

for some constant C_{op} *independent of* h*, i.e. discrete Lagrangian* L_d *has error* $\mathscr{O}\left(h^{n+1}\right)$*, and hence the discrete Hamiltonian flow map has error* $\mathscr{O}\left(h^{n+1}\right)$*.*

Proof. First, we rewrite both the exact discrete Lagrangian and the Galerkin discrete Lagrangian:

$$\left|L_d^E\left(q_0, q_h, h\right) - L_d^G\left(q_0, q_h, h\right)\right|$$

$$
\begin{aligned}
&= \left| \int_0^h L(\bar{q}(t), \dot{\bar{q}}(t))\,\mathrm{d}t - \mathscr{G}\left(L\left(\tilde{q}_n(t), \dot{\tilde{q}}_n(t)\right)\right) \right| \\
&= \left| \int_0^h L(\bar{q}(t), \dot{\bar{q}}(t))\,\mathrm{d}t - h\sum_{j=1}^{m} b_j L\left(\tilde{q}_n\left(c_j h\right), \dot{\tilde{q}}_n\left(c_j h\right)\right) \right| \\
&= \left| \int_0^h L(\bar{q}, \dot{\bar{q}})\,\mathrm{d}t - h\sum_{j=1}^{m} b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right|,
\end{aligned}
$$

where in the last line, we have suppressed the t argument, a convention we will continue throughout the proof. Now we introduce the action evaluated on the $\hat{q}_n$ curve, which is an approximation with error $\mathscr{O}(h^n)$ to the exact solution $\bar{q}$:

$$
\begin{aligned}
&\left| \int_0^h L(\bar{q}, \dot{\bar{q}})\,\mathrm{d}t - h\sum_{j=1}^{m} b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right| \\
&= \left| \int_0^h L(\bar{q}, \dot{\bar{q}})\,\mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right)\mathrm{d}t + \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right)\mathrm{d}t - h\sum_{j=1}^{m} b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right| \\
&\leq \left| \int_0^h L(\bar{q}, \dot{\bar{q}})\,\mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right)\mathrm{d}t \right| \qquad (2.11a) \\
&\qquad + \left| \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right)\mathrm{d}t - h\sum_{j=1}^{m} b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right|. \qquad (2.11b)
\end{aligned}
$$

Considering the first term (2.11a):

$$
\begin{aligned}
\left| \int_0^h L(\bar{q}, \dot{\bar{q}})\,\mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right)\mathrm{d}t \right| &= \left| \int_0^h L(\bar{q}, \dot{\bar{q}}) - L\left(\hat{q}_n, \dot{\hat{q}}_n\right)\mathrm{d}t \right| \\
&\leq \int_0^h \left| L(\bar{q}, \dot{\bar{q}}) - L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \right| \mathrm{d}t.
\end{aligned}
$$

By assumption, all partials of L are continuous on U, and since U is closed and bounded, this implies L is Lipschitz on U. Let L_α denote that Lipschitz constant. Since, again by assumption, $(\bar{q}, \dot{\bar{q}}) \in U$ and $\left(\hat{q}_n, \dot{\hat{q}}_n\right) \in U$, we can rewrite:

$$
\begin{aligned}
\int_0^h \left| L(\bar{q}, \dot{\bar{q}}) - L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \right| \mathrm{d}t &\leq \int_0^h L_\alpha \left| (\bar{q}, \dot{\bar{q}}) - \left(\hat{q}_n, \dot{\hat{q}}_n\right) \right| \mathrm{d}t \\
&\leq \int_0^h L_\alpha C_A h^n \mathrm{d}t \\
&= L_\alpha C_A h^{n+1},
\end{aligned}
$$

where we have made use of the best approximation estimate. Hence,

$$\left| \int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t \right| \leq L_\alpha C_1 h^{n+1}. \tag{2.12}$$

Next, considering the second term (2.11b),

$$\left| \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t - h\sum_{j=1}^m b_j L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right) \right|,$$

since $\tilde{q}_n$, the stationary point of the discrete action, minimizes its action and $\hat{q}_n \in \mathbb{M}^n\left([0,h],Q\right)$,

$$h\sum_{j=1}^m b_j L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right) \leq h\sum_{j=1}^m b_j L\left(\hat{q}_n,\dot{\hat{q}}_n\right) \leq \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t + C_g h^{n+1} \tag{2.13}$$

where the inequalities follow from the assumptions on the order of the quadrature rule. Furthermore,

$$\begin{aligned} h\sum_{j=1}^m b_j L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right) &\geq \int_0^h L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right)\mathrm{d}t - C_g h^{n+1} \\ &\geq \int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - C_g h^{n+1} \\ &\geq \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t - L_\alpha C_A h^{n+1} - C_g h^{n+1}, \end{aligned} \tag{2.14}$$

where the inequalities follow from (2.12), the order of the quadrature rule, and the assumption that $\bar{q}$ minimizes its action. Putting (2.13) and (2.14) together, we can conclude:

$$\left| \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t - h\sum_{j=1}^m b_j L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right) \right| \leq \left(L_\alpha C_A + C_g\right) h^{n+1}. \tag{2.15}$$

Now, combining the bounds (2.12) and (2.15) in (2.11a) and (2.11b), we can conclude

$$\left| L_d^E\left(q_0,q_h,h\right) - L_d^G\left(q_0,q_h,h\right) \right| \leq \left(2L_\alpha C_A + C_g\right) h^{n+1}$$

which, combined with Theorem 2.1.1, establishes the order of the error of the integrator.

□

The above proof establishes a significant convergence result for Galerkin variational integrators, namely that for sufficiently well behaved Lagrangians, Galerkin variational integrators will produce discrete approximate flows that converge to the exact flow as $h \to 0$ with the highest possible order allowed by the approximation space, provided the quadrature rule is of sufficiently high order.

We will discuss assumption 4 in §2.3.3. While in general we cannot assume that stationary points of the action are minimizers, it can be shown that for Lagrangians of the canonical form

$$L(q,\dot{q}) = \dot{q}^T M \dot{q} - V(q),$$

under some mild assumptions on the derivatives of V and the accuracy of the quadrature rule, there always exists an interval $[0,h]$ over which stationary points are minimizers. In §2.3.3 we will show the result extends to the discretized action of Galerkin variational integrators. A similar result was established in Müller and Ortiz [26].

Geometric convergence of spectral variational integrators is not strictly covered under the proof of order optimality. While the above theorem establishes convergence of Galerkin variational integrators by shrinking h, the interval length of each discrete Lagrangian, spectral variational integrators achieve convergence by holding the interval length of each discrete Lagrangian constant and increasing the dimension of the approximation space $\mathbb{M}^n([0,h],Q)$. Thus, for spectral variational integrators, we have the following analogous convergence theorem:

Theorem 2.3.4. (Geometric Convergence of Spectral Variational Integrators) *Given an interval $[0,h]$ and a Lagrangian $L : TQ \to \mathbb{R}$, let $\bar{q}$ be the exact solution to the Euler-Lagrange equations, and $\tilde{q}_n$ be the stationary point of the spectral variational discrete action:*

$$L_d^S(q_0,q_h,n) = \underset{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0,q_n(h)=q_h}}{\operatorname{ext}} \mathbb{S}_d\left(\{q_i\}_{i=1}^n\right)$$

$$= \underset{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q)\\ q_n(0)=q_0, q_n(h)=q_h}}{\operatorname{ext}} h\sum_{j=0}^{m_n} b_{n_j} L\left(q_n\left(c_{n_j}h\right), \dot{q}_n\left(c_{n_j}h\right)\right).$$

If:

1. *there exists constants C_A, K_A, $K_A < 1$, independent of n such that, for each n, there exists a curve $\hat{q}_n \in \mathbb{M}^n\left([0,h], Q\right)$, such that,*

$$\left|(\bar{q},\dot{\bar{q}}) - \left(\hat{q}_n, \dot{\hat{q}}_n\right)\right| \leq C_A K_A^n,$$

2. *there exists a closed and bounded neighborhood $U \subset TQ$, such that, $\left(\bar{q}(t), \dot{\bar{q}}(t)\right) \in U$ and $\left(\hat{q}_n(t), \dot{\hat{q}}_n(t)\right) \in U$ for all t and n, and all partial derivatives of L are continuous on U,*

3. *for the sequence of quadrature rules $\mathscr{G}_n(f) = \sum_{j=1}^{m_n} b_{n_j} f\left(c_{n_j}h\right) \approx \int_0^h f(t)\,dt$, there exists constants C_g, K_g, $K_g < 1$, independent of n such that*

$$\left|\int_0^h L\left(q_n(t), \dot{q}_n(t)\right)dt - h\sum_{j=1}^{m_n} b_{n_j} L\left(q_n\left(c_{n_j}h\right), \dot{q}_n\left(c_{n_j}h\right)\right)\right| \leq C_g K_g^n,$$

   for any $q_n \in \mathbb{M}^n\left([0,h], Q\right)$,

4. *and the stationary points $\bar{q}$, $\tilde{q}_n$ minimize their respective actions,*

then

$$\left|L_d^E\left(q_0, q_1\right) - L_d^S\left(q_0, q_1, n\right)\right| \leq C_s K_s^n \tag{2.16}$$

for some constants C_s, K_s, $K_s < 1$, independent of n, and hence the discrete Hamiltonian flow map has error $\mathscr{O}\left(K_s^n\right)$.

The proof of the above theorem is very similar to that of order optimality, and would be tedious to repeat here. It can be found in the appendix. The main differences between the proofs are the assumption of the sequence of converging functions in the increasingly high-dimensional approximation spaces, and the assumption of a sequence of increasingly high-order quadrature rules. These assumptions are used in the obvious way in the modified proof.

2.3.3 Minimization of the Action

One of the major assumptions made in the convergence theorems (2.3.3) and (2.3.4) is that the the stationary points of both the continuous and discrete actions are minimizers over the interval $[0,h]$. This type of minimization requirement is similar to the one made in the paper on Γ-convergence of variational integrators by Müller and Ortiz [26]. In fact, the results in Müller and Ortiz [26] can easily be extended to demonstrate that for sufficiently well-behaved Lagrangians of the form

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q),$$

where $q \in C^2([0,h],Q)$, there exists an interval $[0,h]$, such that stationary points of the Galerkin action are minimizers.

Theorem 2.3.5. *Consider a Lagrangian of the form*

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$$

where $q \in C^2([0,h],Q)$ *and each component* q^d *of* q, $q^d \in C^2([0,h],Q)$, *is a polynomial of degree at most* s. *Assume* M *is symmetric positive-definite and all second-order partial derivatives of* V *exist, and are continuous and bounded. Then, there exists a time interval* $[0,h]$ *such that stationary points of the discrete action,*

$$\mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right) = h\sum_{j=1}^m b_j\left(\frac{1}{2}\dot{\tilde{q}}_n\left(c_jh\right)^T M\dot{\tilde{q}}_n\left(c_jh\right) - V\left(\tilde{q}_n\left(c_jh\right)\right)\right),$$

on this time interval are minimizers if the quadrature rule used to construct the discrete action is of order at least $2s+1$.

We quickly note that the assumption that each component of q, q^d, is a polynomial of degree *at most s* allows for discretizations where different components of the configuration space are discretized with polynomials of different degrees. This allows for more efficient discretizations where slower evolving components are discretized with lower-degree polynomials than faster evolving ones.

Proof. Let $\tilde{q}_n$ be a stationary point of the discrete action $\mathbb{S}_d(\cdot)$, and let δq be an arbitrary perturbation of the stationary point $\tilde{q}_n$, under the conditions $\delta q^d \in P_{S_d}$, $\delta q(0) = \delta q(h) = 0$, which is uniquely defined by $\left\{\delta q_k^i\right\}_{i=1}^n \subset Q$. Then,

$$\begin{aligned}
&\mathbb{S}_d\left(\left\{q_k^i + \delta q_k^i\right\}_{i=1}^n\right) - \mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right)\\
&= h\sum_j^m b_j\left(\frac{1}{2}\left(\dot{\tilde{q}}_n + \delta\dot{q}\right)^T M\left(\dot{\tilde{q}}_n + \delta\dot{q}\right) - V\left(\tilde{q}_n + \delta q\right)\right) - h\sum_j^m b_j\left(\frac{1}{2}\dot{\tilde{q}}_n^T M\dot{\tilde{q}}_n - V\left(\tilde{q}_n\right)\right)\\
&= h\sum_j^m b_j\left(\frac{1}{2}\left(\dot{\tilde{q}}_n + \delta\dot{q}\right)^T M\left(\dot{\tilde{q}}_n + \delta\dot{q}\right) - V\left(\tilde{q}_n + \delta q\right) - \frac{1}{2}\dot{\tilde{q}}_n^T M\dot{\tilde{q}}_n + V\left(\tilde{q}_n\right)\right).
\end{aligned}$$

Making use of Taylor's remainder theorem, we expand:

$$V\left(\tilde{q}_n + \delta q\right) = V\left(\tilde{q}_n\right) + \nabla V\left(\tilde{q}_n\right)\cdot\delta q + \frac{1}{2}\delta\tilde{q}_n^T R\delta\tilde{q}_n,$$

where $|R_{lm}| \leq \sup_{l,m}\left|\frac{\partial^2 V}{\partial q_l \partial q_m}\right|$. Using this expansion, we rewrite

$$\begin{aligned}
&\mathbb{S}_d\left(\left\{q_k^i + \delta q_k^i\right\}_{i=1}^n\right) - \mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right)\\
&= h\sum_j^m b_j\left(\frac{1}{2}\left(\dot{\tilde{q}}_n + \delta\dot{q}\right)^T M\left(\dot{\tilde{q}}_n + \delta\dot{q}\right) - V\left(\tilde{q}_n\right) - \nabla V\left(\tilde{q}_n\right)\cdot\delta q\right.\\
&\qquad\left. - \frac{1}{2}\delta q^T R\delta q - \left(\frac{1}{2}\dot{\tilde{q}}_n^T M\dot{\tilde{q}}_n + V\left(\tilde{q}_n\right)\right)\right)
\end{aligned}$$

which, given the symmetry in M, rearranges to:

$$\begin{aligned}
&\mathbb{S}_d\left(\left\{q_k^i + \delta q_k^i\right\}_{i=1}^n\right) - \mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right)\\
&= h\sum_j^m b_j\left(\dot{\tilde{q}}_n^T M\delta\dot{q} - \nabla V\left(\tilde{q}_n\right)\cdot\delta q + \frac{1}{2}\delta\dot{q}^T M\delta\dot{q} - \frac{1}{2}\delta q^T R\delta q\right).
\end{aligned}$$

Now, it should be noted that the stationarity condition for the discrete Euler-Lagrange equations is

$$h\sum_{j=1}^{m} b_j \left(\dot{\tilde{q}}_n^T M\delta\dot{q} - \nabla V\left(\tilde{q}_n\right)\cdot\delta q\right) = 0$$

for arbitrary δq, which allows us to simplify the expression to

$$\mathbb{S}_d\left(\left\{q_k^i + \delta q_k^i\right\}_{i=1}^n\right) - \mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right) = h\sum_j^m b_j\left(\frac{1}{2}\delta\dot{q}^T M\delta\dot{q} - \frac{1}{2}\delta q^T R\delta q\right).$$

Now, using the assumption that the partial derivatives of V are bounded,

$$|R_{lm}| \leq \left|\frac{\partial^2 V}{\partial q_l \partial q_m}\right| < C_R,$$

and standard matrix inequalities, we get the inequality:

$$\delta q^T R\delta q \leq \|R\delta q\|_2\|\delta q\|_2 \leq \|R\|_2\|\delta q\|_2^2 \leq \|R\|_F\|\delta q\|_2^2 \leq DC_R\|\delta q\|_2^2 = DC_R\delta q^T\delta q, \tag{2.17}$$

where D is the number of spatial dimensions of Q. Thus

$$h\sum_j^m b_j\left(\frac{1}{2}\delta\dot{q}^T M\delta\dot{q} - \frac{1}{2}\delta q^T R\delta q\right) \geq h\sum_j^m b_j\left(\frac{1}{2}\delta\dot{q}^T M\delta\dot{q} - \frac{1}{2}DC_R\delta q^T\delta q\right).$$

Because M is symmetric positive-definite, there exists $m > 0$ such that $x^T Mx \geq mx^T x$ for any x. Hence,

$$h\sum_j^m b_j\left(\frac{1}{2}\delta\dot{q}^T M\delta\dot{q} - \frac{1}{2}DC_R\delta q^T\delta q\right) \geq h\sum_j^m b_j\left(\frac{1}{2}m\delta\dot{q}^T\delta\dot{q} - \frac{1}{2}DC_R\delta q^T\delta q\right).$$

Now, we note that since each component of δq is a polynomial of degree at most s, $\delta q^T\delta q$ and $\delta\dot{q}^T\delta\dot{q}$ are both polynomials of degree less than or equal to $2s$. Since our quadrature rule is of order $2s+1$, the quadrature rule is exact, and we can rewrite

$$h\sum_j^m b_j\left(\frac{1}{2}m\delta\dot{q}^T\delta\dot{q} - \frac{1}{2}DC_R\delta q^T\delta q\right) = \frac{1}{2}\int_0^h m\delta\dot{q}^T\delta\dot{q} - DC_R\delta q^T\delta q\,\mathrm{d}t$$

$$= \frac{1}{2}\left(\int_0^h m\delta\dot{q}^T\delta\dot{q}\mathrm{d}t - \int_0^h DC_R\delta q^T\delta q\mathrm{d}t\right).$$

From here, we note that $\delta q \in H_0^1([0,h],Q)$, and make use of the Poincaré inequality to conclude

$$\frac{1}{2}\left(\int_0^h m\delta\dot{q}^T\delta\dot{q}\mathrm{d}t - \int_0^h nC_R\delta q^T\delta q\mathrm{d}t\right) \geq \frac{1}{2}\left(m\frac{\pi^2}{h^2}\int_0^h \delta q^T\delta q\mathrm{d}t - DC_R\int_0^h \delta q^T\delta q\mathrm{d}t\right)$$
$$= \frac{1}{2}\left(\frac{m\pi^2}{h^2} - DC_R\right)\int_0^h \delta q^T\delta q\mathrm{d}t.$$

Since $\int_0^h \delta q^T\delta q\mathrm{d}t > 0$,

$$\mathbb{S}_d\left(\left\{q_k^i + \delta q_k^i\right\}_{i=1}^n\right) - \mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right) \geq \frac{1}{2}\left(\frac{m\pi^2}{h^2} - DC_R\right)\int_0^h \delta q^T\delta q\mathrm{d}t > 0$$

so long as $h < \sqrt{\frac{m\pi^2}{DC_R}}$. □

2.3.4 Convergence of Galerkin Curves and Noether Quantities

Galerkin Curves

In order to construct the one-step method, spectral variational integrators determine a curve,

$$\tilde{q}_n(t) = \sum_{i=1}^n q_k^i\phi_i(t),$$

which satisfies

$$\tilde{q}_n(t) = \underset{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q) \\ q_n(0)=q_k,q_n(h)=q_{k+1}}}{\operatorname{argmin}} h\sum_{j=1}^m b_j L\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right).$$

Evaluating this curve at h defines the next step of the one-step method, $q_{k+1} = \tilde{q}_n(h)$, but the curve itself has many desirable properties which makes it a good continuous approximation to the true solution of the Euler Lagrange equations $\bar{q}(t)$. In this section, we will examine some of the favorable properties of $\tilde{q}_n(t)$, hereafter referred to as the

Galerkin curve.

However, before discussing the properties of the Galerkin curve, it is useful review the different curves with which we are working. We have already defined the Galerkin curve, $\tilde{q}_n(t)$, and we will also be making use of the local solution to the Euler-Lagrange equations $\bar{q}(t)$, where

$$\bar{q}(t) = \underset{\substack{q(t)\in C^2([0,h],Q) \\ q(0)=q_k, q(h)=q_{k+1}}}{\operatorname{argmin}} \int_0^h L(q(t), \dot{q}(t))\, \mathrm{d}t.$$

However, while for each interval $\bar{q}$ satisfies the Euler-Lagrange equations exactly, it is not the exact solution of the Euler-Lagrange equations globally, as $q_k \neq \Phi_{kh}(q_0, \dot{q}_0)$, where $\Phi_t(q_0, \dot{q}_0)$ is the flow of the Euler-Lagrange vector field. This is particularly important when discussing invariants, where the invariants of $\bar{q}$ remain constant within a time-step, but not from time-step to time-step.

The first property of the Galerkin curve that we will examine is its rate of convergence to the true flow of the Euler-Lagrange vector field. There are two general sources of error that affect the convergence of these curves, the first being the accuracy to which the curves approximate the local solution to the Euler-Lagrange equations over the interval $[0,h]$ with the boundary (q_k, q_{k+1}), and the second being the accuracy of the boundary conditions (q_k, q_{k+1}) as approximations to a true sampling of the exact flow. Numerical experiments will show that often the second source of error dominates the first, causing the Galerkin curves to converge at the same rate as the one-step map. However, the accuracy to which the Galerkin curves approximate the true minimizers independent of the error of the boundary can also be established under appropriate assumptions about the action. Two theorems which establish this convergence are presented below.

Before we state the theorems, we quickly recall the definitions of the *Sobolev Norm* $\|\cdot\|_{W^{1,p}([0,h])}$,

$$\|f\|_{W^{1,p}([0,h])} = \left(\|f\|^p_{L^p([0,h])} + \|\dot{f}\|^p_{L^p([0,h])}\right)^{\frac{1}{p}} = \left(\int_0^h |f|^p\, \mathrm{d}t + \int_0^h |\dot{f}|^p\, \mathrm{d}t\right)^{\frac{1}{p}}.$$

We will make extensive use of this norm when examining convergence of Galerkin curves.

Theorem 2.3.6. (Geometric Convergence of Galerkin Curves with n-Refinement) *Under the same assumptions as Theorem 2.3.4, if at $\bar{q}$, the action is twice Frechet differentiable, and if the second Frechet derivative of the action $D^2\mathfrak{S}(\cdot)[\cdot,\cdot]$ is coercive in a neighborhood U of $\bar{q}$, that is,*

$$D^2\mathfrak{S}(\nu)[\delta q,\delta q] \geq C_f \left\|\delta q\right\|^2_{W^{1,1}([0,h])},$$

for all curves $\delta q \in H^1_0([0,h],Q)$ and all $\nu \in U$, then the curves which minimize the discrete action converge to the true solution geometrically with n-refinement with respect to $\|\cdot\|_{W^{1,1}([0,h])}$. Specifically, if the discrete Hamiltonian flow map has error $\mathscr{O}(K_s^n)$, $K_s < 1$, then the Galerkin curves have error $\mathscr{O}\left(\sqrt{K_s}^n\right)$.

Proof. We start with the bound (2.16) given at the end of Theorem 2.3.4,

$$\left|L_d^E(q_k,q_{k+1}) - L_d^S(q_k,q_{k+1},n)\right| \leq C_s K_s^n$$

and expand using the definitions of $L_d^E(q_k,q_{k+1})$ and $L_d^S(q_k,q_{k+1},n)$, as well as the assumed accuracy of the quadrature rule $\mathscr{G}_n$ to derive

$$\begin{aligned}
C_s K_s^n &\geq \left|L_d^E(q_k,q_{k+1}) - L_d^S(q_k,q_{k+1},n)\right| &(2.18)\\
&= \left|\int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n\left(c_{n_j}h\right),\dot{\tilde{q}}_n\left(c_{n_j}h\right)\right)\right| \\
&\geq \left|\int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - \int_0^h L(\tilde{q}_n,\dot{\tilde{q}}_n)\,\mathrm{d}t\right| - C_g K_g^n &(2.19)\\
&= \left|\mathfrak{S}(\tilde{q}_n) - \mathfrak{S}(\bar{q})\right| - C_g K_g^n
\end{aligned}$$

which implies:

$$(C_s + C_g) K_s^n \geq \left|\mathfrak{S}(\tilde{q}_n) - \mathfrak{S}(\bar{q})\right|$$

because $K_s \geq K_g$, (see the proof of Theorem 2.3.4 in the appendix). Using this inequality, we make use of a Taylor expansion of $\mathfrak{S}(\tilde{q}_n)$,

$$\mathfrak{S}(\tilde{q}_n) = \mathfrak{S}(\bar{q}) + \mathrm{D}\mathfrak{S}(\bar{q})[\tilde{q}_n - \bar{q}] + \frac{1}{2}\mathrm{D}^2\mathfrak{S}(\nu)[\tilde{q}_n - \bar{q}, \tilde{q}_n - \bar{q}],$$

for some $\nu \in U$, to see that

$$\begin{aligned}(C_s + C_g) K_s^n &\geq |\mathfrak{S}(\tilde{q}_n) - \mathfrak{S}(\bar{q})| \\ &= \left| \mathfrak{S}(\bar{q}) + \mathrm{D}\mathfrak{S}(\bar{q})[\tilde{q}_n - \bar{q}] + \frac{1}{2}\mathrm{D}^2\mathfrak{S}(\bar{q})[\tilde{q}_n - \bar{q}, \tilde{q}_n - \bar{q}] - \mathfrak{S}(\bar{q}) \right|.\end{aligned}$$

But

$$\begin{aligned} D\mathfrak{S}(\bar{q})[\tilde{q}_n - \bar{q}] &= \int_0^h \frac{\partial L}{\partial q}(\bar{q}, \dot{\bar{q}})(\tilde{q}_n - \bar{q}) + \frac{\partial L}{\partial \dot{q}}(\bar{q}, \dot{\bar{q}})\left(\dot{\tilde{q}}_n - \dot{\bar{q}}\right) \mathrm{d}t \\ &= \int_0^h \left(\frac{\partial L}{\partial q}(\bar{q}, \dot{\bar{q}}) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}(\bar{q}, \dot{\bar{q}}) \right) \cdot (\tilde{q}_n - \bar{q}) \,\mathrm{d}t \\ &= 0, \end{aligned}$$

because $\tilde{q}_n(0) = \bar{q}(0)$ and $\tilde{q}_n(h) = \bar{q}(h)$ by definition (note that this implies $(\tilde{q}_n - \bar{q}) \in H_0^1([0,h], Q)$). Then

$$\begin{aligned} (C_s + C_g) K_s^n &\geq \left| \mathrm{D}^2\mathfrak{S}(\nu)[\tilde{q}_n - \bar{q}, \tilde{q}_n - \bar{q}] \right| \\ &\geq C_f \|\tilde{q}_n - \bar{q}\|^2_{W^{1,1}([0,h])} \\ C\sqrt{K_s}^n &\geq \|\tilde{q}_n - \bar{q}\|^2_{W^{1,1}([0,h])} \end{aligned}$$

where $C = \frac{C_s + C_g}{C_f}$. □

This result shows that Galerkin curves converge to the true solution geometrically with n-refinement, albeit with a larger geometric constant, and hence a slower rate. By simply replacing the bounds (2.18) and (2.19) from Theorem 2.3.4 with those from Theorem 2.3.3 and the term $C_s K_s^n$ with $C_{op} h^p$, an identical argument shows that Galerkin curves converge at half the optimal rate with h-refinement.

Theorem 2.3.7. (Convergence of Galerkin Curves with h-Refinement) *Under the same assumptions as Theorem 2.3.3, if at $\bar{q}$, the action is twice Frechet differentiable, and* *if the second Frechet derivative of the action* $D^2\mathscr{S}(\cdot)[\cdot,\cdot]$ *is coercive with a constant* C_f *independent of* h *in a neighborhood* U *of* $\bar{q}$*, for all curves* $\delta q \in H_0^1([0,h],Q)$*, then if the discrete Lagrange map has error* $\mathscr{O}\left(h^{p+1}\right)$*, the Galerkin curves have error at most* $\mathscr{O}\left(h^{\frac{p+1}{2}}\right)$ *in* $\|\cdot\|_{W^{1,1}([0,h])}$*. If* C_f *is a function of* h*, this bound becomes* $\mathscr{O}\left(C_f(h)^{-1}h^{\frac{p+1}{2}}\right)$*.*

Like the requirement that the stationary points of the actions are minimizers, the requirement that the second Frechet derivative of the action is coercive may appear quite strong at first. Again, the coercivity will depend on the properties of the Lagrangian L, but we can establish that for Lagrangians of the canonical form,

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q),$$

there exists a time step $[0,h]$ over which the action is coercive on $H_0^1([0,h],Q)$.

Theorem 2.3.8. (Coercivity of the Action) *For Lagrangian of the form*

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q),$$

where M *is symmetric positive-definite, and the second derivatives of* $V(q)$ *are bounded, there exists an interval* $[0,h]$ *over which the action is coercive over* $H_0^1([0,h],Q)$*, that is,*

$$D^2\mathfrak{S}(\nu)[\delta q,\delta q] \geq C_f \|\delta q\|^2_{W^{1,1}([0,h])},$$

for any $\delta q \in H_0^1([0,h],Q)$ *and any* $\nu \in C^2([0,h],Q)$*.*

Proof. First, we note that if

$$\mathfrak{S}(\nu) = \int_0^h \frac{1}{2}\dot{\nu}^T M\dot{\nu} - V(\nu),$$

then

$$\mathrm{D}^2\mathfrak{S}(\nu)[\delta q,\delta q] = \int_0^h \delta\dot{q}^T M\delta\dot{q} - \delta q^T H(\nu)\,\delta q\,\mathrm{d}t$$

$$= \int_0^h \delta\dot{q}^T M \delta\dot{q} \mathrm{d}t - \int_0^h \delta q^T H(\nu) \delta q \mathrm{d}t$$

where $H(\nu)$ is the Hessian of $V(\nu)$ at the point ν. Since M is symmetric positive-definite, and the second derivatives of $V(\cdot)$ are bounded, then there exists C_r and m such that:

$$\begin{aligned} \int_0^h \delta\dot{q}^T M \delta\dot{q} \mathrm{d}t &\geq \int_0^h m \delta\dot{q}^T \delta\dot{q} \mathrm{d}t \\ \int_0^h \delta q^T H(\nu) \delta q \mathrm{d}t &\leq \int_0^h D C_r \delta q^T \delta q \mathrm{d}t, \end{aligned} \tag{2.20}$$

(see (2.17) for a derivation of (2.20)). Hence,

$$\begin{aligned} \mathrm{D}^2 \mathfrak{S}(\nu)[\delta q, \delta q] &\geq \int_0^h m \delta\dot{q}^T \delta\dot{q} \mathrm{d}t - \int_0^h D C_r f^T f \mathrm{d}t \\ &= \frac{1}{2} m \int_0^h \delta\dot{q}^T \delta\dot{q} \mathrm{d}t + \frac{1}{2} m \int_0^h \delta\dot{q}^T \delta\dot{q} \mathrm{d}t - D C_r \int_0^h \delta q^T \delta q \mathrm{d}t. \end{aligned} \tag{2.21}$$

Considering the last two terms in (2.21), and noting that $\delta q \in H_0^1([0,h], Q)$, we make use of the Poincaré inequality to derive:

$$\begin{aligned} \frac{1}{2} m \int_0^h \delta\dot{q}^T \delta\dot{q} \mathrm{d}t - D C_r \int_0^h \delta q^T \delta q \mathrm{d}t &\geq \frac{m\pi^2}{2h^2} \int_0^h \delta q^T \delta q \mathrm{d}t - n C_r \int_0^h \delta q^T \delta q \mathrm{d}t \\ &\geq \left(\frac{m\pi^2}{2h^2} - D C_r \right) \int_0^h \delta q^T \delta q \mathrm{d}t. \end{aligned} \tag{2.22}$$

Thus, substituting (2.22) in for the last two terms of (2.21), we conclude:

$$\begin{aligned} &\mathrm{D}^2 \mathfrak{S}(q, \dot{q})[\delta q, \delta q] \\ &\geq \left(\frac{m\pi^2}{2h^2} - D C_r \right) \int_0^h \delta q^T \delta q \mathrm{d}t + \frac{m}{2} \int_0^h \delta\dot{q}^T \delta\dot{q} \mathrm{d}t \\ &\geq \min \left(\frac{m}{2}, \left(\frac{m\pi^2}{2h^2} - D C_r \right) \right) \left(\int_0^h \delta q^T \delta q \mathrm{d}t + \int_0^h \delta\dot{q}^T \delta\dot{q} \mathrm{d}t \right) \\ &= \min \left(\frac{m}{2}, \left(\frac{m\pi^2}{2h^2} - D C_r \right) \right) \left(\|\delta q\|_{L^2([0,h])}^2 + \|\delta\dot{q}\|_{L^2([0,h])}^2 \right), \end{aligned}$$

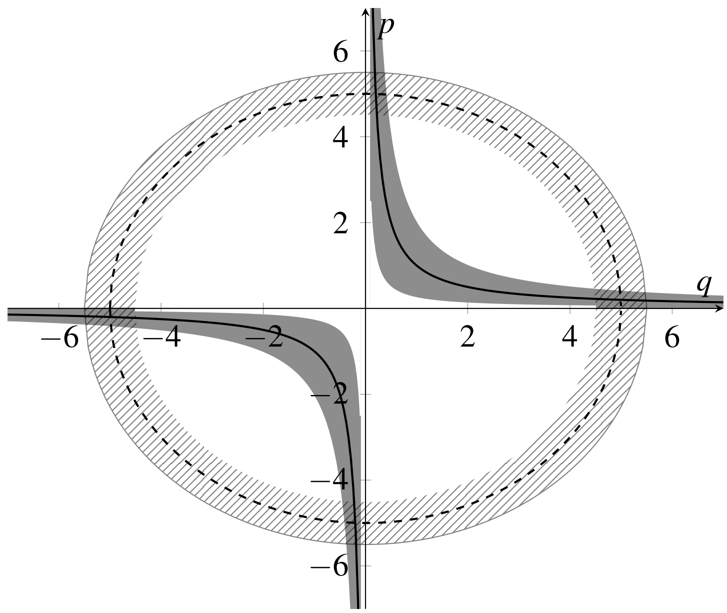


Figure 2.2: Conserved and approximately conserved Noether quantities and the resulting constrained solution space. Suppose that both $p^T q = 1$ and $p^2 + q^2 = 5$ were conserved quantities for a certain Lagrangian. Then the solutions of the Euler-Lagrange equations would be constrained to the intersections of these two constant surfaces in phase space; in the above diagram, this is the intersection of the dashed and solid lines. If these quantities were conserved up to a fixed error along a numerical solution, then the numerical solution would be constrained to the intersection of the shaded regions in the above figure. The constraint of the numerical solution to these regions is what leads to the many excellent qualities of variational integrators.

and making use of Hölder's inequality, we see that $\|\delta q\|_{L^2([0,h])} \geq h^{-\frac{1}{2}} \|\delta q\|_{L^1([0,h])}$, thus

$$\begin{aligned}
&\mathrm{D}^2\mathfrak{S}(q,\dot{q})[\delta q,\delta q] \\
&\geq \min\left(\frac{m}{2},\left(\frac{m\pi^2}{2h^2} - DC_r\right)\right)\left(h^{-1}\|\delta q\|^2_{L^1([0,h])} + h^{-1}\|\delta \dot{q}\|^2_{L^1([0,h])}\right) \\
&\geq \min\left(\frac{m}{2h}, h^{-1}\left(\frac{m\pi^2}{2h^2} - hDC_r\right)\right)\frac{1}{2}\left(\|\delta q\|_{L^1([0,h])} + \|\delta \dot{q}\|_{L^1([0,h])}\right)^2 \\
&= \min\left(\frac{m}{4h}, (2h)^{-1}\left(\frac{m\pi^2}{2h^2} - DC_r\right)\right)\|\delta q\|^2_{W^{1,1}([0,h])}
\end{aligned}$$

which establishes the coercivity result. □

Noether Quantities

One of the great advantages of using variational integrators for problems in geometric mechanics is that by construction they have a rich geometric structure which helps lead to excellent long term and qualitative behavior. An important geometric feature of variational integrators is the preservation of discrete Noether quantities, which are invariants that are derived from symmetries of the action. These are analogous to the more familiar Noether quantities of geometric mechanics in the continuous case. We quickly recall Noether's theorem in both the discrete and continuous case, which will also help define the notation used throughout the proofs that follow. The proofs of both these theorems can be found in Hairer et al. [11].

Theorem 2.3.9. (Noether's Theorem) *Consider a system with Hamiltonian $H(p,q)$ and Lagrangian $L(q,\dot{q})$. Suppose $\{g_s : s \in \mathbb{R}\}$ is a one-parameter group of transformations which leaves the Lagrangian invariant. Let*

$$a(q) = \left.\frac{d}{ds}\right|_{s=0} g_s(q)$$

be defined as the vector field with flow $g_s(q)$, referred to as the infinitesimal generator, and define the canonical momentum

$$p = \frac{\partial L}{\partial \dot{q}}(q,\dot{q}).$$

Then

$$I(p,q) = p^T a(q)$$

is a first integral of the Hamiltonian system.

Theorem 2.3.10. (Discrete Noether's Theorem) *Suppose the one-parameter group of transformations leaves the discrete Lagrangian $L_d(q_k, q_{k+1})$ invariant for all (q_k, q_{k+1}). Then:*

$$p_{k+1}^T a(q_{k+1}) = p_k^T a(q_k)$$

where

$$
\begin{aligned}
p_k &= -D_1 L_d\left(q_k, q_{k+1}\right), \\
p_{k+1} &= D_2 L_d\left(q_k, q_{k+1}\right).
\end{aligned}
$$

For the remainder of this section, we will refer to $I(q,p)$ as the *Noether quantity* and $p_n^T a(q_n) = p_{n+1}^T a(q_{n+1})$ as the *discrete Noether quantity*.

For Galerkin variational integrators, it is possible to bound the error of the Noether quantities along the Galerkin curve from the behavior of the analogous discrete Noether quantities of the discrete problem and, more importantly, this bound is independent of the number of time steps that are taken in the numerical integration. This is significant because it offers insight into the excellent behavior of spectral variational integrators even over long periods of integration.

The proof of convergence and near preservation of Noether quantities is broken into three major parts. First, we note that on step k of a numerical integration the discrete Noether quantity arises from a function of the Galerkin curve and the initial point of the one-step map (q_{k-1}, q_k), and that a bound exists for the difference of this discrete Noether quantity evaluated on the Galerkin curve and evaluated on the local exact solution to the Euler-Lagrange equations $\bar{q}$. Second, we show that a bound exists for the difference of the discrete Noether quantity on the local exact solution of the Euler-Lagrange equations and the value of the Noether quantity of the local exact solution, which is conserved along the flow of the Euler-Lagrange vector field. Finally, we show that under certain smoothness conditions, there exists a point-wise bound between the Noether quantity evaluated on the Galerkin curve and the Noether quantity evaluated on the local exact solution. Thus, we establish a point-wise bound between the Noether quantity evaluated on the Galerkin curve and the discrete Noether quantity, and a bound between the discrete Noether quantity and the Noether quantity, which leads to a point-wise bound between the Noether quantity evaluated on the Galerkin curve, and the Noether quantity which is conserved along the global flow of the Euler-Lagrange vector field.

Throughout this section we will make the simplifying assumptions that

$$\tilde{q}_n = \sum_{i=1}^{n} q_k^i \phi_i$$

where $q_k^1 = q_k$, and thus

$$\frac{\partial \tilde{q}_n}{\partial q_k} = \phi_1.$$

This assumption significantly simplifies the analysis.

We begin by bounding the discrete Noether quantity by a function of the local exact solution of the Euler-Lagrange equations.

Lemma 2.3.3. (Bound on Discrete Noether Quantity) *Define the Galerkin Noether map as:*

$$I_d\left(q\left(t\right), q_k\right) = -\left(h \sum_{j=1}^{n} b_j \left[\frac{\partial L}{\partial q}\left(q, \dot{q}\right) \phi_1 + \frac{\partial L}{\partial \dot{q}}\left(q, \dot{q}\right) \dot{\phi}_1\right]\right)^T a\left(q_k\right)$$

and note that the discrete Noether quantity is given by

$$I_d\left(\tilde{q}_n, q_k\right) = p_n^T a\left(q_k\right).$$

Assuming the quadrature accuracy of Theorem (2.3.4) with n-refinement and Theorem (2.3.3) with h-refinement, if $\frac{\partial L}{\partial q}\left(q, \dot{q}\right)$, $\frac{\partial L}{\partial \dot{q}}\left(q, \dot{q}\right)$ *and* $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}$ *are Lipschitz continuous,* $\|\phi_1\|_{L^\infty([0,h])}$ *is bounded with n refinement, and* $\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])}$ *is bounded below by the quadrature error, then*

$$\begin{aligned} &\left|I_d\left(\tilde{q}_n, q_k\right) - I_d\left(\bar{q}, q_k\right)\right| \\ &\leq C\left|a\left(q_k\right)\right| \left(\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])} + \|\tilde{q}_n - \bar{q}\|_{L^\infty([0,h])} + \left\|\dot{\tilde{q}}_n - \dot{\bar{q}}\right\|_{L^\infty([0,h])}\right) \end{aligned}$$

for some C independent of n and h.

Proof. We begin by expanding the definitions of the discrete Noether quantity:

$$\left|I_d\left(\tilde{q}_n, q_k\right) - I_d\left(\bar{q}, q_k\right)\right|$$

$$= \left| h \left(\sum_{j=1}^{m} b_j \left[\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \phi_1 + \frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \dot{\phi}_1 \right] \right)^T a(q_k) \right.$$
$$\left. - \left(h \sum_{j=1}^{m} b_j \left[\frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \phi_1 + \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \dot{\phi}_1 \right] \right)^T a(q_k) \right|$$
$$= \left| \left(h \sum b_j \left[\left(\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 - \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \dot{\phi}_1 \right] \right)^T a(q_k) \right|$$
$$\leq \left| h \sum_{j=1}^{m} b_j \left[\left(\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 - \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \dot{\phi}_1 \right] \right| |a(q_k)|.$$

Now we introduce the function $e_q(\cdot, \cdot)$ which gives the error of the quadrature rule, and thus

$$|I_d(\tilde{q}_n, q_k) - I_d(\bar{q}, q_k)|$$
$$\leq \left| \int_0^h \left(\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 - \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \dot{\phi}_1 \mathrm{d}t \right.$$
$$\left. + e_q \left(\tilde{q}_n - \bar{q}, \dot{\tilde{q}}_n - \dot{\bar{q}} \right) \right| |a(q_k)|.$$

Integrating by parts, we get:

$$|I_d(\tilde{q}_n, q_k) - I_d(\bar{q}, q_k)|$$
$$\leq \left| \int_0^h \left(\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 \mathrm{d}t \right.$$
$$\left. + \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 \Big|_0^h + e_q \left(\tilde{q}_n - \bar{q}, \dot{\tilde{q}}_n - \dot{\bar{q}} \right) \right| |a(q_k)|.$$

Introducing the Lipschitz constants L_1 for $\frac{\partial L}{\partial q}$, L_2 for $\frac{\partial L}{\partial \dot{q}}$, and L_3 for $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}$,

$$|I_d(\tilde{q}_n, q_k) - I_d(\bar{q}, q_k)|$$
$$\leq \left(\int_0^h (L_1 + L_3) \left| \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \left(\bar{q}, \dot{\bar{q}} \right) \right| |\phi_1| \mathrm{d}t + 2L_2 \left(\|\phi_1\|_{L^\infty([0,h])} \right) \right.$$
$$\left. \left(\|\tilde{q}_n - \bar{q}\|_{L^\infty([0,h])} + \left\|\dot{\tilde{q}}_n - \dot{\bar{q}}\right\|_{L^\infty([0,h])} \right) + e_q \left(\tilde{q}_n - \bar{q}, \dot{\tilde{q}}_n - \dot{\bar{q}} \right) \right) |a(q_k)|$$

$$\leq (L_1+L_3)\,\|\phi_1\|_{L^\infty([0,h])}\,|a(q_k)|\left(\int_0^h \left|\left(\tilde{q}_n,\dot{\tilde{q}}_n\right)-\left(\bar{q},\dot{\bar{q}}\right)\right|\mathrm{d}t\right)$$
$$+2L_2\left(\|\phi_1\|_{L^\infty([0,h])}\right)|a(q_k)|\left(\|\tilde{q}_n-\bar{q}\|_{L^\infty([0,h])}+\left\|\dot{\tilde{q}}_n-\dot{\bar{q}}\right\|_{L^\infty([0,h])}\right)$$
$$+e_q\left(\tilde{q}_n-\bar{q},\dot{\tilde{q}}_n-\dot{\bar{q}}\right)|a(q_k)|\,.$$

We now make the simplification that the quadrature error $\left|e_q(\cdot,\cdot)\right|$ serves as a lower bound for $\|\tilde{q}_n-\bar{q}\|_{W^{1,1}([0,h])}$. While this may not strictly hold, all of our estimates on the convergence for $\tilde{q}_n$ imply this bound, and hence it is a reasonable simplification for establishing convergence in this case. Now, note that $\|\phi_1\|_{L^\infty([0,h])}$ is invariant under h rescaling, and let

$$C=\max(L_1+L_3,2L_2)\,\|\phi_1\|_{L^\infty([0,h])}+1$$

to get

$$|I_d(\tilde{q}_n,q_k)-I_d(\bar{q},q_k)|$$
$$\leq C\,|a(q_k)|\left(\|\tilde{q}_n-\bar{q}\|_{W^{1,1}([0,h])}+\|\tilde{q}_n-\bar{q}\|_{L^\infty([0,h])}+\left\|\dot{\tilde{q}}_n-\dot{\bar{q}}\right\|_{L^\infty([0,h])}\right)$$

which establishes the result. □

Lemma 2.3.3 establishes a bound between the discrete Noether quantity and $I_d(\bar{q},q_k)$. The next step is to establish a bound between $I_d(\bar{q},q_k)$ and the Noether quantity.

Lemma 2.3.4. (Error Between Discrete Noether Quantity and True Noether Quantity) *Assume that $\phi_1(0)=1$ and $\phi_1(h)=0$, and that the sequence $\{|a(q_k)|\}_{k=1}^N$ is bounded by a constant C_a which is independent of N. Let*

$$\bar{p}(t)=\frac{\partial L}{\partial \dot{q}}(\bar{q}(t),\dot{\bar{q}}(t))\,.$$

Once again, let the error of the quadrature rule be given by $e_q(\cdot,\cdot)$. Then

$$\left|I^d(\bar{q},q_k)-I(\bar{p}(t),\bar{q}(t))\right|\leq C_a\left|e_q(\bar{q},\dot{\bar{q}})\right|$$

for any $t \in [0,h]$.

Proof. We note that since $\bar{q}$ solves the Euler-Lagrange equations exactly, $I(\bar{p}(t), \bar{q}(t))$ is a conserved quantity along the flow, so it suffices to show the inequality holds for $t = 0$. We begin by expanding:

$$\begin{aligned}
&\left|I^d(\bar{q}, q_k) - I(\bar{p}(0), \bar{q}(0))\right| \\
&= \left|-h\left(\sum_{j=1}^{m} b_j \frac{\partial L}{\partial q}(\bar{q}, \dot{\bar{q}})\phi_1 + \frac{\partial L}{\partial \dot{q}}(\bar{q}, \dot{\bar{q}})\dot{\phi}_1\right)^T a(q_k) - \bar{p}(0)^T a(\bar{q}(0))\right| \\
&= \left|-\left(\int_0^h \frac{\partial L}{\partial q}(\bar{q}, \dot{\bar{q}})\phi_1 + \frac{\partial L}{\partial \dot{q}}(\bar{q}, \dot{\bar{q}})\dot{\phi}_1 \mathrm{d}t + e_q(\bar{q}, \dot{\bar{q}})\right)^T a(q_k) - \bar{p}(0)^T a(\bar{q}(0))\right| \\
&= \left|-\left(\int_0^h \left(\frac{\partial L}{\partial q}(\bar{q}, \dot{\bar{q}}) - \frac{d}{d\mathrm{t}}\frac{\partial L}{\partial \dot{q}}(\bar{q}, \dot{\bar{q}})\right)\phi_1 \mathrm{d}t + \frac{\partial L}{\partial \dot{q}}(\bar{q}(h), \dot{\bar{q}}(h))\phi_1(h)\right.\right. \\
&\qquad \left.\left. - \frac{\partial L}{\partial \dot{q}}(\bar{q}(0), \dot{\bar{q}}(0))\phi_1(0) + e_q(\bar{q}, \dot{\bar{q}})\right)^T a(q_k) - \bar{p}(0)^T a(\bar{q}(0))\right|
\end{aligned}$$

Since $\bar{q}(t)$ solves the Euler-Lagrange equations, $\phi_1(0) = 1$ and $\phi_1(h) = 0$, and $\bar{q}(0) = q_k$,

$$\begin{aligned}
&\left|I^d(\bar{q}, q_k) - I(\bar{p}(0), \bar{q}(0))\right| \\
&= \left|\left(\frac{\partial L}{\partial \dot{q}}(\bar{q}(0), \dot{\bar{q}}(0))\right)^T a(q_k) + \left(e_q(\bar{q}, \dot{\bar{q}})\right)^T a(q_k) - \bar{p}(0)^T a(q_k)\right| \\
&= \left|(\bar{p}(0))^T a(q_k) + \left(e_q(\bar{q}, \dot{\bar{q}})\right)^T a(q_k) - (\bar{p}(0))^T a(q_k)\right| \\
&= \left|e_q(\bar{q}, \dot{\bar{q}})^T a(q_k)\right| \\
&\leq \left|e_q(\bar{q}, \dot{\bar{q}})\right| |a(q_k)| \\
&\leq C_a \left|e_q(\bar{q}, \dot{\bar{q}})\right|
\end{aligned}$$

which yields the desired bound. □

Once again, if we assume that the quadrature error serves as a lower bound for the Sobolev error, combining the bounds from (2.3.3) and (2.3.4) yields:

$$|I_d(\tilde{q}_n, q_k) - I(\bar{p}(t), \bar{q}(t))|$$

$$\leq 2CC_a \left(\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])} + \|\tilde{q}_n - \bar{q}\|_{L^\infty([0,h])} + \left\|\dot{\tilde{q}}_n - \dot{\bar{q}}\right\|_{L^\infty([0,h])} \right).$$

This bound serves two purposes; the first is to establish a bound between the discrete Noether quantity and the Noether quantity computed on the local exact solution $\bar{q}$. The second is to establish a bound between the discrete Noether quantity after one step and the Noether quantity computed on the initial data:

$$\begin{aligned} &|I_d(\tilde{q}_n, q_1) - I(p(0), q(0))| \\ &\leq 2CC_a \left(\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])} + \|\tilde{q}_n - \bar{q}\|_{L^\infty([0,h])} + \left\|\dot{\tilde{q}}_n - \dot{\bar{q}}\right\|_{L^\infty([0,h])} \right), \end{aligned}$$

since for (q_1, q_2), $\bar{q}$ is the *global* exact flow of the Euler-Lagrange equations.

The difference between these two bounds is subtle but important; by establishing a bound between the discrete Noether quantity and the Noether quantity associated with the initial conditions, on any step of the method we can bound the error between the discrete Noether quantity and the Noether quantity associated with the *global* exact flow. By establishing the bound between the discrete Noether quantity and the Noether quantity associated with $\bar{q}$ at any step, we can bound the error between the Noether quantity associated with the local exact flow $\bar{q}$ and the true Noether quantity conserved along the global exact flow:

$$\begin{aligned} &|I(\bar{p}(t), \bar{q}(t)) - I(p(0), q(0))| \\ &\leq |I(\bar{p}(t), \bar{q}(t)) - I_d(\tilde{q}_n, q_k)| + |I_d(\tilde{q}_n, q_k) - I(p(0), q(0))| \\ &\leq 4CC_a \left(\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])} + \|\tilde{q}_n - \bar{q}\|_{L^\infty([0,h])} + \left\|\dot{\tilde{q}}_n - \dot{\bar{q}}\right\|_{L^\infty([0,h])} \right) \end{aligned} \tag{2.23}$$

for any $t_0 \in [0,h]$ on any time step k. Because the local exact flow $\bar{q}$ is generated from boundary conditions (q_k, q_{k+1}) which only approximate the boundary conditions of the true flow, there is no guarantee that the Noether quantity associated with $\bar{q}$ will be the same step to step, only that it will be conserved within each time step. However, because there is a bound between the Noether quantity associated with $\bar{q}$ and the discrete Noether quantity at every time step, the discrete Noether quantity and the Noether quantity associated with the exact flow, and because the Noether quantity is conserved point-wise along $\bar{q}$ on each time step, there exists a bound between the Noether quantity associated with each point of the local exact flow and the Noether quantity associated with the true solution.

We finally arrive at the desired result, which is a theorem that bounds the error between the Noether quantity along the Galerkin curve and the true Noether quantity. It is significant because not only does it bound the error of the Noether quantity, but the bound is independent of the number of steps taken, and hence will not grow even for extremely long numerical integrations.

Theorem 2.3.11. (Convergence of Conserved Noether Quantities) *Define*

$$\tilde{p}_n = \frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n\right).$$

Under the assumptions of Lemmas (2.3.3 - 2.3.4), if the Noether map $I(p,q)$ *is Lipschitz continuous in both its arguments, then there exists a constant* C_v *independent N, the number of method steps, such that:*

$$\begin{aligned}&\left|I(p(0),q(0)) - I(\tilde{p}_n(t),\tilde{q}_n(t))\right| \\ &\leq C_v \left(\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])} + \|\tilde{q}_n - \bar{q}\|_{L^\infty([0,h])} + \left\|\dot{\tilde{q}}_n - \dot{\bar{q}}\right\|_{L^\infty([0,h])}\right).\end{aligned}$$

for any $t \in [0,Nh]$.

Proof. We begin by introducing the Noether quantity evaluated at t on the local exact flow, $\bar{q}$:

$$\begin{aligned}&\left|I(p(0),q(0)) - I(\tilde{p}_n(t),\tilde{q}_n(t))\right| \\ &\leq \left|I(\tilde{p}_n(t),\tilde{q}_n(t)) - I(\bar{p}(t),\bar{q}(t))\right| + \left|I(\bar{p}(t),\bar{q}(t)) - I(p(0),q(0))\right|.\end{aligned} \tag{2.24}$$

Considering the first term in (2.24), let L_4 be the Lipschitz constant for $I(\cdot,\cdot)$. Then

$$\begin{aligned}&\left|I(\tilde{p}_n(t),\tilde{q}_n(t)) - I(\bar{p}(t),\bar{q}(t))\right| \\ &\leq L_4 \left|(\tilde{p}_n(t),\tilde{q}_n(t)) - (\bar{p}(t),\bar{q}(t))\right| \\ &\leq L_4 \left(\left|\tilde{p}_n(t) - \bar{p}(t)\right| + \left|\tilde{q}_n(t) - \bar{p}(t)\right|\right) \\ &= L_4 \left(\left|\frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n(t),\dot{\tilde{q}}_n(t)\right) - \frac{\partial L}{\partial \dot{q}}\left(\bar{q}(t),\dot{\bar{q}}(t)\right)\right| + \left|\tilde{q}_n(t) - \bar{q}(t)\right|\right)\end{aligned}$$

$$\leq L_4\left(L_2\left|\dot{\tilde{q}}_n(t)-\dot{\bar{q}}(t)\right|+(L_2+1)\left|\tilde{q}_n(t)-\bar{q}(t)\right|\right)$$
$$\leq L_4(L_2+1)\left(\|\tilde{q}_n-\bar{q}\|_{L^\infty([0,h])}+\left\|\dot{\tilde{q}}_n-\dot{\bar{q}}\right\|_{L^\infty([0,h])}\right). \tag{2.25}$$

The second term in (2.24) is exactly the bound given by (2.23) and thus combining (2.25) and (2.23) in (2.24) and defining $C_v=4CC_a+L(L_2+1)$, we have:

$$\left|I(\tilde{p}_n(t),\tilde{q}_n(t))-I(\bar{p}(t),\bar{q}(t))\right|$$
$$\leq C_v\left(\|\tilde{q}_n-\bar{q}\|_{W^{1,1}([0,h])}+\|\tilde{q}_n-\bar{q}\|_{L^\infty([0,h])}+\left\|\dot{\tilde{q}}_n-\dot{\bar{q}}\right\|_{L^\infty([0,h])}\right)$$

which completes the result. □

The convergence of the Noether quantity evaluated on the Galerkin curve to that of the true solution is hampered by one issue. While Theorems (2.3.6) and (2.3.7) provide estimates for convergence in the Sobolev norm $\|\cdot\|_{W^{1,1}([0,h])}$, Theorem (2.3.11) requires estimates in the L^∞ norm. We establish a bound for $\|\tilde{q}_n(t)-\bar{q}(t)\|_{L^\infty([0,h])}$, but it is much more difficult to establish a general estimate for $\left\|\dot{\tilde{q}}_n(t)-\dot{\bar{q}}(t)\right\|_{L^\infty([0,h])}$.

Lemma 2.3.5. (Bound on L^∞ Norm from Sobolev Norm) *For any $t\in[0,h]$, the following bound holds:*

$$|q(t)|\leq\max\left(\frac{1}{h},1\right)\|q\|_{W^{1,1}([0,h])}$$

and thus

$$\|q\|_{L^\infty([0,h])}\leq\max\left(\frac{1}{h},1\right)\|q\|_{W^{1,1}([0,h])}.$$

Proof. This is a basic extension of the arguments from Lemma A.1. in Larsson and Thomée [15], generalizing the lemma from the interval $[0,1]$ to an interval of arbitrary length, $[0,h]$. We note that for any $t,s\in[0,h]$, $q(t)=q(s)+\int_s^t\dot{q}(u)\,\mathrm{d}u$. Thus:

$$|q(t)|\leq|q(s)|+\int_0^h|\dot{q}(u)|\,\mathrm{d}u$$
$$\leq|q(s)|+\|\dot{q}\|_{L^1([0,h])}.$$

Now, we integrate with respect to s:

$$\int_0^h |q(t)|\,\mathrm{d}s \leq \int_0^h |q(s)|\,\mathrm{d}s + \int_0^h \|\dot{q}\|_{L^1([0,h])}\,\mathrm{d}s$$
$$h\,|q(t)| \leq \left(\|q\|_{L^1([0,h])} + h\,\|\dot{q}\|_{L^1([0,h])}\right).$$

which yields the desired result. □

Under certain assumptions about the behavior of $\dot{\tilde{q}}_n - \dot{\bar{q}}$, it is possible to establish bounds on the point-wise error of $\dot{\tilde{q}}_n$ from the Sobolev error $\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])}$. For example, if the length of time that the error is within a given fraction of the max error is proportional to the length of the interval $[0,h]$, i.e. there exists C_1, C_2 independent of h: i.e.,

$$m\left(\left\{t \,\middle|\, \left\|\left(\dot{\tilde{q}}_n(t) - \dot{\bar{q}}(t)\right)\right\| \geq C_1 \left\|\dot{\tilde{q}}_n(t) - \dot{\bar{q}}(t)\right\|_\infty \right\}\right) \geq C_2 h,$$

where m is the Lebesgue measure, then it can easily be seen that:

$$\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])} \geq \int_0^h \left\|\dot{\tilde{q}}_n(t) - \dot{\bar{q}}(t)\right\|\mathrm{d}t \geq C_1 C_2 h \left\|\dot{\tilde{q}}_n - \dot{\bar{q}}\right\|_{L^\infty([0,h])}.$$

While we will not establish here that the $\dot{\tilde{q}}_n$ converges in the L^∞ norm with the same rate that the Galerkin curve converges in the Sobolev norm, our numerical experiments will show that the Noether quantities tend to converge at the same rate as the Galerkin curve.

2.4 Numerical Experiments

To support the results in this paper, as well as to investigate the efficiency and stability of spectral variational integrators, several numerical experiments were conducted by applying spectral variational techniques to well-known variational problems. For each problem, the spectral variational integrator was constructed using a Lagrange interpolation polynomials at n Chebyshev points with a Gauss quadrature rule at $2n$ points. Convergence of both the one-step map and the Galerkin curves was measured using the ℓ^∞ and L^∞ norms respectively, although we record them on the same axis using labeled

L^∞ error in a slight abuse of notation. The experiments strongly support the results of this paper, and suggest topics for further investigation.

2.4.1 Harmonic Oscillator

The first and simplest numerical experiment conducted was the harmonic oscillator. Starting from the Lagrangian,

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

where $q \in \mathbb{R}$, the induced spectral variational discrete Euler-Lagrange equations are linear. Choosing the large time step $h = 20$ over 100 steps yields the expected geometric convergence, and attains very high accuracy, as can be seen in Figure 2.3. In addition, the max error of the energy converges geometrically, see Figure 2.4, and does not grow over the time of integration, see Figure 2.5.

2.4.2 N-body Problems

We now turn our attention towards Kepler N-body problems, which are both good benchmark problems and are interesting in their own right. The general form of the Lagrangian for these problems is

$$L(q,\dot{q}) = \frac{1}{2}\sum_{i=1}^{N}\dot{q}_i^T M\dot{q}_i + G\sum_{i=1}^{N}\sum_{j=0}^{i-1}\frac{m_i m_j}{\|q_i - q_j\|},$$

where $q_i \in \mathbb{R}^D$ is the center of mass for body i, G is a gravitational constant, and m_i is a mass constant associated with the body described by q_i.

2-Body Problem

The first experiment we will examine is the choice of parameters $D = 2$, $m_1 = m_2 = 1$. Centering the coordinate system at q_1, we choose $q_2(0) = (0.4, 0)$, $\dot{q}_2(0) = (0, 2)$, which has a known closed form solution which is a stable closed elliptical orbit with eccentricity 0.6. Knowing the closed form solution allows us to examine the rate of

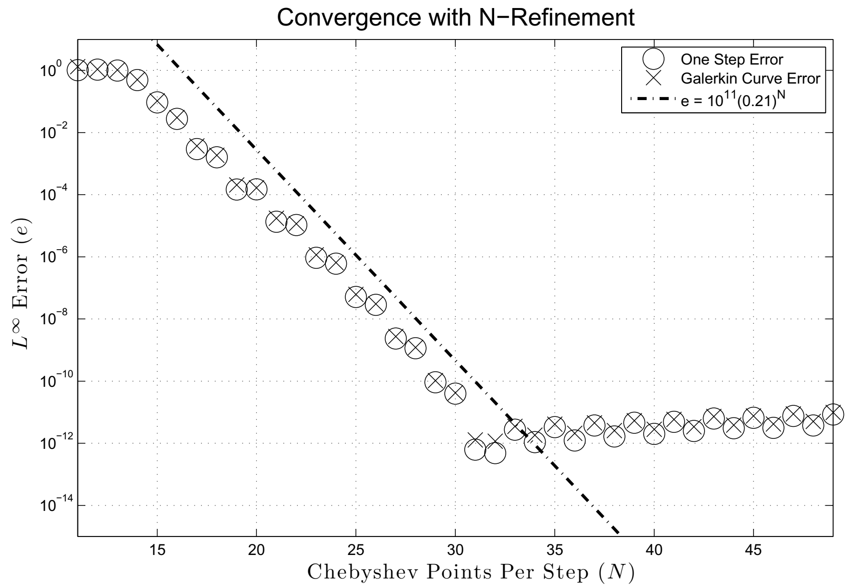


Figure 2.3: Geometric convergence of the spectral variational integration of the harmonic oscillator for 100 steps at step size $h = 20.0$.

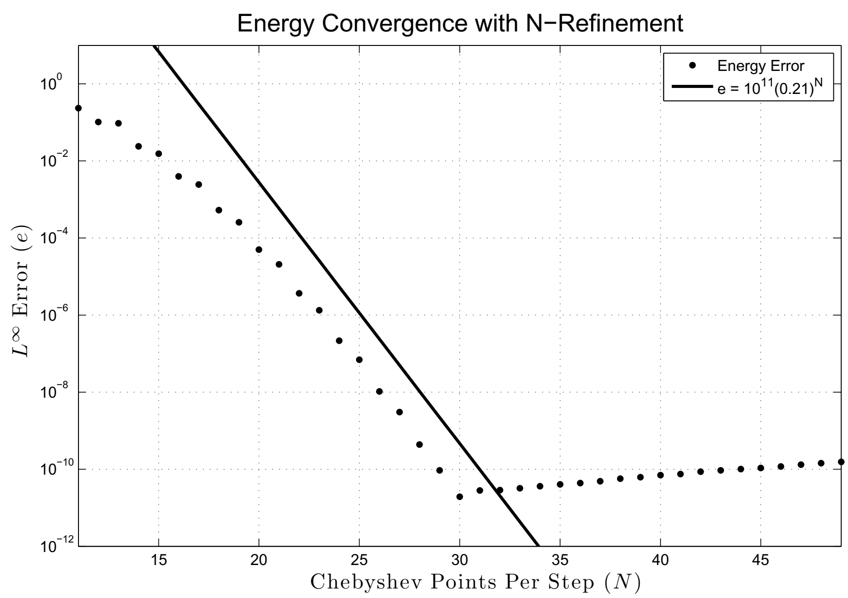


Figure 2.4: Geometric convergence of the energy error of the spectral variational integration of the harmonic oscillator for 100 steps at step size $h = 20.0$.

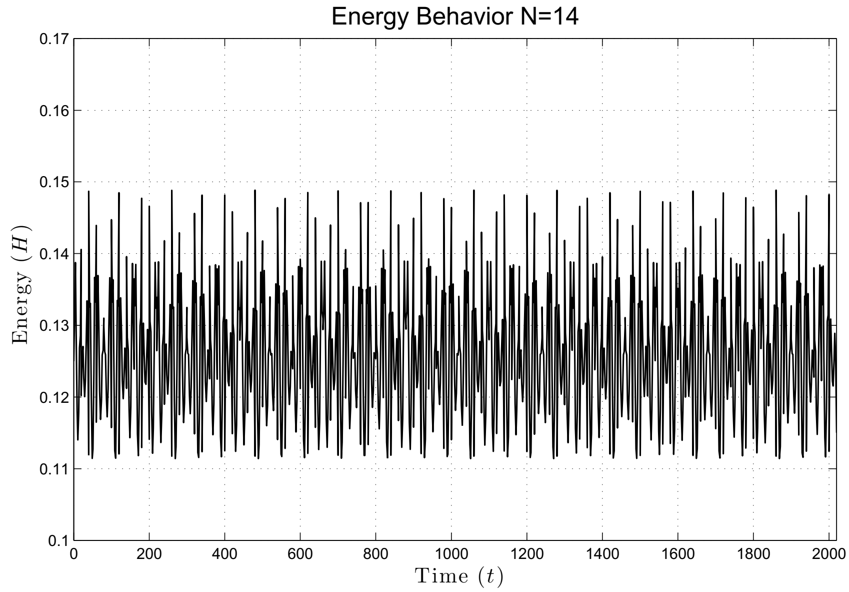


Figure 2.5: Energy Stability of the spectral variational integration of the harmonic oscillator. This energy was computed for the integration using $n = 14$ for step size $h = 20.0$.

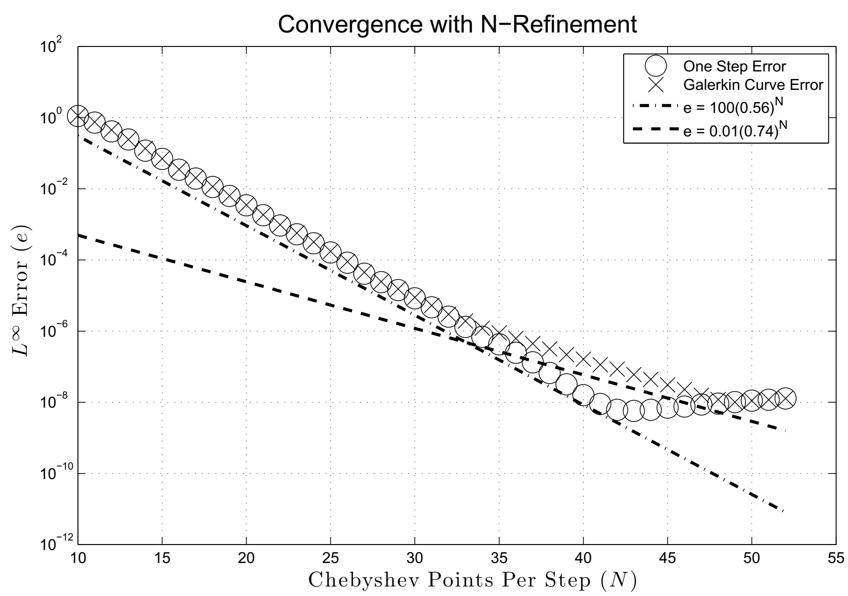


Figure 2.6: Geometric convergence of the Kepler 2-body problem with eccentricity 0.6 over 100 steps of $h = 2.0$. Note that around $n = 32$, the error for the Galerkin curves becomes $\mathscr{O}(0.74^n)$, while the error for the one-step map is always $\mathscr{O}(0.56^n)$.

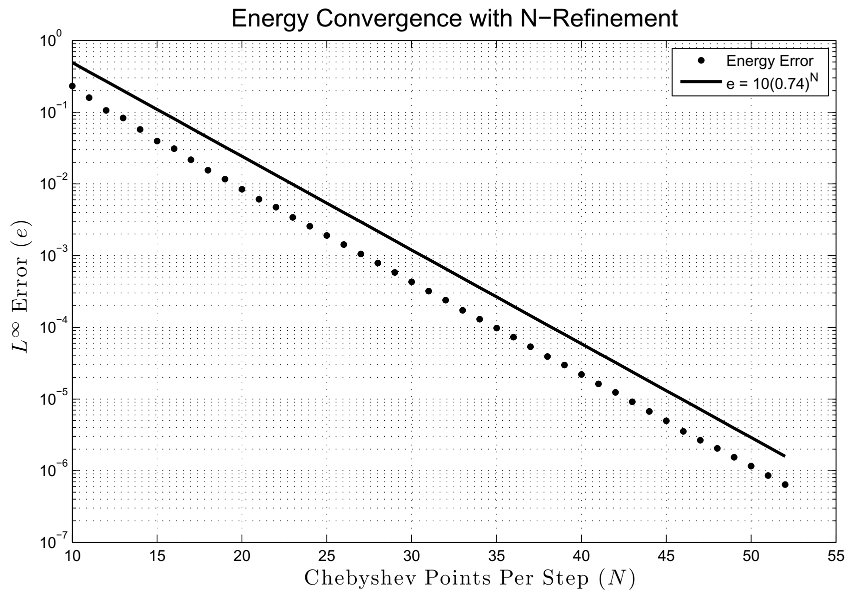


Figure 2.7: Geometric convergence of the energy error of Kepler 2-body problem with eccentricity 0.6 over 100 steps of $h = 2.0$. Note that the error is $\mathcal{O}(0.74^n)$, the same as it was for the Galerkin curves.

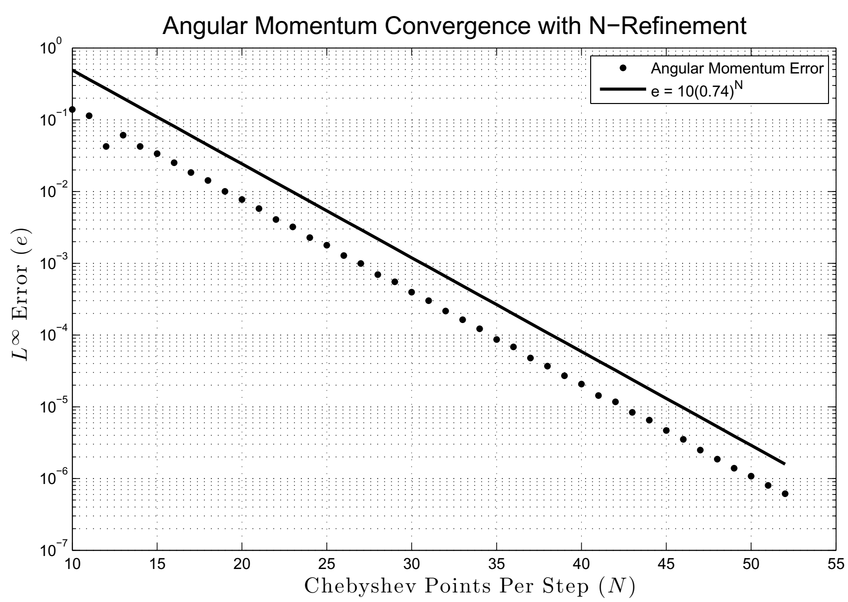


Figure 2.8: Geometric convergence of the angular momentum of the Kepler 2-body problem with eccentricity 0.6 over 100 steps of $h = 2.0$. Again, the error is of the same order as it was the Galerkin curves.

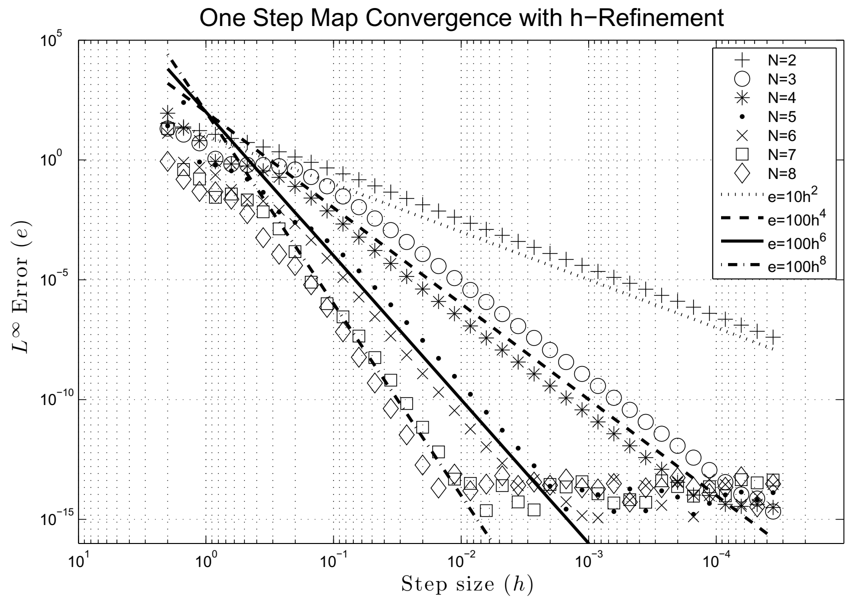


Figure 2.9: Order optimal convergence of the Kepler 2-body problem with eccentricity 0.6 over 10 steps with h refinement. Note our bound is not sharp, as the error is $\mathscr{O}\left(h^{2\left\lceil \frac{n}{2} \right\rceil}\right)$, where $\lceil \cdot \rceil$ is the ceiling function.

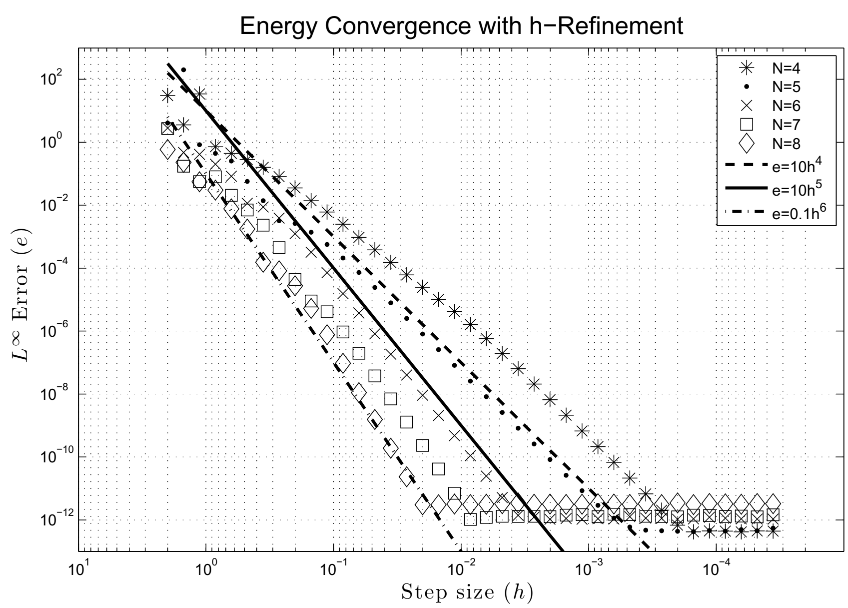


Figure 2.10: Convergence of the Kepler 2-body problem energy with eccentricity 0.6 over 10 steps with h refinement.

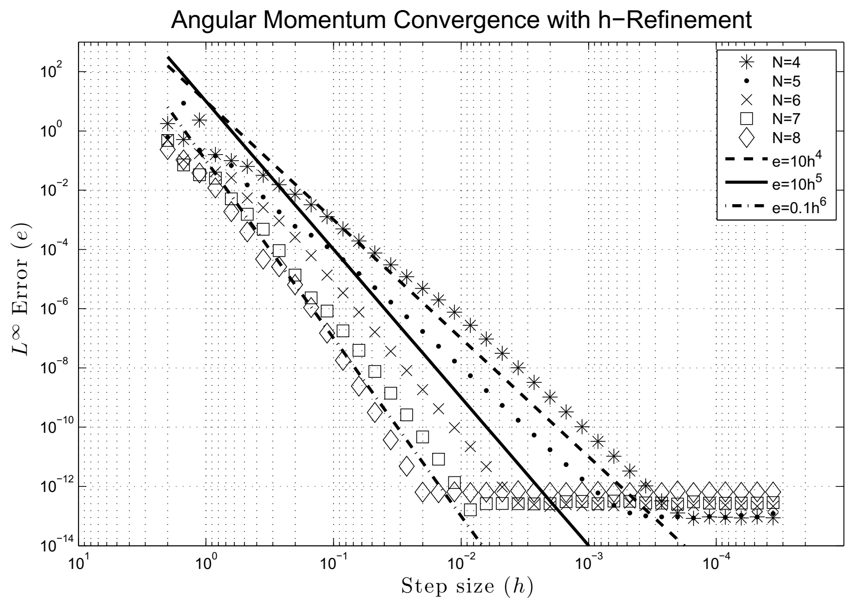


Figure 2.11: Convergence of the angular momentum Kepler 2-body problem with eccentricity 0.6 over 100 steps with h refinement.

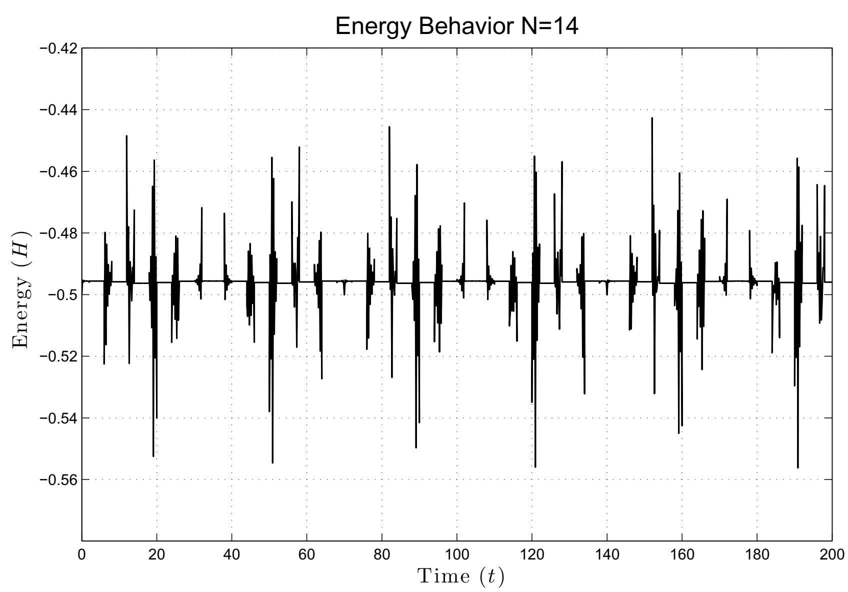


Figure 2.12: Stability of energy for Kepler 2-body problem. This solution was computed with parameters $n = 8$ and $h = 2.0$.

convergence to the true solution, and when solved with the large time step $h = 2.0$, over 100 steps, the error of the one-step map is $\mathscr{O}(0.56^n)$ with n-refinement and $\mathscr{O}\left(h^{2\lceil \frac{n}{2} \rceil}\right)$ with h-refinement, as can be seen in Figure 2.6 and Figure 2.9, respectively. The numerical evidence suggests that our bound for the one-step map with h-refinement is not sharp, as the convergence of the one-step map is always even. Interestingly, it is also possible to observe the different convergence rates of the one-step map and the Galerkin curves with n-refinement, as eventually the Galerkin curves have error approximately $\mathscr{O}(0.74^n)$ while the one-step map has error approximately $\mathscr{O}(0.56^n)$, and $\sqrt{0.56} \approx 0.7483$. However, it appears that the error from the one step map dominates until very high choices of n, and thus it is difficult to observe the error of the Galerkin curves directly with h-refinement, round off error becomes a problem before the error of the Galerkin curves does for smaller choices of n.

The N-body Lagrangian is invariant under the action of $\mathrm{SO}(D)$, which yields the conserved Noether quantity of angular momentum. For the two body problem this is:

$$L(q, \dot{q}) = q_x \dot{q}_y - q_y \dot{q}_x$$

where $q = (q_x, q_y)$. Numerical experiments show that the error of the angular momentum does not grow with the number of steps taken in the integration, Figure 2.13, but that the error is of the same order as the error of Galerkin curve with n-refinement in Figure 2.8. With h-refinement, the angular momentum appears to have error $\mathscr{O}\left(h^{\frac{n}{2}} + 2\right)$ in Figure 2.11. This is interesting because the theoretical bound on the error of the Galerkin curves is $\mathscr{O}\left(h^{\frac{n}{2}}\right)$, and the error of the Noether quantities is theoretically a factor $C(h)$ times the error of the Galerkin curves, where C is the factor that arises in the proof of the convergence of the conserved Noether quantities. Numerical experiments suggest C is $\mathscr{O}\left(h^2\right)$ for this problem, but that the Galerkin curves do converge at a rate of $\mathscr{O}\left(h^{\frac{n}{2}}\right)$, which is consistent with of the Galerkin curve error estimate. While this evidence is not conclusive, it is suggestive that the error analysis provides a plausible bound. A careful analysis of the factor C would be an interesting direction for further investigation.

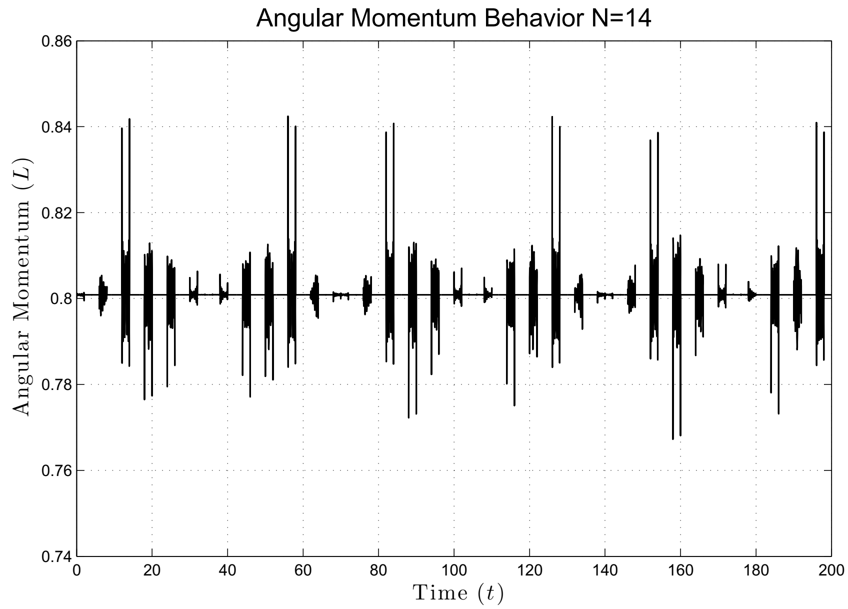


Figure 2.13: Stability of angular momentum for the Kepler 2-body problem. This solution was computed with parameters $n = 8$ and $h = 2.0$.

The Solar System

To illustrate the excellent stability proprieties of spectral variational integrators, we let $D = 3$, $N = 10$, and use the velocities, positions, and masses of the sun, 8 planet, and the dwarf planet Pluto on January 1, 2000 (as provided by the JPL Solar System ephemeris [27]) as initial configuration parameters for the Kepler system. Taking 100 time steps of $h = 100$ days, the $N = 25$ spectral variational integrator produces a highly stable flow in Figure 2.14. It should be noted that orbits are closed, stable, and exhibit almost none of the "precession" effects that are characteristic of symplectic integrators, even though the time step is larger than the orbital period of Mercury. Additionally, considering just the outer solar system (Jupiter, Saturn, Uranus, Neptune, Pluto), and aggregating the inner solar system (Sun, Mercury, Venus, Earth, Mars) to a point mass, an $N = 25$ spectral variational integrator taking 100 time steps $h = 1825$ days (5 year steps) produces the orbital flow seen in Figure 2.15. Again, these are highly stable, precession free orbits. As can be clearly seen, the spectral variational integrator produces extremely stable flows, even for very large time steps.

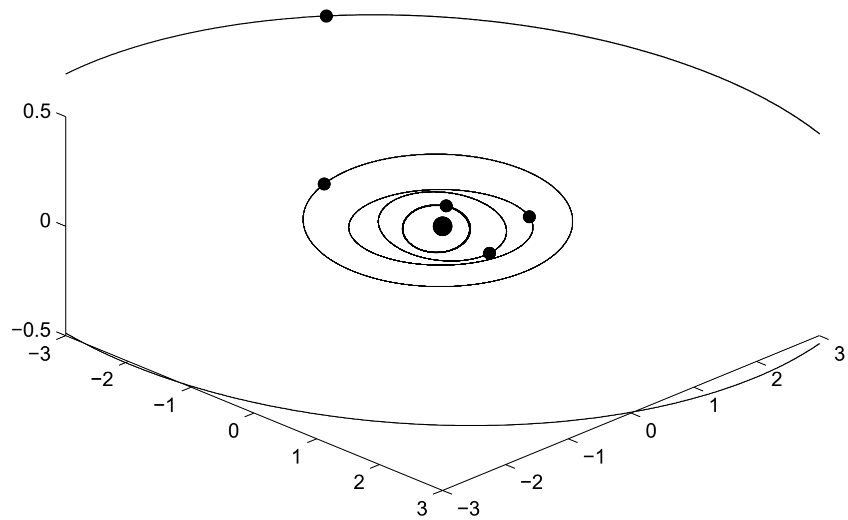


Figure 2.14: Orbital diagram for the inner Solar System produced by the spectral variational integrator using all 8 planets, the Sun, and Pluto with 100 time steps with $h = 100$ days.

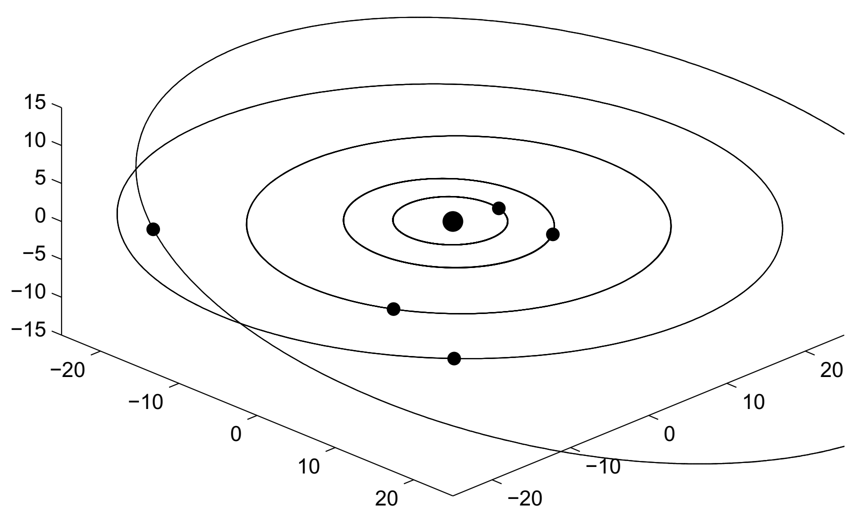


Figure 2.15: Orbital diagram for the outer Solar System produced by the spectral variational integrator using the 4 outer planets, Pluto, with the Sun and 4 inner planets aggregated to a point with 100 time steps at $h = 1825$ days.

2.5 Conclusions and Future Work

In this paper a new numerical method for variational problems was introduced, specifically a symplectic momentum-preserving integrator that exhibits geometric convergence to the true flow of a system under the appropriate conditions. These integrators were constructed under the general framework of Galerkin variational integrators, and made use of the global function paradigm common to many different spectral methods.

Additionally, a general convergence theorem was established for Galerkin type variational integrators, establishing the important result that, under suitable hypotheses, Galerkin variational integrators will inherit the optimal order of convergence permitted by the underlying approximation space used in their construction. This result provides a powerful tool for both constructing and analyzing variational integrators, it provides a methodology for constructing methods of very high order of accuracy, and it also establishes order of convergence for methods that can be viewed as Galerkin variational integrators. It was shown that from the one step map, a continuous approximation to the solution of the Euler-Lagrange equations can be easily recovered over each time step. The error of these continuous approximations was shown to be related to the error of the one step map. Furthermore, the Noether quantities along this continuous approximation approximate the true Noether quantity up to a small error which does not grow with the number of steps taken. It was also shown that the error of the Noether quantities converges to zero with n or h refinement at a predictable rate.

In addition to the convergence results, another interesting feature of spectral variational integrators is the construction of very high order methods that remain stable and accurate using time steps that are orders of magnitude larger than can be tolerated by traditional integrators. The trade off is that the computational effort required to compute each time step is also orders of magnitude larger than that of other methods, which are a major trade off in terms of the practicality of spectral variational integrators. However, a mitigating factor of this trade off is that the approach of solving a short sequence of large problems, as opposed to a large sequence of small problems, lends itself much better to parallelization and computational acceleration. The literature on methods for acceleration of the construction and solution of structured systems of linear and nonlinear problems for PDE problems is extensive, and it is likely that such methods could be applied to spectral variational integrators to greatly improve their computation cost.

2.5.1 Future Work

Future directions for this work are numerous. Because of generality of the construction of Galerkin variational integrators, there exists many possible directions of further exploration.

Lie Group Spectral Variational Integrators

Following the approach of Leok and Shingel [20] or Bou-Rabee and Marsden [1], it is relatively straight forward to extend spectral variational integrators to Lie groups using natural charts. A systematic investigation of the resulting Lie group methods, including convergence and near conservation of Noether quantities, would be a natural extension of the work done here.

Novel Variational Integrators

The power of the Galerkin variational framework is its high flexibility in the choice of approximation spaces and quadrature rules used to construct numerical methods. This flexibility allows for the construction of novel methods specifically tailored to certain applications. One immediate example is the use of periodic functions to construct methods for detecting choreographies in Kepler problems, which would be closely related to methods already used with great success to detect choreographies. Another interesting application would be the use of high order polynomials to develop integrators for high-order Lie group problems, such as the construction of Riemannian splines, which has a variety of applications in motion planning. Enriching traditional polynomial approximation spaces with highly oscillatory functions could be used to develop methods for problems with dynamics evolving on radically different time scales, which are also very challenging for traditional numerical methods.

Multisymplectic Variational Integrators

Multisymplectic geometry (see Marsden et al. [24]) has become an increasingly popular framework for extending much of the geometric theory from classical Lagrangian mechanics to Lagrangian PDEs. The foundations for a discrete theory have been laid, and there have been significant results achieved in geometric techniques for structured problems such as elasticity, fluid mechanics, non-linear wave equations, and computational electromagnetism. However, there is still significant work to be done in the areas of construction of numerical methods, analysis of discrete geometric structure, and especially error analysis. Galerkin type methods have become a standard method in classical numerical PDE methods, such as Finite-Element Methods, Spectral, and Pseudospectral methods. The variational Galerkin framework could provide a natural framework for extending these classical methods to structure preserving geometric methods for PDEs, and the analysis of such methods will rely on the notion of the boundary Lagrangian (see Vankerschaver et al. [30]), which is the PDE analogue of the exact discrete Lagrangian.

2.6 Appendix

2.6.1 Proofs of Geometric Convergence of Spectral Variational Integrators

As stated in §2.3.2, it can be shown that spectral variational integrators converge geometrically to the true flow associated with a Lagrangian under the appropriate conditions. The proof is similar to that of order optimality, and is offered below.

However, before we offer a proof of the theorem, we must establish a result that extends Theorem 2.1.1. Specifically, we must show:

Theorem 2.6.1. (Extension of Theorem 2.1.1 to Geometric Convergence) *Given a regular Lagrangian L and corresponding Hamiltonian H, the following are equivalent for a discrete Lagrangian $L_d(q_0, q_1, n)$:*

1. *there exist a positive constant K, where $K < 1$, such that the discrete Hamiltonian map for $L_d(q_0, q_h, n)$ has error $\mathscr{O}(K^n)$,*

2. *there exists a positive constant K, where* $K<1$*, such that the discrete Legendre transforms of* $L_d(q_0,q_h,n)$ *have error* $\mathscr{O}(K^n)$*,*

3. *there exists a positive constant K, where* $K<1$*, such that* $L_d(q_0,q_h,n)$ *approximates the exact discrete Lagrangian* $L_d^E(q_0,q_h,h)$ *with error* $\mathscr{O}(K^n)$*.*

The proof of this theorem is a simple modification of the proof of Theorem 2.1.1, and is included here for completeness. For details, the interested reader is referred to Marsden and West [23].

Proof. Since we are assuming that the time step h is being held constant, we will suppress it as an argument to the exact discrete Lagrangian, writing $L_d^E(q_0,q_h)$ for $L_d^E(q_0,q_h,h)$. First, we will assume that $L_d(q_0,q_h,n)$ approximates $L_d(q_0,q_h)$ with error $\mathscr{O}(K^n)$ and show this implies the discrete Legendre transforms have error $\mathscr{O}(K^n)$. By assumption, if $L_d(q_0,q_h,n)$ has error $\mathscr{O}(K^n)$, there exists a function which is smooth in its first two arguments $e_v: Q\times Q\times\mathbb{N}\to\mathbb{R}$ such that:

$$L_d(q_0,q_h,n)=L_d^E(q_0,q_h)+K^n e_v(q_0,q_h,n),$$

with $|e_v(q_0,q_h,n)|\leq C_v$ on U, some compact subset of Q. Taking derivatives with respect to the first argument yields:

$$\mathbb{F}^-L_d^n(q_0,q_h)=\mathbb{F}^-L_d^E(q_0,q_h)+K^nD_1e_v(q_0,q_h,n),$$

and with respect to the second yields:

$$\mathbb{F}^+L_d^n(q_0,q_h)=\mathbb{F}^+L_d^E(q_0,q_h)+K^nD_2e_v(q_0,q_h,n).$$

Since e_v is smooth and bounded over the compact set U, so are D_1e_v and D_2e_v, yielding that the discrete Legendre transforms have error $\mathscr{O}(K^n)$. Now, to show that if the discrete Legendre transforms have error $\mathscr{O}(K^n)$, the discrete Lagrangian has error $\mathscr{O}(K^n)$, we write:

$$e_v(q_0,q_h,n)=\frac{1}{K^n}\left[L_d(q_0,q_h,n)-L_d^E(q_0,q_h)\right],$$

$$D_1 e_v(q_0, q_h, n) = \frac{1}{K^n}\left[\mathbb{F}^- L_d(q_0, q_h, n) - \mathbb{F}^- L_d^E(q_0, q_h)\right],$$
$$D_2 e_v(q_0, q_h, n) = \frac{1}{K^n}\left[\mathbb{F}^+ L_d(q_0, q_h, n) - \mathbb{F}^+ L_d^E(q_0, q_h)\right].$$

Since $D_1 e_v$ and $D_2 e_v$ are smooth and bounded on a compact set, this implies there exists a function $d(n)$ such that

$$\|e_v(q(0), q(h), n) - d(n)\| \leq C_v,$$

for some constant C_v. This shows that the discrete Lagrangian is equivalent to a discrete Lagrangian with error $\mathscr{O}(K^n)$. We note that the equivalence is a consequence of the fact that one can add a function of h or n to any discrete Lagrangian and the resulting discrete Euler-Lagrange equations and discrete Legendre Transforms are unchanged, hence the function $d(n)$.

To show the equivalence of the discrete Hamiltonian map having error $\mathscr{O}(K^n)$ and the discrete Legendre transforms having error $\mathscr{O}(K^n)$, we recall expressions for the discrete Hamiltonian map for the discrete Lagrangian L_d and exact discrete Lagrangian L_d^E:

$$F_{L_d} = \mathbb{F}^+ L_d \circ \left(\mathbb{F}^- L_d\right)^{-1},$$
$$F_{L_d^E} = \mathbb{F}^+ L_d^E \circ \left(\mathbb{F}^- L_d^E\right)^{-1}.$$

Now, we make use of the following consequence of the implicit function theorem: If we have smooth functions g_1, g_2 and the sequences of functions $\{f_{1_n}\}_{n=1}^{\infty}$, $\{f_{2_n}\}_{n=1}^{\infty}$, $\{e_{1_n}\}_{n=1}^{\infty}$ and $\{e_{2_n}\}_{n=1}^{\infty}$ such that

$$f_{1_n}(x) = g_1(x) + K^n e_{1_n}(x),$$
$$f_{2_n}(x) = g_2(x) + K^n e_{2_n}(x),$$

where $\sup\{\|e_{1_n}\|\}_{n=1}^{\infty} < C_1$ and $\sup\{\|e_{2_n}\|\}_{n=1}^{\infty} < C_2$ on compact sets, then

$$f_{2_n}(f_{1_n}(x)) = g_2(g_1(x)) + K^n e_{12_n}(x) \tag{2.26}$$
$$f_{1_n}^{-1}(y) = g_1^{-1}(y) + K^n \bar{e}_{1_n}(y) \tag{2.27}$$

for some sequences of functions $\{e_{12_n}\}_{n=1}^{\infty}$, $\{\bar{e}_{1_n}\}_{n=1}^{\infty}$ where $\sup\{\|e_{12_n}\|\}_{n=1}^{\infty} < C_3$ and $\sup\{\|\bar{e}_{1_n}\|\}_{n=1}^{\infty} < C_4$ on compact sets.

It follows from (2.26) and (2.27) that if the discrete Legendre transforms have error $\mathscr{O}(K^n)$, the discrete Hamiltonian map does as well. Finally, if we have a discrete Hamiltonian map has error $\mathscr{O}(K^n)$, we use the identity

$$\left(\mathbb{F}^{-}L_d\right)^{-1}(q_0, p_0) = (q_0, \pi_Q \circ F_{L_d}(q_0, p_0))$$

where π_Q is the projection map $\pi_Q : (q, p) \to q$ and (2.27) to see that $\mathbb{F}^{-}L_d$ is has error $\mathscr{O}(K^n)$, and the identity:

$$\mathbb{F}^{+}L_d = F_{L_d} \circ \mathbb{F}^{-}L_d,$$

along with (2.26) to establish that $F^{+}L_d$ also has error $\mathscr{O}(K^n)$, which completes the proof. □

This simple extension is a critical tool for establishing the geometric convergence of spectral variational integrators, and leads to the following theorem concerning the accuracy of spectral variational integrators.

Theorem 2.6.2. (Geometric Convergence of Spectral Variational Integrators) *Given an interval* $[0,h]$ *and a Lagrangian* $L : TQ \to \mathbb{R}$, *let* $\bar{q}$ *be the exact solution to the Euler-Lagrange equations, and* $\tilde{q}_n$ *be the stationary point of the spectral variational discrete action*

$$\begin{aligned} L_d^S(q_0, q_h, n) &= \operatorname*{ext}_{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} \mathbb{S}_d\left(\{q_i\}_{i=1}^n\right) \\ &= \operatorname*{ext}_{\substack{q_n(t)\in\mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} h\sum_{j=0}^{m_n} b_{n_j} L\left(q_n\left(c_{n_j}h\right), \dot{q}_n\left(c_{n_j}h\right)\right). \end{aligned}$$

If:

1. *there exists constants* C_A, K_A, $K_A < 1$, *independent of* n, *such that, for each* n, *there*

exists a curve $\hat{q}_n \in \mathbb{M}^n([0,h],Q)$ *such that,*

$$\left\|(\bar{q},\dot{\bar{q}}) - \left(\hat{q}_n,\dot{\hat{q}}_n\right)\right\| \leq C_A K_A^n,$$

2. *there exists a closed and bounded neighborhood* $U \subset TQ$ *such that* $\left(\bar{q}(t),\dot{\bar{q}}(t)\right) \in U$ *and* $\left(\hat{q}_n(t),\dot{\hat{q}}_n(t)\right) \in U$ *for all* t *and* n*, and all partial derivatives of* L *are continuous on* U*,*

3. *for the sequence of quadrature rules* $\mathscr{G}_n(f) = h\sum_{j=1}^{m_n} b_{n_j} f\left(c_{n_j}h\right) \approx \int_0^h f(t)\,dt$ *there exists constants* C_g*,* K_g*,* $K_g < 1$*, independent of* n *such that:*

$$\left|\int_0^h L\left(q_n(t),\dot{q}_n(t)\right)dt - h\sum_{j=1}^{m_n} b_{n_j} L\left(q_n\left(c_{n_j}h\right),\dot{q}_n\left(c_{n_j}h\right)\right)\right| \leq C_g K_g^n,$$

   for any $q_n \in \mathbb{M}^n([0,h],Q)$*,*

4. *and the stationary points* $\bar{q}$*,* $\tilde{q}_n$ *minimize their respective actions,*

then

$$\left|L_d^E(q_0,q_1) - L_d^S(q_0,q_1,n)\right| \leq C_s K_s^n \tag{2.28}$$

for some constants C_s, K_s*,* $K_s < 1$*, independent of* n*, and hence the discrete Hamiltonian flow map has error* $\mathscr{O}\left(K_s^n\right)$*.*

Proof. As before, we rewrite both the exact discrete Lagrangian and the spectral discrete Lagrangian:

$$\begin{aligned}
\left|L_d^E(q_0,q_1) - L_d^S(q_0,q_1,n)\right| &= \left|\int_0^h L\left(\bar{q}(t),\dot{\bar{q}}(t)\right)\mathrm{d}t - \mathscr{G}_n\left(L\left(\tilde{q}_n(t),\dot{\tilde{q}}_n(t)\right)\right)\right| \\
&= \left|\int_0^h L\left(\bar{q}(t),\dot{\bar{q}}(t)\right)\mathrm{d}t - h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n\left(c_{n_j}h\right),\dot{\tilde{q}}_n\left(c_{n_j}h\right)\right)\right| \\
&= \left|\int_0^h L\left(\bar{q},\dot{\bar{q}}\right)\mathrm{d}t - h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right)\right|,
\end{aligned}$$

with suppression of the t argument. We introduce the action evaluated on the curve $\hat{q}_n$:

$$\begin{aligned}
&\left|\int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right)\right| \\
&= \left|\int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t + \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t - h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right)\right| \\
&\leq \left|\int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t\right| + \left|\int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t - h\sum_{j=1}^{m} b_{n_j} L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right)\right|. \quad (2.29)
\end{aligned}$$

Considering the first term in (2.29):

$$\begin{aligned}
\left|\int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t\right| &= \left|\int_0^h L(\bar{q},\dot{\bar{q}}) - L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t\right| \\
&\leq \int_0^h \left|L(\bar{q},\dot{\bar{q}}) - L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\right|\mathrm{d}t.
\end{aligned}$$

By assumption, all partials of L are continuous on U, and since U is closed and bounded, this implies L is Lipschitz on U, so let L_α denote the Lipschitz constant. Since, again by assumption, $(\bar{q},\dot{\bar{q}}) \in U$ and $\left(\hat{q}_n,\dot{\hat{q}}_n\right) \in U$, we can obtain:

$$\begin{aligned}
\int_0^h \left|L(\bar{q},\dot{\bar{q}}) - L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\right|\mathrm{d}t &\leq \int_0^h L_\alpha \left|(\bar{q},\dot{\bar{q}}) - \left(\hat{q}_n,\dot{\hat{q}}_n\right)\right|\mathrm{d}t \\
&\leq \int_0^h L_\alpha C_A K_A^n \mathrm{d}t \\
&= hL_\alpha C_A K_A^n.
\end{aligned}$$

Hence,

$$\left|\int_0^h L(\bar{q},\dot{\bar{q}})\,\mathrm{d}t - \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t\right| \leq hL_\alpha C_A K_A^n. \quad (2.30)$$

Next, considering the second term in (2.29),

$$\left|\int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right)\mathrm{d}t - \sum_{j=1}^{m} hb_{n_j} L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right)\right|,$$

since $\tilde{q}_n$ minimizes its action,

$$h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right) \leq h\sum_{j=1}^{m_n} b_{n_j} L\left(\hat{q}_n,\dot{\hat{q}}_n\right) \leq \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right) \mathrm{d}t + C_g K_g^n \tag{2.31}$$

where the inequalities follow from the assumptions on the order of the quadrature rule and (2.30). Furthermore,

$$\begin{aligned} h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right) &\geq \int_0^h L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right) \mathrm{d}t - C_g K_g^n \geq \int_0^h L\left(\bar{q},\dot{\bar{q}}\right) \mathrm{d}t - C_g K_g^n \\ &\geq \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right) \mathrm{d}t - hL_\alpha C_A K_A^n - C_g K_g^n, \end{aligned} \tag{2.32}$$

where the inequalities follow from (2.30), the order of the sequence of quadrature rules, and the assumption that $\bar{q}$ minimizes its action. Putting (2.31) and (2.32) together, we can conclude:

$$\left| \int_0^h L\left(\hat{q}_n,\dot{\hat{q}}_n\right) \mathrm{d}t - h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n,\dot{\tilde{q}}_n\right) \right| \leq \left(hL_\alpha C_A + C_g\right) K_s^{-n}. \tag{2.33}$$

where $K_s = \max\left(K_A, K_g\right)$. Now, combining the bounds (2.30) and (2.33) in (2.29), we can conclude

$$\left| L_d^E\left(q_0,q_1\right) - L_d^S\left(q_0,q_1,n\right) \right| \leq \left(2hL_\alpha C_A + C_g\right) K_s^{-n}$$

which, combined with Theorem 2.6.1, establishes the rate of convergence. □

Acknowledgements

This work was supported in part by NSF CAREER Award DMS-1010687, NSF Grants CMMI-1029445, and DMS-1065972.

References

[1] N. Bou-Rabee and J. E. Marsden. Hamilton-Pontryagin integrators on Lie groups. I. Introduction and structure-preserving properties. *Found. Comput. Math.*, 9(2): 197–219, 2009.

[2] N. Bou-Rabee and H. Owhadi. Stochastic variational integrators. *IMA J. Numer. Anal.*, 29(2):421–443, 2009.

[3] N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.*, 48(1):278–297, 2010.

[4] J. P. Boyd. *Chebyshev and Fourier spectral methods*. Dover Publications Inc., Mineola, NY, second edition, 2001.

[5] J. Cortés and S. Martínez. Non-holonomic integrators. *Nonlinearity*, 14(5):1365–1392, 2001.

[6] G. Dal Maso. *An introduction to* Γ*-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[7] D. Estep and D. French. Global error control for the continuous Galerkin finite element method for ordinary differential equations. *RAIRO Modél. Math. Anal. Numér.*, 28(7):815–852, 1994.

[8] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[9] Y. N. Fedorov and D. V. Zenkov. Discrete nonholonomic LL systems on Lie groups. *Nonlinearity*, 18(5):2211–2241, 2005.

[10] R. Fetecau, J. E. Marsden, M. Ortiz, and M. West. Nonsmooth Lagrangian mechanics and variational collision integrators. *SIAM Journal on Applied Dynamical Systems*, 2(3):381–416, 2003.

[11] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[12] N. Hale and L. N. Trefethen. New quadrature formulas from conformal maps. *SIAM J. Numer. Anal.*, 46(2):930–948, 2008.

[13] B. L. Hulme. One-step piecewise polynomial Galerkin methods for initial value problems. *Math. Comp.*, 26:415–426, 1972.

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

[15] S. Larsson and V. Thomée. *Partial differential equations with numerical methods*, volume 45 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 2003.

[16] T. Lee, M. Leok, and N. H. McClamroch. Lie group variational integrators for the full body problem. *Comput. Methods Appl. Mech. Engrg.*, 196(29-30):2907–2924, 2007.

[17] T. Lee, M. Leok, and N. H. McClamroch. Lagrangian mechanics and variational integrators on two-spheres. *Internat. J. Numer. Methods Engrg.*, 79(9):1147–1174, 2009.

[18] M. Leok. Generalized Galerkin variational integrators: Lie group, multiscale, and pseudospectral methods. (preprint, arXiv:math.NA/0508360), 2004.

[19] M. Leok and T. Shingel. Prolongation–collocation variational integrators. *IMA J. Numer. Anal.*, 32(3):1194–1216, 2012.

[20] M. Leok and T. Shingel. General techniques for constructing variational integrators. *Frontiers of Mathematics in China*, 7(2):273–303, 2012. (Special issue on computational mathematics, invited paper).

[21] M. Leok and J. Zhang. Discrete Hamiltonian variational integrators. *IMA J. Numer. Anal.*, 31(4):1497–1532, 2011.

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

[23] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001.

[24] J. E. Marsden, G. W. Patrick, and S. Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Comm. Math. Phys.*, 199(2):351–395, 1998.

[25] R. McLachlan and M. Perlmutter. Integrators for nonholonomic mechanical systems. *J. Nonlinear Sci.*, 16(4):283–328, 2006.

[26] S. Müller and M. Ortiz. On the Γ-convergence of discrete dynamics and variational integrators. *J. Nonlinear Sci.*, 14(3):279–296, 2004.

[27] NASA. JPL Horizons system online solar system data and ephemeris, September 2012. URL http://ssd.jpl.nasa.gov/?horizons.

[28] G. W. Patrick and C. Cuell. Error analysis of variational integrators of unconstrained Lagrangian systems. *Numer. Math.*, 113(2):243–264, 2009.

[29] L. N. Trefethen. *Spectral methods in MATLAB*, volume 10 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[30] J. Vankerschaver, C. Liao, and M. Leok. Generating functionals and Lagrangian PDEs. *Journal of Nonlinear Science*, 2011. (submitted, arXiv:1111.0280 [math-ph]).

Chapter 2, in full, has been submitted for publication of the material as it may appear in Numerische Mathematik, 2012, James Hall and Melvin Leok, Springer, 2012. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Lie Group Galerkin Variational Integrators

Lie Group Galerkin Variational Integrators

James Hall and Melvin Leok

Abstract. We present a new class of high-order variational integrators on Lie groups. We show that these integrators are symplectic, momentum preserving, and can be constructed to be of arbitrarily high-order, or can be made to converge geometrically. Furthermore, these methods are stable and accurate for very large time steps. We demonstrate the construction of one such variational integrator for the rigid body, and discuss how this construction could be generalized to other related Lie group problems. We close with several numerical examples which demonstrate our claims, and discuss further extensions of our work.

3.1 Introduction

There is a deep and elegant geometric structure underlying the dynamics of many mechanical systems. Conserved quantities, such as the energy, momentum, and symplectic form offer insight into this structure, and through this, we obtain an understanding of the behavior of these systems that goes beyond what is conventionally available. Conservation laws reveal much about the stability and long term behavior of a system, and can even characterize the entire dynamics of a system when a sufficient number of them exist. Hence, there has been much recent interest in the field of geometric mechanics, which seeks to understand this structure using differential geometric and symmetry techniques.

From this geometric mechanics framework, it is possible to formulate numerical methods which respect much of the geometry of mechanical systems. There are a variety of approaches for constructing such methods, often known as *structure-preserving methods*, including projection methods, splitting methods, symplectic Runge-Kutta methods, B-series expansion methods, to name a few. An extensive introduction can be found in Hairer et al. [14]. One of the powerful frameworks, discrete mechanics, approaches the construction of numerical methods by developing much of the theory of geometric mechanics from a discrete standpoint. This approach has proven highly effective for constructing methods for problems in Hamiltonian and Lagrangian mechanics, specifically because these type of problems arise from a variational principle. Methods that make use of a variational principle and the framework of discrete mechanics are referred to as *variational integrators*, and they have many favorable geometric properties, including conservation of the symplectic form and momentum.

A further advantage of variational integrators is that it is often straightforward to analyze the error of these methods. This has led turn to the development of high-order variational integrators, which can be constructed so that they converge very quickly. In Hall and Leok [16], such integrators for vector space problems were presented and analyzed. It was shown that such integrators can be arbitrarily high-order or even exhibit geometric convergence. Furthermore, these integrators are stable and accurate even with extremely large time steps, and using them it is easy to reconstruct highly accurate continuous approximations to the dynamics of the system of interest.

In this paper, we present an extension of that work to Lie group methods. Lie group methods are of particular interest in science and engineering applications. It can be shown that many problems of interest, from the dynamics of rigid bodies to the behavior of incompressible fluids, evolve on Lie groups. Furthermore, if a traditional numerical method is applied to a problem with dynamics in a Lie group, the approximate solution will typically depart from the Lie group, destroying a critical structural property of the solution. Our work gives a general framework for constructing methods which will always evolve in the Lie group and which will share many of the desirable properties of the vector space type methods. Specifically, we will be able to construct methods of arbitrarily high-order and with geometric convergence, and we will be able to reconstruct high quality continuous approximations from these methods.

Lie group methods have a rich history and remain the subject of significant interest. An extensive introduction can be found in Iserles et al. [18], which provides an excellent exposition of both the motivation for Lie group methods and many of the techniques used on Lie groups. Likewise, Celledoni and Owren [7], provide a very helpful general introduction to Lie group methods for the rigid body, which is a prototypical example of an interesting Lie group problem. In this paper, we provide a thorough example of the construction of our method for the rigid body, as this approach can easily be extended to other interesting problems. More recently Bogfjellmo and Marthinsen [1] investigated the construction of high-order symplectic Lie group integrators from a discrete Hamilton-Pontryagin principle, and Burnett et al. [6] described applications of such high-order Lie group discretizations, such as interpolation in $SO(3)$.

Galerkin variational integrators were proposed in Marsden and West [29], and expanded on by Leok [24]. The concept of a Galerkin Lie group integrator was proposed in Leok [24] and expanded in Leok and Shingel [26]. Our work expands upon this by generalizing both the diffeomorphisms used to construct the natural charts and the approximation spaces used to construct the curve on the Lie group, and establishing convergence results and properties of both the discrete solution and the continuous approximation.

3.1.1 Discrete Mechanics

Since we are working from the perspective of discrete mechanics, we will take a moment to review the fundamentals of the theory here. This will only be a brief summary, and extensive exposition of the theory can be found in Marsden and West [29].

Consider a configuration manifold, Q, which describes the configuration of a mechanical system at a given point in time. In discrete mechanics, the fundamental object is the *discrete Lagrangian*, $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$. The discrete Lagrangian can be viewed as an approximation to the *exact discrete Lagrangian* L_d^E, where the L_d^E is defined to be the action of the Lagrangian on the solution of the Euler-Lagrange equations over a short time interval:

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

The discrete Lagrangian gives rise to a discrete action sum, which can be viewed as an approximation to the action over a long time interval:

$$\mathbb{S}\left(\{q_k\}_{k=1}^n\right) = \sum_{k=0}^{n-1} L_d\left(q_k, q_{k+1}\right) \approx \int_{t_0}^{t_n} L\left(q, \dot{q}\right) \mathrm{d}t,$$

and requiring stationarity of this discrete action sum subject to fixed endpoint conditions q_0, q_n, gives rise to the *discrete Euler-Lagrange equations*:

$$D_1 L_d\left(q_k, q_{k+1}, h\right) + D_2 L_d\left(q_{k-1}, q_k, h\right) = 0, \tag{3.1}$$

where D_i denotes partial differentiation of a function with respect to the i-th argument. Given a point (q_{k-1}, q_k), these equations implicitly define an update map $F_{L_d} : (q_{k-1}, q_k) \to (q_k, q_{k+1})$, which approximates the solution of the Euler-Lagrange equations for the continuous system. A numerical method which uses the update map F_{L_d} to construct numerical solutions to ODEs is referred to as a *variational integrator*.

The power of discrete mechanics is derived from the discrete variational structure. Since the update map F_{L_d} is induced from a discrete analogue of the variational principle, much of the geometric structure from continuous mechanics can be extended to discrete mechanics. The discrete Lagrangian gives rise to *discrete Legendre Transforms* $\mathbb{F}L^{\pm} : Q \times Q \to T^*Q$:

$$\begin{aligned}
\mathbb{F}L_d^+ (q_k, q_{k+1}) &= (q_{k+1}, D_2 L_d (q_k, q_{k+1})), \\
\mathbb{F}L_d^- (q_k, q_{k+1}) &= (q_k, -D_1 L_d (q_k, q_{k+1})),
\end{aligned}$$

which lead to the extension of other classical geometric structures. It is important to note that, while there are two different discrete Legendre transforms, (3.1) guarantees that $\mathbb{F}L_d^- (q_k, q_{k+1}) = \mathbb{F}L_d^+ (q_{k-1}, q_k)$, and thus they can be used interchangeably when defining the discrete geometric structure. By their construction, variational integrators induce a discrete symplectic form, i.e. $\Omega_{L_d} = (\mathbb{F}^{\pm} L_d)^* \Omega$ which is conserved by the update map $F_{L_d}^* \Omega_{L_d} = \Omega_{L_d}$, and a discrete analogue of Noether's Theorem, which states that if a discrete Lagrangian is invariant under a diagonal group action on (q_k, q_{k+1}), it induces a discrete momentum map $J_{L_d} = \left(\mathbb{F}L_d^{\pm}\right)^* J$, which is preserved under the update map: $F_{L_d}^* J_{L_d} = J_{L_d}$. The existence of these discrete geometric conservation laws gives a systematic framework to construct powerful numerical methods which preserve structure.

The discrete Legendre transforms also allow us to define an update map through phase space $\tilde{F}_{L_d} : T^*Q \to T^*Q$,

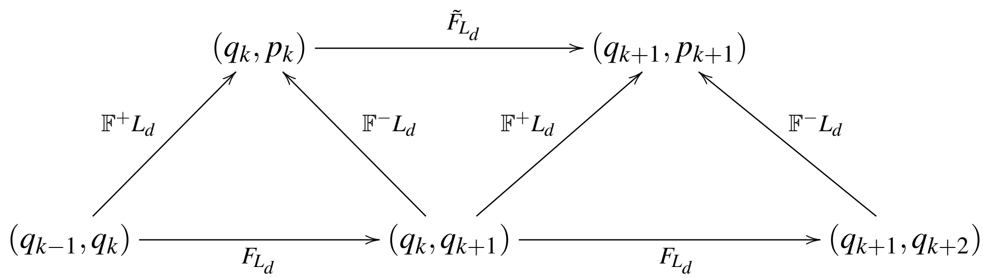
$$\tilde{F}_{L_d} (q_k, p_k) = (q_{k+1}, p_{k+1}),$$

which is given by

$$\tilde{F}_{L_d} (q_k, p_k) = \mathbb{F}^+ L_d \left(\left(\mathbb{F}^- L_d \right)^{-1} (q_k, p_k) \right),$$

known as the Hamiltonian flow map. As long as the discrete Lagrangian is sufficiently smooth, the Hamiltonian flow map and the Lagrangian flow map are compatible, and the geometric structure of discrete flow can be understood from either perspective, just as in the continuous theory.

The following commutative diagram illustrates the relationship between the discrete Legendre transforms, the Lagrangian flow map, the Hamiltonian flow map, and the discrete Lagrangian.



A further consequence of the discrete mechanics framework is that it provides a natural mechanism for analyzing the order of accuracy of a variational integrator. Specifically, it can be shown that the variational integrator induced by the exact discrete Lagrangian produces an exact sampling of the true flow. Based on this, we have the following theorem which is critical for the error analysis of variational integrators:

Theorem 3.1.1. Variational Order Analysis (Theorem 2.3.1 of Marsden and West [29]). *If a discrete Lagrangian L_d approximates the exact discrete Lagrangian L_d^E to order p, i.e.* $L_d(q_0, q_1, h) = L_d^E(q_0, q_1, h) + \mathscr{O}\left(h^{p+1}\right)$*, then the variational integrator induced by L_d is order p accurate.*

This theorem allows for greatly simplified *a priori* error estimates of variational integrators, and is a fundamental tool for the development and analysis of high-order variational integrators.

3.2 Construction

3.2.1 General Galerkin Variational Integrators

Lie group Galerkin variational integrators are an extension of Galerkin variational integrators to Lie groups. As such, we will briefly review the construction of general Galerkin variational integrators.

The driving idea behind Galerkin variational integrators is approaching the construction of a discrete Lagrangian as the approximation of a variational problem. We know from discrete mechanics that the exact discrete Lagrangian $L_d^E : Q \times Q \times \mathbb{R} \to \mathbb{R}$,

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

induces a variational integrator that produces an exact sampling of the true flow, and the accuracy with which a variational integrator approximates the true solution is the same as the accuracy to which the discrete Lagrangian used to construct it approximates the exact discrete Lagrangian. Hence, to construct a highly accurate discrete Lagrangian, we construct a discrete approximation

$$\begin{aligned} L_d^G(q_0, q_1, h) &= \underset{\substack{q_n(t) \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_1}}{\operatorname{ext}} h \sum_{j=1}^{m} b_j L\left(q_n\left(c_j h\right), \dot{q}_n\left(c_j h\right)\right) \\ &\approx \underset{\substack{q(t) \in C^2([0,h],Q) \\ q(0)=q_0, q(h)=q_1}}{\operatorname{ext}} \int_0^h L(q, \dot{q})\, \mathrm{d}t \end{aligned}$$

by replacing the function space $C^2([0,h], Q)$ with a finite-dimensional subspace

$$\mathbb{M}^n([0,h], Q) \subset C^2([0,h], Q)$$

and the integral with a quadrature rule, $h\sum_{j=1}^{m} b_j f\left(c_j h\right) \approx \int_0^h f \mathrm{d}t$. Finding the extremizer of the discrete action is computationally feasible, and by computing this extremizer we can construct the variational integrator that results from the discrete Lagrangian. Because this approach of replacing the function space $C^2([0,h], Q)$ with a finite-dimensional subspace is inspired by Galerkin methods for partial differential equations, we refer to variational integrators constructed in this way as *Galerkin variational integrators*.

In Hall and Leok [16], we studied the Galerkin variational integrator construction on vector spaces. Specifically, we obtained several significant results, including that Galerkin variational integrators over linear spaces are in a certain sense order-optimal, and that by enriching the function space $\mathbb{M}^n([0,h], Q)$, as opposed to shortening the time step h, we can construct variational integrators that converge geometrically. Furthermore, we established that it is easy to recover a continuous approximation to the trajectory over the time step $[0,h]$, and that the convergence of this continuous approximation is related to the rate of convergence of the variational integrator. Finally, we established an error bound on Noether quantities evaluated on this continuous approximation which is independent of the number of steps taken.

3.2.2 Lie Group Galerkin Variational Integrators

The construction and analysis in Hall and Leok [16] relied on the linear structure of the spaces involved. At their heart, Galerkin variational integrators make use of a Galerkin curve

$$\tilde{q}_n(t) = \sum_{i=1}^{n} q^i \phi_i(t)$$

for some set of points $\left\{q^i\right\}_{i=1}^{n} \subset Q$ and basis functions $\{\phi_i\}_{i=1}^{n}$. While for linear spaces, $\tilde{q}_n(t) \in Q$ for any choice of t, in nonlinear spaces this will not be the case. However, when Q is a Lie group, it is possible to extend this construction in a way that keeps the curve $\tilde{q}_n(t)$ in Q.

Natural Charts

To generalize Galerkin variational integrators to Lie groups, we will make use of the linear nature of the Lie algebra associated with the Lie group. Specifically, given a Lie group G and its associated Lie algebra $\mathfrak{g}$, we choose a local diffeomorphism $\Phi : \mathfrak{g} \to G$. Then, given a set of points in the Lie group $\{g_i\}_{i=1}^{n} \subset G$ and a set of associated interpolation times t_i, we can construct an interpolating curve $g : G^n \times \mathbb{R} \to G$ such that $g\left(\{g_i\}_{i=1}^{n}, t_i\right) = g_i$, given by

$$g\left(\{g_i\}_{i=1}^{n}, t\right) = L_{g_1}\Phi\left(\sum_{i=1}^{n} \Phi^{-1}\left(L_{g_1^{-1}} g_i\right)\phi_i(t)\right),$$

where L_g is the left group action of g and $\phi_i(t)$ is the Lagrange interpolation polynomial for t_i. A key feature of this type of curve is that is *Lie group equivariant*, that is, $g\left(\left\{L_{g_0} g_i\right\}_{i=1}^{n}, t\right) = L_{g_0} g\left(\{g_i\}_{i=1}^{n}, t\right)$, as we shall show in the following lemma.

Lemma 3.2.1. *The curve $g\left(\{g_i\}_{i=1}^n, t\right)$ is Lie group equivariant.*

Proof. The proof is a direct calculation.

$$\begin{aligned}
g\left(\left\{L_{g_0} g_i\right\}_{i=1}^n, t\right) &= L_{L_{g_0} g_1} \Phi\left(\sum_{i=1}^n \Phi^{-1}\left(L_{\left(L_{g_0} g_1\right)^{-1}} L_{g_0} g_i\right) \phi_i(t)\right) \\
&= L_{g_0} L_{g_1} \Phi\left(\sum_{i=1}^n \Phi^{-1}\left(L_{g_1^{-1}} L_{g_0^{-1}} L_{g_0} g_i\right) \phi_i(t)\right) \\
&= L_{g_0} L_{g_1} \Phi\left(\sum_{i=1}^n \Phi^{-1}\left(L_{g_1^{-1}} g_i\right) \phi_i(t)\right) \\
&= L_{g_0} g\left(\{g_i\}_{i=1}^n, t\right).
\end{aligned}$$

□

This property will be important for ensuring that the Lie group Galerkin discrete Lagrangian inherits the symmetries of the continuous Lagrangian; these inherited symmetries give rise to the structure-preserving properties of the resulting variational integrator.

Throughout this paper, we will consider the function spaces composed of curves of this form. We note that $\Phi^{-1}\left(L_{g_1^{-1}} g_i\right) \in \mathfrak{g}$, and for any $\xi \in \mathfrak{g}$, $L_{g_1}\Phi(\xi) \in G$, so we can construct interpolation curves on the group in terms of interpolation curves in the Lie algebra. In light of this, we define

$$\begin{aligned}
&\mathbb{GM}^n\left(g_0 \times [0,h], G\right) \\
&:= \left\{ g\left(\left\{\xi^i\right\}_{i=1}^n, t\right) \;\middle|\; g\left(\left\{\xi^i\right\}_{i=1}^n, t\right) = L_{g_0} \Phi\left(\sum_{i=1}^n \xi^i \phi_i(t)\right), \xi^i \in \mathfrak{g} \right\}
\end{aligned}$$

where $\{\phi_i(t)\}_{i=1}^n$ forms the basis for a finite-dimensional approximation space in $\mathbb{R}$, for example, Lagrange interpolation polynomials, which is what we will use in our explicit construction in §3.4 and numerical examples in §3.5. We refer to the space of finite dimensional curves in the Lie algebra as

$$\mathbb{M}^n\left([0,h], \mathfrak{g}\right) := \left\{ \xi(t) \;\middle|\; \xi(t) = \sum_{i=1}^n \xi^i \phi_i(t), \xi^i \in \mathfrak{g}, \phi_i : [0,h] \to \mathbb{R} \right\}.$$

Because we are identifying every point in a neighborhood of the Lie group with a point in the Lie algebra, which is a vector space, it is natural to think of this construction as choosing a set of coordinates for a neighborhood in the Lie group. Thus, we can consider this construction as choosing a chart for a neighborhood of the Lie group, and because it makes use of the "natural" relationship between the Lie group G, its Lie algebra $\mathfrak{g}$, and the tangent space of the Lie group TG, we call the function $\varphi_{g_0} : G \to \mathfrak{g}$, $\varphi_{g_0}(\cdot) = \Phi^{-1}\left(L_{g_0^{-1}}(\cdot)\right)$ a "natural chart."

Discrete Lagrangian

Now that we have introduced a Lie group approximation space, we can define a compatible discrete Lagrangian for Lie group problems. We take a similar approach to the construction for vector spaces; we construct an approximation to the action of the Lagrangian over $[0,h]$ by replacing $C^2([0,h],G)$ with a finite-dimensional approximation space and the integral with a quadrature rule, and then compute its extremizer. Specifically, given a Lagrangian on the tangent space of a Lie group $L : TG \to \mathbb{R}$, the associated Lie group Galerkin discrete Lagrangian is defined to be:

$$L_d(g_k, g_{k+1}) = \operatorname*{ext}_{\substack{g_n(t) \in \mathbb{GM}^n(g_k \times [0,h],G) \\ g_n(0)=g_k, g_n(h)=g_{k+1}}} h \sum_{j=1}^{m} b_j L\left(g_n(c_j h), \dot{g}_n(c_j h)\right).$$

Internal Stage Discrete Euler-Poincaré Equations

This discrete Lagrangian involves solving an optimization problem, namely: find $\tilde{g}_n(t) \in \mathbb{GM}^n(g_k \times [0,h],G)$ such that $\tilde{g}_n(0) = g_k$, $\tilde{g}_n(h) = g_{k+1}$, and

$$h \sum_{j=1}^{m} b_j L\left(\tilde{g}_n(c_j h), \dot{\tilde{g}}_n(c_j h)\right) = \operatorname*{ext}_{\substack{g_n(t) \in \mathbb{GM}^n(g_k \times [0,h],G) \\ g_n(0)=g_k, g_n(h)=g_{k+1}}} h \sum_{j=1}^{m} b_j L\left(g_n(c_j h), \dot{g}_n(c_j h)\right). \tag{3.2}$$

While this problem can be solved using standard methods of numerical optimization, it is also possible to reduce it to a root finding problem. Since each curve $\tilde{g}_n(t) \in \mathbb{GM}^n(g_0 \times [0,h],G)$ is parametrized by a finite number of Lie algebra points $\{\xi^i\}_{i=1}^n$, by taking discrete variations of the discrete Lagrangian with respect to these points, we can derive stationarity conditions for the extremizer. Specifically, if we denote

$$\xi(t) = \sum_{i=1}^{n} \xi^i \phi_i(t)$$

then a straightforward computation reveals the stationarity condition:

$$h\sum_{j=1}^{m} b_j \left(\mathbf{D}_1 L \circ \mathbf{D}_{\Phi(\xi(c_j h))} L_{g_k} \circ \mathbf{D}_{\xi(c_j h)} \Phi \circ \left(\sum_{i=1}^{n} \mathbf{D}_{\xi^i} \xi\left(c_j h\right) \cdot \delta\xi^i \right) \right.$$
$$+ \mathbf{D}_2 L \circ \mathbf{D}_{\left(\Phi\circ\xi(c_j h), \mathbf{D}_{\dot{\xi}(c_j h)} \Phi\circ\dot{\xi}(c_j h)\right)} \mathbf{D}_{\Phi\circ\xi(c_j h)} L_{g_k}$$
$$\left. \circ \mathbf{D}_{(\xi(c_j h), \dot{\xi}(c_j h))} \Phi \circ \left(\sum_{i=1}^{n} \mathbf{D}_{\xi^i} \dot{\xi}\left(c_j h\right) \cdot \delta\xi^i \right) \right) = 0$$

for arbitrary $\left\{\delta\xi^i\right\}_{i=2}^{n-1}$. Using standard calculus of variations arguments, this reduces to

$$h\sum_{j=1}^{m} b_j \left(\mathbf{D}_1 L \circ \mathbf{D}_{\Phi(\xi(c_j h))} L_{g_k} \circ \mathbf{D}_{\xi(c_j h)} \Phi \circ \mathbf{D}_{\xi^i} \xi\left(c_j h\right) \cdot \delta\xi^i \right.$$
$$+ \mathbf{D}_2 L \circ \mathbf{D}_{\left(\Phi\circ\xi(c_j h), \mathbf{D}_{\dot{\xi}(c_j h)} \Phi\circ\dot{\xi}(c_j h)\right)} \mathbf{D}_{\Phi\circ\xi(c_j h)} L_{g_k}$$
$$\left. \circ \mathbf{D}_{(\xi(c_j h), \dot{\xi}(c_j h))} \Phi \circ \mathbf{D}_{\xi^i} \dot{\xi}\left(c_j h\right) \cdot \delta\xi^i \right) = 0$$

for $i = 2, .., n-1$ (note that the sum of the Lie algebra elements has disappeared). Now using the linearity of one-forms, we can collect terms to further simplify this expression to

$$h\sum_{j=1}^{m} b_j \left(\left[\mathbf{D}_1 L \circ \mathbf{D}_{\Phi(\xi(c_j h))} L_{g_k} \circ \mathbf{D}_{\xi(c_j h)} \Phi \circ \mathbf{D}_{\xi^i} \xi\left(c_j h\right) \right. \right.$$
$$+ \mathbf{D}_2 L \circ \mathbf{D}_{\left(\Phi\circ\xi(c_j h), \mathbf{D}_{\dot{\xi}(c_j h)} \Phi\circ\dot{\xi}(c_j h)\right)} \mathbf{D}_{\Phi\circ\xi(c_j h)} L_{g_k}$$
$$\left. \left. \circ \mathbf{D}_{(\xi(c_j h), \dot{\xi}(c_j h))} \Phi \circ \mathbf{D}_{\xi^i} \dot{\xi}\left(c_j h\right) \right] \cdot \delta\xi^i \right) = 0$$

for $i=2,...,n-1$. Since $\delta\xi^i$ is arbitrary, this implies that

$$h\sum_{j=1}^{m} b_j \Bigg(\mathbf{D}_1 L \circ \mathbf{D}_{\Phi(\xi(c_jh))} L_{g_k} \circ \mathbf{D}_{\xi(c_jh)} \Phi \circ \mathbf{D}_{\xi^i} \xi\left(c_jh\right) \tag{3.3}$$

$$+\mathbf{D}_2 L \circ \mathbf{D}_{\left(\Phi\circ\xi(c_jh), \mathbf{D}_{\dot{\xi}(c_jh)} \Phi\circ\dot{\xi}(c_jh)\right)} \mathbf{D}_{\Phi\circ\xi(c_jh)} L_{g_k} \tag{3.4}$$

$$\circ\, \mathbf{D}_{(\xi(c_jh),\dot{\xi}(c_jh))} \Phi \circ \mathbf{D}_{\xi^i} \dot{\xi}\left(c_jh\right) \Bigg) = 0$$

for $i=2,...,n-1$. These equations, which we shall refer to as the *internal stage discrete Euler-Poincaré equations*, combined with the standard momentum matching condition,

$$D_2 L_d\left(g_{k-1}, g_k\right) + D_1 L_d\left(g_k, g_{k+1}\right) = 0, \tag{3.5}$$

which we will discuss in more detail in the §3.2.2, can be easily solved with an iterative nonlinear equation solver. The result is a curve $\tilde{g}_n(t)$ which satisfies condition (3.2). The next step of the one-step map is given by $\tilde{g}_n(h) = g_{k+1}$, which gives the variational integrator.

It should be noted that while the internal stage discrete Euler-Poincaré equations can be computed by deriving all of the various differentials in the chosen coordinates, it is often much simpler to form the discrete action

$$\mathbb{S}_d\left(\left\{\xi^i\right\}_{i=1}^{n}\right) = h\sum_{j=1}^{m} b_j L\left(L_{g_k}\Phi\left(\sum_{i=1}^{n}\xi^i\phi_i\left(c_jh\right)\right), \frac{\mathrm{d}}{\mathrm{d}t}\left(L_{g_k}\Phi\left(\sum_{i=1}^{n}\xi^i\phi_i\left(c_jh\right)\right)\right)\right)$$

explicitly and then compute the stationarity conditions directly in coordinates, rather than a step by step computation of the different maps in (3.3). This is the approach we take when deriving the integrator for the rigid body in §3.4, and it appears to be the much simpler approach in this case. However, the two approaches are equivalent, so if done carefully either will suffice to give the internal stage Euler-Poincaré equations.

Momentum Matching Condition

A difficulty in the derivation of the discrete Euler-Poincaré equations is the computation of the discrete momentum terms

$$\begin{aligned} p_{k,k+1}^{-} &= -D_1 L_d\left(g_k, g_{k+1}\right) \\ p_{k-1,k}^{+} &= D_2 L_d\left(g_{k-1}, g_k\right) \end{aligned}$$

which are used in the discrete Euler-Poincaré equations (3.5),

$$D_1 L_d\left(g_k, g_{k+1}\right) + D_2 L_d\left(g_{k-1}, g_k\right) = 0$$

or

$$p_{k-1,k}^{+} = p_{k,k+1}^{-}.$$

The difficulty arises because the discrete Lagrangian makes use of a local left trivialization. Through the local charts, we reduce the discrete Lagrangian to a function of algebra elements, and because the corresponding group elements are recovered through a complicated computation, working with the group elements directly to compute the discrete Euler-Poincaré equations is difficult. Because of this, to compute the discrete Euler-Poincaré equations, it is more natural to think of the discrete Lagrangian as a function of two Lie algebra elements. If we define a discrete Lagrangian on the Lie algebra $\hat{L}_d : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ as

$$\hat{L}_d\left(\xi_k, \xi_{k+1}\right) = \underset{\substack{g_n(t) \in \mathbb{GM}^n(g_k \times [0,h], G) \\ \Phi^{-1}\left(L_{g_k^{-1}} g_n(0)\right) = \xi_k, \Phi^{-1}\left(L_{g_k^{-1}} g_n(h)\right) = \xi_{k+1}}}{\operatorname{ext}} h \sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right), \dot{g}_n\left(c_j h\right)\right)$$

and compare it to the discrete Lagrangian on the Lie group,

$$L_d\left(g_k, g_{k+1}\right) = \underset{\substack{g_n(t) \in \mathbb{GM}^n(g_k \times [0,h], G) \\ g_n(0) = g_k, g_n(h) = g_{k+1}}}{\operatorname{ext}} h \sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right), \dot{g}_n\left(c_j h\right)\right)$$

it can be seen that there is a simple one-to-one correspondence through the natural charts between points in $G \times G$ and points in $\mathfrak{g} \times \mathfrak{g}$, and that if

$$\left(\Phi^{-1}\left(L_{g_k^{-1}} g_0\right), \Phi^{-1}\left(L_{g_k^{-1}} g_1\right)\right) = (\xi_0, \xi_1),$$

then

$$L(g_0, g_1) = \hat{L}_d(\xi_0, \xi_1).$$

Hence, for every sequence $\{g_k\}_{k=1}^{N}$, there exists a unique sequence $\{\xi_k\}_{i=1}^{N}$ such that

$$\sum_{k=1}^{N-1} L_d(g_k, g_{k+1}) = \sum_{k=1}^{N-1} \hat{L}_d(\xi_k, \xi_{k+1}), \tag{3.6}$$

and vice versa. Thus, we can find the sequence $\{g_k\}_{k=1}^{N}$ that makes the sum on the left hand side of (3.6) stationary by finding the sequence $\{\xi_k\}_{k=1}^{N}$ that makes the sum on the right hand side of (3.6) stationary.

It can easily be seen that the stationarity condition of the action sum on the right is

$$D_2\hat{L}_d(\xi_{k-1}, \xi_k) + D_1\hat{L}_d(\xi_k, \xi_{k+1}) = 0,. \tag{3.7}$$

However, from the definition of $\hat{L}_d$, this implicitly assumes that (ξ_{k-1}, ξ_k) and (ξ_k, ξ_{k+1}) are in the same natural chart. Unfortunately, in our construction (ξ_{k-1}, ξ_k) and (ξ_k, ξ_{k+1}) are in different natural charts. This is because the construction of the Lie group interpolating curve

$$g(t) = L_{g_\beta}\Phi\left(\sum_{i=1}^{n} \xi_k^i \phi_i(t)\right)$$

requires the choice of a base point for the natural chart $g_\beta \in G$. If a consistent choice of base point was made for each time step, then the above equations could be directly computed without difficulty. However, because many natural chart functions contain coordinate singularities, our construction uses a different base point, and thus a different natural chart, at each time step. Specifically, on the interval $[kh,(k+1)h]$, we choose $g_\beta = g_k$ and define

$$g(t) = L_{g_k}\Phi\left(\sum_{i=1}^{n}\xi_k^i\phi_i(t)\right).$$

Thus

$$g(t) = L_{g_{k-1}}\Phi\left(\sum_{j=1}^{n}\xi_{k-1}^i\phi_i(t)\right), t\in[(k-1)h,kh]$$
$$g(t) = L_{g_k}\Phi\left(\sum_{j=1}^{n}\xi_k^i\phi_i(t)\right), t\in[kh,(k+1)h],$$

where we now denote internal stage points ξ_k^i with the subscript k to denote in which interval they occur. While the change in natural chart is expedient for the construction, it creates a difficulty for the computation of the discrete Euler-Poincaré equations, in that now we are using discrete Lagrangians with different natural charts for the different time steps, and hence we cannot compute the discrete Euler-Poincaré equations using (3.7). This problem can be resolved by expressing $g(t)$, and hence (ξ_{k-1},ξ_k) and (ξ_k,ξ_{k+1}), in the same natural chart for $t\in[(k-1)h,(k+1)h]$. Rewriting

$$\begin{aligned} g_n(t) &= L_{g_k}\Phi\left(\sum_{i=1}^{n}\xi_k^i\phi_i(t)\right) && t\in[kh,(k+1)h], \\ &= L_{g_{k-1}}\Phi\left(\Phi^{-1}\left(L_{g_{k-1}^{-1}}L_{g_k}\Phi\left(\sum_{i=1}^{n}\xi_k^i\phi_i(t)\right)\right)\right) && t\in[kh,(k+1)h], \end{aligned}$$

(note that $g_n(t)$ is still in $\mathbb{GM}^n(g_k\times[0,h],G)$), and defining

$$\lambda(t) = \Phi^{-1}\left(L_{g_{k-1}^{-1}}L_{g_k}\Phi\left(\sum_{i=1}^{n}\xi_k^i\phi_i(t)\right)\right) = \Phi^{-1}\left(L_{\Phi(\xi_k)}\Phi\left(\sum_{i=1}^{n}\xi_k^i\phi_i(t)\right)\right)$$

we can reexpress the discrete Lagrangian as

$$\tilde{L}_d(\lambda_k,\lambda_{k+1}) = \underset{\substack{g_n(t)\in\mathbb{GM}^n(g_k\times[0,h],G)\\ \Phi^{-1}\left(L_{g_{k-1}^{-1}}g_n(0)\right)=\lambda_k,\Phi^{-1}\left(L_{g_{k-1}^{-1}}g_n(h)\right)=\lambda_{k+1}}}{\operatorname{ext}} h\sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right),\dot{g}_n\left(c_j h\right)\right).$$

Note that if $L_{g_{k-1}}\Phi(\lambda_k) = L_{g_k}\Phi(\xi_k)$ and $L_{g_{k-1}}\Phi(\lambda_{k+1}) = L_{g_k}\Phi(\xi_{k+1})$ then

$$\hat{L}_d(\xi_k,\xi_{k+1}) = \tilde{L}_d(\lambda_k,\lambda_{k+1}).$$

Furthermore, $(\lambda_k,\lambda_{k+1})$ are in the same chart as (ξ_{k-1},ξ_k), and hence the discrete Euler-Poincaré equations are

$$D_2\hat{L}_d(\xi_{k-1},\xi_k) + D_1\tilde{L}_d(\lambda_k,\lambda_{k+1}) = 0.$$

It remains to compute λ_k as a function of ξ_k. If we consider the definition of $\lambda(t)$, then

$$\lambda_k = \lambda(0) = \Phi^{-1}\left(L_{\Phi(\xi_k)}\Phi\left(\sum_{i=1}^{n}\xi_k^i\phi_i(0)\right)\right)$$

and

$$\xi_k = \Phi^{-1}\left(L_{g_k^{-1}}g_n(0)\right) = \Phi^{-1}\left(L_{\Phi(\xi_k)^{-1}}\Phi(\lambda_k)\right). \tag{3.8}$$

This is simply a change of coordinates, and hence computing the discrete Euler Lagrange equations amounts to using the change of coordinates map to transform the algebra elements into the same chart. Thus,

$$\begin{aligned} D_2\hat{L}_d(\xi_{k-1},\xi_k) &= \frac{\partial L_d}{\partial \xi_k} \\ D_1\tilde{L}_d(\lambda_k,\lambda_{k+1}) &= \frac{\partial L_d}{\partial \xi_k}\frac{\partial \xi_k}{\partial \lambda_k} \end{aligned} \tag{3.9}$$

where (3.8) can be used to compute $\frac{\partial \xi_k}{\partial \lambda_k}$. An explicit example is presented in section §3.4.

There are several features of this computation that should be noted. First, since we are considering specific choices of natural charts, we may think of ξ_k and λ_k as corresponding to a specific coordinate choice, and hence it is natural to use standard partial derivatives as opposed to coordinate free notation. Second, because λ_k is a function of ξ_k, which is in turn a function of ξ_k^i, this is still a root finding problem over ξ_k^i, and hence may be solved concurrently with the internal stage Euler-Poincaré equations (3.3).

3.3 Convergence

Thus far, we have discussed the construction of Lie group Galerkin variational integrators. Now we will prove several theorems related to their convergence. Unlike traditional integrators, we will achieve convergence in two different ways; the first will be the standard shortening of the time step $[0,h]$, which we refer to as h-refinement. In practice, we refer to methods that achieve convergence through h-refinement as Lie group Galerkin variational integrators, after the method used to construct them. The second is by enriching the function space $\mathbb{GM}^n\left(g_0 \times [0,h], G\right)$ and holding the time step $[0,h]$ constant. Because enriching $\mathbb{GM}^n\left(g_0 \times [0,h], G\right)$ involves increasing the number of basis functions, and hence the value of n, we refer to this as n-refinement. Because this approach of enriching the function space is inspired by classical spectral methods, as in Trefethen [36], when we use n-refinement to achieve convergence we will refer to the the resulting method as a *Lie group spectral variational integrator*.

3.3.1 Geometric and Optimal Convergence

Naturally, the goal of applying the spectral paradigm to the construction of Galerkin variational integrators is to construct methods which achieve geometric convergence. In this section, we will prove that under certain assumptions about the behavior of the Lagrangian and the approximation space, Lie group spectral variational integrators achieve geometric convergence. Additionally, the argument that establishes geometric convergence can be easily modified to show that the convergence of Lie group Galerkin integrators is, in a certain sense, optimal.

The proof of the rate of convergence Galerkin Lie group variational integrators is superficially similar to the proof of the rate of convergence of Galerkin variational integrators, which was established in [16]. The specific major difference is the need to quantify the error between two different curves on the Lie group. Unlike a normed vector space, there may not be a simple method of quantifying this error. For the moment, we will avoid this difficulty by simply assuming that the error between two curves that share a common point in a Lie group can be characterized through the error between curves in the Lie algebra. Specifically, we will make the following "natural chart conditioning" assumption:

$$e_g\left(L_{g_0}\Phi\left(\xi\left(t\right)\right), L_{g_0}\Phi\left(\eta\left(t\right)\right)\right) \leq C_G \left\langle \xi\left(t\right)-\eta\left(t\right), \xi\left(t\right)-\eta\left(t\right)\right\rangle^{\frac{1}{2}} \tag{3.10}$$

$$e_a\left(\frac{\mathrm{d}}{\mathrm{d}t}L_{g_0}\Phi\left(\xi\left(t\right)\right), \frac{\mathrm{d}}{\mathrm{d}t}L_{g_0}\Phi\left(\eta\left(t\right)\right)\right) \leq C_{\mathfrak{g}} \left\langle \dot{\xi}\left(t\right)-\dot{\eta}\left(t\right), \dot{\xi}\left(t\right)-\dot{\eta}\left(t\right)\right\rangle^{\frac{1}{2}} + C_{\mathfrak{g}}^{G} \left\langle \xi\left(t\right)-\eta\left(t\right), \xi\left(t\right)-\eta\left(t\right)\right\rangle^{\frac{1}{2}} \tag{3.11}$$

for some functions $e_g(\cdot,\cdot)$ and $e_a(\cdot,\cdot)$, which are chosen to measure the error in the Lie group and tangent bundle of the Lie group, respectively, and for some choice of Riemannian metric $\langle\cdot,\cdot\rangle$ on the Lie algebra. It is important to note that while the error function may be chosen to be the length of the geodesic curve that connects $L_{g_k}\Phi(\xi)$ and $L_{g_k}\Phi(\eta)$, there are other valid choices. This will greatly simplify error calculations; for example, in §3.4 we choose the error function to be the matrix two-norm, $\|\cdot\|_2$, which is quickly and easily computed and will obey this inequality for the Riemannian metric we use.

Optimal Convergence

We will begin by proving optimal convergence of Lie group Galerkin variational integrators. In this case, we take "optimal" to mean that the Lie group Galerkin variational integrator will converge at the same rate as the best possible approximation in the approximation space used to construct it.

Theorem 3.3.1. *Given an interval $[0,h]$, and a Lagrangian $L : TG \to \mathbb{R}$, suppose that $\bar{g}(t)$ solves the Euler-Lagrange equations on that interval exactly. Furthermore, suppose* *that the exact solution* $\bar{g}(t)$ *falls within the range of the natural chart, that is:*

$$\bar{g}(t) = L_{g_k}\Phi(\bar{\eta}(t))$$

for some $\bar{\eta}(t) \in C^2([0,h],\mathfrak{g})$. *For the function space* $\mathbb{GM}^n(g_0 \times [0,h], G)$ *and the quadrature rule* $\mathcal{G}$, *define the Galerkin discrete Lagrangian* $L_d^G(g_0, g_1, h) \to \mathbb{R}$ *as*

$$\begin{aligned} L_d^G(g_0,g_1,h) &= \underset{\substack{g_n(t)\in\mathbb{GM}^n(g_0\times[0,h],G)\\ g_n(0)=g_0, g_n(h)=g_1}}{\operatorname{ext}} h\sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right), \dot{g}_n\left(c_j h\right)\right) \\ &= h\sum_{h=1}^{m} b_j L\left(\tilde{g}_n\left(c_j h\right), \dot{\tilde{g}}_n\left(c_j h\right)\right) \end{aligned} \tag{3.12}$$

where $\tilde{g}_n(t)$ *is the extremizing curve in* $\mathbb{GM}^n(g_0 \times [0,h], G)$. *If:*

1. *there exists an approximation* $\hat{\eta}_n \in \mathbb{M}^n([0,h],\mathfrak{g})$ *such that,*

$$\begin{aligned} \langle \bar{\eta}(t) - \hat{\eta}_n(t), \bar{\eta}(t) - \hat{\eta}_n(t)\rangle^{\frac{1}{2}} &\leq C_A h^n \\ \left\langle \dot{\bar{\eta}}(t) - \dot{\hat{\eta}}_n(t), \dot{\bar{\eta}}(t) - \dot{\hat{\eta}}_n(t)\right\rangle^{\frac{1}{2}} &\leq C_{\mathfrak{A}} h^n, \end{aligned}$$

   for some constants $C_A \geq 0$ *and* $C_{\mathfrak{A}} \geq 0$ *independent of h,*

2. *the Lagrangian L is Lipschitz in the chosen norms in both its arguments, that is:*

$$|L(g_1,\dot{g}_1) - L(g_2,\dot{g}_2)| \leq L_\alpha\left(e_g(g_1,g_2) + e_a(\dot{g}_1,\dot{g}_2)\right),$$

3. *the chart function* Φ *is well-conditioned in* $e_g(\cdot,\cdot)$ *and* $e_a(\cdot,\cdot)$, *that is (3.10) and (3.11) hold,*

4. *for the quadrature rule* $\mathcal{G}(f) = h\sum_{j=1}^{m} b_j f\left(c_j h\right) \approx \int_0^h f(t)\,dt$, *there exists a constant* $C_g \geq 0$ *such that,*

$$\left|\int_0^h L(g_n(t), \dot{g}_n(t))\,dt - h\sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right), \dot{g}_n\left(c_j h\right)\right)\right| \leq C_g h^{n+1}$$

   for any $g_n(t) = L_{g_0}\Phi(\xi(t))$ *where* $\xi(t) \in \mathbb{M}^n([0,h],\mathfrak{g})$,

5. *the stationary points of the discrete action and the continuous action are minimizers,*

then the variational integrator induced by $L_d^G(g_0,g_1)$ *has error* $\mathscr{O}\left(h^{n+1}\right)$.

Proof. We begin by rewriting the exact discrete Lagrangian and the Galerkin discrete Lagrangian:

$$\left|L_d^E(g_0,g_1,h)-L_d^G(g_0,g_1,h)\right| = \left|\int_0^h L(\bar{g},\dot{\bar{g}})\,\mathrm{d}t - h\sum_{j=1}^m b_j L\left(\tilde{g}_n\left(c_jh\right),\dot{\tilde{g}}_n\left(c_jh\right)\right)\right|,$$

where we have introduced $\tilde{g}_n(t)$, which is the stationary point of the local Galerkin action (3.12). We introduce the solution in the approximation space which takes the form $\hat{g}_n(t)=L_{g_k}\Phi(\hat{\eta}_n(t))$, and compare the action on the exact solution to the action on this solution:

$$\begin{aligned}\left|\int_0^h L(\bar{g},\dot{\bar{g}})\,\mathrm{d}t-\int_0^h L\left(\hat{g}_n,\dot{\hat{g}}_n\right)\mathrm{d}t\right| &= \left|\int_0^h L(\bar{g},\dot{\bar{g}})-L\left(\hat{g}_n,\dot{\hat{g}}_n\right)\mathrm{d}t\right| \\ &\leq \int_0^h\left|L(\bar{g},\dot{\bar{g}})-L\left(\hat{g}_n,\dot{\hat{g}}_n\right)\right|\mathrm{d}t.\end{aligned}$$

Now, we use the Lipschitz assumption to establish the bound

$$\begin{aligned}&\int_0^h\left|L(\bar{g},\dot{\bar{g}})-L\left(\hat{g}_n,\dot{\hat{g}}_n\right)\right|\mathrm{d}t \\ &\leq \int_0^h L_\alpha\left(e_g\left(\bar{g},\hat{g}_n\right)+e_a\left(\dot{\bar{g}},\dot{\hat{g}}_n\right)\right)\mathrm{d}t \\ &= \int_0^h L_\alpha\left(e_g\left(L_{g_k}\Phi(\bar{\eta}),L_{g_k}\Phi(\hat{\eta}_n)\right)\right. \\ &\qquad\left.+e_a\left(D_{\Phi(\bar{\eta})}L_{g_0}D_{\bar{\eta}}\Phi\left(\dot{\bar{\eta}}\right),D_{\Phi(\hat{\eta}_n)}L_{g_0}D_{\hat{\eta}_n}\Phi\left(\dot{\hat{\eta}}_n\right)\right)\right)\mathrm{d}t,\end{aligned}$$

and the chart conditioning assumptions to see

$$\begin{aligned}\int_0^h\left|L(\bar{g},\dot{\bar{g}})-L\left(\hat{g}_n,\dot{\hat{g}}_n\right)\right|\mathrm{d}t &\leq \int_0^h L_\alpha\left(C_G\langle\bar{\eta}-\hat{\eta}_n,\bar{\eta}-\hat{\eta}_n\rangle^{\frac{1}{2}}+C_{\mathfrak{g}}\left\langle\dot{\bar{\eta}}-\dot{\hat{\eta}}_n,\dot{\bar{\eta}}-\dot{\hat{\eta}}_n\right\rangle^{\frac{1}{2}}+\right. \\ &\qquad\left.C_{\mathfrak{g}}^G\langle\bar{\eta}-\hat{\eta}_n,\bar{\eta}-\hat{\eta}_n\rangle^{\frac{1}{2}}\right)\mathrm{d}t \\ &\leq \int_0^h L_\alpha\left(C_GC_Ah^n+C_{\mathfrak{g}}C_{\mathfrak{A}}h^n+C_{\mathfrak{g}}^GC_Ah^n\right)\mathrm{d}t\end{aligned}$$

$$= L_\alpha \left(\left(C_G + C_{\mathfrak{g}}^G \right) C_A + C_{\mathfrak{g}} C_{\mathfrak{A}} \right) h^{n+1}.$$

This establishes a bound between the action evaluated on the exact discrete Lagrangian and the optimal solution in the approximation space, $\hat{g}_n$. Considering the Galerkin discrete action,

$$\begin{aligned}
h \sum_{j=1}^{m} b_j L\left(\tilde{g}_n, \tilde{g}_n\right) &\leq h \sum_{j=1}^{m} b_j L\left(\hat{g}_n, \dot{\hat{g}}_n\right) \\
&\leq \int_0^h L\left(\hat{g}_n, \dot{\hat{g}}_n\right) \mathrm{d}t + C_g h^{n+1} \\
&\leq \int_0^h L\left(\bar{g}, \dot{\bar{g}}\right) \mathrm{d}t + C_g h^{n+1} + L_\alpha \left(\left(C_G + C_{\mathfrak{g}}^G \right) C_A + C_{\mathfrak{g}} C_{\mathfrak{A}} \right) h^{n+1} \quad (3.13)
\end{aligned}$$

where we have used the assumption that the Galerkin approximation $\tilde{g}_n$ minimizes the Galerkin discrete action and the assumption on the accuracy of the quadrature. Now, using the fact that $\bar{g}(t)$ minimizes the action and that $\mathbb{GM}^n\left(g_0 \times [0,h], G\right) \subset C^2\left([0,h], G\right)$,

$$\begin{aligned}
h \sum_{j=1}^{m} b_j L\left(\tilde{g}_n, \dot{\tilde{g}}_n\right) &\geq \int_0^h L\left(\tilde{g}_n, \dot{\tilde{g}}_n\right) \mathrm{d}t - C_g h^{n+1} \\
&\geq \int_0^h L\left(\bar{g}, \bar{g}\right) \mathrm{d}t - C_g h^{n+1} \quad (3.14)
\end{aligned}$$

Combining inequalities (3.13) and (3.14), we see that,

$$\begin{aligned}
\int_0^h L\left(\bar{g}, \dot{\bar{g}}\right) \mathrm{d}t - C_g h^{n+1} &\leq h \sum_{j=1}^{m} b_j L\left(\tilde{g}_n, \dot{\tilde{g}}_n\right) \\
&\leq \int_0^h L\left(\bar{g}, \dot{\bar{g}}\right) \mathrm{d}t + C_g h^{n+1} + L_\alpha \left(\left(C_G + C_{\mathfrak{g}}^G \right) C_A + C_{\mathfrak{g}} C_{\mathfrak{A}} \right) h^{n+1}
\end{aligned}$$

which implies

$$\left| \int_0^h L\left(\bar{g}, \dot{\bar{g}}\right) \mathrm{d}t - h \sum_{j=1}^{m} L\left(\tilde{g}_n, \dot{\tilde{g}}_n\right) \right| \leq \left(C_g + L_\alpha \left(\left(C_G + C_{\mathfrak{g}}^G \right) C_A + C_{\mathfrak{g}} C_{\mathfrak{A}} \right) \right) h^{n+1} \quad (3.15)$$

The left hand side of (3.15) is exactly $\left| L_d^E\left(g_0, g_1, h\right) - L_d^G\left(g_0, g_1, h\right) \right|$, and thus

$$\left| L_d^E\left(g_0, g_1, h\right) - L_d^G\left(g_0, g_1, h\right) \right| \leq C_{op} h^{n+1}.$$

where

$$C_{op} = C_g + L_\alpha \left(\left(C_G + C_{\mathfrak{g}}^G \right) C_A + C_{\mathfrak{g}} C_{\mathfrak{A}} \right).$$

This states that the Galerkin discrete Lagrangian approximates the exact discrete Lagrangian with error $\mathscr{O}\left(h^{n+1}\right)$, and by Theorem (3.1.1) this further implies that the Lagrangian update map, and hence the Lie group Galerkin variational integrator has error $\mathscr{O}\left(h^{n+1}\right)$. □

Geometric Convergence

Under similar assumptions, we can demonstrate that Lie group spectral variational integrators will converge geometrically with n-refinement, that is, enrichment of the function space $\mathbb{GM}^n\left(g_0 \times [0,h], G\right)$ as opposed to the shortening of the time step, h.

Theorem 3.3.2. *Given an interval* $[0,h]$*, and a Lagrangian* $L : TG \to \mathbb{R}$*, suppose that* $\bar{g}(t)$ *solves the Euler-Lagrange equations on that interval exactly. Furthermore, suppose that the exact solution* $\bar{g}(t)$ *falls within the range of the natural chart, that is:*

$$\bar{g}(t) = L_{g_k} \Phi\left(\bar{\eta}(t)\right)$$

for some $\bar{\eta} \in C^2\left([0,h], \mathfrak{g}\right)$*. For the function space* $\mathbb{M}^n\left([0,h], \mathfrak{g}\right)$ *and the quadrature rule* $\mathscr{G}$*, define the Galerkin discrete Lagrangian* $L_d^G\left(g_0, g_1, h\right) \to \mathbb{R}$ *as*

$$\begin{aligned} L_d^G\left(g_0, g_1, h\right) &= \underset{\substack{g_n(t) \in \mathbb{GM}^n(g_0 \times [0,h], G) \\ g_n(0) = g_0, g_n(h) = g_1}}{\operatorname{ext}} h \sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right), \dot{g}_n\left(c_j h\right)\right) \\ &= h \sum_{h=1}^{m} b_j L\left(\tilde{g}_n\left(c_j h\right), \dot{\tilde{g}}_n\left(c_j h\right)\right) \end{aligned} \tag{3.16}$$

where $\tilde{g}_n(t)$ *is the extremizing curve in* $\mathbb{GM}^n\left(g_0 \times [0,h], G\right)$*. If:*

1. *there exists an approximation* $\hat{\eta}_n \in \mathbb{M}^n\left([0,h], \mathfrak{g}\right)$ *such that,*

$$\begin{aligned} \left\langle \bar{\eta} - \hat{\eta}_n, \bar{\eta} - \hat{\eta}_n \right\rangle^{\frac{1}{2}} &\leq C_A K_A^n \\ \left\langle \dot{\bar{\eta}} - \dot{\hat{\eta}}_n, \dot{\bar{\eta}} - \dot{\hat{\eta}}_n \right\rangle^{\frac{1}{2}} &\leq C_{\mathfrak{A}} K_A^n, \end{aligned}$$

for some constants $C_A \geq 0$ *and* $C_{\mathfrak{A}} \geq 0$, $0 < K_A < 1$ *independent of n,*

2. *the Lagrangian L is Lipschitz in the chosen error norm in both its arguments, that is:*

$$\left|L(g_1, \dot{g}_1) - L(g_2, \dot{g}_2)\right| \leq L_\alpha \left(e_g(g_1, g_2) + e_a(\dot{g}_1, \dot{g}_2)\right)$$

3. *the chart function* Φ *is well conditioned in* $e_g(\cdot,\cdot)$ *and* $e_a(\cdot,\cdot)$*, that is (3.10) and (3.11) hold,*

4. *there exists a sequence of quadrature rules* $\{\mathscr{G}_n\}_{n=1}^{\infty}$,

$$\mathscr{G}_n(f) = h \sum_{j=1}^{m_n} b_{n_j} f\left(c_{n_j} h\right) \approx \int_0^h f(t)\, dt,$$

and there exists a constant $0 < K_g < 1$ *independent of n such that,*

$$\left| \int_0^h L(g_n(t), \dot{g}_n(t))\, dt - h \sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right), \dot{g}_n\left(c_j h\right)\right) \right| \leq C_g K_g^n$$

for any $g_n(t) = L_{g_0} \Phi(\xi(t))$ *where* $\xi(t) \in \mathbb{M}^n([0,h], \mathfrak{g})$,

5. *the stationary points of the discrete action and the continuous action are minimizers,*

then the variational integrator induced by $L_d^G(g_0, g_1)$ *has error* $\mathscr{O}(K^n)$.

The proof for this theorem is very similar to that for Theorem 3.3.1, using the modified assumptions in the obvious way. It would be tedious to repeat it here, but it has been included in the appendix for completeness.

These proofs may seem quite strong in their assumptions. However, as we shall see in §3.4, for many Lagrangians, there are many reasonable choices of function spaces, natural chart functions, quadrature rules and error norms such that the assumptions are satisfied. We will specifically examine Lagrangians over $SO(3)$ of the form:

$$L\left(R, \dot{R}\right) = \operatorname{tr}\left(\dot{R}^T R J_d R^T \dot{R}\right) - V(R),$$

which is the rigid body under the influence of a potential. We will show that for Lie group Galerkin variational integrators, stationary points of the discrete action are minimizers under a certain time step restriction. In addition we will give a specific construction of a Lie group Galerkin variational integrator for this type of problem, and demonstrate the expected convergence on several example problems.

3.3.2 Stationary Points are Minimizers

A major assumption in both Theorem 3.3.1 and Theorem 3.3.2 is that the stationary point of the discrete action is a minimizer. While in general this may not hold, we can show that given a time step restriction on h, that this condition holds for problems on $SO(3)$ for Lagrangians of the form

$$L\left(R,\dot{R}\right) = \operatorname{tr}\left(\dot{R}^T R J_d R^T \dot{R}\right) - V(R).$$

This includes a broad range of problems. Furthermore, we establish a similar result for problems in vector space in Hall and Leok [16], and it may be possible to combine these two results to include a large class of problems, including those that evolve on the special Euclidean group $SE(3) = \mathbb{R}^3 \ltimes SO(3)$.

Lemma 3.3.1. *Consider a Lagrangian on $SO(3)$ of the form*

$$L\left(R,\dot{R}\right) = tr\left(\dot{R}^T R J_d R^T \dot{R}\right) - V(R).$$

If a Lie group Galerkin variational integrator is constructed with $\{\phi_i\}_{i=1}^n$ forming the basis for polynomials of degree $n+1$ and the quadrature rule is of order at least $2n+2$, then the stationary points of the discrete action are minimizers.

Proof. We begin by noting that we can identify every element of $\mathfrak{so}(3)$, the Lie algebra associated with $SO(3)$, with an element of $\mathbb{R}^3$ using the *hat map* $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$,

$$\widehat{\begin{pmatrix} a \\ b \\ c \end{pmatrix}} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}. \tag{3.17}$$

Hence, it is natural to consider the discrete action as a function on $H^1\left([0,h],\mathbb{R}^3\right)$,

$$\mathbb{S}_d\left(\xi(t),\dot{\xi}(t)\right)=h\sum_{j=1}^{m}b_jL\left(L_{g_k}\Phi\left(\hat{\xi}(c_jh)\right),\frac{d}{\mathrm{d}t}L_{g_k}\Phi\left(\hat{\xi}(c_jh)\right)\right).$$

Let $\xi_s(t)$ be the stationary point of $\mathbb{S}_d$. Now, consider a perturbation to $\xi_s(t)$, $\xi_s(t)+\delta\xi(t)$. Since $\xi_s(t)$ is the extremizer over curves $\xi(t)$ subject to the constraints $\xi(0)=\xi_0$, $\xi(h)=\xi_1$, we know $\delta\xi(0)=0$ and $\delta\xi(h)=0$, but it is otherwise arbitrary. Hence, we consider an arbitrary perturbation $\delta\xi(t)\in H_0^1\left([0,h],\mathbb{R}^3\right)$. Since $\mathbb{S}_d$ is a function on $H^1\left([0,h],\mathbb{R}^3\right)$, we can Taylor expand around the stationary point:

$$\begin{aligned}\mathbb{S}_d\left(\xi_s+\delta\xi,\dot{\xi}_s+\delta\dot{\xi}\right)=\mathbb{S}_d\left(\xi_s,\dot{\xi}_s\right)+D\mathbb{S}_d\left(\xi_s,\dot{\xi}_s\right)\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\\+\frac{1}{2}D^2\mathbb{S}_d(\eta,\dot{\eta})\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\end{aligned}$$

where $\eta(t)=\lambda(t)\xi_0(t)+(1-\lambda(t))\delta\xi(t)$ for some $\lambda(t):[0,h]\to[0,1]$ and $D\mathbb{S}_d$, $D^2\mathbb{S}_d$ are the first and second Frechet derivative of $\mathbb{S}_d$, respectively. Thus

$$\begin{aligned}\mathbb{S}_d\left(\xi_s+\delta\xi,\dot{\xi}_s+\delta\dot{\xi}\right)-\mathbb{S}_d\left(\xi_s,\dot{\xi}_s\right)=D\mathbb{S}_d\left(\xi_s,\dot{\xi}_s\right)\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\\+\frac{1}{2}D^2\mathbb{S}_d(\eta,\dot{\eta})\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\left[\left(\delta\xi,\delta\dot{\xi}\right)\right].\end{aligned}$$

Now, note that

$$D\mathbb{S}_d\left(\xi_s,\dot{\xi}_s\right)\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]=0$$

is exactly the stationarity conditions for the internal stage discrete Euler-Poincaré equations. Thus,

$$\mathbb{S}_d\left(\xi_s+\delta\xi,\dot{\xi}_s+\delta\dot{\xi}\right)-\mathbb{S}_d\left(\xi_s,\dot{\xi}_s\right)=\frac{1}{2}D^2\mathbb{S}_d(\eta,\dot{\eta})\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\left[\left(\delta\xi,\delta\dot{\xi}\right)\right].$$

We will examine $D^2\mathbb{S}_d$. The second Frechet derivative of the discrete action is given by

$$D^2\mathbb{S}_d\left(\xi,\dot{\xi}\right)\left[\left(\delta\xi_a,\delta\dot{\xi}_a\right)\right]\left[\left(\delta\xi_b,\delta\dot{\xi}_b\right)\right]$$

$$= h \sum_{j=1}^{m} b_j \nabla^2 L\left(\xi, \dot{\xi}\right) \left[\left(\delta\xi_a\left(c_j h\right), \delta\dot{\xi}_a\left(c_j h\right)\right)\right] \left[\left(\delta\xi_b\left(c_j h\right), \delta\dot{\xi}_b\left(c_j h\right)\right)\right].$$

In order to examine the second Frechet derivative, we must examine the Hessian of the Lagrangian. We will do this term-wise. The Lagrangian has the form

$$L\left(\xi, \dot{\xi}\right) = K\left(\xi, \dot{\xi}\right) - V\left(\xi\right)$$

where

$$K\left(\xi\left(t\right), \dot{\xi}\left(t\right)\right) = \dot{R}\left(\xi\left(t\right)\right)^T R\left(\xi\left(t\right)\right) J_d R\left(\xi\left(t\right)\right)^T \dot{R}\left(\xi\left(t\right)\right).$$

is the kinetic energy and V is the potential energy. Considering K, note that

$$R\left(\xi\left(t\right)\right)^T \dot{R}\left(\xi\left(t\right)\right) = \Phi\left(\xi\left(t\right)\right)^T \nabla\Phi\left(\xi\left(t\right)\right) \dot{\xi}\left(t\right)$$

and hence as a function of $\dot{\xi}\left(t\right)$,

$$K\left(\dot{\xi}\left(t\right)\right) = \dot{\xi}\left(t\right)^T \nabla\Phi\left(\xi\left(t\right)\right)^T \Phi\left(\xi\left(t\right)\right) J_d \Phi\left(\xi\left(t\right)\right)^T \nabla\Phi\left(\xi\left(t\right)\right) \dot{\xi}\left(t\right).$$

J_d is a diagonal matrix with (J_1, J_2, J_3) on the diagonal, and because $\Phi\left(\xi\left(t\right)\right)$ is an orthogonal matrix, $\Phi\left(\xi\left(t\right)\right) J_d \Phi\left(\xi\left(t\right)\right)^T$ has the eigenvalues (J_1, J_2, J_3). Furthermore, $\Phi\left(\cdot\right)$ is a diffeomorphism, which implies $\nabla\Phi\left(\cdot\right)$ is non-singular, so

$$\begin{aligned}
&\dot{\xi}\left(t\right)^T \nabla\Phi\left(\xi\left(t\right)\right)^T \Phi\left(\xi\left(t\right)\right) J_d \Phi\left(\xi\left(t\right)\right)^T \nabla\Phi\left(\xi\left(t\right)\right) \dot{\xi}\left(t\right) \\
&\geq J_{\min} \left\| \nabla\Phi\left(\xi\left(t\right)\right) \dot{\xi}\left(t\right) \right\|_2^2 \\
&\geq J_{\min} \left|\sigma_{\min}\left(t\right)\right| \left\| \dot{\xi}\left(t\right) \right\|_2^2
\end{aligned}$$

where $J_{\min} = \min\left(\{J_1, J_2, J_3\}\right)$ and $\sigma_{\min}\left(t\right)$ is the singular value of $\nabla\Phi\left(\dot{\xi}\left(t\right)\right)$ with the smallest magnitude. Since $\left|\sigma_{\min}\left(t\right)\right|$ is a continuous function of t and $\left|\sigma_{\min}\left(t\right)\right| > 0$ for all t over the compact interval $[0, h]$, there exists a constant $C_\sigma > 0$ such that $\left|\sigma_{\min}\left(t\right)\right| > C_\sigma$

for all $t \in [0,h]$. Finally, we note

$$\frac{\partial^2 K}{\partial \dot{\xi}^2}(\eta(t),\dot{\eta}(t))\left[\delta\dot{\xi}_a\right]\left[\delta\dot{\xi}_b\right]$$
$$= 2\delta\dot{\xi}_a^T \nabla\Phi(\eta(t))^T \Phi(\eta(t)) J_d \Phi(\eta(t))^T \nabla\Phi(\eta(t))\,\delta\dot{\xi}_b,$$

and hence

$$\frac{\partial^2 K}{\partial \dot{\xi}^2}(\eta(t),\dot{\eta}(t))\left[\delta\dot{\xi}\right]\left[\delta\dot{\xi}\right] \geq 2J_{\min}C_\sigma \delta\dot{\xi}^T\delta\dot{\xi}. \tag{3.18}$$

Now, considering the full term $\nabla^2 K(\eta(t),\dot{\eta}(t))\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]$, we see that:

$$\begin{aligned}
&\nabla^2 K(\eta(t),\dot{\eta}(t))\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\left[\left(\delta\xi,\delta\dot{\xi}\right)\right] \\
&= \frac{\partial^2 K}{\partial \xi^2}(\eta(t),\dot{\eta}(t))[\delta\xi][\delta\xi] + 2\frac{\partial^2 K}{\partial \xi \partial \dot{\xi}}(\eta(t),\dot{\eta}(t))[\delta\xi]\left[\delta\dot{\xi}\right] \\
&\qquad + \frac{\partial^2 K}{\partial \dot{\xi}^2}(\eta(t),\dot{\eta}(t))\left[\delta\dot{\xi}\right]\left[\delta\dot{\xi}\right],
\end{aligned} \tag{3.19}$$

where we have made use of the symmetry of mixed second derivatives. K is smooth in all of its components, and hence there exists $C_m > 0$, $C_d > 0$ such that

$$\frac{\partial^2 K}{\partial \xi \partial \dot{\xi}}(\eta(t),\dot{\eta}(t))[\delta\xi]\left[\delta\dot{\xi}\right] \geq -C_m \delta\xi^T\delta\dot{\xi} \tag{3.20}$$

$$\frac{\partial^2 K}{\partial \xi^2}(\eta(t),\dot{\eta}(t))[\delta\xi][\delta\xi] \geq -C_d \delta\xi^T\delta\xi. \tag{3.21}$$

Defining $C_p = 2J_{\min}C_\sigma$, inserting (3.18), (3.20), and (3.21) into (3.19) gives

$$\begin{aligned}
&\nabla^2 K(\eta(t),\dot{\eta}(t))\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\left[\left(\delta\xi,\delta\dot{\xi}\right)\right] \\
&\geq C_p \delta\dot{\xi}(t)^T\delta\dot{\xi}(t) - C_m\delta\dot{\xi}(t)^T\delta\xi(t) - C_d\delta\xi(t)^T\delta\xi(t) \\
&= \frac{C_p}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t) + \frac{C_p}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t) - C_m\delta\dot{\xi}(t)^T\delta\xi(t) \\
&\qquad - C_d\delta\xi(t)^T\delta\xi(t).
\end{aligned}$$

Completing the square, we see that

$$\begin{aligned}
&\frac{C_p}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t)+\frac{C_p}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t)-C_m\delta\dot{\xi}(t)^T\delta\xi(t)-C_d\delta\xi(t)^T\delta\xi(t)\\
=&\frac{C_p}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t)+\left(\frac{\sqrt{C_p}}{\sqrt{2}}\delta\dot{\xi}(t)-\frac{\sqrt{2}C_m}{2\sqrt{C_p}}\delta\xi(t)\right)^T\left(\frac{\sqrt{C_p}}{\sqrt{2}}\delta\dot{\xi}(t)-\frac{\sqrt{2}C_m}{2\sqrt{C_p}}\delta\xi(t)\right)\\
&\qquad-\left(C_d+\frac{C_m^2}{2C_p}\right)\delta\xi(t)^T\delta\xi(t)\\
=&\frac{C_p}{2}\left\|\delta\dot{\xi}(t)\right\|_2^2+\left\|\frac{\sqrt{C_p}}{\sqrt{2}}\delta\dot{\xi}(t)-\frac{\sqrt{2}C_m}{2\sqrt{C_p}}\delta\xi(t)\right\|_2^2-\left(C_d+\frac{C_m^2}{2C_p}\right)\|\delta\xi(t)\|_2^2.
\end{aligned}$$

Making use of the trivial bound that for any $a,b\in\mathbb{R}^3$, $\|a-b\|_2^2\geq 0$, we see

$$\begin{aligned}
&\frac{C_p}{2}\left\|\delta\dot{\xi}(t)\right\|_2^2+\left\|\frac{\sqrt{C_p}}{\sqrt{2}}\delta\dot{\xi}(t)-\frac{\sqrt{2}C_m}{2\sqrt{C_p}}\delta\xi(t)\right\|_2^2-\left(C_d+\frac{C_m^2}{2C_p}\right)\|\delta\xi(t)\|_2^2\\
\geq&\frac{C_p}{2}\left\|\delta\dot{\xi}(t)\right\|_2^2-\left(C_d+\frac{C_m^2}{2C_p}\right)\|\delta\xi(t)\|_2^2\\
=&C_{\dot{\xi}}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t)-C_{\xi}\delta\xi(t)^T\delta\xi(t)
\end{aligned}$$

for constants $C_{\dot{\xi}}>0$, $C_{\xi}>0$, where

$$\begin{aligned}
C_{\dot{\xi}}&=\frac{C_p}{2}\\
C_{\xi}&=C_d+\frac{C_m^2}{2C_p}.
\end{aligned}$$

This bound allows us to conclude

$$\nabla^2K(\eta(t),\dot{\eta}(t))\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\left[\left(\delta\xi,\delta\dot{\xi}\right)\right]\geq C_{\dot{\xi}}\dot{\xi}(t)^T\dot{\xi}(t)-C_{\xi}\xi(t)^T\xi(t). \tag{3.22}$$

We now turn our attention to the potential term, $V(R(\xi(t)))$. Since V and $R(\cdot)$ are both smooth we know that the second partial derivatives of $V(R(\cdot))$ are bounded, and since V does not depend on $\dot{\xi}(t)$,

$$\nabla^2V(R(\eta(t)))\left[\left(\delta\xi(t),\delta\dot{\xi}(t)\right)\right]\left[\left(\delta\xi(t),\delta\dot{\xi}(t)\right)\right]\leq C_V\delta\xi(t)^T\delta\xi(t) \tag{3.23}$$

for a constant C_V. Thus, combining (3.22) and (3.23), we can bound $\nabla^2\mathbb{S}_d$,

$$\nabla^2\mathbb{S}_d\left(\eta\left(t\right),\dot{\eta}\left(t\right)\right)\left[\left(\delta\xi\left(t\right),\delta\dot{\xi}\left(t\right)\right)\right]\left[\left(\delta\xi\left(t\right),\delta\dot{\xi}\left(t\right)\right)\right]\geq$$
$$h\sum_{j=1}^{m}b_jC_{\dot{\xi}}\delta\dot{\xi}\left(c_jh\right)^T\delta\dot{\xi}\left(c_jh\right)-\left(C_\xi+C_V\right)\delta\xi\left(c_jh\right)^T\delta\xi\left(c_jh\right).$$

By assumption, $\delta\xi\left(t\right)$ and $\delta\dot{\xi}\left(t\right)$ are polynomials of degree at most $n+1$, so we know that $\delta\xi\left(t\right)^T\delta\xi\left(t\right)$ and $\delta\dot{\xi}\left(t\right)^T\delta\dot{\xi}\left(t\right)$ are polynomials of degree at most $2n+2$, which implies the quadrature rule is exact, and thus

$$h\sum_{j=1}^{m}b_jC_{\dot{\xi}}\delta\dot{\xi}\left(c_jh\right)^T\delta\dot{\xi}\left(c_jh\right)-\left(C_\xi+C_V\right)\delta\xi\left(c_jh\right)^T\delta\xi\left(c_jh\right)= \qquad (3.24)$$
$$C_{\dot{\xi}}\int_0^h\delta\dot{\xi}\left(t\right)^T\delta\dot{\xi}\left(t\right)\mathrm{d}t-\left(C_\xi+C_V\right)\int_0^h\delta\xi\left(t\right)^T\delta\xi\left(t\right)\mathrm{d}t.$$

$\delta\xi\left(t\right)\in H_0^1\left(\left[0,h\right],\mathbb{R}^3\right)$, so we can apply Poincaré's inequality to see

$$\begin{aligned}&C_{\dot{\xi}}\int_0^h\delta\dot{\xi}\left(t\right)^T\delta\dot{\xi}\left(t\right)\mathrm{d}t-\left(C_\xi+C_V\right)\int_0^h\delta\xi\left(t\right)^T\delta\xi\left(t\right)\mathrm{d}t\\
\geq&\frac{C_{\dot{\xi}}\pi^2}{h^2}\int_0^h\delta\xi\left(t\right)^T\delta\xi\left(t\right)\mathrm{d}t-\left(C_\xi+C_V\right)\int_0^h\delta\xi\left(t\right)^T\delta\xi\left(t\right)\mathrm{d}t\\
=&\left(\frac{C_{\dot{\xi}}\pi^2}{h^2}-\left(C_\xi+C_V\right)\right)\int_0^h\xi\left(t\right)^T\xi\left(t\right)\mathrm{d}t\end{aligned}$$

which is positive so long as $h<\sqrt{\frac{C_{\dot{\xi}}\pi^2}{C_\xi+C_V}}$. Thus, given that $h<\sqrt{\frac{C_{\dot{\xi}}\pi^2}{C_\xi+C_V}}$, for arbitrary $\left(\delta\xi\left(t\right),\delta\dot{\xi}\left(t\right)\right)$

$$\mathbb{S}_d\left(\xi_s\left(t\right)+\delta\xi\left(t\right),\dot{\xi}_s\left(t\right)+\delta\dot{\xi}\left(t\right)\right)-\mathbb{S}_d\left(\xi_s\left(t\right),\dot{\xi}_s\left(t\right)\right)>0$$

which demonstrates that $\left(\xi_s\left(t\right),\dot{\xi}_s\left(t\right)\right)$ minimizes the action. □

It should be noted that the only use of the assumption that the approximation space is polynomials of order at least n is when we use the order of the quadrature rule to change the quadrature to the exact integral (3.24). Thus, this proof can easily be generalized to other approximation spaces, so long as the quadrature rule used is exact for the product of any two elements of the approximation space and the product of any two derivatives of the elements of the approximation space.

3.3.3 Convergence of Galerkin Curves

Lie group Galerkin variational integrators require the construction of a curve

$$\tilde{g}_n(t) \in \mathbb{GM}^n(g_k \times [0,h], G)$$

such that

$$\tilde{g}_n(t) = \underset{\substack{g_n(t)\in\mathbb{GM}^n(g_k\times[0,h],G)\\ g_n(0)=g_k, g_n(h)=g_{k+1}}}{\operatorname{argext}} \sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right), \dot{g}_n\left(c_j h\right)\right).$$

This curve, which we shall refer to as the *Galerkin curve*, is a finite-dimension approximation to the true solution of the Euler-Poincaré equations over the interval $[0,h]$. For the one-step map, we are only concerned with the right endpoint of the Galerkin curve, as

$$g_{k+1} = g_n(h).$$

However, the curve itself has excellent approximation properties as a continuous approximation to the solution of the Euler-Poincaré equations over the interval $[0,h]$. Because Lie group Galerkin variational integrators are capable of taking very large time steps, the dynamics during these time steps may be of interest, and hence the quality of the approximation by these Galerkin curves is also of particular interest.

Ideally, these curves would have the same order of error as the one-step map. Unfortunately, we can only establish error estimates with lower orders of approximation. We established similar results in the vector space case, see Hall and Leok [16], and observed that at high enough accuracy, there is indeed greater error in the Galerkin curve than the one-step map. However, when comparing these curves to the true solution, typically the error introduced by the inaccuracies in (g_k, g_{k+1}) dominates the error from the Galerkin curve, and thus this lower rate of convergence is not observable in practice.

Before we formally establish the rates of convergence for the Galerkin curves, we will briefly review the norms we will use in our theorems and proofs. First, recall the L_p norm for functions over the interval $[0,h]$ given by

$$\|f\|_{L^p([0,h])} = \left(\int_0^h |f|^p \, \mathrm{d}t \right)^{\frac{1}{p}}$$

and next, the Sobolev norm $\|\cdot\|_{W^{1,p}([0,h])}$ for functions on the interval $[0,h]$, given by:

$$\|f\|_{W^{1,p}([0,h])} = \left(\|f\|^p_{L^p([0,h])} + \|\dot{f}\|^p_{L^p([0,h])} \right)^{\frac{1}{p}}.$$

Also, note that for curves $\xi(t) \in \mathfrak{g}$, $|\xi(t)| = \langle \xi(t), \xi(t) \rangle^{\frac{1}{2}}$. We will make extensive use of these definitions in the next three theorems.

Theorem 3.3.3. *Under the same assumptions as Theorem 3.3.2, consider the action as a function of the local left trivialization of the Lie group curve and its derivative,*

$$\mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right) = \int_0^h L\left(L_g \Phi(\bar{\eta}(t)), \frac{d}{dt} L_g \Phi(\bar{\eta}(t)) \right) dt,$$

where $L_g\Phi(\bar{\eta}(t))$ satisfies the Euler-Poincaré equations exactly. If at $\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right)$ the action $\mathfrak{S}_{\mathfrak{g}}(\cdot,\cdot)$ is twice Frechet differentiable and the second Frechet derivative is coercive in variations of the Lie algebra, that is,

$$\left| D^2 \mathfrak{S}_{\mathfrak{g}}\left(\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right)\right) \left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right] \left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right] \right| \geq C_f \|\delta\xi(t)\|^2_{W^{1,1}([0,h])}$$

for all $\delta\xi(t) \in H^1_0([0,h], \mathfrak{g})$, then if the one-step map has error $\mathscr{O}(K^n)$, the Galerkin curves have error $\mathscr{O}\left(\sqrt{K}^n\right)$ in Sobolev norm $\|\cdot\|_{W^{1,1}([0,h])}$.

Proof. We start with the bound (3.56), given at the end of the proof of Theorem 3.3.2 in the appendix,

$$\left| L^E_d(g_k, g_{k+1}, h) - L^G_d(g_k, g_{k+1}, n) \right| \leq C_s K^n_s,$$

and expand using the definitions of $L_d^E(g_k, g_{k+1}, h)$ and $L_d^G(g_k, g_{k+1}, n)$,

$$\begin{aligned}
C_s K_s^n &\geq \left| L_d^E(g_k, g_{k+1}, h) - L_d^G(g_k, g_{k+1}, n) \right| \\
&\geq \left| \int_0^h L\left(L_{g_k}\Phi(\bar{\eta}(t)), \frac{\mathrm{d}}{\mathrm{d}t} L_{g_k}\Phi(\bar{\eta}(t)) \right) \mathrm{d}t \right. \\
&\qquad \left. - \sum_{j=1}^m b_j L\left(L_{g_k}\Phi(\tilde{\eta}(t)), \frac{\mathrm{d}}{\mathrm{d}t} L_{g_k}\Phi(\tilde{\eta}(t)) \right) \right| \\
&\geq \left| \int_0^h L\left(L_{g_k}\Phi(\bar{\eta}(t)), \frac{\mathrm{d}}{\mathrm{d}t} L_{g_k}\Phi(\bar{\eta}(t)) \right) \mathrm{d}t \right. \\
&\qquad \left. - \int_0^h L\left(L_{g_k}\Phi(\tilde{\eta}(t)), \frac{\mathrm{d}}{\mathrm{d}t} L_{g_k}\Phi(\tilde{\eta}(t)) \right) \mathrm{d}t \right| - C_g K_g^n \\
&= \left| \mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right) - \mathfrak{S}_{\mathfrak{g}}\left(\tilde{\eta}(t), \dot{\tilde{\eta}}(t)\right) \right| - C_g K_g^n,
\end{aligned}$$

and since $K_g \leq K_s$, this implies

$$(C_s + C_g) K_s^n \geq \left| \mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right) - \mathfrak{S}_{\mathfrak{g}}\left(\tilde{\eta}(t), \dot{\tilde{\eta}}_n(t)\right) \right|. \tag{3.25}$$

We now Taylor expand around the exact solution $\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right)$

$$\mathfrak{S}_{\mathfrak{g}}\left(\tilde{\eta}(t), \dot{\tilde{\eta}}(t)\right) \tag{3.26}$$
$$= \mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right) + D\mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right)\left[\bar{\eta}(t) - \tilde{\eta}(t), \dot{\bar{\eta}}(t) - \dot{\tilde{\eta}}(t)\right] \tag{3.27}$$
$$+ \frac{1}{2} D^2\mathfrak{S}_{\mathfrak{g}}(\nu(t), \dot{\nu}(t))\left[\left(\bar{\eta}(t) - \tilde{\eta}(t), \dot{\bar{\eta}}(t) - \dot{\tilde{\eta}}(t)\right)\right]\left[\left(\bar{\eta}(t) - \tilde{\eta}(t), \dot{\bar{\eta}}(t) - \dot{\tilde{\eta}}(t)\right)\right],$$

where $\nu(t)$ is a curve in $\mathfrak{g}$. Now, note that $D\mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right) = 0$ is exactly the stationarity condition of the Euler-Poincaré equations. Thus, inserting (3.27) into (3.25) yields

$$(C_s + C_g) K_s^n \tag{3.28}$$
$$\geq \frac{1}{2}\left| D^2\mathfrak{S}_{\mathfrak{g}}(\nu(t), \dot{\nu}(t))\left[\left(\bar{\eta}(t) - \tilde{\eta}(t), \dot{\bar{\eta}}(t) - \dot{\tilde{\eta}}(t)\right)\right]\left[\left(\bar{\eta}(t) - \tilde{\eta}(t), \dot{\bar{\eta}}(t) - \dot{\tilde{\eta}}(t)\right)\right] \right|$$
$$\geq \frac{C_f}{2} \left\| \bar{\eta}(t) - \tilde{\eta}(t) \right\|_{W^{1,1}([0,h])}^2 \tag{3.29}$$

where we have made use of the coercivity of the second derivative of the action. Simplifying (3.29) yields

$$\sqrt{\frac{2\left(C_s+C_g\right)}{C_f}}\sqrt{K_s}^n \geq \left\|\bar{\eta}\left(t\right)-\tilde{\eta}\left(t\right)\right\|_{W^{1,1}([0,h])},$$

which establishes convergence in the Sobolev norm. □

Just as we proved an order optimality theorem, Theorem 3.3.1, that was analogous to the geometric convergence theorem, Theorem 3.3.2, we can establish an analogous convergence theorem for Galerkin curves with h-refinement.

Theorem 3.3.4. *Under the same assumptions as Theorem 3.3.1, consider the action as a function of the local left trivialization of the Lie group curve and its derivative,*

$$\mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta},\dot{\bar{\eta}}\right)=\int_0^h L\left(L_g\Phi\left(\bar{\eta}\right),\frac{d}{dt}L_g\Phi\left(\bar{\eta}\right)\right)dt.$$

If at $\left(\bar{\eta},\dot{\bar{\eta}}\right)$ the action $\mathfrak{S}_{\mathfrak{g}}\left(\cdot,\cdot\right)$ is twice Frechet differentiable and the second Frechet derivative is coercive in variations of the Lie algebra as in Theorem 3.3.3, then if the one-step map has error $\mathscr{O}\left(h^{n+1}\right)$, then the Galerkin curves have error $\mathscr{O}\left(h^{\frac{n+1}{2}}\right)$ in the Sobolev norm $\left\|\cdot\right\|_{W^{1,1}([0,h])}$.

The proof Theorem 3.3.4 is nearly identical to that of Theorem 3.3.3, the only difference being that the bounds containing K_s^n are replaced with bounds containing h^{n+1} in the obvious way.

Like the assumption that the stationary point of the discrete action is a minimizer in Theorems 3.3.2 and 3.3.1, the assumption that the second Frechet derivative of the action is coercive might seem quite strong. However, we can show that for Lagrangians on $SO\left(3\right)$ of the form

$$L\left(R,\dot{R}\right)=\operatorname{tr}\left(\dot{R}^TRJ_dR^T\dot{R}\right)-V\left(R\right),$$

the second Frechet derivative of the action is coercive, subject to a time-step restriction on h.

Lemma 3.3.2. *For Lagrangians on $SO(3)$ of the form*

$$L\left(R,\dot{R}\right) = tr\left(\dot{R}^T R J_d R^T \dot{R}\right) - V(R),$$

there exists a $C > 0$ such that for $h < C$, the second Frechet derivative of $\mathfrak{S}_{\mathfrak{g}}(\cdot,\cdot)$ at $\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right)$ is coercive on the interval $[0,h]$.

Proof. First, we note that for this Lagrangian

$$\begin{aligned}
&D^2\mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right)\left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right]\left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right] \\
&= \int_0^h \nabla^2 L\left(L_{g_k}\Phi(\bar{\eta}(t)), \frac{\mathrm{d}}{\mathrm{d}t}L_{g_k}\Phi(\bar{\eta}(t))\right)\left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right]\left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right]\mathrm{d}t
\end{aligned}$$

From the proof of Lemma 3.3.1, we know that

$$\begin{aligned}
&\nabla^2 L\left(L_{g_k}\Phi(\bar{\eta}(t)), \frac{\mathrm{d}}{\mathrm{d}t}L_{g_k}\Phi(\bar{\eta}(t))\right)\left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right]\left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right] \\
&\geq C_{\dot{\xi}}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t) - \left(C_{\xi} + C_V\right)\delta\xi(t)^T\delta\xi(t) \\
&= \frac{C_{\dot{\xi}}}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t) + \frac{C_{\dot{\xi}}}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t) - \left(C_{\xi} + C_V\right)\delta\xi(t)^T\delta\xi(t),
\end{aligned}$$

and hence

$$\begin{aligned}
&D^2\mathfrak{S}_{\mathfrak{g}}\left(\bar{\eta}(t), \dot{\bar{\eta}}(t)\right)\left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right]\left[\left(\delta\xi(t), \delta\dot{\xi}(t)\right)\right] \\
&\geq \int_0^h \frac{C_{\dot{\xi}}}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t)\,\mathrm{d}t + \int_0^h \frac{C_{\dot{\xi}}}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t) - \left(C_{\xi} + C_V\right)\delta\xi(t)^T\delta\xi(t)\,\mathrm{d}t.
\end{aligned} \tag{3.30}$$

Applying Poincaré's inequality, we see that

$$\begin{aligned}
&\int_0^h \frac{C_{\dot{\xi}}}{2}\delta\dot{\xi}(t)^T\delta\dot{\xi}(t)\,\mathrm{d}t - \int_0^h \left(C_{\xi} + C_V\right)\delta\xi(t)^T\delta\xi(t)\,\mathrm{d}t \\
&\geq \frac{C_{\dot{\xi}}\pi^2}{2h^2}\int_0^h \delta\xi(t)^T\delta\xi(t)\,\mathrm{d}t - \left(C_{\xi} + C_V\right)\int_0^h \delta\xi(t)^T\delta\xi(t)\,\mathrm{d}t \\
&= \left(\frac{C_{\dot{\xi}}\pi^2}{2h^2} - \left(C_{\xi} + C_V\right)\right)\int_0^h \delta\xi(t)^T\delta\xi(t)\,\mathrm{d}t.
\end{aligned} \tag{3.31}$$

Replacing the last two terms in (3.30) with (3.31), we see

$$\begin{aligned}
&D^2\mathfrak{S}_{\mathfrak{g}}\left(\left(\bar{\eta},\dot{\bar{\eta}}\right)\right)\left[\left(\delta\xi\left(t\right),\delta\dot{\xi}\left(t\right)\right)\right]\left[\left(\delta\xi\left(t\right),\delta\dot{\xi}\left(t\right)\right)\right]\\
&\geq \frac{C_{\dot{\xi}}}{2}\int_0^h \delta\dot{\xi}\left(t\right)^T\delta\dot{\xi}\left(t\right)\mathrm{d}t+\left(\frac{C_{\dot{\xi}}\pi^2}{2h^2}-\left(C_{\xi}+C_V\right)\right)\int_0^h \delta\xi\left(t\right)^T\delta\xi\left(t\right)\mathrm{d}t\\
&=\frac{C_{\dot{\xi}}}{2}\left\|\delta\dot{\xi}\left(t\right)\right\|^2_{L^2([0,h])}+\left(\frac{C_{\dot{\xi}}\pi^2}{2h^2}-\left(C_{\xi}+C_V\right)\right)\left\|\delta\xi\left(t\right)\right\|^2_{L^2([0,h])}.
\end{aligned}$$

We now apply Hölder's inequality

$$\|fg\|_{L^1([0,h])}\leq\|f\|_{L^2([0,h])}\|g\|_{L^2([0,h])}$$

to derive the bounds

$$\begin{aligned}
\left\|\delta\dot{\xi}\left(t\right)\right\|_{L^1([0,h])}&\leq\sqrt{h}\left\|\delta\dot{\xi}\left(t\right)\right\|_{L^2([0,h])}\\
\left\|\delta\xi\left(t\right)\right\|_{L^1([0,h])}&\leq\sqrt{h}\left\|\delta\xi\left(t\right)\right\|_{L^2([0,h])},
\end{aligned}$$

and hence,

$$\begin{aligned}
&D^2\mathfrak{S}_{\mathfrak{g}}\left(\left(\bar{\eta},\dot{\bar{\eta}}\right)\right)\left[\left(\delta\xi\left(t\right),\delta\dot{\xi}\left(t\right)\right)\right]\left[\left(\delta\xi\left(t\right),\delta\dot{\xi}\left(t\right)\right)\right]\\
&\geq\frac{C_{\dot{\xi}}}{2h}\left\|\delta\dot{\xi}\left(t\right)\right\|^2_{L^1([0,h])}+\frac{1}{h}\left(\frac{C_{\dot{\xi}}\pi^2}{2h^2}-\left(C_{\xi}+C_V\right)\right)\left\|\delta\xi\left(t\right)\right\|^2_{L^1([0,h])}\\
&\geq\min\left(\frac{C_{\dot{\xi}}}{2h},\frac{1}{h}\left(\frac{C_{\dot{\xi}}\pi^2}{2h^2}-\left(C_{\xi}+C_V\right)\right)\right)\left(\left\|\delta\xi\left(t\right)\right\|^2_{L^1([0,h])}+\left\|\delta\dot{\xi}\left(t\right)\right\|^2_{L^1([0,h])}\right)\\
&\geq\min\left(\frac{C_{\dot{\xi}}}{2h},\frac{1}{h}\left(\frac{C_{\dot{\xi}}\pi^2}{2h^2}-\left(C_{\xi}+C_V\right)\right)\right)\left(\frac{1}{2}\right)\left(\left\|\delta\xi\left(t\right)\right\|_{L^1([0,h])}+\left\|\delta\dot{\xi}\left(t\right)\right\|_{L^1([0,h])}\right)^2\\
&\geq\min\left(\frac{C_{\dot{\xi}}}{4h},\frac{1}{2h}\left(\frac{C_{\dot{\xi}}\pi^2}{2h^2}-\left(C_{\xi}+C_V\right)\right)\right)\left\|\delta\xi\left(t\right)\right\|^2_{W^{1,1}([0,h])}
\end{aligned}$$

which establishes the required coercivity result so long as $0<h<\sqrt{\frac{C_{\dot{\xi}}\pi^2}{2\left(C_{\xi}+C_V\right)}}$. □

3.4 Cayley Transform Based Method on the Special Orthogonal Group

Because the construction of a Lie group Galerkin variational integrator can be involved, we will provide an example of an integrator based on the Cayley transform for the rigid body on $SO(3)$ and related problems. We will first construct the method and then verify that it satisfies the hypotheses of Theorems 3.3.1 and 3.3.2, and in §3.5 we will demonstrate numerically that it exhibits the expected convergence.

Additionally, discretizing the rigid body amounts to discretizing a kinetic energy term that can be used in many different applications. It appears that discretizing the kinetic energy term of the rigid body is more painstaking than the potential term, so we provide a detailed description so that others will not have to repeat the derivation of this discretization for future applications.

3.4.1 Free Rigid Body

The Lagrangian:

$$L\left(R,\dot{R}\right) = \operatorname{tr}\left(\dot{R}^T R J_d R^T \dot{R}\right) \tag{3.32}$$

$$J_d = \frac{1}{2}\operatorname{tr}[J]\, I_{3\times 3} - J$$

$$J = \operatorname{tr}[J_d]\, I_{3\times 3} - J_d, \tag{3.33}$$

where $R \in SO(3)$ and J are the moments of inertia in the reference coordinate frame, gives rise to the equations of motion for the rigid body. The rigid body has a rich geometric structure, which is discussed in Lee et al. [23], Celledoni and Owren [7], and Marsden and Ratiu [31]. In addition to being an interesting example of a non-canonical Lagrangian system, it is a standard model problem for discretization for numerical methods on Lie groups, and an overview of integrators applied to the rigid body can be found in Hairer et al. [14].

3.4.2 Construction

To construct the Lie group Galerkin variational integrator, we have to choose:

1. a map $\Phi(\cdot) : \mathfrak{so}(3) \to SO(3)$,
2. a finite dimensional function space $\mathbb{M}^n([0,h],\mathfrak{g})$, and
3. a quadrature rule,

and to complete the error analysis, we must also choose

1. a metric on $\mathfrak{so}(3)$ $\langle \cdot, \cdot \rangle$,
2. error functions $e_g(\cdot,\cdot)$ and $e_a(\cdot,\cdot)$.

For our construction, we will make use of the Cayley transform for our map $\Phi(\cdot)$ and Lagrange interpolation polynomials through $\mathfrak{so}(3)$ for the finite-dimensional function space $\mathbb{M}^n([0,h],\mathfrak{g})$, that is,

$$\mathbb{M}^n([0,h],\mathfrak{g}) = \left\{ \xi(t) \,\middle|\, \xi(t) = \sum_{i=1}^{n} \widehat{q^i \phi_i(t)}, q^i \in \mathbb{R}^3 \right\},$$

where $\hat{\cdot}$ is that hat map described by (3.17) and $\phi_i(t)$ is the Lagrange interpolation polynomial for t_i. For the error analysis we will choose:

$$\begin{aligned} \langle \hat{\eta}, \hat{\nu} \rangle &= \eta^T \nu, \\ e_g(G_1, G_2) &= \|G_1 - G_2\|_2 \\ e_a(\hat{\eta}, \hat{\nu}) &= \|\hat{\eta} - \hat{\nu}\|_2, \end{aligned}$$

for arbitrary $G_1, G_2 \in SO(3)$ and $\eta, \nu \in \mathbb{R}^3$, where the $\|\cdot\|_2$ norm is understood as arising from the $\|\cdot\|_2$ from the embedding space $R^{3\times 3}$. We will discuss these below, and elaborate on the motivation for these choices in our construction.

The Cayley Transform

To construct our Lie group Galerkin variational Integrator, we will make use of the Cayley Transform, $\Phi(\cdot) : \mathfrak{so}(3) \to SO(3)$ which is given by:

$$\Phi(q) = (I - Q)(I + Q)^{-1}.$$

The reader should note that we are using an unscaled version of the Cayley transform, but for the purposes of constructing the natural chart, different versions of the Cayley transform should result in equivalent methods. Furthermore, the choice of the Cayley transform for the integrator is certainly not necessary; different choices of maps, such as the exponential map, would result in equally valid methods. We make use of the Cayley transform simply because it is easy to manipulate and compute, is its own inverse, and because it satisfies our chart conditioning assumptions, as we will establish shortly.

Lemma 3.4.1. *For $\eta, \nu \in \mathfrak{so}(3)$, so long as*

$$2\|\eta\|_2 + \|\nu\|_2 \leq < 1, \tag{3.34}$$

the natural chart constructed by the Cayley transform locally satisfies chart conditioning assumption, that is:

$$\|\Phi(\eta) - \Phi(\nu)\|_2 \leq C_G \langle \eta - \nu, \eta - \nu \rangle^{\frac{1}{2}}$$
$$\left\| D_\eta \Phi(\dot{\eta}) - D_\nu \Phi(\dot{\nu}) \right\|_2 \leq C_{\mathfrak{g}} \langle \eta - \nu, \eta - \nu \rangle^{\frac{1}{2}} + C_{\mathfrak{g}}^G \langle \dot{\eta} - \dot{\nu}, \dot{\eta} - \dot{\nu} \rangle^{\frac{1}{2}}.$$

If $\|\eta - \nu\|_2 < \varepsilon$, assumption (3.34) can be relaxed to

$$\|\eta\|_2 + \varepsilon \leq C_{con} < 1.$$

Proof. Throughout the proof of this lemma, we will make extensive use of two inequalities. The first is the bound:

$$\left\| (I + E)^{-1} \right\|_p \leq C_{con} \left(1 - \|E\|_p \right)^{-1}, \tag{3.35}$$

if $\|E\|_p < 1$, and the second is the bound:

$$\left\|(A+E)^{-1} - A^{-1}\right\|_p \le \|E\|_p \left\|A^{-1}\right\|_p^2 \left(1 - \left\|A^{-1}E\right\|_p\right)^{-1} \tag{3.36}$$

which generalizes (3.35). We begin with

$$\begin{aligned}
&\|\Phi(\eta) - \Phi(\nu)\|_2 \\
&= \left\|(I-\eta)(I+\eta)^{-1} - (I-\nu)(I+\nu)^{-1}\right\|_2 \\
&= \left\|(I-\eta)(I+\eta)^{-1} - (I-\eta)(I+\nu)^{-1} + (I-\eta)(I+\nu)^{-1} - (I-\nu)(I+\nu)^{-1}\right\|_2 \\
&= \left\|(I-\eta)\left[(I+\eta)^{-1} - (I+\nu)^{-1}\right] + [(I-\eta) - (I-\nu)](I+\nu)^{-1}\right\|_2 \\
&= \left\|(I-\eta)\left[(I+\eta)^{-1} - (I+\nu)^{-1}\right] + [\nu-\eta](I+\nu)^{-1}\right\|_2 \\
&\le \left\|(I-\eta)\left[(I+\eta)^{-1} - (I+\nu)^{-1}\right]\right\|_2 + \left\|[\nu-\eta](I+\nu)^{-1}\right\|_2.
\end{aligned} \tag{3.37}$$

Considering the term $[\nu-\eta](I+\nu)^{-1}$, we make use of (3.35) to see

$$\begin{aligned}
\left\|[\nu-\eta](I+\nu)^{-1}\right\|_2 &\le \|\nu-\eta\|_2 \left\|(I+\nu)^{-1}\right\|_2 \\
&\le (1 - \|\nu\|_2)^{-1} \|\eta-\nu\|_2.
\end{aligned} \tag{3.38}$$

Next, considering the term $\left\|(I-\eta)\left[(I+\eta)^{-1} - (I+\nu)^{-1}\right]\right\|_2$,

$$\begin{aligned}
\left\|(I-\eta)\left[(I+\eta)^{-1} - (I+\nu)^{-1}\right]\right\|_2 &\le \|I-\eta\|_2 \left\|(I+\eta)^{-1} - (I+\nu)^{-1}\right\|_2 \\
&= \|I-\eta\|_2 \left\|(I+\nu+(\eta-\nu))^{-1} - (I+\nu)^{-1}\right\|_2.
\end{aligned} \tag{3.39}$$

Applying (3.36), with $E = \eta - \nu$ and $A = I + \nu$,

$$\begin{aligned}
&\left\|(I+\nu+(\eta-\nu))^{-1} - (I+\nu)^{-1}\right\|_2 \\
&\le \|\eta-\nu\|_2 \left\|(I+\nu)^{-1}\right\|_2^2 \left(1 - \left\|(I+\nu)^{-1}(\eta-\nu)\right\|_2\right)^{-1}.
\end{aligned} \tag{3.40}$$

But

$$1-\left\|(I+\nu)^{-1}(\eta-\nu)\right\|_2 \geq 1-\left\|(I+\nu)^{-1}\right\|_2 \|(\eta-\nu)\|_2$$
$$\geq 1-(1-\|\nu\|_2)^{-1}\|(\eta-\nu)\|_2$$

which implies

$$\left(1-\left\|(I+\nu)^{-1}(\eta-\nu)\right\|_2\right)^{-1} \leq \left(1-(1-\|\nu\|_2)^{-1}\|(\eta-\nu)\|_2\right)^{-1}$$

and

$$\left\|(I+\nu)^{-1}\right\|_2^2 \leq (1-\|\nu\|_2)^{-2},$$

so

$$\begin{aligned}
&\left\|(I+\nu)^{-1}\right\|_2^2 \left(1-\left\|(I+\nu)^{-1}(\eta-\nu)\right\|_2\right)^{-1} \\
&\leq (1-\|\nu\|_2)^{-2}\left(1-(1-\|\nu\|_2)^{-1}\|(\eta-\nu)\|_2\right)^{-1} \\
&= (1-\|\nu\|_2)^{-1}((1-\|\nu\|_2)-\|\eta-\nu\|_2)^{-1} \\
&= (1-\|\nu\|_2)^{-1}(1-\|\nu\|_2-\|\eta-\nu\|_2)^{-1}. \qquad (3.41)
\end{aligned}$$

The triangle inequality gives

$$\|\eta-\nu\|_2 \leq \|\eta\|_2+\|\nu\|_2$$

and thus

$$\begin{aligned}
1-\|\eta\|_2-\|\eta-\nu\|_2 &\geq 1-2\|\eta\|_2-\|\nu\|_2 \\
(1-\|\eta\|_2-\|\eta-\nu\|_2)^{-1} &\leq (1-2\|\eta\|_2-\|\nu\|_2)^{-1}. \qquad (3.42)
\end{aligned}$$

So applying (3.42) to (3.41) gives,

$$\left\|(I+\nu)^{-1}\right\|_2^2\left(1-\left\|(I+\nu)^{-1}(\eta-\nu)\right\|_2\right)^{-1} \leq (1-\|\nu\|_2)^{-1}(1-2\|\eta\|_2-\|\nu\|_2)^{-1}, \tag{3.43}$$

then applying (3.43) to (3.40) gives,

$$\left\|(I+\nu+(\eta-\nu))^{-1}-(I+\nu)^{-1}\right\|_2 \leq \|\eta-\nu\|_2(1-\|\nu\|_2)^{-1}(1-2\|\eta\|_2-\|\nu\|_2)^{-1}, \tag{3.44}$$

and finally applying (3.44) to (3.39) yields

$$\begin{aligned}
&\left\|(I-\eta)\left[(I+\eta)^{-1}-(I+\nu)^{-1}\right]\right\|_2 \\
&\leq \|I-\eta\|_2(1-\|\nu\|_2)^{-1}(1-2\|\eta\|_2-\|\nu\|_2)^{-1}\|\eta-\nu\|_2 \\
&\leq (1-\|\eta\|_2)^{-1}(1-\|\nu\|_2)^{-1}(1-2\|\eta\|_2-\|\nu\|_2)^{-1}\|\eta-\nu\|_2
\end{aligned} \tag{3.45}$$

Substituting (3.38) and (3.45) into (3.37), we see

$$\begin{aligned}
&\|\Phi(\eta)-\Phi(\nu)\|_2 \\
&\leq \left[(1-\|\nu\|_2)^{-1}+(1-\|\eta\|_2)^{-1}(1-\|\nu\|_2)^{-1}(1-2\|\eta\|_2-\|\nu\|_2)^{-1}\right]\|\eta-\nu\|_2.
\end{aligned}$$

Hence, so as long as $2\|\eta\|+\|\nu\| \leq C_{con} < 1$

$$\|\Phi(\eta)-\Phi(\nu)\|_2 \leq C_G\|\eta-\nu\|_2.$$

where

$$C_G = \left[(1-\|\nu\|_2)^{-1}+(1-\|\eta\|_2)^{-1}(1-\|\nu\|_2)^{-1}(1-2\|\eta\|_2-\|\nu\|_2)^{-1}\right].$$

If we make the stronger assumption that $\|\eta-\nu\|_2 < \varepsilon$, we can weaken the assumption to simply $\|\eta\|_2+\varepsilon \leq C_{con} < 1$ and $\|\nu\|_2+\varepsilon \leq C_{con} < 1$. As we expect the error between the approximate and exact solutions in the Lie algebra to be orders of magnitude smaller than the magnitude of the Lie algebra elements, this is a reasonable assumption to make.

Next, to examine $\left\|D_{\eta}\Phi(\dot{\eta})-D_{\nu}\Phi(\dot{\nu})\right\|_2$, we consider the definition

$$D_X\Phi(Y)=-Y(I+X)^{-1}-(I-X)(I+X)^{-1}Y(I+X)^{-1}.$$

Using this definition,

$$\begin{aligned}
&\left\|D_{\eta}\Phi(\dot{\eta})-D_{\nu}\Phi(\dot{\nu})\right\|_2\\
&=\Big\|-\dot{\eta}(I+\eta)^{-1}-(I-\eta)(I+\eta)^{-1}\dot{\eta}(I+\eta)^{-1}+\\
&\qquad\qquad \dot{\nu}(I+\nu)^{-1}+(I-\nu)(I+\nu)^{-1}\dot{\nu}(I+\nu)^{-1}\Big\|_2\\
&\leq\left\|\dot{\nu}(I+\nu)^{-1}-\dot{\eta}(I+\eta)^{-1}\right\|_2+\\
&\qquad\qquad \left\|(I-\nu)(I+\nu)^{-1}\dot{\nu}(I+\nu)^{-1}-(I-\eta)(I+\eta)^{-1}\dot{\eta}(I+\eta)^{-1}\right\|_2.
\end{aligned}$$

Considering

$$\begin{aligned}
&\left\|\dot{\nu}(I+\nu)^{-1}-\dot{\eta}(I+\eta)^{-1}\right\|_2\\
&=\left\|\dot{\nu}(I+\nu)^{-1}-\dot{\eta}(I+\nu)^{-1}+\dot{\eta}(I+\nu)^{-1}-\dot{\eta}(I+\eta)^{-1}\right\|_2\\
&=\left\|(\dot{\nu}-\dot{\eta})(I+\nu)^{-1}+\dot{\eta}\left[(I+\nu)^{-1}-(I+\eta)^{-1}\right]\right\|_2\\
&\leq\left\|(I+\nu)^{-1}\right\|_2\|\dot{\eta}-\dot{\nu}\|_2+\|\dot{\eta}\|_2\left\|(I+\nu)^{-1}-(I+\eta)^{-1}\right\|_2\\
&\leq(1-\|\nu\|_2)^{-1}\|\dot{\eta}-\dot{\nu}\|_2+\|\dot{\eta}\|_2\left((1-\|\nu\|_2)(1-\|\nu\|_2-\|\nu-\eta\|_2)\right)^{-1}\|\eta-\nu\|_2,
\end{aligned}$$

where we have made use of (3.44) to bound $\left\|(I+\nu)^{-1}-(I+\eta)^{-1}\right\|_2$. Now, considering the second term, we first note,

$$\begin{aligned}
&\left\|(I-\nu)(I+\nu)^{-1}\dot{\nu}(I+\nu)^{-1}-(I-\eta)(I+\eta)^{-1}\dot{\eta}(I+\eta)^{-1}\right\|_2\\
&=\left\|\Phi(\nu)\dot{\nu}(I+\nu)^{-1}-\Phi(\eta)\dot{\eta}(I+\eta)^{-1}\right\|_2.
\end{aligned}$$

Using this, we see

$$\begin{aligned}
&\left\|\Phi(\nu)\dot{\nu}(I+\nu)^{-1}-\Phi(\eta)\dot{\eta}(I+\eta)^{-1}\right\|_2\\
&=\Big\|\Phi(\nu)\dot{\nu}(I+\nu)^{-1}-\Phi(\nu)\dot{\nu}(I+\eta)^{-1}
\end{aligned}$$

$$+\,\Phi(\nu)\,\dot{\nu}\,(I+\eta)^{-1}-\Phi(\eta)\,\dot{\eta}\,(I+\eta)^{-1}\Big\|_2 .$$

$$\le \left\|\Phi(\nu)\,\dot{\nu}\,(I+\nu)^{-1}-\Phi(\nu)\,\dot{\nu}\,(I+\eta)^{-1}\right\|_2$$

$$+\left\|\Phi(\nu)\,\dot{\nu}\,(I+\eta)^{-1}-\Phi(\eta)\,\dot{\eta}\,(I+\eta)^{-1}\right\|_2 \tag{3.46}$$

For the first term in (3.46),

$$\begin{aligned}
&\left\|\Phi(\nu)\,\dot{\nu}\,(I+\nu)^{-1}-\Phi(\nu)\,\dot{\nu}\,(I+\eta)^{-1}\right\|_2\\
&=\left\|\Phi(\nu)\,\dot{\nu}\left[(I+\nu)^{-1}-(I+\eta)^{-1}\right]\right\|_2\\
&\le \|\Phi(\nu)\|_2\,\|\dot{\nu}\|_2\left\|(I+\nu)^{-1}-(I+\eta)^{-1}\right\|_2\\
&\le \|\dot{\nu}\|_2\left((1-\|\nu\|_2)\,(1-\|\nu\|_2-\|\nu-\eta\|_2)\right)^{-1}\|\eta-\nu\|_2 .
\end{aligned} \tag{3.47}$$

where we once again have made use of (3.44) to bound $\left\|(I+\nu)^{-1}-(I+\eta)^{-1}\right\|_2$ and the fact that $\Phi(\nu)$ is orthogonal to set $\|\Phi(\nu)\|_2=1$. Now, considering the second term in (3.46),

$$\begin{aligned}
\left\|\Phi(\nu)\,\dot{\nu}\,(I+\eta)^{-1}-\Phi(\eta)\,\dot{\eta}\,(I+\eta)^{-1}\right\|_2 &= \left\|(\Phi(\nu)\,\dot{\nu}-\Phi(\eta)\,\dot{\eta})\,(I+\eta)^{-1}\right\|_2\\
&\le \|\Phi(\nu)\,\dot{\nu}-\Phi(\eta)\,\dot{\eta}\|_2\left\|(I+\eta)^{-1}\right\|_2
\end{aligned} \tag{3.48}$$

and additionally,

$$\begin{aligned}
\|\Phi(\nu)\,\dot{\nu}-\Phi(\eta)\,\dot{\eta}\|_2 &= \|\Phi(\nu)\,\dot{\nu}-\Phi(\eta)\,\dot{\nu}+\Phi(\eta)\,\dot{\nu}-\Phi(\eta)\,\dot{\eta}\|_2\\
&\le \|\Phi(\nu)-\Phi(\eta)\|_2\,\|\dot{\nu}\|_2+\|\Phi(\eta)\|_2\,\|\dot{\nu}-\dot{\eta}\|_2\\
&\le C_G\,\|\dot{\nu}\|_2\,\|\eta-\nu\|_2+\|\dot{\eta}-\dot{\nu}\|_2 .
\end{aligned} \tag{3.49}$$

Combining (3.47), (3.48), (3.49) in (3.46) yields

$$\left\|D_\eta\Phi(\dot{\eta})-D_\nu\Phi(\dot{\nu})\right\|_2 \le C_{\mathfrak{g}}\,\|\eta-\nu\|_2+C_{\mathfrak{g}}^{G}\,\|\dot{\eta}-\dot{\nu}\|_2$$

with constants

$$C_{\mathfrak{g}} = (1-\|\nu\|_2 - \|\nu-\eta\|_2)^{-1}\left(\|\dot{\eta}\|_2(1-\|\eta\|_2)^{-1} + \|\dot{\nu}\|_2(1-\|\nu\|_2)^{-1}\right) + C_G\|\dot{\nu}\|_2$$
$$C_{\mathfrak{g}}^G = 1 + (1-\|\nu\|_2)^{-1}.$$

To complete the proof of the lemma, we need to establish a bound on the matrix two norm from the metric on the Lie algebra. For arbitrary algebra element ξ, standard vector and matrix norm equivalences yield

$$\left\|\hat{\xi}\right\|_2 \leq \sqrt{3}\left\|\hat{\xi}\right\|_1 \leq \sqrt{3}\|\xi\|_1 \leq 3\|\xi\|_2 = 3\left\langle\hat{\xi},\hat{\xi}\right\rangle^{\frac{1}{2}}$$

which completes the proof. □

It should be noted that as $\|\nu\|_2$ or $\|\eta\|_2$ approaches 1, C_G, $C_{\mathfrak{g}}$ and $C_{\mathfrak{g}}^G$ increase without bound. This amounts to a time step restriction for the method; if the configuration changes too dramatically during the time step, the chart will become poorly conditioned and the numerical solution will degrade. However, as long as $\|\nu\|_2 \leq C_{con}$ and $\|\eta\|_2 \leq C_{con}$ on each time step for some $C_{con} < 1$ which is independent of the number of the time step, these constants will remain bounded and the natural chart will be well conditioned.

Choice of Basis Functions

The final feature of the construction of our Cayley transform Lie group Galerkin variational integrator is the choice of function space $\mathbb{M}^n([0,h],\mathfrak{g})$ for approximation of curves in the Lie algebra. Since the curves in the Lie algebra $\mathfrak{so}(3)$ that we use have a natural correspondence with curves in $\mathbb{R}^3$ through the hat map, constructing these curves reduces to choosing an approximation space for curves in $\mathbb{R}^3$.

We make the choice of polynomials of degree at most $n+1$ for $\mathbb{M}^n([0,h],\mathfrak{g})$. We choose polynomials because approximation theory and particularly the theory of spectral numerical methods, see Trefethen [36], tells us that polynomials have excellent convergence under both h and n refinement to smooth curves, and in particular, analytic curves. For the basis functions $\{\phi_i(t)\}_{i=1}^n$, we choose $\phi_i(t)$ to be the Lagrange interpolation polynomial for the i-th of n Chebyshev points rescaled to the interval $[0,h]$, that is

$$\phi_i(t) = \frac{\prod_{j=1,j\neq i}^{n}\left(t-t_j\right)}{\prod_{j=1,j\neq i}^{n}\left(t_i-t_j\right)}$$

for $t_i = \frac{h}{2}\cos\left(\frac{i\pi}{n}\right)+\frac{h}{2}$. While our convergence theory does not depend on the choice of polynomial basis, there are two major benefits for this choice of basis functions. The first is that these polynomials interpolate 0 and h, which greatly simplifies the computation of $D_1\left(R_k,R_{k+1}\right)$ and $D_2\left(R_{k-1},R_k\right)$. The second is that this choice of basis function leads to methods which are more stable than other choices of interpolation points, most likely because of the excellent stability properties of the interpolation polynomials that are constructed from them. The interested reader is referred to Trefethen [36] and Boyd [5] and the references therein for more details on spectral numerical methods.

Choice of Quadrature Rule

The final selection we must make when constructing the integrator is a choice of quadrature rule. We choose to use Gaussian quadrature, mostly because this quadrature rule is optimally accurate in the number of points and because it is simple to compute higher order Gaussian quadrature points and weights by solving a small eigenvalue problem. However, it is possible to use other rules, and we make no claim that our choice is the best for our choice of parameters.

3.4.3 Discrete Euler Poincaré Equations

While in §3.2.2 we presented a general form of the internal stage discrete Euler-Poincaré equations in coordinate-free notation, direct construction of these equations is probably not the easiest way to formulate a numerical method. This is because it requires the computation and composition of many different functions, some of which may be complicated (for example, working out $\mathbf{D}_{\alpha,\dot{\alpha}}\mathbf{D}_{\alpha}\Phi\left(\mathbf{D}_{\mathrm{q}_i}\alpha,\mathbf{D}_{\mathrm{q}_i}\dot{\alpha}\right)$ for the Cayley transform is straightforward, but also slightly obnoxious). An alternative approach, to which we alluded in §3.2.2, is to compute the discrete action in coordinates, and then explicitly compute the stationarity conditions for this discrete action. We do this here for the rigid body equations.

For the construction of the Lie group Galerkin variational integrator for the rigid body, we make use of the following functions:

$$\begin{aligned} R_n\left(\left\{\xi^i\right\}_{i=1}^n, t\right) &= R_k \Phi\left(\sum_{i=1}^n \hat{\xi}^i \phi_i(t)\right) \\ L\left(R(t), \dot{R}(t)\right) &= \operatorname{tr}\left(\dot{R}(t)^T R(t) J_d R(t)^T \dot{R}(t)\right) \\ L_d\left(R_k, R_{k+1}\right) &= \underset{\substack{R_n(t)\in\mathbb{GM}^n(R_k\times[0,h],G) \\ R_n(0)=R_k, R_n(0)=R_{k+1}}}{\operatorname{ext}} h \sum_{j=1}^m b_j L\left(R_n\left(c_j h\right), \dot{R}_n\left(c_j h\right)\right) \end{aligned} \tag{3.50}$$

where $\xi^i \in \mathbb{R}^3$. Since the curve $R_n(t)$ is a function on n points in $\mathbb{R}^3$, denoting $\xi^i = \left(\xi_a^i, \xi_b^i, \xi_c^i\right)$, we can write (3.50) as

$$\begin{aligned} L_d\left(R_k, R_{k+1}\right) = \underset{\xi^0=0, \hat{\xi}^n=\Phi^{-1}\left(R_k^T R_{k+1}\right)}{\operatorname{ext}} h \sum_{j=1}^m b_j \frac{2}{\left(1+\left\|\xi\left(c_j h\right)\right\|_2^2\right)^2} & \left(I_1 \varphi\left(\xi_c, \xi_a, \xi_b\right)^2\right. \\ & \left. + I_2 \varphi\left(\xi_b, \xi_c, \xi_a\right)^2 + I_3 \varphi\left(\xi_a, \xi_b, \xi_c\right)^2\right) \end{aligned}$$

where

$$\varphi\left(\xi_a, \xi_b, \xi_c\right) = \dot{\xi}_a\left(c_j h\right) + \xi_b\left(c_j h\right) \dot{\xi}_c\left(c_j h\right) - \xi_c\left(c_j h\right) \dot{\xi}_b\left(c_j h\right),$$

$\varphi\left(\xi_c, \xi_a, \xi_b\right)$ and $\varphi\left(\xi_b, \xi_c, \xi_a\right)$ are defined analogously, and

$$\begin{aligned} I_i &= \sum_{j\neq i} \left(J_d\right)_{jj} \\ \xi_x(t) &= \sum_i^n \xi_x^i \phi_i(t) \\ \xi(t) &= \left(\xi_a(t), \xi_b(t), \xi_c(t)\right). \end{aligned}$$

Forming the action sum as a function of the ξ^i,

$$\mathbb{S}_d\left(\left\{\xi^i\right\}_{i=1}^n\right) = h\sum_{j=1}^m b_j \frac{2}{\left(1+\left\|\xi\left(c_j h\right)\right\|_2^2\right)^2}\left(I_1\varphi\left(\xi_c,\xi_a,\xi_b\right)^2 + I_2\varphi\left(\xi_b,\xi_c,\xi_a\right)^2 + I_3\varphi\left(\xi_a,\xi_b,\xi_c\right)^2\right)$$

computing its variational derivative from ξ^i directly and setting it equal to 0,

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathbb{S}_d\left(\left\{\xi^i+\varepsilon\delta\xi^i\right\}_{i=1}^n\right)\bigg|_{\varepsilon=0} = 0,$$

gives the internal stage discrete Euler-Poincaré equations,

$$\begin{aligned} & h\sum_{j=1}^m b_j 4\left(1+\|\xi\|_2^2\right)^{-2} \\ & \Big[\left(I_3\varphi\left(\xi_c,\xi_a,\xi_b\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_c,\xi_a,\xi_b\right)\xi_a\phi_i+\dot{\xi}_b\phi_i-\xi_b\dot{\phi}_i\right)+ \\ & \left(I_2\varphi\left(\xi_b,\xi_c,\xi_a\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_b,\xi_c,\xi_a\right)\xi_a\phi_i+\xi_c\dot{\phi}_i-\dot{\xi}_c\phi_i\right)+ \\ & \left(I_1\varphi\left(\xi_a,\xi_b,\xi_c\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_a,\xi_b,\xi_c\right)\xi_a\phi_i+\dot{\phi}_i\right)\Big] = 0 \end{aligned} \tag{3.51a}$$

$$\begin{aligned} & h\sum_{j=1}^m b_j 4\left(1+\|\xi\|_2^2\right)^{-2} \\ & \Big[\left(I_3\varphi\left(\xi_c,\xi_a,\xi_b\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_c,\xi_a,\xi_b\right)\xi_b\phi_i+\xi_a\dot{\phi}_i-\dot{\xi}_a\phi_i\right)+ \\ & \left(I_2\varphi\left(\xi_b,\xi_c,\xi_a\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_b,\xi_c,\xi_a\right)\xi_b\phi_i+\dot{\phi}_i\right)+ \\ & \left(I_1\varphi\left(\xi_a,\xi_b,\xi_c\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_a,\xi_b,\xi_c\right)\xi_b\phi_i+\dot{\xi}_c\phi_i-\xi_c\dot{\phi}_i\right)\Big] = 0 \end{aligned} \tag{3.51b}$$

$$\begin{aligned} & h\sum_{j=1}^m b_j 4\left(1+\|\xi\|_2^2\right)^{-2} \\ & \Big[\left(I_3\varphi\left(\xi_c,\xi_a,\xi_b\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_c,\xi_a,\xi_b\right)\xi_c\phi_i+\dot{\phi}_i\right)+ \\ & \left(I_2\varphi\left(\xi_b,\xi_c,\xi_a\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_b,\xi_c,\xi_a\right)\xi_c\phi_i+\dot{\xi}_a\phi_i-\xi_a\dot{\phi}_i\right)+ \\ & \left(I_1\varphi\left(\xi_a,\xi_b,\xi_c\right)\right)\left(-2\left(1+\|\xi\|_2^2\right)^{-1}\varphi\left(\xi_a,\xi_b,\xi_c\right)\xi_c\phi_i+\xi_b\dot{\phi}_i-\dot{\xi}_b\phi_i\right)\Big] = 0, \end{aligned} \tag{3.51c}$$

for $i = 2, \ldots, n-1$, and where we have suppressed the t argument on all of our functions. Solving these equations, along with the condition $\xi^1 = 0$ and the discrete Euler-Poincaré equations

$$D_1 L_d\left(R_k, R_{k+1}\right) + D_2 L_d\left(R_{k-1}, R_k\right) = 0, \tag{3.52}$$

which we will discuss in §3.4.3, yields $\tilde{R}(t)$, the stationary point of the discrete action. Using this stationary point, computing $\tilde{R}(h) = R_{k+1}$ gives us the next step of our one-step map.

Momentum Matching

As we mentioned in our general derivation of the discrete Euler-Poincaré equations, (3.52) must be treated with care. We described in an expedient method for computing $D_1 L_d\left(R_k, R_{k+1}\right)$ so that the result is compatible with our change of natural charts. We will provide an explicit example below.

We already know the expression for $D_1 L_d\left(R_k, R_{k+1}\right)$ for the coordinates in the current natural chart, the vector of the form (3.51a) – (3.51c), with $i = 1$. This is the map $\frac{\partial L_d}{\partial \xi_k}$ described in §3.2.2. Now, we need to compute an expression for λ_k and $\frac{\partial \xi_k}{\partial \lambda_k}$. Given $\xi_0 = \left(\xi_a^0, \xi_b^0, \xi_c^0\right)$ and $\xi_k = \left(\xi_a, \xi_b, \xi_c\right)$, we compute λ by

$$\hat{\lambda} = \Phi^{-1}\left(\Phi\left(\hat{\xi}_0\right)\Phi\left(\hat{\xi}_k\right)\right)$$

which gives in coordinates $\lambda = \left(\lambda_a, \lambda_b, \lambda_c\right)$,

$$\begin{aligned}
\lambda_a &= \frac{-\xi_a - \xi_a^0 + \xi_c \xi_b^0 - \xi_b \xi_c^0}{-1 + \xi_a^0 \xi_a + \xi_b^0 \xi_b + \xi_c^0 \xi_c} \\
\lambda_b &= \frac{-\xi_b - \xi_b^0 + \xi_a \xi_c^0 - \xi_c \xi_a^0}{-1 + \xi_a^0 \xi_a + \xi_b^0 \xi_b + \xi_c^0 \xi_c} \\
\lambda_c &= \frac{-\xi_c - \xi_c^0 + \xi_b \xi_a^0 - \xi_a \xi_b^0}{-1 + \xi_a^0 \xi_a + \xi_b^0 \xi_b + \xi_c^0 \xi_c}.
\end{aligned}$$

Now, we recompute ξ_k in terms of λ,

$$\hat{\xi}_k = \Phi^{-1}\left(\left(\Phi\left(\hat{\xi}_0\right)\right)^{-1}\Phi\left(\hat{\lambda}\right)\right)$$

which, when expressed in coordinates $\xi_k = (\xi_a, \xi_b, \xi_c)$, gives

$$\begin{aligned}
\xi_a &= \frac{\lambda_a - \xi_a^0 + \lambda_c\xi_b^0 - \lambda_b\xi_c^0}{1+\lambda_a\xi_a^0+\lambda_b\xi_b^0+\lambda_c\xi_c^0} \\
\xi_b &= \frac{\lambda_b - \xi_b^0 + \lambda_a\xi_c^0 - \lambda_c\xi_a^0}{1+\lambda_a\xi_a^0+\lambda_b\xi_b^0+\lambda_c\xi_c^0} \\
\xi_c &= \frac{\lambda_c - \xi_c^0 + \lambda_b\xi_a^0 - \lambda_a\xi_b^0}{1+\lambda_a\xi_a^0+\lambda_b\xi_b^0+\lambda_c\xi_c^0}.
\end{aligned}$$

So, to compute $D_1L_d(R_k, R_{k+1}) = \left(\frac{\partial L_d}{\partial \lambda_a}, \frac{\partial L_d}{\partial \lambda_a}, \frac{\partial L_d}{\partial \lambda_a}\right)$, we can take the easily computed expression $\frac{\partial L_d}{\partial \xi_k}$ and apply a change of coordinates computation,

$$\begin{aligned}
\frac{\partial L_d}{\partial \lambda_a} &= \frac{\partial L_d}{\partial \xi_a}\frac{\partial \xi_a}{\partial \lambda_a} + \frac{\partial L_d}{\partial \xi_b}\frac{\partial \xi_b}{\partial \lambda_a} + \frac{\partial L_d}{\partial \xi_c}\frac{\partial \xi_c}{\partial \lambda_a} \\
\frac{\partial L_d}{\partial \lambda_b} &= \frac{\partial L_d}{\partial \xi_a}\frac{\partial \xi_a}{\partial \lambda_b} + \frac{\partial L_d}{\partial \xi_b}\frac{\partial \xi_b}{\partial \lambda_b} + \frac{\partial L_d}{\partial \xi_c}\frac{\partial \xi_c}{\partial \lambda_b} \\
\frac{\partial L_d}{\partial \lambda_c} &= \frac{\partial L_d}{\partial \xi_a}\frac{\partial \xi_a}{\partial \lambda_c} + \frac{\partial L_d}{\partial \xi_c}\frac{\partial \xi_b}{\partial \lambda_c} + \frac{\partial L_d}{\partial \xi_c}\frac{\partial \xi_c}{\partial \lambda_c},
\end{aligned}$$

which is the momentum matching condition expressed so that it is compatible with the change of natural charts.

3.5 Numerical Experiments

Thus far, we have discussed the construction of Lie group Galerkin variational integrators, and established bounds on their rate of convergence. We will now turn to several numerical examples to demonstrate that our methods behave in practice as our theory predicts.

3.5.1 Cayley Transform Method for the Rigid Body

In §3.4 we have discussed in great detail a specific construction of a Lie group Galerkin variational integrators for the free rigid body based on the Cayley transform. Based on the convergence results from Theorems 3.3.1 and 3.3.2, we would expect our construction to converge geometrically with n-refinement and optimally with h-refinement.

Using MATLAB, we implemented the Lie group Galerkin variational integrator described in §3.4, using a finite-difference Newton method as a root finder. We used the parameters

$$\begin{aligned} J_d &= \operatorname{diag}(1.3, 2.1, 1.2) \\ R(0) &= I \\ R^T(0)\dot{R}(0) &= \left(2.0, \widehat{-1.9}, 1.0\right)^T. \end{aligned}$$

To establish convergence, we first computed a numerical solution using a low-order splitting method with a very small time step, and once we established that the Lie group Galerkin variational integrator's solution and the splitting method's numerical solutions agreed, we used a Lie group Galerkin variational integrator solution with $n = 26$ and $h = 0.5$ as a high-order approximation to the exact solution, and established convergence to this solution. We made this choice of parameters for our approximate exact solution because it appeared that for this choice of parameters, the residual from the nonlinear solver was the dominant source of error, and neither h nor n refinement improved our numerical solution.

The results, which are summarized in Figures 3.1 – 3.4, establish the rates of convergence predicted in Theorems 3.3.1 and 3.3.2. For n-refinement, we see that our integrator did indeed achieve geometric convergence, as can be seen in Figure 3.1. However, unlike the vector space method (see Hall and Leok [16]), we did not observe the difference in convergence rates of the continuous approximation and the one-step map. We suspect this is because until very high accuracy is achieved, the inaccurate boundary conditions due to the one-step map error dominates the continuous approximation error, and the threshold at which the continuous approximation error is greater than the one-step error is related to the time step. While we can take extremely large time steps in the vector space case, in the Lie group case the time step length is limited by the natural chart, and hence we never observe the lower convergence rate of the continuous approximation. We explore convergence with h-refinement, see Figure 3.2, and observe the optimal rate of convergence for our construction for even n. However, for odd n, we see convergence at a rate of $n-1$. We do not have a clear explanation for this.

Now, considering the geometric invariants related to the rigid body, we see that the Cayley transform based method has excellent conservation properties. Figure 3.4 shows one of the classic depictions of geometric invariants for the rigid body, that is the intersection of the two hypersurfaces in momentum space given by the two geometric invariants $C(y) = \frac{1}{2}\sum_{i=1}^{3} y_i^2$ and $H(y) = \frac{1}{2}\sum_{i=1}^{3} I_i^{-1} y_i^2$ where y is the angular momentum of the rigid body. These invariants correspond to the norm of the body fixed angular momentum and the energy, respectively. Discussions of these invariants, and comparable behavior of other methods can be found in Marsden and Ratiu [31] and Hairer et al. [14] (specifically, see Hairer et al. [14] for a comparison to other numerical methods). Our method has nearly perfect conservation of these invariants.

Additionally, while it is not perfectly conserved, the energy behavior of our method is oscillatory and remains bounded even for very long integration times, as can be seen in Figure 3.3. This type of behavior is typical for variational integrators, and can be understood in terms of backwards error analysis.

3.5.2 Cayley Transform Method for the 3D Pendulum

For a second numerical experiment, we examine the 3D pendulum. The 3D pendulum is the rigid body with one point fixed and under the influence of gravity, and its Lagrangian is:

$$\begin{aligned}
L\left(R,\dot{R}\right) &= \frac{1}{2}\operatorname{tr}\left(\dot{R}^T R J_d R^T \dot{R}\right) + mge_3^T R\rho \\
J_d &= \operatorname{diag}\left(1, 2.8, 2\right) \\
\rho &= \left(0,0,1\right)^T
\end{aligned}$$

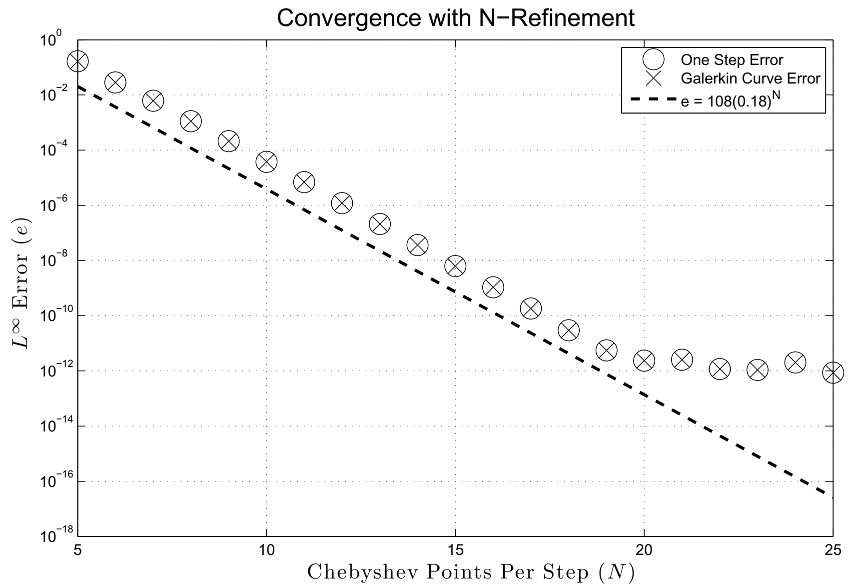


Figure 3.1: Geometric convergence of the Lie group spectral variational integrator based on the Cayley transform for the rigid body. We use a constant step-size $h = 0.5$. Note that the Galerkin curves have the same error as the one-step map, even though they have a theoretical lower rate of convergence.

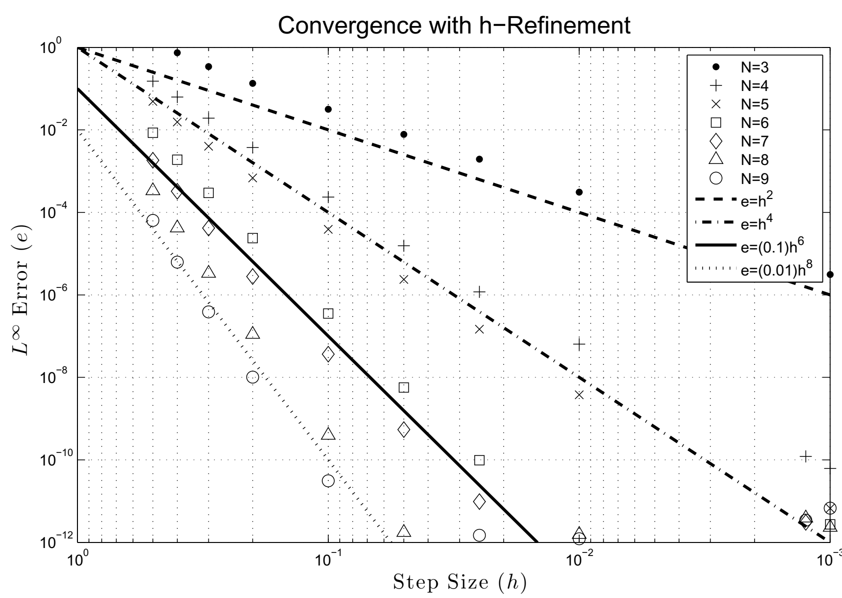


Figure 3.2: Order optimal convergence of the Lie group Galerkin variational integrator based on the Cayley transform for the rigid body.

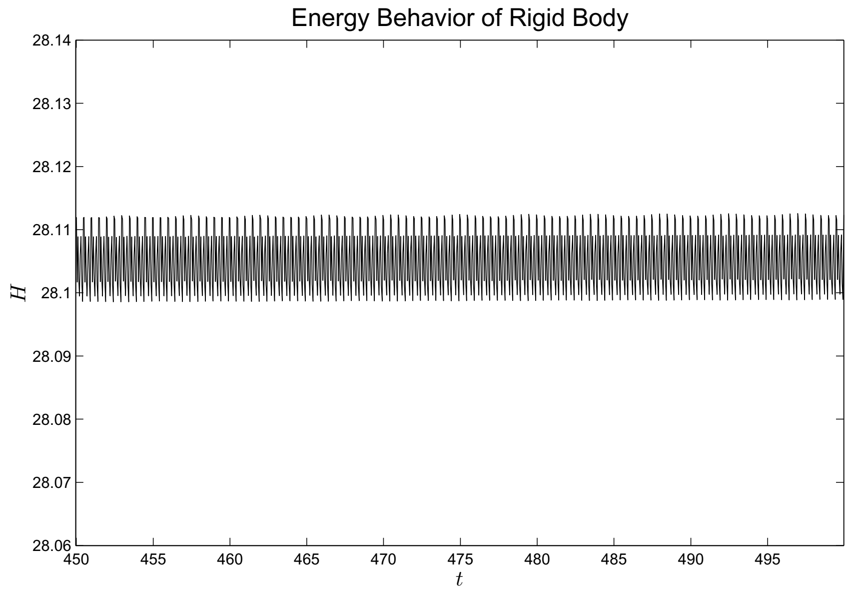


Figure 3.3: Energy behavior of the Lie group Galerkin variational integrator based on the Cayley transform for the rigid body. This is from a simulation starting at $t_0 = 0.0$, and using the parameters $n = 12$, $h = 0.5$ for the integrator.

where ρ is center of mass for $R = I$, m is the mass of the pendulum and g is the gravitational constant. We consider two sets of initial conditions, the first,

$$R(0) = I$$
$$R(0)^T \dot{R}(0) = \widehat{(0.5, -0.5, 0.4)^T},$$

which is a slight perturbation from stable equilibrium, and the second

$$R(0) = \operatorname{diag}(-1, 1, -1)$$
$$R(0)^T \dot{R}(0) = \widehat{(0.5, -0.5, 0.4)^T}$$

which is the pendulum slightly perturbed from its unstable equilibrium.

We construct the variational integrator for this system using the Cayley transform. This involves adding the term $V(\xi(t)) = mge_3^T L_{g_k} \Phi(\xi(t)) \rho$ to the discrete action in equations (3.50), and finding the stationarity conditions of this new discrete action, which gives us the new internal stage discrete Euler-Poincaré equations. These have the same form as equations (3.51a) – (3.51c), with added terms for the potential.

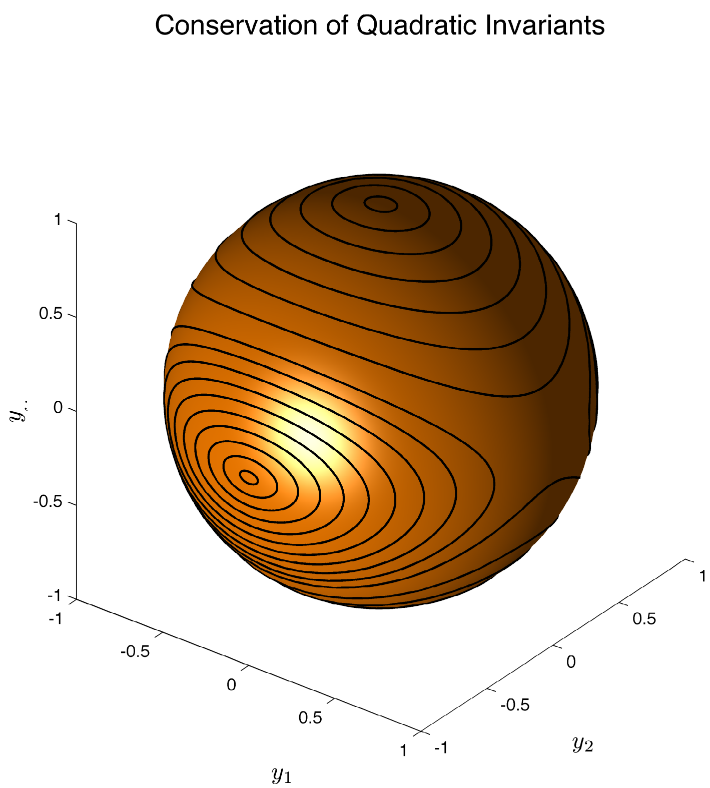


Figure 3.4: Conserved quantities for the Lie group Galerkin variational integrator based on the Cayley transform for the rigid body. This is from a series of computations using the parameters $n = 8$, and $h = 0.5$, from a variety of initial conditions. Note that the trajectories computed by the Lie group Galerkin variational integrators, which are the black curves, lie on the intersections of $\sum_{i=1}^{3} y_i^2 = 1$, $\sum_{i=1}^{3} I_i^{-1} y_i^2 = 2$, which are the norm of angular momentum and energy, respectively.

For the first set of initial data, which are near the stable equilibrium, we see exactly the expected convergence with both h and N refinement, as is illustrated in Figures 3.5 and 3.6. Furthermore, we see bounded oscillatory energy behavior over the length of the integration, as in Figure 3.7.

For the second set of initial data, this system evolves chaotically, so convergence of individual trajectories is not of great interest. What is more important is the conservation of geometric invariants as the system evolves. As can be seen from Figures 3.8 and 3.9, the energy of the system is nearly conserved, even with very aggressive time stepping. Of particular note is that even though there are many steps where the solution undergoes a change that approaches the limit on the conditioning of the natural chart, the energy error remains small.

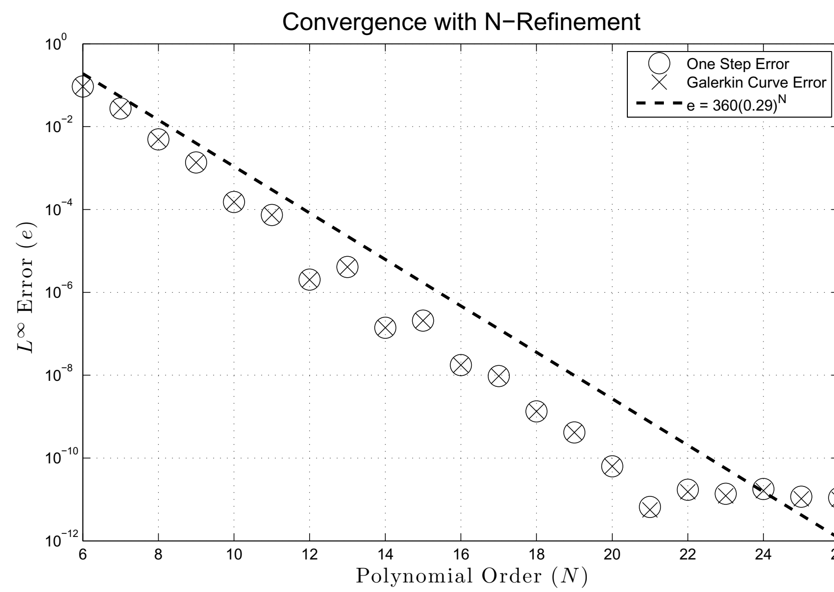


Figure 3.5: Geometric convergence of the Lie group spectral variational integrator based on the Cayley transform for the 3D pendulum for a small perturbation from the stable equilibrium. We use the time step $h = 0.5$. Note that, once again, the Galerkin curves have the same error as the one-step map.

3.6 Conclusions and Future Work

In this paper, we have presented a new numerical method for Lagrangian problems on Lie groups. Specifically, we used a Galerkin construction to create variational

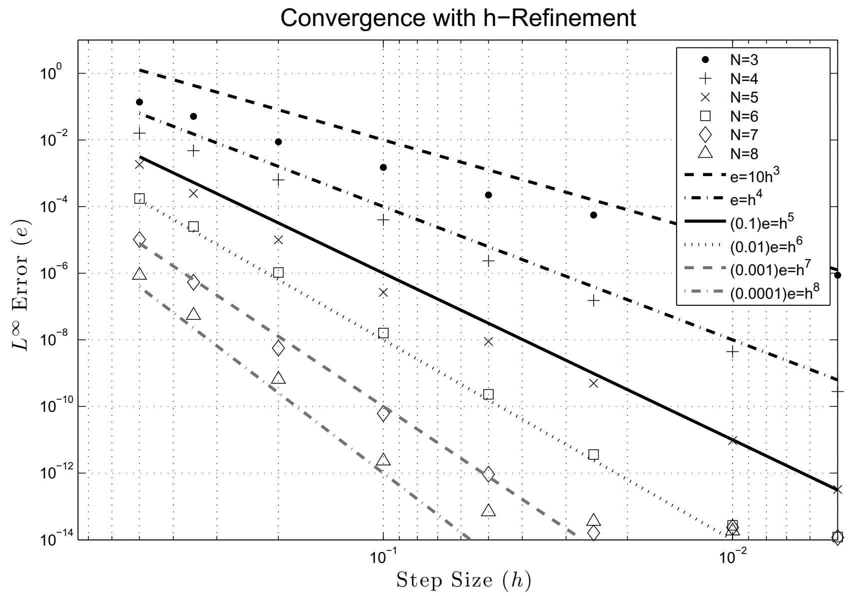


Figure 3.6: Order optimal convergence of Lie group Galerkin variational integrator based on the Cayley transform for the 3D pendulum for a small perturbation from the stable equilibrium. Note that we have almost exactly order optimal convergence.

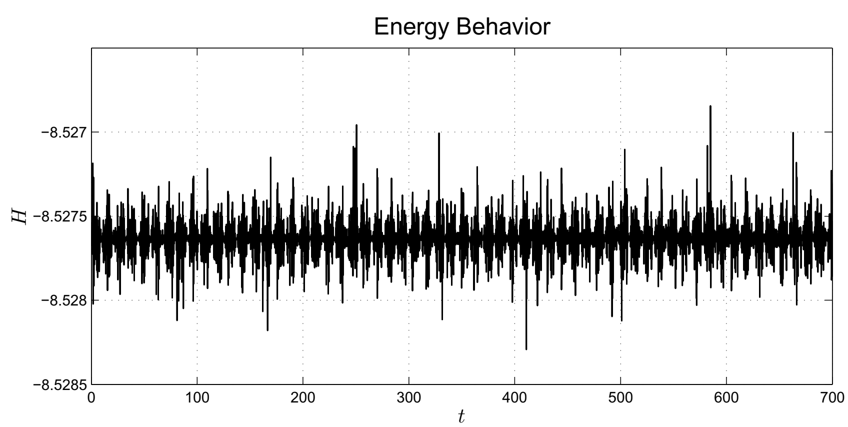


Figure 3.7: Energy behavior of the Lie group Galerkin variational integrator based on the Cayley transform for the 3D pendulum for a small perturbation from the stable equilibrium. This is the behavior of an integrator constructed with parameters $n = 8$, step size $h = 1.5$. Note that the error is both small and oscillatory, but not increasing.

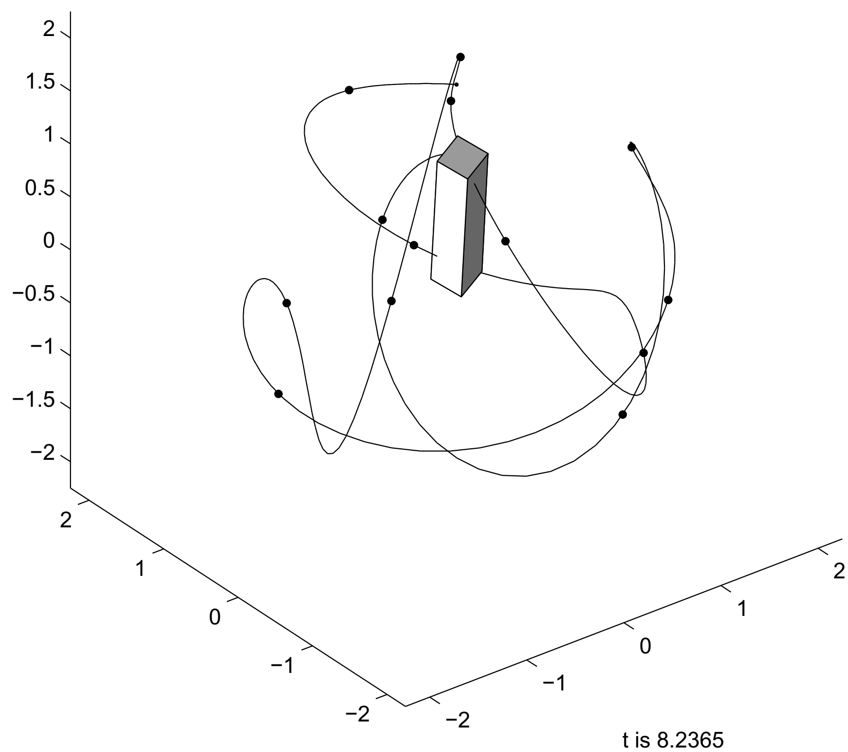


Figure 3.8: Dynamics of the numerical simulation of the 3D pendulum constructed from a Lie group Galerkin variational integrator. These dynamics were constructed from an integrator with $n = 20$, $h = 0.6$. The black dots each represent a single step of the one-step map, and the solid lines are the Galerkin curves. Note that some of the steps are almost through an angle of length π, which is the limit of the conditioning of the natural chart.

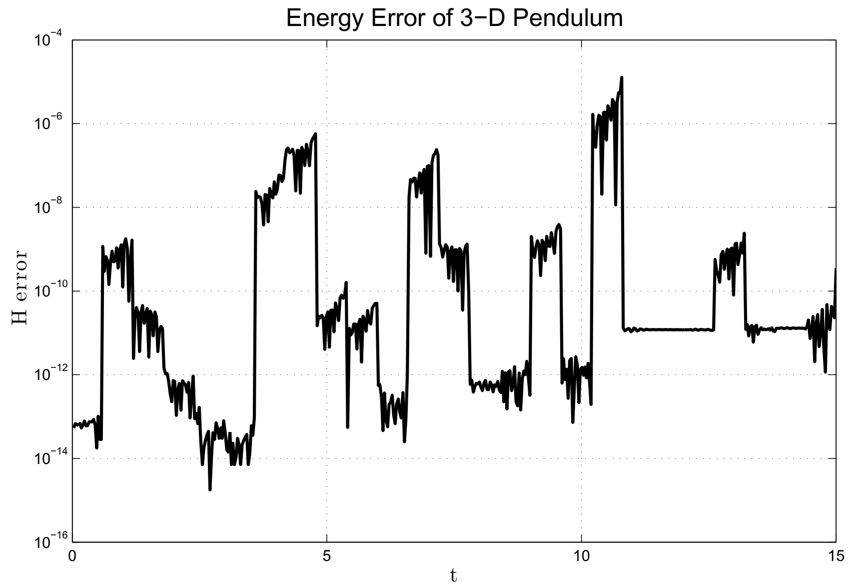


Figure 3.9: Energy error of the dynamics depicted in Figure 3.8. The large jumps in error are associated with time steps that almost exceed the conditioning of the natural chart.

integrators of arbitrarily high-order, and also Lie group spectral variational integrators, which converge geometrically. We demonstrated that in addition to inheriting the excellent geometric properties common to all variational integrators, which include conservation of the symplectic form, and conservation of momentum, that such integrators also are extremely stable even for large time steps, can be adapted for a large class of problems, and yield highly accurate continuous approximations to the true trajectory of the system.

We also gave an explicit example of a Lie group Galerkin variational integrator constructed using the Cayley transform. Using this construction, we demonstrated the expected rates of convergence on two different example problems, the rigid body and the 3D pendulum. We also showed that these methods both have excellent energy and momentum conservation properties. Additionally, we provided explicit expressions for the internal stage discrete Euler-Poincaré equations for the free rigid body, which form the foundation of a numerical method for a variety of problems.

3.6.1 Future Work

Symplectic integrators continue to be an area of interest, and there has been considerable success in developing high-order structure-preserving methods and applying such methods to relevant problems. While we have developed a significant amount of the theory of Lie group Galerkin variational integrators, there is considerable future work to be done.

Choice of Natural Charts

In our construction, we chose the Cayley transform to construct our natural chart. While this choice made the derivation of the resulting integrator simpler, it also introduced a limitation on the conditioning of the natural chart. A possible extension of our framework would be constructions based on natural charts constructed from other functions. An obvious choice is the exponential map, which was the choice of chart function used in earlier works that proposed this construction. A comparison of the behavior of integrators constructed from other choices of natural chart functions would be interesting further work.

Novel Variational Integrators

One of the attractive features of our work is that we establish an optimality result for arbitrary approximation spaces. Because of this, our results hold for a variety of different possible constructions of variational integrators. It would be interesting to investigate the behavior of variational integrators constructed from novel approximation spaces, such as wavelets, or for variational integrators that make use of specialized function spaces, such as spaces that include both high and low frequency functions for problems with components that evolve on different time scales.

Larger classes of Problems

In this paper, we have focused most of our attention on the rigid body and problems that evolve on $SO(3)$. However, there are many examples of Lie group problems that evolve on other spaces. Our analysis suggests that the Galerkin approach would be effective for these problems. It would be interesting to examine Galerkin variational integrators for problems that evolve on other Lie groups, and apply our methods to other interesting applications.

Parallel Implementation and Computational Efficiency

Perhaps our method's greatest flaw is that it requires the solution of a large number of nonlinear equations at every time step. This problem is further exasperated by the fact that assembling the Newton matrix at every time step requires the repeated application of a high-order quadrature rule. While the fact that our method is stable even for very large time steps helps to overcome this computational difficulty, it would be interesting to see how much our method could be accelerated by assembling our Newton matrix in parallel.

Multisymplectic Variational Integrators

Multisymplectic geometry has become an increasingly popular framework for extending much of the geometric theory from classical Lagrangian mechanics to Lagrangian PDEs. The foundations for a discrete theory have been laid, and there have been significant results achieved in geometric techniques for structured problems such as elasticity, fluid mechanics, non-linear wave equations, and computational electromagnetism. However, there is still significant work to be done in the areas of construction of numerical methods, analysis of discrete geometric structure, and especially error analysis. Galerkin type methods have become a standard in classical numerical PDE methods, popular examples include Finite-Element, Spectral, and Pseudospectral methods. The variational Galerkin framework could provide a natural framework for extending these classical methods to structure-preserving geometric methods for PDEs.

3.7 Appendix

In §3.3.1, we stated Theorem 3.3.2 but did not provide a proof. This is because the proof is essentially the same as that for optimal convergence, with slight and obvious modifications. For completeness, we will provide the proof here.

Theorem 3.7.1. *Given an interval $[0,h]$, and a Lagrangian $L : TG \to \mathbb{R}$, suppose that $\bar{g}(t)$ solves the Euler-Lagrange equations on that interval exactly. Furthermore, suppose that the exact solution $\bar{g}(t)$ falls within the range of the natural chart, that is:*

$$\bar{g}(t) = L_{g_k}\Phi(\bar{\eta}(t))$$

for some $\bar{\eta} \in C^2([0,h],\mathfrak{g})$. For the function space $\mathbb{M}^n([0,h],\mathfrak{g})$ and the quadrature rule $\mathscr{G}$, define the Galerkin discrete Lagrangian $L_d^G(g_0,g_1) \to \mathbb{R}$ as

$$\begin{aligned} L_d^G(g_0,g_1,n) &= \operatorname*{ext}_{\substack{g_n(t)\in\mathbb{GM}^n(g_0\times[0,h],G)\\ g_n(0)=g_0,g_n(h)=g_1}} h\sum_{j=1}^{m} b_j L\left(g_n\left(c_j h\right),\dot{g}_n\left(c_j h\right)\right) \\ &= h\sum_{h=1}^{m} b_j L\left(\tilde{g}_n\left(c_j h\right),\dot{\tilde{g}}_n\left(c_j h\right)\right) \end{aligned} \tag{3.53}$$

where $\tilde{g}_n(t)$ is the extremizing curve in $\mathbb{GM}^n(g_0\times[0,h],G)$. If:

1. *there exists an approximation $\hat{\eta}_n \in \mathbb{M}^n([0,h],\mathfrak{g})$ such that,*

$$\begin{aligned} \langle\bar{\eta}-\hat{\eta}_n,\bar{\eta}-\hat{\eta}_n\rangle^{\frac{1}{2}} &\leq C_A K_A^n \\ \left\langle\dot{\bar{\eta}}-\dot{\hat{\eta}}_n,\dot{\bar{\eta}}-\dot{\hat{\eta}}_n\right\rangle^{\frac{1}{2}} &\leq C_{\mathfrak{A}} K_A^n, \end{aligned}$$

   for some constants $C_A \geq 0$ and $C_{\mathfrak{A}} \geq 0$, $0 < K_A < 1$ independent of n,

2. *the Lagrangian L is Lipschitz in the chosen error norm in both its arguments, that is:*

$$|L(g_1,\dot{g}_1)-L(g_2,\dot{g}_2)| \leq L_\alpha\left(e_g(g_1,g_2)+e_a(\dot{g}_1,\dot{g}_2)\right)$$

3. *the chart function Φ is well-conditioned in $e_g(\cdot,\cdot)$ and $e_a(\cdot,\cdot)$, that is (3.10) and (3.11) hold,*

4. *there exists a sequence of quadrature rules $\{\mathscr{G}_n\}_{n=1}^{\infty}$,*

$$\mathscr{G}_n(f) = h\sum_{j=1}^{m_n} b_{n_j} f\left(c_{n_j}h\right) \approx \int_0^h f(t)\,dt,$$

and there exists a constant $0 < K_g < 1$ *independent of n such that,*

$$\left| \int_0^h L(g_n(t), \dot{g}_n(t))\, dt - h \sum_{j=1}^{m_n} b_{n_j} L\left(g_n\left(c_{n_j} h\right), \dot{g}_n\left(c_{n_j} h\right)\right) \right| \leq C_g K_g^n$$

for any $g_n(t) = L_{g_0} \Phi(\xi(t))$ *where* $\xi(t) \in \mathbb{M}^n([0,h], \mathfrak{g})$.

5. *the stationary points of the discrete action and the continuous action are minimizers,*

then the variational integrator induced by $L_d^G(g_0, g_1, n)$ *has error* $\mathscr{O}(K_s^n)$ *for some constant* K_s *independent of n,* $0 < K_s < 1$.

Proof. We begin by rewriting the exact discrete Lagrangian and the Galerkin discrete Lagrangian:

$$\left| L_d^E(g_0, g_1, n) - L_d^G(g_0, g_1, h) \right| = \left| \int_0^h L(\bar{g}, \dot{\bar{g}})\, \mathrm{d}t - h \sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{g}_n\left(c_{n_j} h\right), \dot{\tilde{g}}_n\left(c_{n_j} h\right)\right) \right|,$$

where we have introduced $\tilde{g}_n$, which is the stationary point of the local Galerkin action (3.53). We introduce the solution in the approximation space which takes the form $\hat{g}_n(t) = L_{g_k} \Phi(\hat{\eta}(t))$, and compare the action on the exact solution to the action on this solution:

$$\begin{aligned}
\left| \int_0^h L(\bar{g}, \dot{\bar{g}})\, \mathrm{d}t - \int_0^h L\left(\hat{g}_n, \dot{\hat{g}}_n\right) \mathrm{d}t \right| &= \left| \int_0^h L(\bar{g}, \dot{\bar{g}}) - L\left(\hat{g}_n, \dot{\hat{g}}_n\right) \mathrm{d}t \right| \\
&\leq \int_0^h \left| L(\bar{g}, \dot{\bar{g}}) - L\left(\hat{g}_n, \dot{\hat{g}}_n\right) \right| \mathrm{d}t.
\end{aligned}$$

Now, we use the Lipschitz assumption to establish the bound

$$\begin{aligned}
&\int_0^h \left| L(\bar{g}, \dot{\bar{g}}) - L\left(\hat{g}_n, \dot{\hat{g}}_n\right) \right| \mathrm{d}t \\
&\leq \int_0^h L_\alpha \left(e_g(\bar{g}, \hat{g}_n) + e_a\left(\dot{\bar{g}}, \dot{\hat{g}}_n\right) \right) \mathrm{d}t \\
&= \int_0^h L_\alpha \left(e_g\left(L_{g_k}\Phi(\bar{\eta}), L_{g_k}\Phi(\hat{\eta}_n)\right) + \right. \\
&\qquad \left. e_a\left(D_{\Phi(\bar{\eta})} L_{g_0} D_{\bar{\eta}} \Phi\left(\dot{\bar{\eta}}\right), D_{\Phi(\hat{\eta}_n)} L_{g_0} D_{\hat{\eta}_n} \Phi\left(\dot{\hat{\eta}}_n\right) \right) \right) \mathrm{d}t,
\end{aligned}$$

and the chart conditioning assumptions to see

$$\begin{aligned}\int_0^h \left|L(\bar{g},\dot{\bar{g}}) - L\left(\hat{g}_n,\dot{\hat{g}}_n\right)\right| \mathrm{d}t &\leq \int_0^h L_\alpha \Bigg(C_G \langle \bar{\eta} - \hat{\eta}_n, \bar{\eta} - \hat{\eta}_n \rangle^{\frac{1}{2}} + C_{\mathfrak{g}} \left\langle \dot{\bar{\eta}} - \dot{\hat{\eta}}_n, \dot{\bar{\eta}} - \dot{\hat{\eta}}_n \right\rangle^{\frac{1}{2}} + \\ &\qquad C_{\mathfrak{g}}^G \langle \bar{\eta} - \hat{\eta}_n, \bar{\eta} - \hat{\eta}_n \rangle^{\frac{1}{2}} \Bigg) \mathrm{d}t \\ &\leq \int_0^h L_\alpha \left(C_G C_A K_A^n + C_{\mathfrak{g}} C_{\mathfrak{A}} K_A^n + C_{\mathfrak{g}}^G C_A K_A^n \right) \mathrm{d}t \\ &= L_\alpha \left(\left(C_G + C_{\mathfrak{g}}^G \right) C_A + C_{\mathfrak{g}} C_{\mathfrak{A}} \right) K_A^n.\end{aligned}$$

This establishes a bound between the action evaluated on the exact discrete Lagrangian and the optimal solution in the approximation space. Considering the Galerkin discrete action,

$$\begin{aligned}h\sum_{j=1}^{m_n} b_{n_j} L(\tilde{g}_n,\tilde{g}_n) &\leq h\sum_{j=1}^{m_n} b_{n_j} L\left(\hat{g}_n,\dot{\hat{g}}_n\right) \\ &\leq \int_0^h L\left(\hat{g}_n,\dot{\hat{g}}_n\right)\mathrm{d}t + C_g K_g^n \\ &\leq \int_0^h L(\bar{g},\dot{\bar{g}})\,\mathrm{d}t + C_g K_g^n + L_\alpha \left(\left(C_G + C_{\mathfrak{g}}^G \right) C_A + C_{\mathfrak{g}} C_{\mathfrak{A}} \right) K_A^n \qquad (3.54)\end{aligned}$$

where we have used the assumption that the Galerkin approximation minimizes the Galerkin discrete action and the assumption on the accuracy of the quadrature. Now, using the fact that $\bar{g}(t)$ minimizes the action and that $\mathbb{GM}^n(g_0 \times [0,h], G) \subset C^2([0,h], G)$,

$$\begin{aligned}h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{g}_n,\dot{\tilde{q}}_n\right) &\geq \int_0^h L\left(\tilde{g}_n,\dot{\tilde{g}}_n\right)\mathrm{d}t - C_g K_g^n \\ &\geq \int_0^h L(\bar{g},\bar{g})\,\mathrm{d}t - C_g K_g^n. \qquad (3.55)\end{aligned}$$

Combining inequalities (3.54) and (3.55), we see that,

$$\begin{aligned}\int_0^h L(\bar{g},\dot{\bar{g}})\,\mathrm{d}t - C_g K_g^n &\leq h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{g}_n,\dot{\tilde{g}}_n\right) \\ &\leq \int_0^h L(\bar{g},\dot{\bar{g}})\,\mathrm{d}t + C_g K_g^n + L_\alpha \left(\left(C_G + C_{\mathfrak{g}}^G \right) C_A + C_{\mathfrak{g}} C_{\mathfrak{A}} \right) K_A^n\end{aligned}$$

which implies

$$\left|\int_0^h L(\bar{g},\dot{\bar{g}})\,\mathrm{d}t - h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{g}_n,\dot{\tilde{g}}_n\right)\right| \leq \left(C_g + L_\alpha\left(\left(C_G + C_{\mathfrak{g}}^G\right)C_A + C_{\mathfrak{g}} C_{\mathfrak{A}}\right)\right)K_s^n \quad (3.56)$$

where $K_s = \max\left(K_A, K_g\right)$. The left hand side of (3.56) is exactly

$$\left|L_d^E\left(g_0,g_1,h\right) - L_d^G\left(g_0,g_1,n\right)\right|,$$

and thus

$$\left|L_d^E\left(g_0,g_1,h\right) - L_d^G\left(g_0,g_1,n\right)\right| \leq \left(C_g + L_\alpha\left(\left(C_G + C_{\mathfrak{g}}^G\right)C_A + C_{\mathfrak{g}} C_{\mathfrak{A}}\right)\right)K_s^n.$$

This states that the Galerkin discrete Lagrangian approximates the exact discrete Lagrangian with error $\mathscr{O}\left(K_s^n\right)$, and by Theorem (3.1.1) this further implies that the Lagrangian update map has error $\mathscr{O}\left(K_s^n\right)$. □

References

[1] G. Bogfjellmo and H. Marthinsen. Symplectic Lie group methods. *ArXiv e-prints*, March 2013.

[2] N. Bou-Rabee and J. E. Marsden. Hamilton-Pontryagin integrators on Lie groups. I. Introduction and structure-preserving properties. *Found. Comput. Math.*, 9(2): 197–219, 2009.

[3] N. Bou-Rabee and H. Owhadi. Stochastic variational integrators. *IMA J. Numer. Anal.*, 29(2):421–443, 2009.

[4] N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.*, 48(1):278–297, 2010.

[5] J. P. Boyd. *Chebyshev and Fourier spectral methods*. Dover Publications Inc., Mineola, NY, second edition, 2001.

[6] C. L. Burnett, D. D. Holm, and D. M. Meier. Geometric integrators for higher-order mechanics on Lie groups. *ArXiv e-prints*, December 2011.

[7] Elena Celledoni and Brynjulf Owren. Lie group methods for rigid body dynamics and time integration on manifolds. *Computer Methods in Applied Mechanics and Engineering*, 192(3 - 4):421 – 438, 2003. ISSN 0045-7825. doi:

10.1016/S0045-7825(02)00520-0. URL http://www.sciencedirect.com/science/article/pii/S0045782502005200.

[8] J. Cortés and S. Martínez. Non-holonomic integrators. *Nonlinearity*, 14(5):1365–1392, 2001.

[9] G. Dal Maso. *An introduction to* Γ*-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[10] D. Estep and D. French. Global error control for the continuous Galerkin finite element method for ordinary differential equations. *RAIRO Modél. Math. Anal. Numér.*, 28(7):815–852, 1994.

[11] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[12] Y. N. Fedorov and D. V. Zenkov. Discrete nonholonomic LL systems on Lie groups. *Nonlinearity*, 18(5):2211–2241, 2005.

[13] R. Fetecau, J. E. Marsden, M. Ortiz, and M. West. Nonsmooth Lagrangian mechanics and variational collision integrators. *SIAM Journal on Applied Dynamical Systems*, 2(3):381–416, 2003.

[14] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[15] N. Hale and L. N. Trefethen. New quadrature formulas from conformal maps. *SIAM J. Numer. Anal.*, 46(2):930–948, 2008.

[16] J. Hall and M. Leok. Spectral Variational Integrators. *ArXiv e-prints*, November 2012.

[17] B. L. Hulme. One-step piecewise polynomial Galerkin methods for initial value problems. *Math. Comp.*, 26:415–426, 1972.

[18] Arieh Iserles, Hans Z. Munthe-Kaas, Syvert P. NÃÿrsett, and Antonella Zanna. Lie-group methods. *Acta Numerica 2000*, 9:215–365, 0 2000. ISSN 1474-0508.

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

[20] S. Larsson and V. Thomée. *Partial differential equations with numerical methods*, volume 45 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 2003.

[21] T. Lee, M. Leok, and N. H. McClamroch. Lie group variational integrators for the full body problem. *Comput. Methods Appl. Mech. Engrg.*, 196(29-30):2907–2924, 2007.

[22] T. Lee, M. Leok, and N. H. McClamroch. Lagrangian mechanics and variational integrators on two-spheres. *Internat. J. Numer. Methods Engrg.*, 79(9):1147–1174, 2009.

[23] Taeyoung Lee, N.H. McClamroch, and M. Leok. A lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3d pendulum. In *Control Applications, 2005. CCA 2005. Proceedings of 2005 IEEE Conference on*, pages 962–967, 2005. doi: 10.1109/CCA.2005.1507254.

[24] M. Leok. Generalized Galerkin variational integrators: Lie group, multiscale, and pseudospectral methods. (preprint, arXiv:math.NA/0508360), 2004.

[25] M. Leok and T. Shingel. Prolongation–collocation variational integrators. *IMA J. Numer. Anal.*, 32(3):1194–1216, 2012.

[26] M. Leok and T. Shingel. General techniques for constructing variational integrators. *Frontiers of Mathematics in China*, 7(2):273–303, 2012. (Special issue on computational mathematics, invited paper).

[27] M. Leok and J. Zhang. Discrete Hamiltonian variational integrators. *IMA J. Numer. Anal.*, 31(4):1497–1532, 2011.

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

[29] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001.

[30] J. E. Marsden, G. W. Patrick, and S. Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Comm. Math. Phys.*, 199(2):351–395, 1998.

[31] J.E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. Texts in Applied Mathematics. Springer, 1999. ISBN 9780387986432. URL http://books.google.com/books?id=I2gH9ZIs-3AC.

[32] R. McLachlan and M. Perlmutter. Integrators for nonholonomic mechanical systems. *J. Nonlinear Sci.*, 16(4):283–328, 2006.

[33] S. Müller and M. Ortiz. On the Γ-convergence of discrete dynamics and variational integrators. *J. Nonlinear Sci.*, 14(3):279–296, 2004.

[34] NASA. JPL Horizons system online solar system data and ephemeris, September 2012. URL http://ssd.jpl.nasa.gov/?horizons.

[35] G. W. Patrick and C. Cuell. Error analysis of variational integrators of unconstrained Lagrangian systems. *Numer. Math.*, 113(2):243–264, 2009.

[36] L. N. Trefethen. *Spectral methods in MATLAB*, volume 10 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[37] J. Vankerschaver, C. Liao, and M. Leok. Generating functionals and Lagrangian PDEs. *Journal of Nonlinear Science*, 2011. (submitted, arXiv:1111.0280 [math-ph]).

Chapter 3, in full, is currently being prepared for submission for publication of the material. James Hall and Melvin Leok. The dissertation author was the primary investigator and author of this material.

Chapter 4

Conclusions

4.1 Summary of Work

This dissertation presented a new framework for developing high order structure preserving numerical methods, using tools from spectral methods, Galerkin methods, and discrete mechanics. Techniques from these fields were used to construct structure preserving methods that share some of the best properties of these theories; specifically, highly stable, symplectic and momentum preserving methods that have optimal convergence. A brief summary of the major results of this dissertation are summarized below, and precise statements of these results can be found in Chapters 2 and 3.

1. If Q is a vector space, for Lagrangians of the form

   $$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q),$$

   as long as the quadrature rule is sufficiently accurate, Galerkin variational integrators will converge at the same rate as the best approximation in the approximation space used to construct them.

2. If Q is a vector space, if the stationary points of the action for a certain Lagrangian over the interval $[0,h]$ are minimizers, as long as the quadrature rule is sufficiently accurate, Galerkin variational integrators will converge at the same rate as the best approximation in the approximation space used to construct them.

3. If Q is a vector space, for Lagrangians of the form

   $$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q),$$

   then the Galerkin curves $q_n(t)$ which are the stationary points of the local discrete action converge to the true flow with error $\mathscr{O}\left(h^{\frac{p}{2}}\right)$ if the Galerkin variational integrator has error $\mathscr{O}(h^p)$ and with error $\mathscr{O}\left(\sqrt{K^n}\right)$ if the Galerkin variational integrator has error $\mathscr{O}(K^n)$.

4. If Q is $SO(3)$, for Lagrangians of the form

   $$L(g,\dot{q}) = \dot{g}^T g J_d g^T \dot{g} - V(g),$$

as long as the quadrature rule is sufficiently accurate and the natural chart is well conditioned, Lie group Galerkin variational integrators will converge at the same rate as the best approximation in the approximation space used to construct them.

5. If Q is a Lie group, if the stationary points of the action for a certain Lagrangian over the interval $[0,h]$ are minimizers, as long as the quadrature rule is sufficiently accurate and the natural chart is well conditioned, Lie group Galerkin variational integrators will converge at the same rate as the best approximation in the approximation space used to construct them.

6. If Q is a Lie group G, as long as the action of the Lagrangian over $[0,h]$ is coercive, then the Galerkin curves $g(t)$ will converge to the true flow with error $\mathscr{O}\left(h^{\frac{p}{2}}\right)$ if the Galerkin variational integrator has error $\mathscr{O}(h^p)$ and with error $\mathscr{O}\left(\sqrt{K}^n\right)$ if the Galerkin variational integrator has error $\mathscr{O}(K^n)$,

While the Galerkin approach used here has been proposed before, this is the first work that explores a general error analysis for these types of methods. Previous works have either developed error theories that are specific to certain constructions, or have not discussed the error analysis of these methods at all. Furthermore, in this dissertation the Galerkin construction and the general error theory were extended to Lie group problems. The cumulative result is a general framework for constructing high quality numerical methods for problems with structure, and a rigorous error analysis for this framework.

4.2 Future Applications

One of the surprising and powerful features of Galerkin variational integrators is that they are capable of producing high quality solutions even with very large time-steps. For example, when computing a numerical simulation of the Solar System, these methods can produce highly accurate simulations even using time-steps which are greater than the orbital period of Mercury. Furthermore, because time-step shortening is not necessary to achieve convergence, it is possible to produce solutions using Galerkin integrators that are more accurate than explicit methods even when the time-step used for the Galerkin integrators is several orders of magnitude times larger than those of the explicit methods. By increasing the efficiency of these methods, highly accurate long term simulations may be achievable with these methods that are simply not possible using lower order methods. Investigating the extent to which these methods can be accelerated remains an avenue of future work.

Additionally, the error analysis for Galerkin variational integrators creates the opportunity for the development of many new and novel methods. By connecting the accuracy of the numerical method to the best possible approximation in an approximation space, this error analysis connects the behavior of Galerkin variational integrators to approximation theory. If there exists an understanding the behavior of the dynamics of a Lagrangian system acquired through analytic means, Galerkin variational integrators allow for the construction of integrators which are specially designed for specific problems, in addition to the ability to construct high order integrators. Such methods could have applications for problems with dynamics on different scales or problems with periodic behavior. A simple example is using high order polynomials to discretize components with dynamics which evolve on short time scales and low order polynomials to discretize components with dynamics on longer time scales, which reduces the dimension of the discrete problem without sacrificing accuracy. Another interesting direction for future research would be to investigate the behavior of Galerkin variational integrators on such problems, and using the Galerkin framework to construct variational integrators based on novel approximation spaces.