

DEVELOPMENT OF MULTISTEP AND DEGENERATE  
VARIATIONAL INTEGRATORS FOR APPLICATIONS  
IN PLASMA PHYSICS

CHARLES LELAND ELLISON

A DISSERTATION  
PRESENTED TO THE FACULTY  
OF PRINCETON UNIVERSITY  
IN CANDIDACY FOR THE DEGREE  
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE  
BY THE DEPARTMENT OF  
ASTROPHYSICAL SCIENCES  
PROGRAM IN PLASMA PHYSICS  
ADVISERS: HONG QIN, WILLIAM TANG

APRIL 2016

ProQuest Number: 10090246

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10090246

Published by ProQuest LLC (2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

© Copyright by Charles Leland Ellison, 2016.

All rights reserved.

## Abstract

Geometric integrators yield high-fidelity numerical results by retaining conservation laws in the time advance. A particularly powerful class of geometric integrators is symplectic integrators, which are widely used in orbital mechanics and accelerator physics. An important application presently lacking symplectic integrators is the guiding center motion of magnetized particles represented by non-canonical coordinates. Because guiding center trajectories are foundational to many simulations of magnetically confined plasmas, geometric guiding center algorithms have high potential for impact. The motivation is compounded by the need to simulate long-pulse fusion devices, including ITER, and opportunities in high performance computing, including the use of petascale resources and beyond.

This dissertation uses a systematic procedure for constructing geometric integrators - known as variational integration - to deliver new algorithms for guiding center trajectories and other plasma-relevant dynamical systems. These variational integrators are non-trivial because the Lagrangians of interest are *degenerate* - the Euler-Lagrange equations are first-order differential equations and the Legendre transform is not invertible. The first contribution of this dissertation is that variational integrators for degenerate Lagrangian systems are typically *multistep methods*. Multistep methods admit parasitic mode instabilities that can ruin the numerical results. These instabilities motivate the second major contribution: *degenerate variational integrators*. By replicating the degeneracy of the continuous system, degenerate variational integrators avoid parasitic mode instabilities. The new methods are therefore robust geometric integrators for degenerate Lagrangian systems.

These developments in variational integration theory culminate in one-step degenerate variational integrators for non-canonical magnetic field line flow and guiding center dynamics. The guiding center integrator assumes coordinates such that one component of the magnetic field is zero; it is shown how to construct such coordinates for nested magnetic surface configurations. Additionally, collisional drag effects are incorporated in the variational

guiding center algorithm for the first time, allowing simulation of energetic particle thermalization. Advantages relative to existing canonical-symplectic and non-geometric algorithms are numerically demonstrated. All algorithms have been implemented as part of a modern, parallel, ODE-solving library, suitable for use in high-performance simulations.

---

## Acknowledgements

*I imagine these moments as fragments or splinters scattered across a lifetime. If a person could somehow collect them all up and stick them together he would have a perfect hour or even a perfect day. And I think in that hour or day he would be closer to the mystery of what it is to be human. It would be like having a glimpse of heaven.*

—Daniel Tammet, *Born on a Blue Day*

Throughout my graduate tenure I have collected many such “splinters” of cherished moments, each of which can be attributed to the extraordinary individuals with whom I’ve had the fortune of crossing paths. To all those explicitly named herein or otherwise implied: thank you.

To my advisers: thank you for your guidance, wisdom, and selfless investments of time during this stage of my journey toward scientific maturity. Tobin, thank you for inspiring my career in plasma physics. The combination of rational thinking and insatiable curiosity you demonstrated during my time at CU remain my ideal of what it means to be a physicist. Yevgeny, thank you for welcoming me as an eager first year student, providing abundant support for my success, and pushing me to continually improve my capabilities. Hong, you reinforced a life-long lesson that passion is the greatest guide for success, and eagerly supported the value in my ideas. Thank you for the invitations to collaborate with your colleagues in China; I will never forget these trips. Bill, thank you for keeping me focused on achieving results with real, tangible impact. Thank you also for the many enthralling discussions and for being an unwavering advocate on my behalf. John, I am so grateful for your willingness to dive into the technical trenches of this research with me with little more motivation than scientific curiosity. You taught me an optimistic mentality of perseverance,

and it was a joy to see your versatile physics skills at work. All of you have shaped my personal and professional growth in life changing ways.

To my research colleagues at and outside of PPPL: thank you for being delightful to work with and for patiently teaching me physics lessons. Roscoe, thank you for reading my dissertation, supporting my thesis proposal, and for the personal ORBIT lessons. Stewart, thank you for being willing to discuss magnetic equilibria, numerical methods, and anything else troubling me at the time. Gerrit, I had lots of fun discussing SPIRAL and envisioning numerical improvements together. Michael, it was a blast working with you in China and over many Skype sessions; I hope we have not seen the last of either.

To my graduate peers: thank you for being both inspiring colleagues and my dearest friends. The many of you that preceded me - Jess, Jeff, Luc, Laura, Paul, Craig, Josh, Martin, Clayton, et al. - were always happy to lend words of advice and humble encouragement. You led an example of humility and hard work, and instilled in me confidence that I too could survive this academic gauntlet.

To my immediate classmates: you were an *unbelievably* fun group, and I know we will be lifelong friends. Brendan, it was great sharing a European adventure with you, including French bicycle tours, late night conversations at famous attractions, and a spiteful meal at McDonald's. Tyler, you have shared two joys that are closest to my heart: Starcraft and hoppy beers. These spawned many great memories involving one, the other, or both (Barcraft NYC!). Vasily, thank you for enduring our inferior prelims capabilities and for many fun encounters at parties, conferences, and the lab. Matt, you are responsible for many of my most hilarious (and unspeakable) memories, including our prospective visit night, rugby raffle winnings, and multiple things related to cats. Thank you for the laughs. Nik, the good times we've shared are too numerous to count - goat roasts, an adventure to China, Brooklyn beer gardens... none of these would have been the same without you.

To my graduate colleagues who started later than me: you embody a bright future for the program. Jono and Josh, you may be a younger year, but I look up to you both and it's fitting that you beat me to the finish line. Jono, thank you for the New Zealand treats and many laughs. Josh, my head spins every time we have a technical conversation; much of what I have learned was taught by you - mostly about mathematics but also about spicy food. Jake, I'm so glad our "pentarch" tenures overlapped, and will vividly remember your hosting me in San Diego and for epic game nights. Thank you to my colleagues from China - Yao, Lei, Yuan, Chang, Lan, Ge, Qian, - for being great company during eye-opening culinary excursions. Yao, thank you for being a generous host *both* times I went to China. Daniel, thank you for your humility at the blackboard and ruthlessness at the green table - you know the one. Peter, thank you for slip-n-slide, duck hunt, and starting the now famous ping pong league. Seth, I had a blast tackling GeoPIC with you and raging at many 12-PL and otherwise plasma-related events. Brian, to you I entrust the "party office guy", "guy from Colorado", "Hall thruster guy", and "non-native-Chinese-guy-who-speaks-some-Mandarin" titles. Lee - to you I entrust the revered "Lee of the program" title. Jack - to you I entrust the "head brewer" and "summer student speaker" titles. Noah - thank you for being a party officer, computer guy, and for the many lessons at the ping pong table. Jonathan - it was a blast getting crushed by you and your left-handed alter-ego at afternoon games of pong. Eric - thank you for being a fan of Starcraft and Atticus. Denis - thank you for generously welcoming me to Canadian thanksgiving and for heading up epic lunch excursions. And the rest of you - Jacob, Jeff, Charles, Eugene, James, Jon, Scott, Peter J., and all the first years - I carry fond memories of our shared time in the program.

In addition to the student body, I am tremendously grateful for the support offered by the Graduate Program. Barbara, I cannot state enough how reassuring it was to have you a phone call away for any and all concerns about graduate life. The Autism Speaks walks were a blast each year, and I will forever be reminded of you whenever a burnt hot dog is

to be found. I wish you the greatest of adventures in your retirement. Beth, you have been a delightful addition to the program, and I have no doubt you will be equally cherished by the student passing under your watch. Thank you for all of the extra help as I got close to finishing!

To Kelsey: my greatest delight in graduate school may well have been meeting someone who understood me as if our whole lives had been shared up until the day we met. You supported me when I doubted myself, celebrated with me when I had moments of success, and shared countless meaningful moments. I treasure you and the times we've shared.

To my family: I never would have gotten here, neither the first day nor the last, without your ceaseless support. Dad, you set the wheels in motion by incorporating counting in my earliest life lessons, and your work ethic set a high bar for what I expect of myself. Mom, you placed my well-being at top priority from day one, taught me to seize the day, and ingrained in me that laughter is the greatest coping mechanism. Both of you placed education as a top priority for our family, and you taught me that pursuing my passions would lead to success. John, your thirst for learning set a backdrop for many of the lessons I've acquired in this journey, and I look forward to continuing to learn from you across our delightfully distinct paths.

Thank you all.

Dedicated to my mother and father -  
for making many sacrifices to provide many opportunities.

# Contents

Abstract . . . . .	iii
Acknowledgements . . . . .	v
<b>1 Introduction</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 Motivation . . . . .	5
1.3 Contributions . . . . .	12
<b>2 Background</b>	<b>19</b>
2.1 Overview . . . . .	19
2.2 Variational Integrators . . . . .	26
2.2.1 Variational Principles . . . . .	28
2.2.2 Legendre Transforms . . . . .	32
2.2.3 Symplecticity . . . . .	37
2.2.4 Dissipation in Variational Principles . . . . .	43
2.2.5 Correspondence between Continuous and Discrete Mechanics . . . . .	49
2.3 Action Principles in Phase Space . . . . .	55
2.3.1 Hamiltonian Systems on Symplectic Manifolds . . . . .	56
2.3.2 Action Principles in Phase Space . . . . .	58
2.3.3 Phase-Space Lagrange-d'Alembert . . . . .	62

2.3.4	Applications in Plasma Physics . . . . .	63
2.3.5	Variational Integrators for Action Principles in Phase Space . . . . .	76
2.4	Dissertation Scope . . . . .	77
<b>3</b>	<b>Multistep Variational Integrators</b>	<b>78</b>
3.1	Introduction . . . . .	79
3.2	Definitions . . . . .	85
3.3	Stability . . . . .	91
3.3.1	Spectral Analysis . . . . .	91
3.3.2	Backward Error Analysis . . . . .	99
3.4	Structure Preservation . . . . .	118
3.4.1	Multistep Advance . . . . .	119
3.4.2	Underlying One-Step Method . . . . .	123
3.5	Accuracy . . . . .	128
3.6	Parasitic Mode Mitigation . . . . .	137
3.6.1	Backward Error Initialization . . . . .	137
3.6.2	Re-initialization . . . . .	139
3.6.3	Identification of the Underlying One-Step Method . . . . .	142
3.7	Discussion . . . . .	146
3.7.1	Summary . . . . .	146
3.7.2	Prospects for Future Work . . . . .	147
<b>4</b>	<b>Degenerate Variational Integrators</b>	<b>149</b>
4.1	Introduction . . . . .	150
4.2	Definitions . . . . .	154
4.3	Stability Analysis . . . . .	157
4.4	Structure Preservation . . . . .	159

4.5 Accuracy . . . . .	161
4.6 Applications . . . . .	163
4.6.1 Magnetic Field Line Flow . . . . .	164
4.6.2 Non-Canonical Guiding Center Dynamics . . . . .	167
4.7 Discussion . . . . .	171
4.7.1 Summary . . . . .	171
4.7.2 Future Work . . . . .	171
<b>5 Applications in Plasma Physics</b>	<b>174</b>
5.1 MVI for Nearly-Integrable Non-Canonical Systems . . . . .	174
5.1.1 Motivation . . . . .	175
5.1.2 Algorithm . . . . .	176
5.1.3 Results . . . . .	179
5.1.4 Discussion . . . . .	183
5.2 DVI for Magnetic Field Line Flow . . . . .	184
5.2.1 Motivation . . . . .	184
5.2.2 Algorithm . . . . .	185
5.2.3 Numerical Results . . . . .	186
5.2.4 Discussion . . . . .	189
5.3 DVI for Guiding Center Trajectories . . . . .	191
5.3.1 Algorithms . . . . .	193
5.3.2 Numerical Results . . . . .	196
5.3.3 Discussion . . . . .	204
<b>6 Conclusion</b>	<b>208</b>
6.1 Summary . . . . .	208
6.2 Future Work . . . . .	211

<b>A Mathematical Preliminaries</b>	<b>213</b>
A.1 Introduction . . . . .	216
A.2 Tangent Vectors and One-Forms . . . . .	217
A.3 Tensors . . . . .	222
A.3.1 The Metric Tensor . . . . .	222
A.4 Differential Forms . . . . .	225
A.4.1 Wedge Product . . . . .	225
A.4.2 Exterior Derivative . . . . .	229
A.4.3 Symplectic Two-Forms . . . . .	233
A.5 Push-Forward and Pull-Back . . . . .	237
<b>B GEODES Library</b>	<b>241</b>
B.1 Overview . . . . .	241
B.2 Features . . . . .	242
B.2.1 Language . . . . .	242
B.2.2 Style . . . . .	242
B.2.3 Documentation . . . . .	243
B.2.4 Version Control . . . . .	245
B.2.5 Design . . . . .	245
B.2.6 Unit testing . . . . .	246
B.2.7 Input Specification . . . . .	246
B.2.8 Parallelization . . . . .	247
<b>Bibliography</b>	<b>249</b>

# Chapter 1

## Introduction

*Since the fabric of the universe is most perfect, and is the work of a most wise Creator, nothing whatsoever takes place in the universe in which some relation of maximum and minimum does not appear.*

—Leonhard Euler, *Method of Finding Plane Curves that Show Some Property of Maximum and Minimum*

### 1.1 Overview

As physics has advanced in developing predictive mathematical descriptions of nature, an astounding array of phenomena has been described within two theoretical frameworks: those of Lagrangian and Hamiltonian mechanics [1, 2, 3, 4, 5]. Lagrangian mechanics describes motion as maximizing or minimizing a quantity known as “action”, while the dual perspective of Hamiltonian mechanics refers to a mathematical structure known as a “Poisson bracket”. From the large scale motion of astrophysical objects to the small scale jittering of fundamental particles, these same two mathematical formalisms describe their behavior. Even the counter-intuitive discoveries of modern physics admit descriptions within these frameworks - quantum mechanics, general relativity, and quantum field theory all admit

Lagrangian and/or Hamiltonian descriptions. Many of the triumphs of theoretical physics have been to demonstrate how a new natural phenomenon may be described within these two formalisms, or to simultaneously advance the understanding of many dynamical systems by deriving results pertinent to any Lagrangian or Hamiltonian system.

In an era in which numerical modeling grows increasingly prevalent for understanding complex systems, it is a challenge for computational physicists to (i) retain central features of Lagrangian and Hamiltonian mechanics in the numerical models and (ii) leverage the well-developed tools of these formalisms to resolve outstanding problems. The principal obstacle to achieving these objectives is that the numerical dynamics are inherently discrete, whereas the calculus upon which Lagrangian and Hamiltonian mechanics are built relies on infinities and infinitesimals. Remarkable progress has already been made: symplectic integrators can replicate the Hamiltonian character of the orbits of astrophysical objects, accurate over thousands of millennia [6, 7]. Still, much remains to be done.

For instance, one of the most important Hamiltonian systems in plasma physics does not have any integrators that are both symplectic and formulated in terms of the most natural coordinates for the problem. This system is the guiding center motion of magnetized particles in electric and magnetic fields [8]. Symplectic integrators are important for capturing the Hamiltonian character of the guiding center system [6], and formulation in terms of the natural coordinates is critical for computational efficiency. Because the guiding center motion of magnetized particles is fundamental to many simulations of magnetically confined fusion plasmas, efficient new symplectic algorithms for guiding center trajectories have a huge potential for impact.

This dissertation explores how a recently developed approach to systematically constructing symplectic integrators can be applied to the guiding center system. The systematic approach is referred to as “variational integration” [9]. Unlike conventional numerical methods, which make approximations to the equations of motion, variational integration makes

all approximations to the action principle used to derive the equations of motion. Usually, the “variational integrator” constructed by approximating the action principle is equivalent to a symplectic integrator. An important question is then: *What happens when a variational integrator is constructed by approximating the guiding center action principle?*

Interestingly, variational integration of the guiding center system is far from trivial. This is because the guiding center action principle involves a *degenerate* Lagrangian. When a Lagrangian is degenerate, the Legendre transform is not invertible and the Euler-Lagrange equations do not yield a well-defined system of second-order differential equations [4]. The conventional theory of variational integrators *assumes* the Lagrangian is *not* degenerate [9]. In particular, the equivalence between a variational integrator and a symplectic integrator hinges on the Legendre transform being invertible [9]. To develop variational integrators for the guiding center system, it became necessary to describe more thoroughly the consequences of discretizing action principles with degenerate Lagrangians.

The first such consequence, and first contribution of this dissertation, is the discovery that variational integrators constructed from degenerate Lagrangians are typically *multistep methods*. Multistep methods require additional initial conditions than those supplied by the initial value problem, for example two initial conditions may be required for a multistep integration of a first-order differential equation [10, 11]. This excess dependence has significant numerical analysis implications [6, 11, 12]. Studying and explaining these implications in the context of variational integrators merits designating a new class of numerical methods: *multistep variational integrators (MVIs)*. The analysis performed on MVIs reveals the presence of detrimental numerical instabilities known as “parasitic modes”. The realization of the multistep character of variational integrators constructed from degenerate-Lagrangian systems is therefore critical for analyzing existing methods [13, 14, 15, 16, 17, 18, 19] and for developing new ones.

The second major contribution of this dissertation is the discovery that multistep variational integrators are not an inevitable consequence of discretizing action principles with degenerate Lagrangians; a newly discovered alternative is *degenerate variational integrators (DVIs)*. A variational integrator is degenerate when the *discrete* Legendre transform is not invertible. When a variational integrator is degenerate, the set of initial conditions that must be supplied is reduced and parasitic mode instabilities are in turn eliminated. In the context of guiding center algorithms, capturing the degeneracy of the guiding center Lagrangian when constructing a variational integrator is therefore highly desirable. At present, it is not yet known how to construct a degenerate variational integrator for guiding center trajectories in arbitrary magnetic fields. However, it is shown how to construct a degenerate variational guiding center integrator for a wide class of magnetic fields, including those with nested flux surfaces. No symplectic or variational integrators were previously known for this class of fields and in the natural guiding center coordinates. The new degenerate variational integrators are then an important step forward for the guiding center system and many other systems with degenerate Lagrangians.

From a mathematics perspective, the discovery and analysis of multistep and degenerate variational integrators expands our understanding of discrete Lagrangian mechanics and broadens the applicability of variational methods. From a plasma physics perspective, efficient and robust new algorithms are constructed for the time advance of guiding center trajectories. To further increase the practical utility of these new algorithms, this dissertation also shows how dissipative collisional effects may be incorporated in the new variational integrators. Although the dissipative effects violate the Hamiltonian character of the guiding center system, the dissipation is slow relative to the other characteristic timescales of the problem. It is therefore important that the dissipation exhibited by a simulation is caused by the *physical* dissipation sources and not an ambiguous combination of physical and numerical dissipation. Fair, equal-CPU-expense comparisons are performed between the new

variational guiding center algorithms and industry standard techniques. The new algorithms are shown to be computationally competitive for modeling energetic particle thermalization, collisional drag effects included. These results make a compelling case that careful retention of mathematical properties in plasma physics algorithms yields high performance methods, appropriate for simulating the flagship fusion experiments of today and tomorrow.

## 1.2 Motivation

The contributions - and therefore motivations - of this dissertation are interdisciplinary in nature, spanning applied mathematics and plasma physics. In brief, the intent is to develop and apply tools from applied mathematics to deliver effective new algorithms for computational plasma physics. More specifically, the goal is to apply *geometric* methods to uncover algorithms that retain fundamental qualitative characteristics in numerical approximations of the dynamics. From an applied mathematics perspective, success of this goal expands the theory of discrete mechanics and broadens the applicability of variational methods. From a plasma physics perspective, improved algorithms are obtained for energetic test particle calculations and for the particle advances central to many kinetic simulations.

A major theme of modern physics is that the behavior of natural phenomena is dictated by *geometry* [2, 3, 4, 20]. A “geometric” property may be understood as independent of the particular coordinates used to represent said property. Because the choice of coordinates is at the whim of the physicist, the most fundamental properties of physical systems are geometric. Examples of these fundamental properties include symmetries of the equations of motion and, correspondingly, quantities that are invariant along a solution to these equations. These invariants of motion - which can include energy, momentum, and angular momentum - strongly influence the resulting dynamics. For instance, one can deduce

whether a Hamiltonian system will exhibit chaos by determining how many invariants of motion are present in the system [21].

When numerically modeling dynamical systems, the qualitative behavior of the simulated result will depend on which, if any, geometric properties are retained by the numerical approximation. Numerical methods that capture some geometric property of the dynamical system are deemed *geometric integrators* [6, 7]. Geometric integrators may be further categorized by the specific properties they possess; there exists extensive work on energy-conserving [22, 23], momentum-conserving [24, 25, 26], and volume-preserving [27, 28, 29, 30] integrators, for example. An especially powerful class of geometric integrators is *symplectic integrators* [6, 7, 9, 27, 31, 32, 33, 34, 35, 36, 37]. Symplectic integrators are also volume-preserving, often conserve momentum, and typically exhibit bounded energy error. Hamiltonian systems have long been known to be *symplectic* [2, 3, 9] - a property relating to the preservation of areas in phase space. If one approximates the solution to a Hamiltonian system using a symplectic integrator, then the numerical trajectory will itself be Hamiltonian<sup>1</sup> [6]. Such a strong statement cannot be made based on the conservation of energy, momentum or phase-space volume alone. Important benefits of symplectic integrators (subject to technical caveats) include: bounded energy error for time-independent systems, retention of KAM surfaces in the phase portrait, reduced rates of error accumulation along integrable trajectories, and time-computed averages along chaotic trajectories that are nearby to the true value [6, 7]. Because of these benefits, certain Hamiltonian phenomena such as Arnold diffusion [21, 38, 39, 40] may be reliably captured only by using symplectic integrators [41, 42]. The practical importance of symplectic integration is further evidenced by its extensive use in, for instance, celestial mechanics [43, 44] and accelerator physics [35, 45].

One important application domain presently lacking symplectic algorithms is the guiding center motion of magnetized particles described as a Hamiltonian system in “non-canonical”

---

<sup>1</sup>More rigorously, it will be *nearby* the solution of some Hamiltonian system in the sense that the asymptotic series obtained via backward error analysis is Hamiltonian at every truncation order.

coordinates. This dynamical system lies at the heart of several cornerstone numerical tools in plasma physics. For instance, so-called “test particle” simulations evolve energetic particle distributions in prescribed electric and magnetic fields, typically using the guiding center drift description [46, 47, 48, 49, 50, 51, 52, 53]. Active research topics addressed by test particle calculations include confinement loss induced by error fields and MHD modes [54, 55, 56], heating efficiency [49], the stabilization/destabilization of tokamak modes [57, 58, 59, 60, 61, 55, 62, 63, 64, 65], and runaway electron dynamics [66, 67]. The advance of guiding center trajectories is also fundamental to particle-based drift-kinetic and gyro-kinetic simulations [68, 69]. These kinetic simulations are widely used for the study of turbulence and transport in tokamaks and stellarators [68, 69, 70, 71, 72]. Because the particles and fields are alternately advanced, the particle advance stage of these simulations is essentially equivalent to that of a test particle calculation. Any algorithmic advances to guiding center trajectories may therefore be deployed in particle-based drift- and gyro-kinetic simulations.

Moving forward, trends in high performance computing suggest that test particle and kinetic models will only increase in capabilities and use. Supercomputer architectures are becoming increasingly parallel, driving the need to use algorithms that require minimal inter-processor communication. Both test particle and kinetic simulations are highly parallelizable; test particle calculations require no global communication, and the particle advance stage of kinetic simulations is highly parallelizable [71]. Recent advances have also reduced the communication requirements for the field advance stage of gyrokinetic simulations [73], suggesting the possibility of gyrokinetic simulations that only use “local” communication - i.e. messages are sent only between adjacent processor nodes. The scalability of these numerical methods therefore positions them to lead plasma physics simulations in an era of increasing opportunity for high performance computing [74]. High performance symplectic algorithms would then not only impact today’s simulations, but would be used for years to come.

If symplectic guiding center algorithms should be so beneficial, a reasonable question is why they are not being used already. The main reason symplectic algorithms are absent from existing guiding center simulations [46, 48, 49, 51, 52, 53] is that existing symplectic integrators are formulated in terms of specific “canonical” coordinates. Textbook symplectic integrators assume the Hamiltonian system is described in coordinates that can be paired into position (usually denoted  $q$ ) and conjugate momentum ( $p$ ) coordinates. The guiding center system, as it is most naturally described, does not admit such a partitioning of the coordinates. Instead, it belongs to a more general class of Hamiltonian systems known as “non-canonical Hamiltonian” or “Poisson” systems. Unfortunately, it remains an outstanding problem in numerical analysis to perform symplectic integration of non-canonical Hamiltonian systems [75]. Although it is known how to *transform* the guiding center coordinates to canonical form [46, 76, 77, 78, 79], there are drawbacks to relying on this transformation for computational purposes. Mainly, the transformation to canonical coordinates incurs computational overhead on every time step, which can be critical when deciding whether it is worthwhile to use a symplectic integrator. The alternative is to use an inexpensive non-symplectic integrator with a high order of local accuracy in the numerical step size. Any time spent solving implicit coordinate transformations can instead be spent taking additional, smaller steps with the high-order method. Although the canonical-coordinate symplectic integrator will *eventually* exhibit superior behavior, it takes *too long* for the behavior to emerge, provided the comparison is performed at equal computational expense.

The disadvantages of relying on canonical coordinates for symplectic guiding center integrators motivates the development of non-canonical symplectic guiding center integrators. When attempting to construct such a non-canonical symplectic integrator, one may choose from a number of existing algorithm development techniques. The most standard technique approximates the time derivative in the equations of motion, using for instance a finite difference method [10, 80, 81, 82, 83]. However, this technique is not very systematic for con-

structing symplectic integrators; it is completely unclear how to obtain the desired property when performing the finite difference, especially in the non-canonical setting. Many canonical symplectic integrators are instead systematically constructed using generating function theory [31, 34, 35, 37], but these functions generate *canonical* transformations, so the existing methods cannot be applied to the non-canonical guiding center system. Another systematic approach involves “splitting” the equations of motion into exactly solvable parts; because the solution of each piece has the desired symplectic property, the composition of the solutions also has this property [84, 85]. Again, unfortunately, it is not clear how to apply a splitting technique to the guiding center system, because the individual drift components comprising the equations of motion do not each admit an analytic solution.

A new alternative for systematically constructing symplectic integrators is a technique known as *variational integration* [9, 86]. Variational integrators approximate time derivatives in the Lagrangian function, and vary the action with respect to the discretized path. By variationally deriving the algorithm, geometric properties are inherited from the continuous system. For instance, a discrete Noether’s theorem identifies invariants corresponding to symmetries, and a discrete symplecticity condition is obtained [9]. Because a variational formulation of the guiding center system has long been known [8, 78], an intriguing possibility is that the recently developed theory of variational integration can be used to obtain non-canonical symplectic algorithms for guiding center trajectories.

Recent work has investigated this intriguing possibility, deriving variational integrators for the non-canonical guiding center system [14, 15, 16, 17]. Initial results were very promising, with the variational guiding center algorithms exhibiting behavior characteristic of symplectic integrators [14, 15, 16]. For instance, a test case using the algorithm of Ref. [16] to model a trapped-particle orbit in an axisymmetric tokamak is shown in Fig. 1.1. At the larger step size, the energy error remains bounded for long times and the trapped-particle orbit remains on a closed trajectory. In fact, the original intent of this dissertation was to

implement this method for test particle calculations by incorporating dissipative effects into the existing algorithm. In the process, a surprising discovery was made: in certain parameter regimes, the variational guiding center algorithms exhibit numerical instabilities that completely overwhelm the solution. This is illustrated in the right portion of Fig. 1.1, in which the numerical step size is decreased relative to the prior demonstration. This discovery spawned several questions that had to be addressed before variational integrators could be widely used for test particle and kinetic simulations, specifically:

- Why do instabilities occur in the variational guiding center algorithms, especially given their conservation properties?
- What can be done to mitigate these instabilities?
- Is it possible to construct variational guiding center integrators without such instabilities?

Depending on the answers to these questions, variational integrators may be deemed more or less well-suited for the integration of guiding center trajectories.

It is therefore the high potential for impact of symplectic guiding center algorithms and the promising initial results of variational integrators for guiding center dynamics that motivate this dissertation. In addition to addressing the above questions, this dissertation seeks to incorporate dissipative collisional effects in the best performing variational guiding center integrators. The performance of the new algorithms is then to be assessed relative to the industry standard techniques, identifying parameter regimes in which the new methods are advantageous. This assessment can then be used to argue in favor (or against) geometric integration of guiding center dynamics in high performance test particle and kinetic simulations.

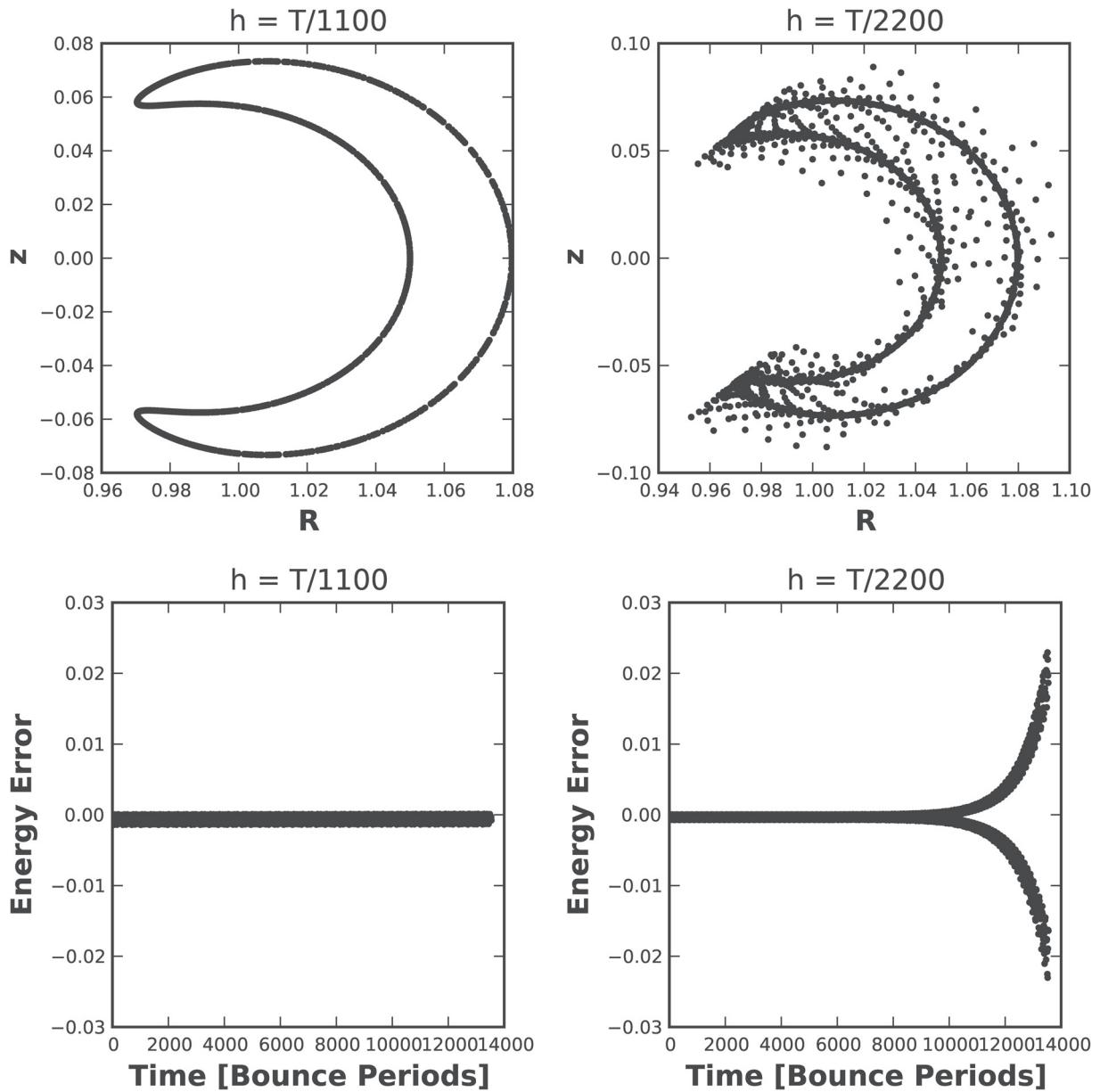


Figure 1.1: Instabilities were discovered in certain parameter regimes of the variational guiding center algorithms presented in Refs. [14, 15, 16]. Here, the algorithm of Ref. [16] is shown to be unstable at *very small* step size  $h$  when modeling the trapped-particle study of Ref. [15]. In these figures,  $T$  denotes the period of the banana orbit,  $R$  and  $z$  are in units of the major radius  $R_0$ , and the energy error is  $\frac{E(t)-E_0}{E_0}$ , where  $E_0$  is the initial energy.

## 1.3 Contributions

In summary, the main contributions of this dissertation in the context of geometric guiding center algorithms may be described as:

- the discovery and explanation of numerical instabilities present in the existing variational integrators for guiding center trajectories,
- the identification of a concise condition under which the instabilities are eliminated,
- the use of this condition to construct new guiding center algorithms without numerical instabilities,
- the first incorporation of dissipative collisional effects into variational guiding center algorithms,
- and the demonstration that the new variational guiding center algorithms are computationally competitive for modeling the thermalization of energetic particles in magnetically confined plasmas.

Although guiding center algorithms are the primary application that motivates these contributions, many of the contributions themselves are general, numerical analysis results. These instabilities are shown not only to occur in guiding center variational algorithms, but in *any* variational algorithm constructed from a degenerate action principle. Similarly, the condition for eliminating the instabilities is phrased in terms of properties of the variational integrator, and is applicable beyond the guiding center system.

To understand these contributions, both in the context of guiding center dynamics and the context of variational integration, it is important to understand how the guiding center Lagrangian differs from those typically considered when developing variational integrators. Toward this end, consider the following two Lagrangians for a charged particle evolving in

an electric field  $E = -\nabla\phi$  and magnetic field  $B = \nabla \times A$ . First, consider the *Lorentz force Lagrangian*, given by:

$$L_{LF}(q, \dot{q}) = \frac{1}{2}m\|\dot{q}\|^2 + \frac{e}{c}A(q) \cdot \dot{q} - e\phi(q). \quad (1.1)$$

The Euler-Lagrange equations corresponding to the Lorentz force Lagrangians give, of course, the Lorentz force law:

$$m\ddot{q} = eE + \frac{e}{c}\dot{q} \times B. \quad (1.2)$$

In particular, a second-order differential equation is obtained, where the second-order derivative results from the  $\|\dot{q}\|^2$  term. Because the Euler-Lagrange equations constitute a well-defined system of second-order ODEs, the Lorentz force Lagrangian is said to be *regular*. In contrast, consider the *guiding center Lagrangian*:

$$L_{GC}(x, u, \dot{x}, \dot{u}) = (A(x) + ub(x)) \cdot \dot{x} - \left( \frac{1}{2}u^2 + \mu\mathcal{B}(x) + \phi(x) \right), \quad (1.3)$$

where  $x$  is the position of the guiding center,  $u$  the parallel velocity,  $\mu$  the magnetic moment,  $b$  the magnetic field unit vector, and  $\mathcal{B} = \|B\|$  the magnitude of the magnetic field. The guiding center description is used when the scale length over which the magnetic and electric fields vary is large relative to the gyroradius and the time variation of the fields is slow relative to the gyrofrequency. In this case, the rapid gyration of the particle about the magnetic field can be averaged over, and only the particle drifts -  $E \times B$ ,  $\nabla\mathcal{B}$ , and curvature drifts - retained in the dynamical description. These drifts are described in the Euler-Lagrange equations obtained from the guiding center Lagrangian:

$$(\nabla \times (A + ub)) \times \dot{x} = -\mu\nabla\mathcal{B} + E + \dot{u}b \quad (1.4)$$

$$b \cdot \dot{x} = u. \quad (1.5)$$

Pertinently, the Euler-Lagrange equations for the guiding center system are *first*-order differential equations. Because the Euler-Lagrange equations are not second order, the guiding center Lagrangian is said to be *degenerate*. In fact, the linear dependence of the guiding center Lagrangian on the velocity coordinates  $(\dot{x}, \dot{u})$  is a defining characteristic of a class of Lagrangians known as “phase-space Lagrangians”. These phase-space Lagrangians are *intended* to give first-order differential equations as the necessary condition for extremizing an action. The emergence of this degeneracy in the guiding center description is reviewed in Sec. 2.3, and the difference between regular and phase-space Lagrangians is illustrated in the upper half of Fig. 1.2.

From a mathematics perspective, the development of variational integrators for guiding center trajectories is then an exercise in variational integrators for degenerate Lagrangian systems. Conventionally, the Lagrangian is assumed to be regular when constructing a variational integrator [9]. Many of the theorems establishing the desirable behavior of variational integrators rely on this assumption. Although there does exist work regarding variational integrators for degenerate Lagrangian systems [13, 87, 88, 89, 90, 91, 18, 92], the numerical instabilities discovered in this dissertation and the proposed solutions are new contributions. In short, when discretizing degenerate Lagrangians (like the guiding center Lagrangian above), two possibilities exist. The first possibility is that the degeneracy is *not* captured in the discrete Lagrangian - it is regular. In this case, as would be appropriate for modeling a second-order differential equation, a numerical trajectory is uniquely specified by the position at *two* instances in time. For phase-space Lagrangians, the initial value problem specifies only the *initial* position; there exists excess freedom in the initial conditions used to start the variational integrator, which may ultimately be considered to be the source of the numerical instabilities depicted in Fig. 1.1. The second possibility is that the degeneracy *is* captured in the discrete Lagrangian - it is degenerate. A satisfying new result is that the numerical instabilities are eliminated whenever the discrete Lagrangian has the same degree

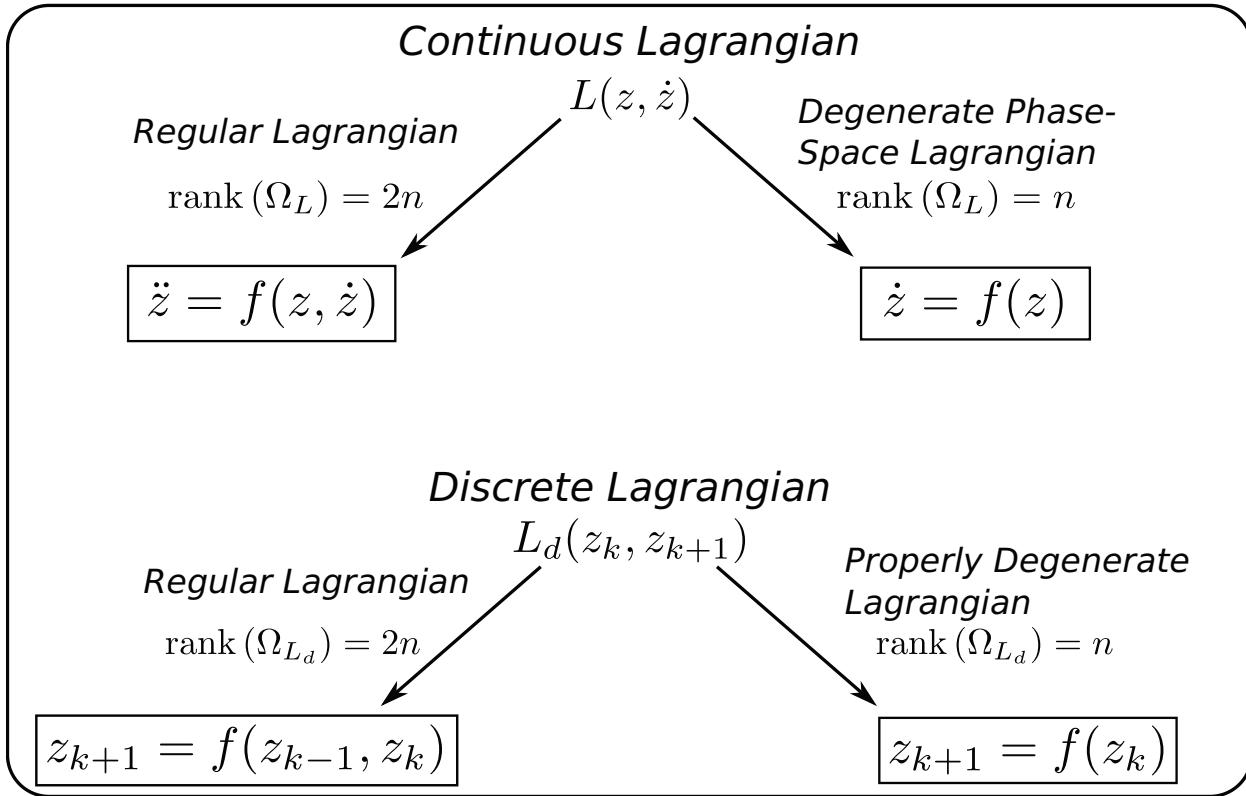


Figure 1.2: Schematic of Lagrangian systems of interest for this dissertation. The regularity of continuous and degenerate Lagrangian systems is determined by the rank of the continuous and discrete symplectic structures,  $\Omega_L$  and  $\Omega_{L_d}$  respectively (defined in Eqs. (2.37) and (2.39)). A regular continuous Lagrangian yields a second-order ODE for the Euler-Lagrange equations. Phase-space Lagrangians, which are degenerate, yield first-order Euler-Lagrange equations. If the discrete Lagrangian is regular, the time advance depends on two previous instants in time. For properly degenerate discrete Lagrangians, the time advance depends only on the state at a single instant in time.

of degeneracy as the phase-space Lagrangian. Discrete Lagrangians satisfying this property are labeled *properly degenerate*. These two possibilities for the discrete system are illustrated in the lower half of Fig. 1.2.

In more detail, Chapter 3 explains the consequences of modeling a (degenerate) phase-space Lagrangian system with a regular discrete Lagrangian. The key realization is that this type of variational integrator belongs to a longstanding class of numerical methods known as *multistep methods* [10, 83, 11, 12]. Multistep methods depend on a larger set of initial

conditions than that specified by the initial value problem; precisely the situation when the discrete Lagrangian is regular while the continuous Lagrangian is degenerate. It is therefore useful to categorize this class of algorithms as *multistep variational integrators (MVIs)*. This new categorization leads naturally to important results on stability and structure-preservation by adapting the existing literature on multistep methods to the variational integrators setting. For instance, the unexpected numerical instabilities in the guiding center variational integrators are in fact a manifestation of *parasitic modes*, long known to be present in any multistep algorithm. Unfortunately, it is determined that these parasitic modes are at best marginally stable for *any* MVI. On a more positive note, MVIs exhibit surprisingly good conservation properties whenever the parasitic modes remain small. It is shown that the “underlying one-step method” [93, 94] of an MVI preserves a symplectic two-form on the original space. This result is surprising because there exist theorems that the underlying one-step method of certain multistep methods *cannot* be symplectic [95, 6]. Contradiction is avoided because the MVI symplectic structure differs from that of the continuous system. The practical implication is that good behavior can be anticipated whenever the parasitic modes remain small. Overall, the developments of this chapter have substantial explanatory power for existing variational algorithms for degenerate Lagrangian systems [13, 14, 15, 16, 96, 18, 19] and offer a new framework for the numerical analysis of multistep methods based on discrete variational principles.

With the intent of eliminating the parasitic mode instabilities inherent to MVIs, Chapter 4 develops *degenerate variational integrators (DVIs)*, which mimic the degeneracy exhibited by the guiding center and any other phase-space Lagrangian system. The key result of the chapter is a concise condition for the elimination of parasitic modes. This condition compares the rank of second-order derivative tensors of the Lagrangian functions, and shows that if the rank is the same for both the continuous and discrete Lagrangians, parasitic modes cannot be present. Ideally, a systematic procedure for satisfying this condition would

be provided. Unfortunately, no such procedure is presently known. However, the condition is easy to check, and inspires deeper consideration of algorithms that *appear* multistep at first glance. Importantly, use of this condition leads to the first one-step (i.e. not multistep) variational algorithms for non-canonical magnetic field line flow and guiding center dynamics. For the magnetic field line integrator, the only restriction is the choice of a gauge such that one of the components of the vector potential is zero. For the guiding center integrator, the vector potential and magnetic field are assumed to have the same two non-zero covariant components. Global coordinates satisfying this condition may be constructed whenever nested magnetic flux surfaces are present. This perspective of retaining degeneracy in the variational integrator is an exciting new approach for developing stable variational integrators for guiding center dynamics, and may eventually yield excellent geometric algorithms for any non-canonical Hamiltonian system.

The existing and newly developed variational integration techniques are deployed in Chapter 5 to obtain new algorithms for dynamical systems in plasma physics. The first application derives a multistep variational integrator for perturbed magnetic field line flow. By using the knowledge of the unperturbed field line flow, one may take large step sizes when constructing the Poincaré section of the perturbed fields. The penultimate application illustrates the long-term numerical fidelity of the degenerate variational integrator for non-canonical magnetic field line flow. This sets the stage for the final application of a degenerate variational integrator for guiding center trajectories, including collisional drag forces. Because the guiding center variational integrator retains the non-canonical coordinate description, the computational expense is reduced by a factor of two from previous symplectic integrators for guiding center dynamics. Consequently, the timescale over which the conservative algorithms are advantageous is greatly reduced when performing an equal computational expense comparison. It is shown that the benefits of the conservative algorithm may be observed over an energetic ion thermalization process, thereby demonstrating

the relevance of variational integrators to experimental simulations of fast particles in magnetically confined fusion plasmas.

All numerical demonstrations performed in this dissertation use a newly developed code called “GEODES” (GEometric ODE Solver). This software has been developed with modern programming practices, including an object-oriented C++ design, version control, unit testing, and GPGPU parallelization. These methods result in a user-friendly, maintainable, and high-performance numerical platform. The library interfaces with the EFIT equilibrium code, and may be easily extended to other magnetic equilibrium software. For details, see Appendix B.

Altogether, this dissertation makes numerical analysis contributions by developing the theory of multistep and degenerate variational integrators, and plasma physics contributions by developing new algorithms for magnetic field line and energetic particle dynamics. The central theme is that failure to capture the degeneracy when discretizing a phase-space action principle has significant consequences for the stability of the resulting variational integrator. This dissertation makes the novel suggestion of replicating the degeneracy in the form of a degenerate variational integrator. These formalisms are put to use in a variety of numerical examples, ultimately demonstrating the practical impact of these algorithms for ITER-relevant test particle calculations. Combined, this dissertation makes a compelling case for geometric integration methods to play a central role in flagship high performance computing efforts in the upcoming years.

# Chapter 2

## Background

*You can't hold no groove if you ain't got no pocket*

—J.D., *A Show of Hands by Victor Wooten*

### 2.1 Overview

This chapter assembles fundamental material from the two topics on which the rest of the dissertation is built: *variational integrators* and *action principles in phase space*. In short, the dissertation explains what happens when one constructs variational integrators from action principles in phase space, so it is important to establish the terminology and fundamental results of these two topics. The section on variational integrators reviews this modern paradigm for numerical analysis, explaining that a discrete formulation of Lagrangian mechanics may be used to derive and analyze geometric algorithms. For greater depth, consult the thorough reviews of Refs. [9, 86]. The second section summarizes phase-space action principles, which utilize a degenerate Lagrangian and appear in several dynamical systems in plasma physics. General references for this material include classical mechanics texts [5, 3, 20], while plasma-specific references include Refs. [8, 97, 98, 78].

To build basic intuition behind the two key topics of this chapter, the central ideas will be informally illustrated here using the Lorentz force dynamical system as a concrete example. First, it will be shown how to construct a simple variational integrator for this system. Then, it will be shown how a “phase-space” Lagrangian can be constructed from the more familiar Lagrangian, and how the variational principle using this phase-space Lagrangian recovers Hamilton’s equations.

Consider then a particle of charge  $e$  and mass  $m$  evolving in an electric field  $E = -\nabla\phi$  and magnetic field  $B = \nabla \times A$ , presumed time-independent for simplicity. Letting  $q$  denote the particle’s position in  $\mathbb{R}^3$ , the Lorentz force Lagrangian is given by Eq. (1.1). The action  $S$  is given by:

$$S(q) = \int_0^T L_{LF}(q(t), \dot{q}(t)) dt, \quad (2.1)$$

where  $[0, T] \subset \mathbb{R}$  is the time interval of interest. *Hamilton’s principle of least action* states that the true trajectory is an extremum of the action functional (Thm. 2.2.1). By the familiar action variation procedure, a trajectory is an extremum of the action functional if and only if it satisfies the *Euler-Lagrange equations* at each instant in time:

$$\frac{\partial L_{LF}}{\partial q}(q, \dot{q}) - \frac{d}{dt} \left( \frac{\partial L_{LF}}{\partial \dot{q}}(q, \dot{q}) \right) = 0. \quad (2.2)$$

For this system, the Euler-Lagrange equations yield the Lorentz force law, as given in Eq. (1.2).

Typically, an algorithm for the Lorentz force system would be obtained by finite differencing the time derivatives appearing in the Lorentz force law Eq. (1.2). To construct a variational integrator, all approximations will be performed on the *action* instead. One

might choose, for instance, a discrete Lagrangian according to:

$$\begin{aligned} L_d(q_k, q_{k+1}) &= L_{LF}(q_k, \frac{q_{k+1} - q_k}{h}) \\ &= \frac{1}{2}m \frac{\|q_{k+1} - q_k\|^2}{h^2} + \frac{e}{c} A(q_k) \cdot \frac{(q_{k+1} - q_k)}{h} - e\phi(q_k), \end{aligned} \quad (2.3)$$

where  $q_k$  denotes a numerical approximation to the solution at time  $t = t_k$ , and  $h$  is a numerical step size. A discrete action may then be defined according to:

$$S_d(q_0, \dots, q_N) = \sum_{k=0}^{N-1} h L_d(q_k, q_{k+1}). \quad (2.4)$$

Because the time was discretized, the discrete action is a summation rather than an integral. The next step is to extremize the action by varying the discrete action with respect to  $q_k$ , for each  $k$  (except the endpoints  $k = 0, N$ ), and requiring the result to be zero:

$$\frac{\partial S_d}{\partial q_k} = h \frac{\partial L_d(q_{k-1}, q_k)}{\partial q_k} + h \frac{\partial L_d(q_k, q_{k+1})}{\partial q_k} = 0. \quad (2.5)$$

This system of equations is known as the *discrete Euler-Lagrange equations*, and is more often [9] written:

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \quad (2.6)$$

where  $D_1$  and  $D_2$  indicate differentiation with respect to the first and second arguments, respectively. The discrete Euler-Lagrange equations may not immediately resemble a finite difference approximation of the Euler-Lagrange equations, but calculation of these equations for the Lorentz force discrete Lagrangian chosen in Eq. (2.3) reveals the correspondence:

$$m \frac{q_{k+1} - 2q_k + q_{k-1}}{h^2} = \frac{e}{c} \nabla A(q_k) \cdot \frac{q_{k+1} - q_k}{h} + \frac{e}{hc} (A(q_{k-1}) - A_j(q_k)) - e \nabla \phi(q_k). \quad (2.7)$$

Here, Cartesian coordinates have been assumed for simplicity of the acceleration term on the left hand side. The term  $(A(q_{k-1}) - A(q_k))$  may be Taylor expanded about  $q_k$  and the result combined with the first term on the right hand side to obtain the correspondence with the  $\dot{q} \times B$  term in the Lorentz force equations.

Reflecting on the Lorentz force variational integrator of Eq. (2.7), first consider the zero magnetic field case  $A = B = 0$ . Upon this simplification, the discrete Euler-Lagrange equations become:

$$m \frac{q_{k+1} - 2q_k + q_{k-1}}{h^2} = -e \nabla \phi(q_k). \quad (2.8)$$

This algorithm is known as the *Störmer* algorithm for a particle evolving in a potential  $V(q) = e\phi(q)$ . The Störmer method exhibits excellent long-term behavior. This good behavior can be explained by showing that the Störmer advance is equivalent to a symplectic integration of Hamilton's equations, in particular it is equivalent to:

$$p_{k+1} = m \frac{q_{k+1} - q_k}{h} \quad (2.9a)$$

$$\frac{p_{k+1} - p_k}{h} = -e \nabla \phi(q_k). \quad (2.9b)$$

This method is known as *symplectic Euler*, and is a first-order accurate in  $h$  symplectic integrator for Hamilton's equations. It is intimately related to the well-known Leapfrog method, which “staggers” the momentum variables  $p_k \rightarrow p_{k+1/2}$ . The Leapfrog method is also known to be symplectic<sup>1</sup>. One might wonder whether this correspondence between the variational integrator and a symplectic integrator is particular to this toy problem or a general feature. Indeed, this is a general feature of variational integrators - one can *always* find an equivalent symplectic integrator, provided the map (Eq. (2.33)) is well-defined (see Thm. 2.2.10). This correspondence is one of the primary benefits of variational integrators - unlike when finite

---

<sup>1</sup>See Sec. A.5 for a tutorial on symplectic maps

differencing the equations of motion, finite differencing the action guarantees the algorithm will be equivalent to a symplectic integrator.

Before moving on from variational integrators, several comments are appropriate regarding the Lorentz force variational integrator of Eq. (2.7) in the case when  $A$  and  $B$  are not equal to zero. Notice that, unlike the Euler-Lagrange equations, the discrete Euler-Lagrange equations (Eq. (2.7)) depend on the vector potential  $A$ . This algorithm would therefore be an unlikely choice when performing a finite difference directly on the Euler-Lagrange equations. One justification for the presence of these terms comes from the fact that, even though the Euler-Lagrange equations do not depend on  $A$ , Hamilton's equations for the Lorentz force system *do*. It is then not too surprising that a method equivalent to a symplectic integrator for Hamilton's equations also depends on the vector potential. Still, this is not an entirely desirable feature of the numerical method because one may need to construct a magnetic vector potential if only supplied with, say, an experimentally determined magnetic field  $B$ .

Now that intuition has been established for variational integrators, the same example will be used to illustrate action principles in phase space. The Lorentz force Lagrangian  $L(q, \dot{q})$  defined in Eq. (1.1) accepts a position  $q$  and a velocity  $\dot{q}$  as its argument. It is said to be a *configuration space Lagrangian*, where the configuration space for this problem is the space in which  $q$  resides, namely,  $\mathbb{R}^3$ . The Euler-Lagrange equations (Eq. (2.18)) for the configuration space Lagrangian provide a three-dimensional system of second-order differential equations on the configuration space. We can instead formulate the dynamics as a six-dimensional system of first-order differential equations on the *phase space*. To do so, we replace the velocity coordinates  $\dot{q}$  with momentum coordinates  $p$  according to the *Legendre transform* (defined more rigorously in Sec. 2.2.2):

$$p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}). \quad (2.10)$$

If the Legendre transform is invertible, then we can determine  $\dot{q}$  as a function of  $p$  and possibly also  $q$ . We can use this to identify a *Hamiltonian* on the  $(q, p)$  phase space according to:

$$H(q, p) = p \cdot \dot{q}(q, \dot{p}) - L(q, \dot{q}(q, p)). \quad (2.11)$$

In the typical presentation for undergraduate mechanics, one would next write down *Hamilton's equations* for the time evolution of this system. Instead, Hamilton's equations can be derived from a variational principle as follows: Let the *phase-space Lagrangian*  $L_{PS}(q, p, \dot{q}, \dot{p})$  be defined according to:

$$L_{PS}(q, p, \dot{q}, \dot{p}) = p \cdot \dot{q} - H(q, p). \quad (2.12)$$

This Lagrangian *looks* like the definition of a Lagrangian as would be obtained using a Hamiltonian and a Legendre transform. Indeed, a configuration-space Lagrangian would be recovered if, in the above expression,  $p$  were a function of  $\dot{q}$  and possibly also  $q$  according to the Legendre transform (applied to the Hamiltonian). For the phase-space Lagrangian, however,  $q$  and  $p$  are considered *independent*; there is no presumed relation between  $p$  and  $\dot{q}$  until informed of one by the equations of motion. Next, *Hamilton's principle in phase space* states that the true path extremizes the phase-space action  $S_{PS}$  given by:

$$S_{PS}(q, p) = \int_0^T L_{PS}(q(t), p(t), \dot{q}(t), \dot{p}(t)) dt. \quad (2.13)$$

The *phase-space Euler-Lagrange equations* emerge from the standard variational procedure, and are given by:

$$\dot{p} = -H_{,q}(q, p) \quad (2.14a)$$

$$\dot{q} = H_{,p}(q, p). \quad (2.14b)$$

That is, we have recovered Hamilton's equations from a variational principle! Notice that this is a  $2n$ -dimensional system of first-order differential equations, where  $n = 3$  for the Lorentz force problem. Because of this, the phase-space Lagrangian is said to be *degenerate*; it's degenerate from the perspective that we typically expect the Euler-Lagrange equations to be second-order differential equations.

The utility of Hamilton's principle in phase-space is that we can perform arbitrary coordinate transformations on the  $2n$ -dimensional  $(q, p)$  phase-space coordinates, while relying on the variational principle to recover the correct equations of motion. For instance, the phase-space Lagrangian for the Lorentz force system is given by:

$$L_{PS}(q, p, \dot{q}, \dot{p}) = p \cdot \dot{q} - \left( \frac{1}{2m} \|p - \frac{e}{c} A\|^2 + e\phi(q) \right). \quad (2.15)$$

From here, we might choose to change variables from positions  $p_j$  to velocities  $v^j$  according to  $v^j = \frac{1}{m} g^{ji} (p_i - \frac{e}{c} A_i(q))$ , where  $g$  is the metric tensor. In terms of these (non-canonical!) coordinates, the phase-space Lagrangian becomes:

$$L_{PS}(q, v, \dot{q}, \dot{v}) = (mv + \frac{e}{c} A(q)) \cdot \dot{q} - \left( \frac{1}{2} m \|v\|^2 + e\phi(q) \right). \quad (2.16)$$

Although the phase-space Euler-Lagrange equations in terms of  $(q, v)$  will not take the form of Eq. (2.14), we have not violated the Hamiltonian character of the system by this transformation. After all, we have only changed coordinates; the dynamics are the same. This point is detailed in Sec. 2.3.1.

Glancing ahead, Chapters 3 and 4 seek to answer the question: *What happens when we construct variational integrators from Hamilton's principle in phase space?* In the example above, the phase-space Lagrangians for Lorentz force dynamics were obtained by starting with a configuration-space Lagrangian and using a Legendre transform. For the guiding center system, *only* a phase-space Lagrangian is known [8]. The guiding center Lagrangian

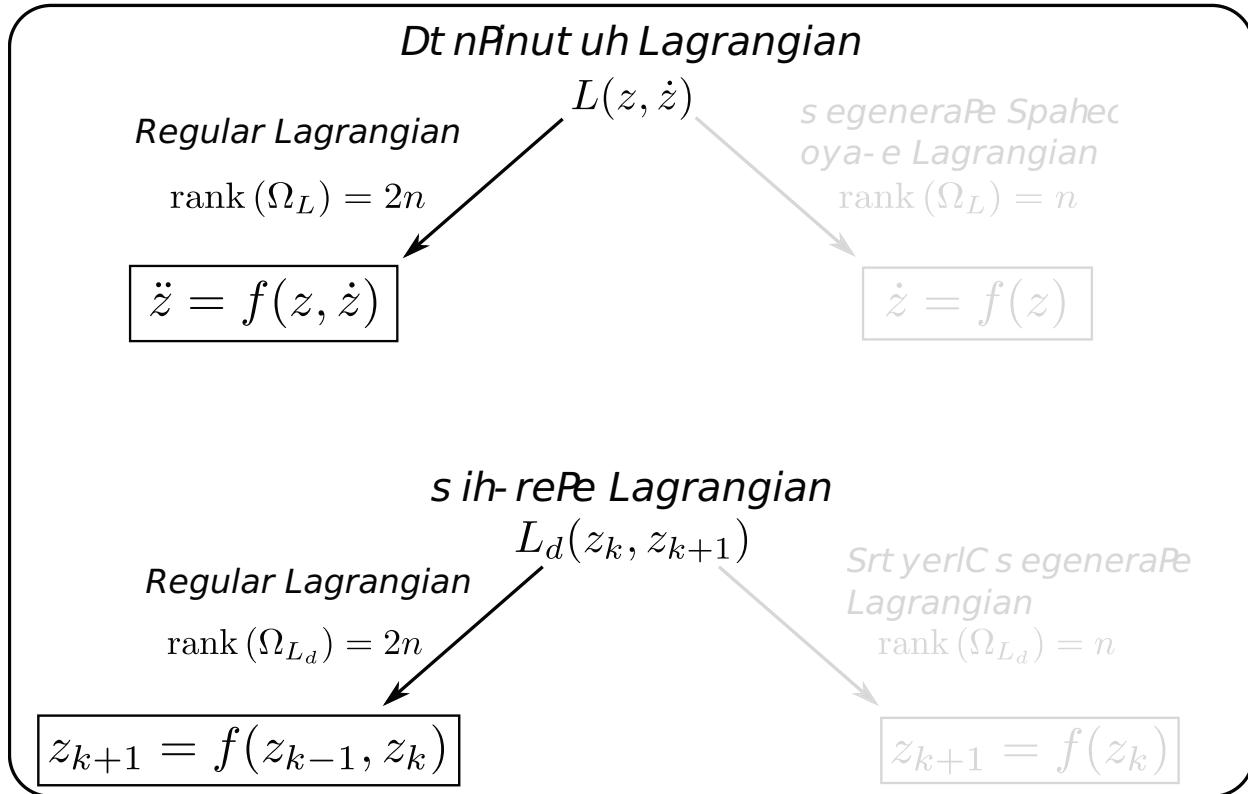
is derived from the Lorentz-force phase-space Lagrangian [8]; one further transforms from  $(q, v)$  coordinates to the guiding center position  $x$  and three additional variables  $(u, \mu, \Theta)$ , where  $u$  is the velocity parallel to the magnetic field,  $\mu$  the magnetic moment, and  $\Theta$  the “gyrophase” - the phase of the charged particle as it gyrates around the magnetic field line. In the guiding center approximation,  $\Theta$  is an ignorable coordinate and its corresponding Noether invariant is the magnetic moment  $\mu$ . This is used to reduce the system to a *four* dimensional system, with coordinates  $(x^1, x^2, x^3, u)$ . The full transformation to the guiding center system is reviewed in Sec. 2.3.4. In terms of variational integration, the new setting emerged as soon as we wrote down a phase-space Lagrangian for the Lorentz force dynamics. At that point, the Lagrangian became degenerate, and many results from the standard theory of variational integrators no longer applied.

For now, the following sections re-visit the two foundational topics - variational integrators and action principles in phase-space - providing rigor and depth to the preceding heuristic discussion.

## 2.2 Variational Integrators

### Section Summary

- Instead of discretizing time derivatives in the equations of motion, variational integrators discretize the path (as a function of time) in the action
- If both the continuous and discrete Lagrangians are regular, then the variational integrator for the Euler-Lagrange equations is equivalent to a symplectic integrator for Hamilton’s equations



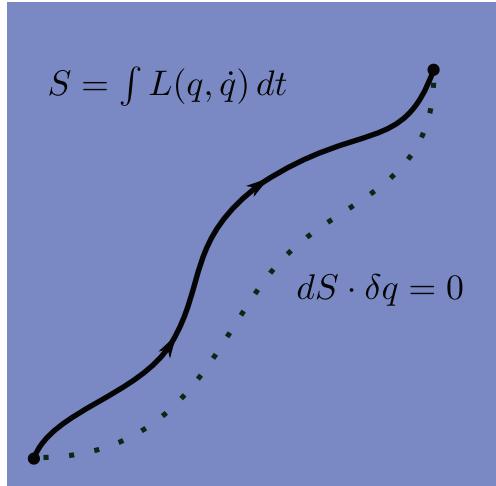
- Symplectic integrators generate trajectories nearby to Hamiltonian systems, yielding many desirable benefits
- Dissipative forces may be incorporated to continuous and discrete variational principles by modifying the statement of the variational principle

The conventional paradigm for numerical analysis considers the equations of motion to be the starting point for constructing algorithms. From a physics perspective, however, the equations of motion are not the starting point but instead derive from a variational principle. Variational integrators strive to take this mathematical structure into account by asserting the *variational principle* to also be the starting point for constructing algorithms. The benefits of adopting such a procedure for the numerical description mimics the benefits

delivered by the variational principle in the continuous setting: coordinate transformations become trivial, conservation laws are guaranteed and straightforward to identify, and unifying features may be observed across seemingly different systems. This section reviews how variational principles may be discretized to construct algorithms while detailing the benefits [9, 6].

### 2.2.1 Variational Principles

#### Continuous Variational Principles



The starting point of a variational principle is the identification of a *configuration space*  $Q$  representing all possible states of the physical system.  $Q$  will be assumed to be a differentiable manifold. The end goal is to identify paths  $q : [0, T] \subset \mathbb{R} \rightarrow Q$  admissible for the dynamical system. Admissible paths satisfy some *equation of motion* governing the system; given an initial condition, one uniquely identifies a corresponding path or “trajectory”  $q$ .

To derive the equations of motion, one identifies a *Lagrangian*  $L : TQ \rightarrow \mathbb{R}$ . The domain of the Lagrangian  $L$  being the tangent bundle of the configuration space  $TQ$  intuitively means  $L$  accepts as its arguments a position and a velocity. Equipped with a Lagrangian, one may define the action  $S$  acting on paths in  $Q$  according to:

$$S(q) = \int_0^T L(q(t), \dot{q}(t)) \, dt. \quad (2.17)$$

Together, these definitions enable the following theorem [1, 2, 3, 4, 5].

**Theorem 2.2.1 (Hamilton's Principle of Least Action)** *Given a configuration space  $Q$ , Lagrangian  $L$ , and action  $S$ , a trajectory  $q$  with  $q(0) = q_0$ ,  $q(T) = q_T$  is a critical point of the action if and only if it satisfies the Euler-Lagrange equations:*

$$\frac{\partial L}{\partial q}(q(t), \dot{q}(t)) - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}}(q(t), \dot{q}(t)) \right) = 0, \quad (2.18)$$

for all  $t \in [0, T]$ .

**Proof** The Euler-Lagrange equations emerge from varying the action with respect to its argument - the path  $q$ :

$$\begin{aligned} dS \cdot \delta q &= \int_0^T \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} dt \\ &= \int_0^T \left( \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q dt + \frac{\partial L}{\partial \dot{q}} \delta q \Big|_0^T, \end{aligned}$$

where integration by parts has been used. Asserting the variations to be zero at the endpoints obtains the Euler-Lagrange equations (2.18) as a necessary and sufficient condition for the path to extremize the action. ■

Assuming the Euler-Lagrange equations are well-defined, notation will be required to discuss solutions of the Euler-Lagrange equations. The *Lagrangian flow map*  $F_L : TQ \times \mathbb{R} \rightarrow TQ$  advances an initial condition  $(q_0, \dot{q}_0) \in TQ$  some amount of time  $t \in \mathbb{R}$  along the solution of the Euler-Lagrange equations, i.e.  $F_L(q_0, \dot{q}_0, t)$  satisfies:

$$\frac{\partial L}{\partial q}(F_L(q_0, \dot{q}_0, t)) - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}}(F_L(q_0, \dot{q}_0, t)) \right) = 0, \quad (2.19)$$

whenever  $t \in [0, T]$ . The *time-t Lagrangian flow map* is denoted  $F_L^t : TQ \rightarrow TQ$  and given by  $F_L^t(q_0, \dot{q}_0) = F_L(q_0, \dot{q}_0, t)$ .

### Discrete Variational Principles

Suppose that, instead of continuous trajectories, one is interested only in the state  $q(t)$  at a finite number of instants in time. It is interesting to propagate this discretization through Lagrangian mechanics from the earliest opportunity. Starting with the same configuration space  $Q$ , consider a discretization of the time interval  $[0, T]$  into  $N+1$  discrete instances  $t_0, t_1, \dots, t_N$  with  $t_0 = 0, t_N = T$ . Also assume the time instances to be equally spaced with *step size*  $h$ . Corresponding to the discretized time trajectory, a discrete path

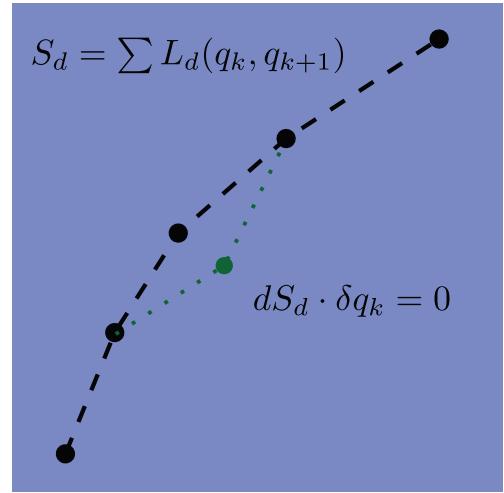
$q_d = (q_0, \dots, q_N)$  with  $q_k \in Q$  may be constructed. The goal of the discrete variational principle will be to identify a condition governing the time advance of this discrete path.

The discrete analog of the Lagrangian is a *discrete Lagrangian*  $L_d : Q \times Q \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $(q_0, q_1, h) \mapsto L_d(q_0, q_1, h)$ . Often, the parametric dependence of  $L_d$  on the numerical step size  $h$  will be implied and not listed in the arguments. Because the path is no longer continuous, the identification of a velocity at each point  $q_k$  in the path is ambiguous. However, note that the dimensions of the spaces  $TQ$  and  $Q \times Q$  are equal, so the domain of the discrete Lagrangian is a plausible choice. From this discrete Lagrangian, a discrete action  $S_d$  may be constructed according to:

$$S_d(q_d) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}). \quad (2.20)$$

From here, the following theorem constitutes a discrete analog of Hamilton's principle of critical action.

**Theorem 2.2.2** *Given a configuration space  $Q$ , discrete Lagrangian  $L_d$ , and a discrete action  $S_d$ , a discrete path  $q_d$  starting at  $q_0$  and ending at  $q_N$  is a critical point of the discrete*



*action if and only if it satisfies the discrete Euler-Lagrange (DEL) equations:*

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \quad (2.21)$$

*for  $k = 1, \dots, N - 1$ . Here, the slot derivative  $D_i$  denotes differentiation with respect to the  $i$ -th argument of a function.*

**Proof** Variation of the discrete action with respect to the discrete trajectory gives:

$$\begin{aligned} dS_d \cdot \delta q_d &= \sum_{k=0}^N \frac{\partial S_d}{\partial q_k} \cdot \delta q_k \\ &= \sum_{k=0}^{N-1} D_1 L_d(q_k, q_{k+1}) \delta q_k + D_2 L_d(q_k, q_{k+1}) \delta q_{k+1} \\ &= \sum_{k=1}^{N-1} (D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1})) \delta q_k + \\ &\quad D_1 L_d(q_0, q_1) \delta q_0 + D_2 L_d(q_{N-1}, q_N) \delta q_N, \end{aligned}$$

where the summation has been re-arranged (a discrete analog of integration by parts!) to collect the variations with respect to a particular  $q_k$ . Asserting the end point variations to be zero and allowing  $\delta q_k$  to be arbitrary identifies the discrete Euler-Lagrange equations (Eq. (2.21)) as the stationarity condition. ■

From the form of the discrete Euler-Lagrange equations, one can infer that specifying an initial condition of the form  $(q_0, q_1)$  identifies  $q_2$  (subject to some solvability conditions to be discussed in Sec. 2.2.2); from there, one can iterate the procedure to obtain the rest of the discrete trajectory. In terms of maps, the discrete analog of the Lagrangian flow map is the

discrete Lagrangian map  $F_{L_d} : Q \times Q \times \mathbb{R} \rightarrow Q \times Q$  and is defined by:

$$D_2 L_d(q_0, q_1) + D_1 L_d(F_{L_d}(q_0, q_1, h)) = 0 \quad (2.22a)$$

$$\pi_1 \circ F_{L_d}(q_0, q_1, h) = q_1, \quad (2.22b)$$

where  $\pi_1 : Q \times Q \rightarrow Q$  is the projection operator, i.e.  $\pi_1(q_0, q_1) = q_0$ . Put simply,  $F_{L_d}$  maps  $(q_0, q_1)$  to  $(q_1, q_2)$  according to the DEL equations. Much like the discrete Lagrangian, the  $h$ -dependence of  $F_{L_d}$  will be implied; when it is necessary to be more explicit the *step size-h discrete Lagrangian map*  $F_{L_d}^h : Q \times Q \rightarrow Q \times Q$  will typically be used.

### 2.2.2 Legendre Transforms

There exist two sides of the proverbial classical mechanics coin: Lagrangian and Hamiltonian mechanics. Where possible, these two formalisms are related via the Legendre transform. This subsection details the Legendre transform in the familiar continuous setting and presents an analog for the discrete setting.

#### Continuous Legendre Transform and Hamiltonian Mechanics

The Legendre transform relates the dynamics on the position-velocity space (tangent bundle)  $TQ$  to dynamics on the phase space (cotangent bundle)  $T^*Q$ . Given a configuration space  $Q$  and a Lagrangian  $L$ , the Legendre transform  $\mathbb{F}L : TQ \rightarrow T^*Q$  is given by:

$$\mathbb{F}L : (q, \dot{q}) \mapsto (q, p) = (q, \frac{\partial L}{\partial \dot{q}}(q, \dot{q})). \quad (2.23)$$

An important technical consideration is whether the Legendre transform is an invertible mapping, namely, whether one can uniquely identify a pair  $(q, \dot{q})$  given values for  $(q, p)$ . A Lagrangian is called *regular* if the Legendre transform is an invertible mapping. Some texts

distinguish “hyper-regular” for the case when the Legendre transform is globally invertible [9], but here I will assume the invertibility is global and only use the term “regular”. A sufficient and necessary condition for the Legendre transform to be locally invertible (and therefore Lagrangian to be locally regular) is that:

$$\det\left(\frac{\partial p}{\partial \dot{q}}\right) = \det\left(\frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}}\right) \neq 0. \quad (2.24)$$

The tensor  $\frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}}$  is referred to as the *Hessian* of the Lagrangian, so one might say that a Lagrangian is regular if and only if the Hessian is non-degenerate. Notice that a Lagrangian that depends on  $\dot{q}$  according to  $\|\dot{q}\|^2$  is globally regular.

An important consequence of regularity is that the Euler-Lagrange equations define a unique second-order differential equation and that the Lagrangian flow map defined in Eq. (2.19) is well-defined [9]. To see the connection between the invertibility of the Legendre transform and the degree to which the Euler-Lagrange equations specify a second-order ODE system, expand the time derivative in the Euler-Lagrange equations (2.18):

$$\frac{\partial L}{\partial q} - \frac{\partial^2 L}{\partial q \partial \dot{q}} \cdot \dot{q} - \frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}} \cdot \ddot{q} = 0. \quad (2.25)$$

If the Hessian is non-degenerate, then  $\frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}}$  can be inverted and the Euler-Lagrange equations uniquely specify  $\ddot{q}$  as a function of  $q, \dot{q}$ . This non-degeneracy is the same condition required to invert the Legendre transform, Eq. (2.24).

Another important consequence of regularity is the ability to identify a Hamiltonian system corresponding to a Lagrangian system. Assuming the Lagrangian is regular, a *Hamiltonian function*  $H : T^*Q \rightarrow \mathbb{R}$  may be identified according to:

$$H(q, p) = p \cdot \dot{q}(q, p) - L(q, \dot{q}(q, p)), \quad (2.26)$$

where  $\dot{q}(q, p)$  is given by inverting the Legendre transform.

Upon identifying a Hamiltonian  $H(q, p)$ , the dynamics evolve according to *Hamilton's equations*:

$$\dot{q} = \frac{\partial H}{\partial p} \quad (2.27a)$$

$$\dot{p} = -\frac{\partial H}{\partial q}. \quad (2.27b)$$

Later, to distinguish from more general representations of Hamilton's equations, this set will be referred to as “Hamilton's equations in canonical coordinates” or “Hamilton's canonical equations”. It is a simple exercise to verify these equations are equivalent to the Euler-Lagrange equations when the Legendre transform is invertible.

Hamilton's equations may be written in terms of a *Hamiltonian vector field*  $X_H : T^*Q \rightarrow T(T^*Q)$  according to:

$$\dot{z} = X_H(z), \quad (2.28)$$

where  $z = (q, p)^T$  and

$$X_H = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial q} \\ \frac{\partial H}{\partial p} \end{pmatrix}. \quad (2.29)$$

The solution to Hamilton's equations is the flow map of this vector field, called the *Hamiltonian flow map*  $F_H : T^*Q \times \mathbb{R} \rightarrow T^*Q$ , or the *time-t Hamiltonian flow map*  $F_H^t : T^*Q \rightarrow T^*Q$ .

### Discrete Legendre Transform

In the discrete setting, there exist two discrete Legendre transforms corresponding to a given discrete Lagrangian  $L_d$ . The discrete Legendre transforms  $\mathbb{F}^+ L_d, \mathbb{F}^- L_d : Q \times Q \rightarrow T^*Q$  are

given by:

$$\mathbb{F}^+ L_d : (q_0, q_1) \mapsto (q_1, p_1) = (q_1, D_2 L_d(q_0, q_1)) \quad (2.30a)$$

$$\mathbb{F}^- L_d : (q_0, q_1) \mapsto (q_0, p_0) = (q_0, -D_1 L_d(q_0, q_1)). \quad (2.30b)$$

The discrete Legendre transforms enable the definition of regularity for discrete Lagrangians. Given a discrete Lagrangian  $L_d$ , is it regular if and only if the discrete Legendre transforms are invertible. Although there exist two discrete Legendre transforms, there is only a single regularity condition:

$$\det(D_1 D_2 L_d(q_0, q_1)) \neq 0. \quad (2.31)$$

If  $\det(D_1 D_2 L_d(q_0, q_1)) = 0$ , the discrete Lagrangian is said to be *degenerate*. Much like regularity in the continuous context, regularity of a discrete Lagrangian ensures that the discrete Lagrangian map is well-defined, that is the discrete Euler-Lagrange equations possess a unique solution  $q_{k+1}$  given a  $(q_{k-1}, q_k)$ .

In terms of identifying Hamiltonian dynamics, there does not exist a common notion of a “discrete Hamiltonian” as identified by a discrete Lagrangian and the discrete Legendre transforms. However, identification of discrete momenta corresponding to a pair of positions allows discussion of discrete trajectories evolving in the phase space  $T^*Q$ . One may cast the discrete Euler-Lagrange equations of Eq. (2.21) in *position-momentum* form:

$$p_k = -D_1 L_d(q_k, q_{k+1}) \quad (2.32a)$$

$$p_{k+1} = D_2 L_d(q_k, q_{k+1}). \quad (2.32b)$$

To verify the equivalence of the position-momentum time advance and that specified by the original discrete Euler-Lagrange equations, consider  $p_k^+$  to be the discrete momentum at time

$t_k$  identified by Eq. (2.32b) and  $p_k^-$  to be the discrete momentum at time  $t_k$  identified by Eq. (2.32a). One would hope that these two momenta are equal so that the two definitions are not contradictory; substituting the definitions:

$$p_k^+ - p_k^- = D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}).$$

So, for trajectories that satisfy the discrete Euler-Lagrange equations, we indeed have  $p_k^+ = p_k^-$ ; the definitions of the discrete momenta are in agreement, and the position-momentum equations are equivalent to the Euler-Lagrange equations.

Using the time advance specified by the position-momentum formulation of the discrete Euler-Lagrange equations, one can identify the *discrete Hamiltonian map*  $F_{H_d} : T^*Q \times \mathbb{R} \rightarrow T^*Q$  according to  $F_{H_d}(q_0, p_0) \mapsto (q_1, p_1)$  such that  $(q_1, p_1)$  satisfy:

$$p_0 = -D_1 L_d(q_0, q_1) \tag{2.33a}$$

$$p_1 = D_2 L_d(q_0, q_1). \tag{2.33b}$$

In practice, one can iterate the discrete Hamiltonian map as an alternative to iterating the discrete Lagrangian map: given  $(q_k, p_k)$ , solve the implicit Eq. (2.32a) for  $q_{k+1}$ , then evaluate the explicit Eq. (2.32b); this constitutes a single step of the discrete Hamiltonian map. Expressed intrinsically, one can define the discrete Hamiltonian map according to:

$$F_{H_d} = \mathbb{F}^- \circ F_{L_d} \circ (\mathbb{F}^-)^{-1}, \tag{2.34}$$

or similarly with  $\mathbb{F}^+$ . The relation is expressed diagrammatically in Fig. 2.1.

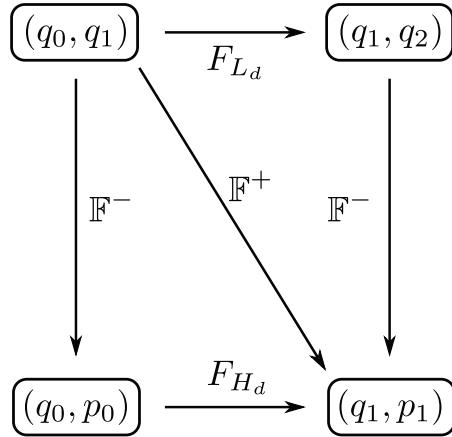


Figure 2.1: Relationship between the discrete Lagrangian map and discrete Hamiltonian map, as related by the discrete Legendre transforms. The diagram commutes; one obtains the same results regardless of the maps chosen to get from one set of variables to the other.

### 2.2.3 Symplecticity

#### Symplecticity in Continuous Mechanics

Beginning the discussion from the Hamiltonian perspective, where a physicist is most likely to first encounter a discussion of symplectic maps, the following theorem states that solutions to Hamilton's equations are symplectic with respect to the “canonical” symplectic structure  $\Omega_c$ . As detailed in Appendix A, a symplectic structure is a non-degenerate anti-symmetric tensor; the tensor components of  $\Omega_c$  may be written:

$$\Omega_c = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad (2.35)$$

where  $I$  is the  $n \times n$  identity matrix. Hopefully the tensor representation lends concreteness to the differential form representation used in the following Theorem.

**Theorem 2.2.3 (The Hamiltonian Flow Map is Symplectic)** *The time- $t$  Hamiltonian flow map  $F_H^t$  of the Hamiltonian vector field  $X_H$  defined in Eq. (2.29) is symplectic*

with respect to the canonical symplectic structure  $\Omega_c$  on  $T^*Q$ , with:

$$\Omega_c = dq^i \wedge dp_i. \quad (2.36)$$

That is,  $(F_H^t)^*\Omega_c = \Omega_c$ , where  $F^*\Omega$  denotes the pull-back of  $\Omega$  by  $F$  (see Sec. A.5).

**Proof** The proof of this statement (following Ref. [6]), proceeds in two steps: First, it will be shown that the pull-back of the symplectic structure is not changing in time. Next, the time  $t = 0$  flow map will be shown to satisfy the statement, so it must hold for *all* time (over which  $F_H^t$  is defined). For a more concise proof, refer to that of Thm. 2.3.2.

Letting  $z = (q, p)^T$  and  $z(t) = F_H^t(z_0, t)$ , consider the time derivative of the pull-back:

$$\frac{d}{dt} \left( \left( \frac{\partial z(t)}{\partial z_0} \right)^T \Omega_c \left( \frac{\partial z(t)}{\partial z_0} \right) \right) = \left( \frac{d}{dt} \frac{\partial z(t)}{\partial z_0} \right)^T \Omega_c \left( \frac{\partial z(t)}{\partial z_0} \right) + \left( \frac{\partial z(t)}{\partial z_0} \right)^T \Omega_c \left( \frac{d}{dt} \frac{\partial z(t)}{\partial z_0} \right).$$

The term  $\frac{d}{dt} \frac{\partial z(t)}{\partial z_0}$  is given by:

$$\frac{d}{dt} \frac{\partial z(t)}{\partial z_0} = \Omega_c^{-1} \frac{\partial^2 H}{\partial z^2}(z(t)) \frac{\partial z(t)}{\partial z_0},$$

This follows from interchanging the order of the derivatives on the left hand side and application of the equations of motion. The time derivative of the pull-back is then confirmed to be zero:

$$\begin{aligned} & \left( \frac{\partial z(t)}{\partial z_0} \right)^T \frac{\partial^2 H}{\partial z^2} (\Omega_c^{-1})^T \Omega_c \left( \frac{\partial z(t)}{\partial z_0} \right) + \left( \frac{\partial z(t)}{\partial z_0} \right)^T \Omega_c \Omega_c^{-1} \frac{\partial^2 H}{\partial z^2} \left( \frac{\partial z(t)}{\partial z_0} \right) = \\ & - \left( \frac{\partial z(t)}{\partial z_0} \right)^T \frac{\partial^2 H}{\partial z^2} \left( \frac{\partial z(t)}{\partial z_0} \right) + \left( \frac{\partial z(t)}{\partial z_0} \right)^T \frac{\partial^2 H}{\partial z^2} \left( \frac{\partial z(t)}{\partial z_0} \right) = 0. \end{aligned}$$

Next, because  $F_H^0(z_0, t) = \text{id}$ , where  $\text{id}$  is the identity map on  $T^*Q$ , then  $(F_H^0)^*\Omega_c = \Omega_c$ .

Therefore, it must hold for all time, and  $F_H^t$  is symplectic. ■

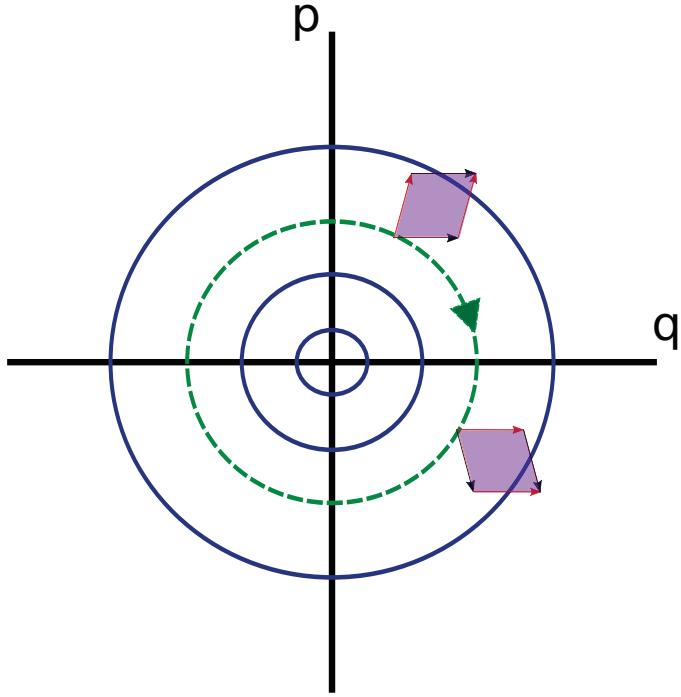


Figure 2.2: Solutions to Hamilton’s equations in canonical coordinates preserve areas in phase space; as two tangent vectors are dragged along with the flow, the area they span remains the same.

More intuitively, Thm. 2.2.3 implies that solutions to Hamilton’s canonical equations preserve phase-space area in the sense that the sum of the oriented areas in the  $p_i$ - $q^i$  planes is the same before and after the map. In a two-dimensional, one degree of freedom system, this reduces to the preservation of phase-space area, as illustrated in Fig. 2.2. See Appendix A for additional detail.

Turning now to the Lagrangian perspective, the Lagrangian flow map  $F_L^t$  is also symplectic, albeit with respect to a less intuitive symplectic structure.

**Theorem 2.2.4 (The Lagrangian Flow Map is Symplectic)** *Given a regular Lagrangian  $L$ , the time- $t$  Lagrangian flow map  $F_L^t$  is symplectic with respect to  $\Omega_L$  on  $TQ$  given by:*

$$\Omega_L = \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j} dq^i \wedge dq^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} d\dot{q}^i \wedge d\dot{q}^j. \quad (2.37)$$

That is,  $(F_L^t)^*\Omega_L = \Omega_L$ .

**Proof** Following Ref. [9], consider the “restricted” action  $\bar{S}$ , restricted in the sense that it acts only on trajectories that satisfy the discrete Euler-Lagrange equations. That is, consider trajectories  $\bar{q}(t) = F_L^t(q_0, \dot{q}_0)$  for some  $(q_0, \dot{q}_0) \in TQ$ , so:

$$\bar{S}(\bar{q}) = \bar{S}(q_0, \dot{q}_0) = \int_0^T L(\bar{q}(t), \dot{\bar{q}}(t)) dt.$$

Taking one exterior derivative (see Sec. A.4.2) of the action,

$$\begin{aligned} d\bar{S} &= \int_0^T \frac{\partial L}{\partial q^i}(\bar{q}(t), \dot{\bar{q}}(t)) \left( \frac{\partial \bar{q}^i(t)}{\partial q_0^j} dq^j + \frac{\partial \bar{q}^i(t)}{\partial \dot{q}_0^j} d\dot{q}^j \right) + \\ &\quad \frac{\partial L}{\partial \dot{q}^i}(\bar{q}(t), \dot{\bar{q}}(t)) \left( \frac{\partial \dot{\bar{q}}^i(t)}{\partial q_0^j} dq^j + \frac{\partial \dot{\bar{q}}^i(t)}{\partial \dot{q}_0^j} d\dot{q}^j \right) dt. \end{aligned}$$

Note that  $dq^j, d\dot{q}^j$  are the basis elements for  $T_{(q_0, \dot{q}_0)}^*(TQ)$ . Integrating by parts and using the fact that  $\bar{q}(t)$  satisfies the Euler-Lagrange equations leaves only the boundary term:

$$d\bar{S} = \left. \frac{\partial L}{\partial \dot{q}^i} \left( \frac{\partial \bar{q}^i}{\partial q_0^j} dq^j + \frac{\partial \bar{q}^i}{\partial \dot{q}_0^j} d\dot{q}^j \right) \right|_0^T,$$

or in terms of the pull-back:

$$d\bar{S} = (F_L^t)^* \frac{\partial L}{\partial \dot{q}^i} dq^i \Big|_0^T = (F_L^T)^* \frac{\partial L}{\partial \dot{q}^i} dq^i - \frac{\partial L}{\partial \dot{q}^i} dq^i.$$

Application of a second exterior derivative and  $d^2\bar{S} = 0$  yield:

$$(F_L^T)^*\Omega_L = \Omega_L,$$

with  $\Omega_L$  as given in Eq. (2.37). ■

The intuition behind the symplecticity of the Lagrangian flow map is that vectors tangent to  $TQ$  have the same areas before and after the advance when the sum is scaled by the non-constant tensor components of Eq. (2.37). Notice that  $\Omega_L$  may also be obtained from  $\Omega_c$  using the Legendre transform, precisely:  $\Omega_L = -\mathbb{F}L^*\Omega_c$ <sup>2</sup>.

### Symplecticity in Discrete Mechanics

In the discrete setting, one of the most important results is that the discrete Hamiltonian map is symplectic in *the same* sense as the continuous Hamiltonian map, as explained by the following theorem.

**Theorem 2.2.5** *The discrete Hamiltonian map  $F_H^d$  is symplectic with respect to the canonical symplectic structure  $\Omega_c$ .*

**Proof** Pursuing a direct calculation of the pull-back, we're interested in whether  $(F_H^d)^*\Omega_c = \Omega_c$ , or:

$$\left( \frac{\partial z_1}{\partial z_0} \right)^T \Omega_c \frac{\partial z_1}{\partial z_0} = \Omega_c,$$

with  $z_1 = (q_1, p_1)$  obtained from  $z_0 = (q_0, p_0)$  according to Eq. (2.33). By implicitly differentiating the position-momentum equations, one can obtain:

$$\frac{\partial z_1}{\partial z_0} = \begin{pmatrix} -(D_2 D_1 L_d)^{-1} (D_1 D_1 L_d) & -(D_2 D_1 L_d)^{-1} \\ -(D_2 D_2 L_d) (D_2 D_1 L_d)^{-1} (D_1 D_1 L_d) + D_1 D_2 L_d & -(D_2 D_2 L_d) (D_2 D_1 L_d)^{-1} \end{pmatrix},$$

with all functions evaluated at  $(q_0, q_1)$ . From this expression, it is straight forward matrix multiplication to verify  $\left( \frac{\partial z_1}{\partial z_0} \right)^T \Omega_c \frac{\partial z_1}{\partial z_0} = \Omega_c$  and therefore  $(F_H^d)^*\Omega_c = \Omega_c$ .  $\blacksquare$

As an alternative proof, we can appeal to generating function theory. It is known that one can perform a canonical coordinate transformation from “old” coordinates  $(q_0, p_0)$  to

---

<sup>2</sup>This minus sign is a bit of a nuisance of a convention. It is chosen to keep the Hamiltonian expressions consistent with e.g. Refs. [9, 4], although note that the minus sign is missing in some of the texts when comparing  $\Omega_L$  and  $\Omega_c$  via pullback.

“new” coordinates  $(q_1, p_1)$  based on a function  $f_1(q_0, q_1)$  and the relations:

$$\begin{aligned} p_0 &= -\frac{\partial f_1}{\partial q_0}(q_0, q_1) \\ p_1 &= \frac{\partial f_1}{\partial q_1}(q_0, q_1). \end{aligned}$$

The function  $f_1$  is referred to as a “type one generating function” [5, 21]. A comparison with the position-momentum formulation of the DEL equations (Eq. (2.33)) identifies the discrete Lagrangian  $L_d(q_0, q_1)$  as a type one generating function. By construction, the transformation given by the discrete Hamiltonian map is symplectic for *any* regular discrete Lagrangian  $L_d$ . ■

This theorem is arguably the most important result on variational integrators: the discrete Hamiltonian map is a symplectic integrator, preserving *the same* symplectic structure as solutions to Hamilton’s equations. The converse is also true, at least locally (c.f. Thm. 2.2.1 of [9]): given a symplectic integrator, there locally exists a type-one generating function, which one might interpret as a discrete Lagrangian function.

From the Lagrangian perspective, the discrete Lagrangian map is also symplectic:

**Theorem 2.2.6 (The discrete Lagrangian map is Symplectic)** *The discrete Lagrangian map is symplectic in the sense that:*

$$F_{L_d}^* \Omega_{L_d} = \Omega_{L_d}, \quad (2.38)$$

where:

$$\Omega_{L_d}(q_0, q_1) = d_1 d_2 L_d(q_0, q_1) = \frac{\partial^2 L_d(q_0, q_1)}{\partial q_0^i \partial q_1^j} dq_0^i \wedge dq_1^j. \quad (2.39)$$

**Proof** Appealing to the discrete variational principle, consider the *restricted discrete action*  $\bar{S}_d$ , restricted to act on paths  $\bar{q}_d$  satisfying the DEL equations:

$$\bar{S}_d(\bar{q}_d) = \bar{S}_d(q_0, q_1) = \sum_{k=0}^{N-1} L_d(F_{L_d}^k(q_0, q_1)) = \sum_{k=0}^{N-1} (F_{L_d}^k)^* L_d(q_0, q_1).$$

Taking an exterior derivative:

$$\begin{aligned} d\bar{S}_d &= \sum_{k=0}^{N-1} (F_{L_d}^k)^* d_1 L_d(q_0, q_1) + (F_{L_d}^k)^* d_2 L_d(q_0, q_1) \\ &= \sum_{k=1}^{N-1} \left[ (F_{L_d}^{k-1})^* d_2 L_d(q_0, q_1) + (F_{L_d}^k)^* d_1 L_d(q_0, q_1) \right] + \\ &\quad d_1 L_d(q_0, q_1) + (F_{L_d}^{N-1})^* L_d(q_0, q_1) \\ &= d_1 L_d(q_0, q_1) + (F_{L_d}^{N-1})^* L_d(q_0, q_1), \end{aligned}$$

where the terms inside the later summand are zero by satisfaction of the DEL equations. Operation by a second exterior derivative combined with  $d^2\bar{S} = 0$  obtains the desired result:

$$F_{L_d}^* d_1 d_2 L_d(q_0, q_1) = -d_2 d_1 L_d(q_0, q_1) = d_1 d_2 L_d(q_0, q_1). \quad \blacksquare$$

#### 2.2.4 Dissipation in Variational Principles

Although the Lagrangian and Hamiltonian formalisms as presented are quite general, it is naive to expect every dynamical system of interest to fit within these frameworks. In particular, although microscopic dynamics may possess a conservative description, one must often average over the shortest timescales, thereby introducing dissipation into the model. The utility of dissipative models for addressing real world problems is irrefutable; as such, one is greatly limited in scope if one cannot include dissipative dynamics in a continuous or discrete model.

Fortunately, established methods exist for extending the Lagrangian and Hamiltonian formalisms to include dissipative effects. In the Lagrangian setting a modification to Hamilton's principle of critical action known as a “Lagrange-d'Alembert variational principle” allows appending dissipation or “external forcing” to an otherwise Lagrangian system. The Legendre transform of such a system identified dissipative terms that must be appended to Hamilton's equations. Of course, these terms will violate certain conservation laws, but the formalisms remain useful by specifying *rates* at which previously invariant quantities dissipate and by facilitating coordinate transformations.

### Continuous Systems with Dissipation

Dissipation may be described by extending the Lagrangian framework as follows:

**Theorem 2.2.7 (Lagrange-d'Alembert Variational Principle)** *Let  $f_L : TQ \rightarrow T^*Q$  be an external forcing function in the sense that trajectories evolve according to the forced Euler-Lagrange equations:*

$$\frac{\partial L}{\partial q}(q, \dot{q}) - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) + f_L(q, \dot{q}) = 0. \quad (2.40)$$

*The forced Euler-Lagrange equations satisfy the Lagrange-d'Alembert variational principle, given by [9]:*

$$\delta \int_0^T L(q(t), \dot{q}(t)) dt + \int_0^T f_L(q(t), \dot{q}(t)) \cdot \delta q dt = 0. \quad (2.41)$$

**Proof** The proof is identical to that of Thm. 2.2.1, with the additional term in the variational principle appending the external forcing term to the Euler-Lagrange equations. ■

The Lagrange-d'Alembert variational principle ensures the addition of the forcing term in the Euler-Lagrange equations by inserting a  $\delta q$  at the appropriate place in the variational principle. Of course, because this additional term does not emerge from performing an

actual *variation*, the additional term will be present when analyzing conservation laws. For instance, the symplecticity proof for Lagrangian flow maps will fail due to the presence of this additional term. This is an appropriate result for systems that are simply not conservative, as in the following example.

### Example - Lorentz Force with Collisional Drag

The Lagrangian for a mass  $m$  particle with charge  $e$  interacting with electric field  $E = -\nabla\phi$  and magnetic field  $B = \nabla \times A$  is given by:

$$L_{LF}(q, \dot{q}) = \frac{1}{2}mg_{ij}(q)\dot{q}^i\dot{q}^j + eA_i(q)\dot{q}^i - e\phi(q). \quad (2.42)$$

If the Lorentz force particle is collisionally slowing down on a background plasma, it is common to model this with a drag force [99, 65, 49]:

$$(f_L(q, \dot{q}))_j = -g_{ij}(q)\nu \frac{v^3 + v_c^3}{v^3} \dot{q}^i. \quad (2.43)$$

Here,  $v^3 = (g_{ij}(q)\dot{q}^i\dot{q}^j)^{3/2}$ ,  $v_c$  is the critical velocity given by:

$$v_c^3 = \frac{3\sqrt{\pi}}{4} \frac{m_e}{m_i} \left( \frac{2k_B T_e}{m_e} \right)^{3/2}, \quad (2.44)$$

and  $\nu$  is the collision frequency:

$$\nu = \frac{n_e Z_{eff} q_e^2 e^3 \log \Lambda}{4\pi \epsilon_0^2 m_i m v_c^3}, \quad (2.45)$$

and the constants have their usual meanings <sup>3</sup>.

---

<sup>3</sup> $m_e$ : electron mass,  $m_i$ : ion mass,  $k_B$ : Boltzmann constant,  $T_e$ : electron temperature,  $n_e$ : plasma density,  $Z_{eff}$ : effective charge of background plasma,  $q_e$ : elementary charge,  $\Lambda = 12\pi\lambda_D^3$ ,  $\lambda_D$ : Debye length,  $\epsilon_0$ : permittivity of free space

The Lagrange-d'Alembert variational principle for the Lorentz force particle subject to collisional drag is given by:

$$\delta \int_0^T L_{LF}(q, \dot{q}) dt - \int_0^T \nu \frac{v^3 + v_c^3}{v^3} g_{ij}(q) \dot{q}^i \delta q^j dt = 0. \quad (2.46)$$

The corresponding forced Euler-Lagrange equations are, as desired:

$$\left( \frac{1}{2} g_{ik,j} - g_{ij,k} \right) m \dot{q}^i \dot{q}^k - mg_{ij} \ddot{q}^i + e(A_{i,j} - A_{j,i}) \dot{q}^i - e\phi_{,j} - \nu \frac{v^3 + v_c^3}{v^3} g_{ij} \dot{q}^i = 0. \quad (2.47)$$

External forcing may similarly be incorporated into canonical Hamiltonian systems by considering *forced Hamilton's equations*:

$$\dot{q} = \frac{\partial H}{\partial p}(q, p) \quad (2.48a)$$

$$\dot{p} = -\frac{\partial H}{\partial q}(q, p) + f_H(q, p), \quad (2.48b)$$

where  $f_H$  is the *Hamiltonian external forcing function*. Where appropriate,  $f_H$  may be obtained from  $f_L$  via the Legendre transform, and the forced Hamilton's equations will be equivalent to the forced Euler-Lagrange equations [9, 86]. As an example, consider formulating the charged particle subject to collisional drag from a Hamiltonian perspective.

### Example - Hamiltonian Lorentz Force with Collisional Drag

The Legendre transform yields, for the Lorentz force problem:

$$\begin{aligned} p_i(q, \dot{q}) &= mg_{ij}(q) \dot{q}^j + eA_i(q) \\ \dot{q}^i(q, p) &= \frac{1}{m} g^{ij}(q) (p_j - eA_j(q)). \end{aligned}$$

The Hamiltonian is then given by:

$$H_{LF}(q, p) = \frac{1}{2m} \|p - eA\|^2 + e\phi, \quad (2.49)$$

and the Hamiltonian external forcing function by:

$$(f_H(q, p))_i = -\frac{\nu}{m} \frac{v_H^3 + v_c^3}{v_H^3} (p_i - eA_i), \quad (2.50)$$

where  $v_H(q, p) = (g_{ij}\dot{q}^i(q, p)\dot{q}^j(q, p))^{3/2}$ .

A natural question at this point is: if the presence of external forcing violates the conservative nature of Lagrangian and Hamiltonian systems, what benefits are retained by the modified framework? More simply: *why* would one want to use these extended descriptions? One of the primary reasons is that the formalisms identify precisely *how* the external forcing impacts the conservation laws. For instance, there exists a modified Noether's theorem describing the impact of a given external forcing on the Noether invariants; it is possible that the invariants remain intact [9]. If the dissipation is only in the  $x$ -direction, the  $y$ -momentum may remain conserved. If the dissipative effects are slow relative to the non-dissipative dynamics, many of the benefits of the Lagrangian and Hamiltonian descriptions will remain largely intact.

### Discrete Systems with Dissipation

**Theorem 2.2.8 (Discrete Lagrange-d'Alembert Variational Principle)** *A discrete external forcing function is a one-form  $f_d : Q \times Q \rightarrow T^*(Q \times Q)$ , that is:*

$$f_d(q_0, q_1) = f_d^-(q_0, q_1) dq_0 + f_d^+(q_0, q_1) dq_1. \quad (2.51)$$

The discrete Lagrange-d'Alembert Variational Principle states that discrete trajectories governed by a discrete Lagrangian  $L_d$  and a discrete external forcing  $f_d$  satisfy:

$$\delta \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) + \sum_{k=0}^{N-1} f_d(q_k, q_{k+1}) \cdot (\delta q_k, \delta q_{k+1}). \quad (2.52)$$

Discrete trajectories satisfying this condition also satisfy the forced discrete Euler-Lagrange equations:

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) + f_d^+(q_{k-1}, q_k) + f_d^-(q_k, q_{k+1}) = 0 \quad (2.53)$$

**Proof** Again, the proof proceeds identically to the unforced case; the dissipative terms appear in the forced discrete Euler-Lagrange equations upon collecting in  $\delta q_k$ . ■

The appropriate modification to the Hamiltonian setting due to these external forces is given by the *forced position-momentum equations*, which take the form:

$$p_k = -D_1 L_d(q_k, q_{k+1}) - f_d^-(q_k, q_{k+1}) \quad (2.54a)$$

$$p_{k+1} = D_2 L_d(q_k, q_{k+1}) + f_d^+(q_k, q_{k+1}). \quad (2.54b)$$

Correspondingly, the *forced Hamiltonian map* updates some  $(q_0, p_0)$  to  $(q_1, p_1)$  according to these equations.

In both the continuous and discrete settings, modifications to the familiar variational principles allows for external forcing and guides modification of statements about conserved quantities. Even though strict conservation of energy or momentum may likely be violated, it remains useful to prescribe or calculate the time rate of change these quantities according to some physical source of dissipation.

### 2.2.5 Correspondence between Continuous and Discrete Mechanics

Up until now, the theory of discrete mechanics has been presented as self-contained, in particular, without recourse to discretizing some continuous system. Of course, the *motivation* for carefully establishing a discrete theory of mechanics is to be able to develop and analyze numerical methods as discrete models of a continuous system. This subsection describes the relationship between continuous and discrete mechanics as encompassed by the framework of variational integration.

#### Accuracy

The first task toward bridging the gap between the continuous and discrete theories is to establish the sense in which a *variational integrator* - defined by the discrete Lagrangian map and determined by a discrete Lagrangian function - approximates the solution to some continuous Euler-Lagrange equations. To do so, one requires a notion of accuracy and precise identification of what a discrete Lagrangian is approximating.

The foundation for definitions of accuracy and relating continuous and discrete systems relies on an “exact discrete Lagrangian”, defined as follows [9]:

**Definition** The *exact discrete Lagrangian*,  $L_d^E : Q \times Q \rightarrow \mathbb{R}$ , is given by:

$$L_d^E(q_0, q_1, h) = \int_0^h L(\bar{q}(t), \dot{\bar{q}}(t)) dt, \quad (2.55)$$

where  $\bar{q}$  is the unique trajectory satisfying  $\bar{q}(0) = q_0$ ,  $\bar{q}(h) = q_1$  and the Euler-Lagrange equations.

Several comments on this definition are warranted: The existence of  $\bar{q}$  connecting  $q_0$  and  $q_1$  over some time  $h$  is guaranteed for sufficiently small  $h$  and sufficiently nearby  $q_0, q_1$  provided

the Lagrangian is regular [9]. Intuitively, regularity of the Lagrangian ensures solutions exist for the initial value problem given by the Euler-Lagrange equations and an initial condition  $(q_0, \dot{q}_0)$ , at least over some time interval. For each  $(q_0, \dot{q}_0)$  there will be a unique  $(q_0, q(h))$ , so the existence for a particular choice of  $(q_0, q_1)$  in the definition of the exact discrete Lagrangian is isomorphic to a choice of  $(q_0, \dot{q}_0)$ . The next explanatory observation for this definition is that the “exact”-ness of  $L_d^E$  refers to its ability to *exactly* sample the flow of the continuous dynamics at the discrete time intervals. This is most readily seen by calculating the discrete Legendre transforms:

$$\begin{aligned} p_0 = -D_1 L_d^E(q_0, q_1) &= -\frac{\partial}{\partial q_0} \int_0^h L(\bar{q}(t), \dot{\bar{q}}(t)) dt \\ &= -\int_0^h \left( \frac{\partial L}{\partial q} \frac{\partial \bar{q}(t)}{\partial q_0} + \frac{\partial L}{\partial \dot{q}} \frac{\partial \dot{\bar{q}}(t)}{\partial q_0} \right) dt \\ &= -\int_0^h \left( \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \frac{\partial \bar{q}(t)}{\partial q_0} dt - \left. \frac{\partial L}{\partial \dot{q}} \frac{\partial \bar{q}}{\partial q_0} \right|_0 \\ &= \frac{\partial L}{\partial \dot{q}}(\bar{q}_0, \dot{\bar{q}}_0) = \bar{p}(0). \end{aligned}$$

A similar calculation verifies:

$$p_1 = \bar{p}(h).$$

Therefore, an advance according to the position-momentum equations using  $L_d^E$  and an advance according to the flow map of the true Hamiltonian system are *equivalent* - concisely:  $F_{H_d}^h = F_H$ . A similar statement may be made about the discrete Lagrangian map; the trajectory of the variational integrator exactly samples the true solution at the discrete time instances with the variational integrator is constructed from the exact discrete Lagrangian  $L_d$ .

Of course, the practical utility of the exact discrete Lagrangian is not the construction of an exactly correct numerical integrator. Determination of the exact discrete Lagrangian requires knowledge of the exact solution over the time intervals of interest. If one knew

the exact solution, a numerical model is not delivering new insight. Instead, the utility of the exact discrete Lagrangian is in enabling a definition of order of some other discrete Lagrangian, in facilitating calculation of such an order by familiar Taylor expansion methods, and in guiding the construction of novel and high-order variational integrators.

Beginning with the first of these applications, the order of a discrete Lagrangian has the following definition:

**Definition** A discrete Lagrangian  $L_d$  is said to be *convergent of order r* if there exists an open subset  $U \subset TQ$  and constants  $C, \bar{h}$  such that

$$\|L_d(q_0, q_1, h) - L_d^E(q_0, q_1, h)\| \leq Ch^{r+1} \quad (2.56)$$

for all initial conditions  $(q, \dot{q}) \in TQ$  and all  $h \leq \bar{h}$  [9].

The order of a discrete Lagrangian may be readily calculated by Taylor expanding both the discrete Lagrangian and the exact discrete Lagrangian about  $t = 0$ , and determining the truncation error as per the usual procedure for a numerical algorithm. Helpfully, the order of the discrete Lagrangian is directly related to the order of the resulting variational integrator, as established by the following theorem as found in e.g. Refs. [9, 86].

**Theorem 2.2.9** *If a discrete Lagrangian of order r, then its discrete Hamiltonian map is an order-r approximation of the Hamiltonian flow map, provided the Lagrangian is regular.*

**Proof** See Ref. [9, 86]. ■

This theorem also merits some clarifying statements. It is indeed the discrete Lagrangian and the discrete Hamiltonian maps that have the *same* order; the same cannot be said in general for the discrete Lagrangian and the Lagrangian flow map because occasionally the Lagrangian flow map is higher-order accurate. For an illustrative example, given  $L(q, \dot{q}) =$

$\frac{1}{2}m\dot{q}^2 - V(q)$ , the following three discrete Lagrangians:

$$\begin{aligned} L_d^1(q_0, q_1) &= hL(q_0, \frac{q_1 - q_0}{h}) \\ L_d^2(q_0, q_1) &= hL(q_1, \frac{q_1 - q_0}{h}) \\ L_d^3(q_0, q_1) &= \frac{h}{2} \left( L(q_0, \frac{q_1 - q_0}{h}) + L(q_1, \frac{q_1 - q_0}{h}) \right), \end{aligned}$$

yield the *same* discrete Euler-Lagrange equations but *different* position-momentum equations. The discrete Euler-Lagrange equations are second-order accurate for all three discrete Lagrangians, whereas the position-momentum equations are first-order accurate for  $L_d^1$  and  $L_d^2$  and second-order accurate for  $L_d^3$ . Lagrangians like this are said to be “weakly equivalent” [9]. Thus, it is only the discrete Hamiltonian map with an order equal to that of the discrete Lagrangian.

As a final application of the exact discrete Lagrangian, constructing numerical algorithms based on discretizing intervals of the action yields novel methods unlikely to emerge from discretizing the equations of motion alone. Certain discretization procedures yield well-known methods, including symplectic Runge-Kutta methods [9, 18]. Symplectic, partitioned Runge-Kutta methods may be constructed by first choosing a polynomial representation for  $q(t)$  then applying Gaussian quadrature to approximate the action interval [100]. An interesting alternative is to use Hermite polynomial representations and apply Euler-Maclaurin quadrature [101]. This latter approach yields novel methods that cannot be expressed as evaluations of the Euler-Lagrange equations [101]. Other discretization tactics include shooting methods [102] and spectral methods [103]. These varied approaches offer an array of potentially new algorithms for testing in practical applications.

## Symplecticity

One of the most important benefits of variational integrators is that they are guaranteed to be equivalent to a symplectic integrator as long as the discrete Lagrangian is consistent (of order  $r \geq 1$ ). Formally:

**Theorem 2.2.10 (Variational Integrators are Symplectic)** *Given a regular Lagrangian  $L$  with an action approximated by a consistent discrete Lagrangian  $L_d$ , the discrete Hamiltonian map of the resulting variational integrator is an order- $r$  symplectic integrator for Hamilton's equations Eq. (2.27).*

**Proof** This theorem is really a corollary of two other theorems: Theorem 2.2.5 shows the discrete Hamiltonian map is symplectic, and Theorem 2.2.9 relates the orders. ■

Because all approximations are performed at the level of the variational principle, truncation error may be introduced at this point without violating one of the most fundamental conservation properties of classical mechanics. Rather than guessing and checking whether a given discretization of Hamilton's equations yields a symplectic integrator, one may systematically discretize the action to a given order and rest assured the discrete Hamiltonian map is a symplectic integrator. This explains how even a naive choice of discrete Lagrangian, as in the Störmer-Verlet example, yielded a well-behaved algorithm.

## Benefits of Symplectic Integrators

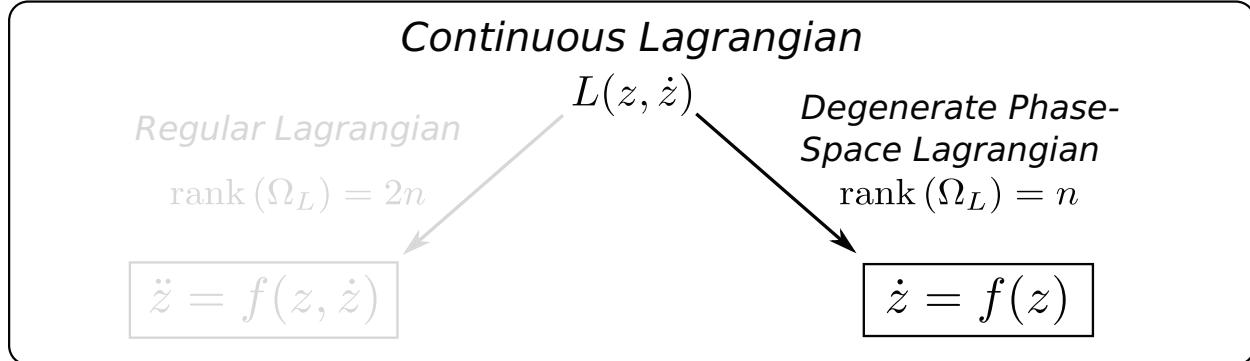
To motivate the use of variational integrators, symplectic integrators offer many benefits for modeling Hamiltonian systems. These benefits all stem from the following fact: *symplectic integrators generate trajectories that are Hamiltonian* [6, 7]. They are Hamiltonian in the following sense: If one performs a backward error analysis of the symplectic integrator, the vector field describing the time evolution of the numerical system is symplectic for every

truncation order. In general, this expansion is an asymptotic series in the numerical step size  $h$ , and must be truncated at optimal order. So, more rigorously, symplectic integrators generate trajectories that are *nearby* to the true solution of some Hamiltonian system. I find it remarkable that, although the introduction of truncation error is inevitable, symplectic integrators manage to remain on the same class of dynamical systems, namely, Hamiltonian systems.

One well-known consequence is that the energy error of symplectic integrators is bounded for exponentially long times for time-independent Hamiltonian systems [6]. This ensues from the Hamiltonian nature of trajectories generated by symplectic integrators; they reside on the constant energy surfaces of the modified Hamiltonian system. Throughout a trajectory, the energy error oscillates as the modified constant energy surfaces interweave the original constant energy surfaces, differing by a small amount proportional to the numerical step size.

Another important consequence of symplectic integrators is the preservation of key features of the phase portrait [7]. Symplectic integrators accurately represent equilibrium points and their eigenspectra, retain KAM tori, and exhibit Hamiltonian chaos. Certain Hamiltonian phenomena, including Arnold diffusion [21, 38, 39, 40], may only be reliably captured by using symplectic integrators [41, 42].

## 2.3 Action Principles in Phase Space



### Section Summary

- Hamiltonian systems are defined by a symplectic structure (non-degenerate anti-symmetric tensor) and a Hamiltonian function
- Phase-space action principles derive Hamilton's equations as the Euler-Lagrange equations of a phase-space Lagrangian
- The phase-space Lagrangian is degenerate
- Important applications of action principles in phase space include magnetic field line flow and the dynamics of guiding center trajectories
- To include dissipation, the Lagrange-d'Alembert variational principle has been extended to phase-space action principles for the first time

Several of the fundamental dynamical systems encountered in plasma physics differ from the conventional Lagrangian descriptions in that the Lagrangian describing their dynamics is *degenerate* [97, 8, 78]. Of primary importance is the guiding center system, which forms the cornerstone of test particle calculations and particle-based drift- and gyro-kinetic

simulations. The guiding center Lagrangian belongs to a class of degenerate Lagrangians categorized as *phase-space Lagrangians*. The corresponding variational principle involving phase-space Lagrangians is referred to as *Hamilton's principle in phase space* or a *phase-space action principle*. This section reviews Hamilton's principle in phase space as a facilitator of non-canonical coordinate transformations, which prove helpful for performing dynamical reductions. To incorporate dissipative effects, the Lagrange-d'Alembert approach is extended to these phase-space action principles. Finally, an illustrative example shows how the degenerate Lagrangian emerges for the guiding center system.

### 2.3.1 Hamiltonian Systems on Symplectic Manifolds

First, a more general notion of Hamiltonian systems is required than that encountered in Eq. (2.27). The scope of Hamiltonian systems considered in this thesis is: *Hamiltonian systems on exact symplectic manifolds*<sup>4</sup>. A symplectic manifold  $(M, \Omega)$  includes a differentiable manifold  $M$  equipped with a non-degenerate two-form  $\Omega$  that is closed, i.e.  $d\Omega = 0$ . If the symplectic structure  $\Omega$  is also exact, then there exists a one-form  $\vartheta$  on  $M$  such that  $\Omega = -d\vartheta$ . Given a Hamiltonian function  $H : M \rightarrow \mathbb{R}$ , one can define a Hamiltonian vector field  $X_H$  on the symplectic manifold according to:

$$\mathbf{i}_{X_H} \Omega = dH, \quad (2.57)$$

where  $\mathbf{i}$  is the interior product [4]. To unwrap this intrinsic expression in terms of coordinates  $z^i$  on  $M$ , recall that  $\Omega = \Omega_{ij} dz^i \wedge dz^j = \Omega_{ij} (dz^i \otimes dz^j - dz^j \otimes dz^i)$  and calculate the interior product to obtain:

$$(\Omega_{ij}(z) - \Omega_{ji}(z)) X_H^i(z) = H_{,j}(z). \quad (2.58)$$

---

<sup>4</sup>The most general class of Hamiltonian systems is those on Poisson manifolds - manifolds equipped with a Poisson bracket. Because a symplectic two-form can be used to construct a Poisson bracket, a symplectic manifold is a Poisson manifold, but the converse is not necessarily true [4].

The non-degeneracy of the symplectic structure  $\Omega$  ensures  $X_H$  is well-defined. Given a Hamiltonian vector field, Hamilton's equations are:

$$\dot{z}(t) = X_H(t). \quad (2.59)$$

A set of coordinates  $z$  on  $M$  is called *canonical* if, in these  $\dim(M) = 2n$  coordinates may be partitioned into  $z = (q^1, \dots, q^n, p_1, \dots, p_n)$  such that  $\Omega$  has the canonical form:

$$\Omega(z) = \Omega_c = dq^i \wedge dp_i. \quad (2.60)$$

Computation of Hamilton's equations using this canonical symplectic structure yields Eq. (2.27). Any *other* coordinates  $z$  on  $M$  are deemed *non-canonical*. Therefore, not all Hamiltonian vector fields take the form Eq. (2.29); the more general notion is represented in Eq. (2.58).

Of course, the *dynamics* are independent of the coordinates used to represent them. Indeed, the intrinsic definition of a Hamiltonian vector field in Eq. (2.57) makes *no* reference to coordinates whatsoever. In the continuous setting, the distinction between canonical and non-canonical coordinates is largely one of convenience. Moreover, given non-canonical coordinates, the *Darboux theorem* guarantees the existence of canonical coordinates, at least locally.

In the discrete setting, however, the presence or absence of canonical coordinates has significant ramifications on the availability of symplectic integration methods. At present, there exist *no known methods for symplectic integration in non-canonical coordinates* [6, 75]. In the literature, an integrator preserving a non-canonical symplectic structure is often referred to as a “Poisson integrator”, but this dissertation will favor the terminology “non-canonical symplectic integrator”.

### 2.3.2 Action Principles in Phase Space

The conventional procedure relating variational principles and Hamiltonian systems appeals to the Legendre transform (see Sec. 2.2.2). Provided the symplectic structure is exact, however, Hamilton's equations may be *directly* derived from a variational principle, as demonstrated by the following theorem.

**Theorem 2.3.1 (Hamilton's Principle of Least Action in Phase Space)** *Consider an exact symplectic manifold  $(M, \Omega)$ ,  $\Omega = -d\vartheta$  with Hamiltonian  $H$ . Let  $z : [0, T] \subset \mathbb{R} \rightarrow M$  denote a path in the Hamiltonian phase space, and let the phase-space Lagrangian  $L_{PS} : TM \rightarrow \mathbb{R}$  and phase-space action  $S_{PS}$  be defined according to:*

$$L_{ps}(z, \dot{z}) = \vartheta_i(z) \dot{z}^i - H(z) \quad (2.61)$$

$$S_{ps}(z) = \int_0^T [\vartheta_i(z(t)) \dot{z}^i(t) - H(z(t))] dt, \quad (2.62)$$

respectively. Then Hamilton's principle of least action in phase space states that a path  $z$  with fixed endpoints is a critical point of the phase-space action if and only if it satisfies the phase-space Euler-Lagrange equations:

$$(\vartheta_{i,j}(z) - \vartheta_{j,i}(z)) \dot{z}^i = H_{,j}(z), \quad (2.63)$$

for all  $t \in [0, T]$ . The phase-space Euler-Lagrange equations describe a Hamiltonian system.

**Proof** This is demonstrated by varying the phase-space action with respect to arbitrary path variations  $\delta z$  with fixed end-points:

$$\begin{aligned} dS_{ps} \cdot \delta z &= \int_0^T \vartheta_{i,j} \dot{z}^i \delta z^j + \vartheta_i \delta \dot{z}^i - H_{,j} \delta z^j dt \\ &= \int_0^T [(\vartheta_{i,j} - \vartheta_{j,i}) \dot{z}^i - H_{,j}] \delta z^j dt + \vartheta_i \delta z^i \Big|_0^T \end{aligned}$$

The fixed endpoint condition states  $\delta z(0) = \delta z(t) = 0$ , so requiring  $dS_{ps} \cdot \delta z = 0$  for arbitrary  $\delta z$  implies the integrand is zero for all  $t \in [0, T]$ , yielding Eq. (2.63). Comparing this result with Eqs. (2.59) and (2.58) verifies that the phase-space Euler-Lagrange equations and Hamilton's equations on an exact symplectic manifold are equivalent. ■

The nomenclature “phase-space” Lagrangian refers to the fact that the domain of the phase-space Lagrangian is the (tangent bundle of) a symplectic manifold, i.e. the “phase space” for a Hamiltonian system. This nomenclature is carried over to the “phase-space” action. Note that it is sometimes, but not always, the case that  $M = T^*Q$ , for some  $Q$ <sup>5</sup>. By “phase space”, in this context, I refer to a symplectic manifold on which a Hamiltonian vector field has been defined, and not necessarily the cotangent bundle  $T^*Q$  of some  $Q$ .

The phase-space action principle is a hybrid of conventional Lagrangian and Hamiltonian mechanics: the Euler-Lagrange equations corresponding to  $L_{PS}$  are Hamilton's equations; the phase-space Lagrangian flow map is a Hamiltonian flow map, i.e.  $F_{L_{PS}} = F_H$ . One consequence of the hybrid approach is a rapid demonstration of the symplecticity of the phase-space Lagrangian and therefore Hamiltonian flow map, even in non-canonical coordinates.

**Theorem 2.3.2 (Phase-Space Lagrangian Flow is Symplectic)** *The Lagrangian flow map of a phase-space Lagrangian  $F_{L_{PS}}$  preserves the symplectic structure  $\Omega$  on  $M$ , i.e.:*

$$F_{L_{PS}}^* \Omega = \Omega. \quad (2.64)$$

**Proof** The proof proceeds by considering a restricted action, like that of Thm. 2.2.4, and arrives at Thm. 2.2.3 as a specialized case. Consider then the restricted phase-space action

---

<sup>5</sup>Only when global canonical coordinates are available on  $M$  and the Hamiltonian is regular, i.e.  $\mathbb{F}H$  is invertible.

$\bar{S}_{PS}$ , restricted to act on trajectories  $\bar{z}$  that satisfy Eq. (2.63):

$$\begin{aligned}\bar{S}_{PS}(\bar{z}) &= \bar{S}_{PS}(z_0) = \int_0^T L_{PS}(\bar{z}(t), \dot{\bar{z}}(t)) dt \\ &= \int_0^T \vartheta_i(\bar{z}(t)) \dot{\bar{z}}^i(t) - H(\bar{z}(t)) dt.\end{aligned}$$

Operating on the restricted action with the exterior derivative:

$$\begin{aligned}d\bar{S}_{PS} &= \int_0^T \vartheta_{i,j} \dot{\bar{z}}^j \frac{\partial \bar{z}^i}{\partial z_0^k} dz^k + \vartheta_i \frac{\partial \bar{z}^i}{\partial z_0^j} dz^j - H_{,j} \frac{\partial \bar{z}^i}{\partial z_0^j} dz^j dt \\ &= \int_0^T [(\vartheta_{i,j} - \vartheta_{j,i}) \dot{\bar{z}}^i - H_{,j}] \frac{\partial \bar{z}^j}{\partial z_0} dz^k dt + \vartheta_i \frac{\partial \bar{z}^i}{\partial z_0^j} dz^j \\ &= F_{L_{PS}}^t \vartheta|_0^T\end{aligned}$$

Performing a second exterior derivative and using  $d^2\bar{S}_{PS} = 0$  obtains:

$$\begin{aligned}F_{L_{PS}}^t \vartheta|_0^T &= -d\vartheta \\ F_{L_{PS}}^t \Omega &= \Omega\end{aligned}$$

with  $\Omega = -d\vartheta = \vartheta_{i,j} dz^i \wedge dz^j$ . ■

The main benefit of the phase-space action principle is the freedom to transform *all* of the phase-space coordinates. In the Lagrangian formalism, one is free to transform the configuration-space coordinates, and the variational principle derives the appropriate equations of motion in these coordinates. In the conventional (regular Lagrangian) setting, this means one can transform  $n$  coordinates whose time evolution is specified by a second-order differential equation. The Hamiltonian formulation of the dynamics places the position and momentum coordinates on equal footing. In the absence of the phase-space variational principle, however, transformations of the Hamiltonian coordinates must remain canonical (or

one must properly transform the symplectic structure or Poisson bracket). The action principle in phase-space places all phase-space coordinates on equal footing, while retaining the ease of coordinate transformations allotted by a variational principle [78].

The cost of the hybrid Lagrangian/Hamiltonian approach embodied by the phase-space action principle is the degeneracy of the Lagrangian. One consequence of this is that the Euler-Lagrange equations do not uniquely specify a second-order differential equation. Of course, because first-order Hamilton's equations are desired as the equations of motion, this is a desired consequence. Another consequence is that  $\Omega_{LPS}$  is degenerate on  $T^*M$ , cf. Eq. (2.37). It is non-degenerate on the lower-dimensional space  $M$ , as again expected from the Hamiltonian formulation. Finally, the Legendre transform  $\mathbb{F}L_{PS}$  is not invertible. Although it is possible to discuss the dynamics as a differential-algebraic system on  $T^*M$  [87, 88, 92], the motivations for doing so are less apparent than in the conventional regular Lagrangian context. One *already* has a Hamiltonian system as the phase-space Euler-Lagrange equations, so  $\mathbb{F}L_{PS}$  should not be expected to recover a familiar Hamiltonian description. In certain cases  $\mathbb{F}H$  may be invertible, but this requires canonical coordinates.

**Example - Canonical Phase-Space Variational Principle** The simplest case of a phase-space action principle is that of a canonical Hamiltonian system. The relevant definitions are:

$$\begin{aligned} z &= (q^1, \dots, q^n, p_1, \dots, p_n) \\ \vartheta(z) &= p_i dq^i \\ \Omega(z) &= dq^i \wedge dp_i \\ L_{PS}(z, \dot{z}) &= p_i \dot{q}^i - H(q, p). \end{aligned}$$

The phase-space Euler-Lagrange equations are given by:

$$\begin{aligned}\dot{q}^i - H_{,p}(q, p) &= 0 \\ -\dot{p}^i - H_{,q}(q, p) &= 0,\end{aligned}$$

which verifies that Hamilton's canonical equations are recovered.

### 2.3.3 Phase-Space Lagrange-d'Alembert

To include dissipation in phase-space action principles, it is necessary to extend the Lagrange-d'Alembert approach. As far as I am aware, this variational principle does not exist elsewhere in the literature.

**Theorem 2.3.3 (Lagrange-d'Alembert Principle in Phase Space)** *Consider an exact symplectic manifold  $(M, \Omega)$  with Hamiltonian  $H$  and coordinates  $z$  on  $M$ . Further, consider an external forcing vector field  $f_{PS}$  on  $M$ . The phase-space Lagrange-d'Alembert variational principle seeks curves  $z$  satisfying:*

$$\delta \int_0^T L_{PS}(z(t), \dot{z}(t)) dt + \int_0^T \Omega_{z(t)}(f_{PS}(z(t)), \delta z(t)) dt = 0. \quad (2.65)$$

A curve  $z$  with fixed endpoints satisfies the phase-space Lagrange-d'Alembert variational principle if and only if it satisfies the forced phase-space Euler-Lagrange equations:

$$\Omega_{ij}(z(t))\dot{z}^i(t) - H_{,j}(z(t)) - \Omega_{ij}(z(t))f_{PS}^i(z(t)) = 0, \quad (2.66)$$

for all  $t \in [0, T]$ . An equivalent expression is the forced Hamilton's equations:

$$\dot{z} = X_{Hf} \quad (2.67a)$$

$$X_{Hf} = X_H + f_{PS}. \quad (2.67b)$$

**Proof** Calculating the variation in the phase-space Lagrange-d'Alembert principle:

$$\begin{aligned} \int_0^T \vartheta_{i,j} \dot{z}^i \delta z^j + \vartheta_i \delta \dot{z}^i - H_{,i} \delta z^i - \Omega_{ij} f_{PS}^i \delta z^j \, dt &= 0 \\ \int_0^T (\Omega_{ij} \dot{z}^i - H_{,j} - \Omega_{ij} f_{PS}^i) \delta z^j \, dt &= 0, \end{aligned}$$

where the fixed endpoints condition has been used to eliminate the boundary terms in the integration by parts. Requiring this to be zero for all  $\delta z^j$  recovers Eq. (2.66). Inverting the tensor associated with the symplectic two-form obtains the vector field formulation of Eq. (2.67). ■

### 2.3.4 Applications in Plasma Physics

#### Non-canonically Hamiltonian Magnetic Field Line Flow

The first application considers magnetic field line flow, cast as a non-canonically Hamiltonian system [97]. Here, the phase-space Lagrangian appears from re-parameterizing a degenerate Lagrangian. The system will later serve as a valuable test case, representing a non-trivial yet low-dimensional Hamiltonian system in non-canonical coordinates.

Consider a non-vanishing magnetic field  $B = (B^1, B^2, B^3)$ . The equations of motion for the field line are given by:

$$\frac{dx^i}{dt} = B^i, \quad i = 1, 2, 3, \quad (2.68)$$

where  $t$  is a time-like variable parameterizing the distance along the field line.

The field line equations satisfy the Euler-Lagrange equations corresponding to the following degenerate (but not phase-space!) Lagrangian:

$$L(x, \dot{x}) = A_i(x)\dot{x}^i, \quad (2.69)$$

where  $A$  is the magnetic vector potential,  $B = \nabla \times A$ . The corresponding Euler-Lagrange equations are given by:

$$(A_{i,j}(x(t)) - A_{j,i}(x(t)))\dot{x}^i = 0, \quad j = 1, 2, 3. \quad (2.70)$$

Strictly speaking, Eq. (2.68) is not equivalent to Eq. (2.70). The latter essentially states: “ $v \times B = 0$ ”, while the former is a well-defined system of first-order differential equations. Put another way, there exists re-parameterization invariance in the action corresponding to the Lagrangian of Eq. (2.69): if  $t \mapsto t(\tau)$ , the Euler-Lagrange equations will remain the same, but in terms of  $\tau$ , regardless of the function  $t(\tau)$ .

One way to “break” this re-parameterization invariance, and simultaneously recover the non-canonical Hamiltonian formulation of magnetic field line flow, is parameterize the motion of two of the coordinates, say  $x^1, x^2$ , in terms of the other, say  $x^3$ . The description will remain valid wherever the component of the magnetic field corresponding to the new time-like parameter remains non-zero. In terms of the Lagrangian, the re-parameterized Lagrangian becomes:

$$L(x^1, x^2, x^{1'}, x^{2'}, x^3) = A_1(x^1, x^2, x^3)x^{1'} + A_2(x^1, x^2, x^3)x^{2'} + A_3(x^1, x^2, x^3), \quad (2.71)$$

where the prime denotes the derivative with respect to  $x^3$ . The corresponding action is:

$$S = \int_{x_0^3}^{x_T^3} (A_1(x^1(x^3), x^2(x^3), x^3)x^{1'} + A_2(x^1(x^3), x^2(x^3), x^3)x^{2'} + A_3(x^1(x^3), x^2(x^3), x^3)) dx^3, \quad (2.72)$$

which may be recognized as a phase-space action with:

$$\vartheta = A_1 dx^1 + A_2 dx^2 \quad (2.73)$$

$$H = -A_3. \quad (2.74)$$

The phase-space Euler-Lagrange equations may be written as:

$$\frac{dx^1}{dx^3} = \frac{B^1}{B^3} \quad (2.75a)$$

$$\frac{dx^2}{dx^3} = \frac{B^2}{B^3}, \quad (2.75b)$$

valid wherever  $B^3 \neq 0$ .

In general, one is allowed to choose any of the coordinates to parameterize the motion of the others. To allow for this in the following summary, let  $i_3$  denote the index labeling the coordinate used as the independent variable, and let the other indices be  $i_1, i_2$ . Then the non-canonical magnetic field line dynamical system may be summarized as:

$$z = (x^{i_1}, x^{i_2}) \quad (2.76a)$$

$$t = x^{i_3} \quad (2.76b)$$

$$\vartheta(z, t) = A_{i_1}(x^{i_1}, x^{i_2}, x^{i_3}) dx^{i_1} + A_{i_2}(x^{i_1}, x^{i_2}, x^{i_3}) dx^{i_2} \quad (2.76c)$$

$$\Omega(z, t) = (A_{i_1, i_2} - A_{i_2, i_1}) dx^{i_1} \wedge dx^{i_2} \quad (2.76d)$$

$$H(z, t) = -A_{i_3}(x^{i_1}, x^{i_2}, x^{i_3}), \quad (2.76e)$$

with phase-space Euler-Lagrange equations:

$$(A_{i_2,i_1} - A_{i_1,i_2}) \dot{x}^{i_2} - (A_{i_1,i_3} - A_{i_3,i_1}) = 0 \quad (2.77a)$$

$$(A_{i_1,i_2} - A_{i_2,i_1}) \dot{x}^{i_1} - (A_{i_2,i_3} - A_{i_3,i_2}) = 0. \quad (2.77b)$$

### Collisionless Guiding Center Dynamics

For many magnetically confined fusion plasmas, the scale length over which the magnetic field varies is much longer than the “gyroradius” - the radius at which a particle orbits a magnetic field line due to the Lorentz force. In this case, the magnetic field is approximately constant with respect to variations in the “gyroangle”  $\Theta$ . Because the gyration of the particle occurs on a much shorter timescale than the cross- and along-field drift motion, it is desirable to average over the rapid gyration and retain only the particle drifts in the description of the dynamical system. This transformation amounts to a dynamical reduction, reducing the six-dimensional charged particle phase space to a four-dimensional space describing the guiding center dynamics. To perform this reduction, coordinate transformations are performed to obtain the gyroangle as an ignorable coordinate in the charged particle Lagrangian. The invariant corresponding to this ignorable coordinate is the magnetic moment  $\mu$ ; ignoring the time evolution of the gyroangle  $\Theta$  and treating  $\mu$  as a constant parameter achieves the desired reduction in dimensionality and the desired reduction in timescales that must be resolved numerically.

In general, the guiding center transformation is an asymptotic expansion ordered by the small parameter that is the ratio of the gyroradius to the magnetic field scale length. Although the transformation may be carried out to arbitrary order [8, 104] typically only the first-order system is retained. Because the asymptotic series is not required to converge, treating the *adiabatic* invariant  $\mu$  as an exact constant can lose stochastic regions present in

the charged particle phase portrait [105]. The limitations of the guiding center approximation are therefore important to keep in mind when considering its validity in various applications.

The collisionless guiding center Lagrangian was first derived in Ref. [8]. The derivation will be reproduced here for convenience and context, appealing also to the presentations in Refs. [57, 78] for clarity and completeness.

Begin with the configuration space Lorentz force Lagrangian for a particle of mass  $m$  and charge  $e$  in an electric field  $E(q, t) = -\nabla\phi(q, t) - \frac{\partial A(q, t)}{\partial t}$  and a magnetic field  $B(q, t) = \nabla \times A(q, t)$ :

$$L(q, \dot{q}, t) = \frac{1}{2}m\|\dot{q}\|^2 + \frac{e}{c}A(q, t) \cdot \dot{q} - e\phi(q, t). \quad (2.78)$$

The first step is to transform to a phase-space Lagrangian by use of the Legendre transform. The canonical phase-space Lagrangian for a Lorentz force particle is given by:

$$L(q, p, \dot{q}, \dot{p}, t) = p \cdot \dot{q} - \left( \frac{1}{2m}\|p - \frac{e}{c}A(q, t)\|^2 + e\phi(q, t) \right), \quad (2.79)$$

where the term in parentheses is the Hamiltonian for a charged particle in canonical coordinates. Next, a non-canonical coordinate transformation is performed:  $(q, p) \mapsto (q, v) = (q, \frac{1}{m}(p - \frac{e}{c}A(q)))$ , yielding a non-canonical phase-space Lagrangian:

$$L(q, v, \dot{q}, \dot{v}, t) = \left( mv + \frac{e}{c}A(q, t) \right) \cdot \dot{q} - \left( \frac{1}{2}m\|v\|^2 + e\phi(q, t) \right). \quad (2.80)$$

The next step is to introduce an ordering parameter  $\epsilon$ , representing terms that are small in the ratio of the gyroradius to the magnetic field scale length. At this point, there exist two potential choices for the ordering of the drifts caused by the electric field. Conventionally, it is assumed that the  $E \times B$  drift is of the same order as the  $\nabla B$  and curvature drifts [8, 57, 15]. This assumption considers the electric field to be small. Alternatively, one can introduce the  $E \times B$  drift at lowest order, capturing the polarization drift (caused by time-dependent

electric fields) at the same order in  $\epsilon$  as the magnetic drifts. This approach is detailed in Ref. [78]. Here, the conventional choice will be made: the  $E \times B$  drift is assumed to be at the same order as the magnetic drifts, and polarization drift will not be captured in the model. Then, using units normalized such that  $e = c = m = 1$ , the ordered Lagrangian is given by:

$$L(q, v, \dot{q}, \dot{v}, t) = (v + \epsilon^{-1} A(q, \epsilon t)) \cdot \dot{q} - \left( \frac{1}{2} \|v\|^2 + \phi(q, \epsilon t) \right), \quad (2.81)$$

where the time dependence of  $A, \phi$  has been presumed to be slow.

After ordering the Lagrangian, a coordinate transformation is made from the particle position  $q$  to the guiding center position  $x$  and from the particle velocity  $(v^1, v^2, v^3)$  to  $(v_{\parallel}, v_{\perp}, \Theta)$ , where  $v_{\parallel}$  is the velocity parallel to the magnetic field,  $v_{\perp}$  the velocity perpendicular to the magnetic field, and  $\Theta$  the gyroangle. To perform these transformations, two sets of basis vectors are helpful, both of which are defined in terms of the magnetic field unit vector  $\hat{b}$ . The first is a “fixed” unit-vector basis  $(\hat{e}_1, \hat{e}_2, \hat{b})$ , with  $\hat{e}_1 \times \hat{e}_2 = \hat{b}$ . Using this fixed basis, one can define a “rotating” unit-vector basis  $(\hat{e}_a, \hat{b}, \hat{e}_c)$  according to:

$$\hat{e}_a = \cos \Theta \hat{e}_1 - \sin \Theta \hat{e}_2 \quad (2.82a)$$

$$\hat{e}_c = -\sin \Theta \hat{e}_1 - \cos \Theta \hat{e}_2. \quad (2.82b)$$

Note that  $\hat{e}_a = \hat{b} \times \hat{e}_c$  and  $\hat{e}_a = \frac{\partial \hat{e}_c}{\partial \Theta}$ . These bases define the transformation from the particle position to the guiding center position according to:

$$q = x + \epsilon \rho = x + \epsilon \frac{v_{\perp}}{\mathcal{B}} \hat{e}_a, \quad (2.83)$$

where  $\rho$  is the gyroradius pointing from the guiding center position to the particle position and  $\mathcal{B} = \|B\|$ . Recall that  $\rho = \frac{v_{\perp}}{\Omega_c}$ , where  $\Omega_c$  is the cyclotron frequency  $\frac{e\mathcal{B}}{mc}$ ; in normalized

units,  $\Omega_c = \mathcal{B}$ . The velocity coordinate transformation is specified by:

$$v = v_{\parallel} \hat{b} + v_{\perp} \hat{e}_c, \quad (2.84)$$

where the  $\Theta$  dependence is introduced in the definition of  $\hat{e}_c$ . In terms of these new variables, the Lagrangian is given by:

$$\begin{aligned} L(x, v_{\parallel}, v_{\perp}, \Theta, \dot{x}, \dot{v}_{\parallel}, \dot{v}_{\perp}, \dot{\Theta}, t) = \\ \left( \frac{1}{\epsilon} A(x + \rho, \epsilon t) + v_{\parallel} \hat{b} + v_{\perp} \hat{e}_c \right) \cdot (\dot{x} + \epsilon \dot{\rho}) - H(x, v_{\parallel}, v_{\perp}, \Theta, t) = \\ \left( \frac{1}{\epsilon} A(x + \rho, \epsilon t) + v_{\parallel} \hat{b} + v_{\perp} \hat{e}_c \right) \cdot (\dot{x} + \epsilon \dot{\rho}) - \left( \frac{1}{2} v_{\parallel}^2 + \frac{1}{2} v_{\perp}^2 + \phi(x + \epsilon \rho, \epsilon t) \right). \end{aligned} \quad (2.85)$$

The next step is to Taylor expand  $A(x + \epsilon \rho, t)$  about  $A(x)$ , obtaining:

$$\begin{aligned} L \approx & \left( \frac{1}{\epsilon} A(x, \epsilon t) + v_{\parallel} \hat{b} + v_{\perp} \hat{e}_c \right) \cdot \dot{x} + A(x, \epsilon t) \cdot \dot{\rho} + \epsilon v_{\parallel} \hat{b} \cdot \dot{\rho} + \epsilon v_{\perp} \hat{e}_c \cdot \dot{\rho} + \\ & (\rho \cdot \nabla) A(x, \epsilon t) \cdot \dot{x} + \epsilon (\rho \cdot \nabla) A(x, \epsilon t) \cdot \dot{\rho} - H(x, v_{\parallel}, v_{\perp}, \Theta, t). \end{aligned} \quad (2.86)$$

From here, it is helpful to evaluate  $\dot{\rho}$ :

$$\frac{d\rho}{dt} \approx \frac{\dot{v}_{\perp}}{\mathcal{B}} \hat{e}_a + \frac{v_{\perp}}{\mathcal{B}} \dot{\Theta} \hat{e}_c, \quad (2.87)$$

obtaining a Lagrangian of the form:

$$\begin{aligned} L \approx & \left( \frac{1}{\epsilon} A + v_{\parallel} \hat{b} + v_{\perp} \hat{e}_c \right) \cdot \dot{x} + \epsilon \frac{v_{\perp}^2}{\mathcal{B}} \dot{\Theta} + \epsilon \frac{v_{\perp} \dot{v}_{\perp}}{\mathcal{B}^2} (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_a + \\ & \epsilon \frac{v_{\perp}^2}{\mathcal{B}^2} \dot{\Theta} (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_c + (\rho \cdot \nabla) A \cdot \dot{x} + A \cdot \dot{\rho} - H(x, v_{\parallel}, v_{\perp}, \Theta, t), \end{aligned} \quad (2.88)$$

where the arguments of  $A(x, \epsilon t)$  are now implied for compactness.

To further simplify the Lagrangian, a *Lagrangian gauge* term will be added. A Lagrangian gauge term is a term that is the total derivative of some function,  $\chi$ . Because  $L$  and  $L + \frac{d\chi}{dt}$  yield the same Euler-Lagrange equations, the dynamics are unaffected by the addition of the Lagrangian gauge term  $\frac{d\chi}{dt}$ . This is analogous to an electromagnetic gauge term  $A \rightarrow A + \nabla\lambda$  for some function  $\lambda$ . A simplification is obtained through the choice of a Lagrangian gauge:

$$\chi(x, v_{\parallel}, v_{\perp}, \Theta) = -A(x, \epsilon t) \cdot \rho \quad (2.89a)$$

$$\frac{d\chi}{dt} = -(\dot{x} \cdot \nabla)A \cdot \rho - \epsilon \frac{\partial A}{\partial t} \cdot \rho - A \cdot \dot{\rho}, \quad (2.89b)$$

yielding:

$$\begin{aligned} L + \frac{d\chi}{dt} &= \left( \frac{1}{\epsilon} A + v_{\parallel} \hat{b} + v_{\perp} \hat{e}_c \right) \cdot \dot{x} + \epsilon \frac{v_{\perp}^2}{\mathcal{B}} \dot{\Theta} + \epsilon \frac{v_{\perp} \dot{v}_{\perp}}{\mathcal{B}^2} (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_a + \\ &\quad \epsilon \frac{v_{\perp}^2}{\mathcal{B}^2} \dot{\Theta} (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_c + (\rho \cdot \nabla) A \cdot \dot{x} - (\dot{x} \cdot \nabla) A \cdot \rho - \epsilon \frac{\partial A}{\partial t} \cdot \rho - \\ &\quad H(x, v_{\parallel}, v_{\perp}, \Theta, t). \end{aligned} \quad (2.90)$$

By using the vector identity:

$$v \times (\nabla \times w) = \nabla w \cdot v - (v \cdot \nabla)w, \quad (2.91)$$

one can determine that:

$$(\rho \cdot \nabla) A \cdot \dot{x} - (\dot{x} \cdot \nabla) A \cdot \rho = -v_{\perp} \hat{e}_c \cdot \dot{x}. \quad (2.92)$$

This fact simplifies the Lagrangian to:

$$L = \left( \frac{1}{\epsilon} A + v_{\parallel} \hat{b} \right) \cdot \dot{x} + \epsilon \frac{v_{\perp}^2}{\mathcal{B}} \dot{\Theta} + \epsilon \frac{v_{\perp} \dot{v}_{\perp}}{\mathcal{B}^2} (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_a + \epsilon \frac{v_{\perp}^2}{\mathcal{B}^2} \dot{\Theta} (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_c - H(x, v_{\parallel}, v_{\perp}, \Theta, t), \quad (2.93)$$

where the  $\epsilon \frac{\partial A}{\partial t} \cdot \rho$  term has been absorbed into the Hamiltonian (no time derivatives of the coordinates are present).

To eliminate the remaining undesired terms, add another Lagrangian gauge term:

$$\chi(x, v_{\parallel}, v_{\perp}, \Theta) = -\epsilon \frac{v_{\perp}^2}{2\mathcal{B}^2} (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_a \quad (2.94)$$

$$\frac{d\chi}{dt} \approx -\epsilon \frac{v_{\perp} \dot{v}_{\perp}}{\mathcal{B}^2} (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_a - \epsilon \frac{v_{\perp}^2}{2\mathcal{B}} \dot{\Theta} ((\hat{e}_c \cdot \nabla) A \cdot \hat{e}_a + (\hat{e}_a \cdot \nabla) A \cdot \hat{e}_c). \quad (2.95)$$

Adding this term to the Lagrangian obtains:

$$L + \frac{d\chi}{dt} = \left( \frac{1}{\epsilon} A + v_{\parallel} \hat{b} \right) \cdot \dot{x} + \epsilon \frac{v_{\perp}^2}{\mathcal{B}} \dot{\Theta} + \epsilon \frac{v_{\perp}^2}{2\mathcal{B}^2} \dot{\Theta} ((\hat{e}_a \cdot \nabla) A \cdot \hat{e}_c - (\hat{e}_c \cdot \nabla) A \cdot \hat{e}_a) - H(x, v_{\parallel}, v_{\perp}, \Theta, t) \quad (2.96)$$

Use of the identity Eq. (2.91) reveals:

$$(\hat{e}_a \cdot \nabla) A \cdot \hat{e}_c - (\hat{e}_c \cdot \nabla) A \cdot \hat{e}_a = -\mathcal{B}, \quad (2.97)$$

which simplifies the Lagrangian to:

$$L = \left( \frac{1}{\epsilon} A + v_{\parallel} \hat{b} \right) \cdot \dot{x} + \epsilon \frac{v_{\perp}^2}{2\mathcal{B}} \dot{\Theta} - H(x, v_{\parallel}, v_{\perp}, \Theta, t). \quad (2.98)$$

The final step is to change from the particle coordinates  $(v_{\parallel}, v_{\perp}, \Theta)$  to  $(u, \mu, \theta)$ , where  $u \approx v_{\parallel}$  is the parallel velocity of the guiding center,  $\mu = \frac{v_{\perp}^2}{2\mathcal{B}}$  is the magnetic moment, and

$\theta \approx \Theta$  is the gyroangle of the averaged guiding center dynamics rather than that of the original particle. These transformations do not introduce any additional terms at this order, yielding *the guiding center Lagrangian*:

$$L_{GC}(x, u, \mu, \theta, \dot{x}, \dot{u}, m\dot{u}, \dot{\theta}) = \\ \left( \frac{e}{c} A(x, t) + m u b(x, t) \right) \cdot \dot{x} + \frac{mc}{e} \mu \dot{\theta} - \left( \frac{m}{2} u^2 + \mu \mathcal{B}(x, t) + e \phi(x, t) \right), \quad (2.99)$$

where  $\epsilon$  has been set to one and physical constants restored in the final result. Also,  $b = \hat{b}$  for simplicity and consistency of notation in future use.

At last, the gyroangle  $\theta$  has been explicitly revealed to be an ignorable coordinate with the magnetic moment as the corresponding Noether invariant. Because  $\theta$  does not appear in the Lagrangian except as  $\dot{\theta}$ , the corresponding Euler-Lagrange equation states:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} = \frac{d}{dt} \left( \frac{mc}{e} \mu \right) = 0. \quad (2.100)$$

The lowest-order magnetic moment is then treated as exactly constant in the guiding center model. Calculating the Euler-Lagrange equation for the variation with respect to  $\mu$  obtains:

$$\dot{\theta} = \frac{e \mod B(x, t)}{mc}, \quad (2.101)$$

recovering the physical result that the gyrophase evolves according to the gyrofrequency. The  $(\mu, \theta)$  dynamics are decoupled from the remaining four dimensions, allowing a dynamical reduction from a six-dimensional phase-space  $(x^1, x^2, x^3, u, \mu, \theta)$  to a four-dimensional phase-space  $(x^1, x^2, x^3, u)$ . In this reduced space, the guiding center Lagrangian is:

$$L_{GC}(x, u, \dot{x}, \dot{u}) = \left( \frac{e}{c} A(x, t) + m u b(x, t) \right) \cdot \dot{x} - \left( \frac{m}{2} u^2 + \mu \mathcal{B}(x, t) + e \phi(x, t) \right), \quad (2.102)$$

where the dependence on  $\mu$  is now interpreted as parametric ( $\mu$  is a constant just like  $m, e, c$ ).

In summary, the guiding center action principle is described by:

$$z = (x^1, x^2, x^3, u) \quad (2.103a)$$

$$\vartheta(z) = (A_i(x) + ub_i(x)) dx^i = A_i^\dagger(x, u) dx^i \quad (2.103b)$$

$$\Omega(z) = A_{i,j}^\dagger dx^i \wedge dx^j + b_i dx^i \wedge du \quad (2.103c)$$

$$H(z) = \frac{u^2}{2} + \mu \mathcal{B}(x) + \phi(x) \quad (2.103d)$$

$$L_{PS}(z, \dot{z}) = (A_i(x) + ub_i(x)) \dot{x}^i - \left( \frac{u^2}{2} + \mu \mathcal{B}(x) + \phi(x) \right). \quad (2.103e)$$

The corresponding guiding center equations of motion are:

$$\left( A_{i,j}^\dagger - A_{j,i}^\dagger \right) \dot{x}^i - b_j \dot{u} - \mu B_{,j} - \phi_{,j} = 0 \quad j = 1, 2, 3 \quad (2.104a)$$

$$b_i \dot{x}^i - u = 0. \quad (2.104b)$$

## Collisional Guiding Center Dynamics

In addition to the collisionless drift motion, test particles evolving in a background plasma experience collisions with other test particles and the electrons and ions that compose the background plasma. If the number of test particles is small relative to the number of particles in the background plasma, a good approximation is to ignore collisions involving two test particles. The interactions between the test particles and the background species may be modeled as a combination of dissipative drag and stochastic kicks. The collision models used here follows common practice for test particle codes, as described in for instance Refs. [49] and [65].

A collisional model describes how the distribution function of a species of particles evolves in time under the influence of collisional processes, as defined by a *collision operator*. The

standard partial differential equation describing such a process for a distribution function  $f(x, v)$  is the *kinetic equation*:

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial x} + \left( \frac{e}{m} E + \frac{e}{mc} v \times B \right) \cdot \frac{\partial f}{\partial v} = C[f], \quad (2.105)$$

where  $C[f]$  is the collision operator  $C$  acting on  $f$ .

For distributions of energetic test particle ions, it is common [49] to use the collision operator:

$$C[f] = \nu_d \frac{\partial}{\partial \lambda} (1 - \lambda^2) \frac{\partial f}{\partial \lambda} - \frac{\nu}{\|v\|^2} \frac{\partial}{\partial \|v\|} (\|v\|^3 + v_c^3) f. \quad (2.106)$$

Here,  $\|v\|$  is the magnitude of the velocity;  $\nu_d = \frac{v_c^3}{2v^3} \nu$  is the pitch angle scattering rate;  $v_c^3 = \frac{3\sqrt{\pi}}{4} \frac{m_e}{m_i} \left( \frac{2kT_e}{m_e} \right)^{3/2}$  is the cube of the critical velocity;  $\lambda = \frac{v_{\parallel}}{\|v\|}$  is the “pitch” of the particle; and  $\nu$  is the collision rate. In SI units, the collision rate is given by:

$$\nu = \frac{n_e Z_{eff} e^2 q_T^2 \log(\Lambda)}{4\pi\epsilon_0^2 m_i m_T v_c^3}, \quad (2.107)$$

with  $n_e$  the electron density of the background plasma;  $Z_{eff}$  the effective charge of the background plasma;  $e$  the fundamental charge;  $q_T$  the charge of the test particle ion;  $\log(\Lambda)$  the “Coulomb logarithm” with  $\Lambda = 12\pi n_e \lambda_D^3$  and  $\lambda_D$  the Debye radius;  $\epsilon_0$  the permittivity of free space;  $m_i$  the mass of the background ion species; and  $m_T$  the mass of the test particle ion.

The collision operator of Eq. (2.106) consists of two terms. The first describes “pitch angle scattering”. Pitch angle scattering is an energy-conserving change to the direction of the particle velocity. The pitch angle variable is convenient when there is a direction of symmetry in the velocity-space distribution, as is the case for magnetized fusion plasmas (with the direction of symmetry being the magnetic field direction). The second term describes drag on the test particle ion due to ion-electron and ion-ion collisions. For a given test particle,

this term amounts to a dissipative term given by:

$$\frac{dv}{dt} = -\nu \frac{\|v\|^3 + v_c^3}{\|v\|^3} v. \quad (2.108)$$

Whereas the pitch angle scattering effect is a stochastic process for the test particle, the drag may be included in the equations of motion as a dissipative process.

For the guiding center dynamical system, the energetic test particle drag effects are given by the vector field:

$$X_f = (0, 0, 0, -\nu \frac{v^3 + v_c^3}{v_c^3} u, -2\nu \frac{v^3 + v_c^3}{v^3} \mu)^T. \quad (2.109)$$

To append this dissipative or “forcing” vector field to the phase-space Euler-Lagrange equations, one may apply a phase-space Lagrange-d’Alembert principle of the form:

$$\begin{aligned} \delta \int_0^T L_{GC}(x, u, \Theta, \mu, \dot{x}, \dot{u}, \dot{\Theta}, \dot{\mu}) dt - \int_0^T (\Omega_{GC})_{ij} X_f^i \delta z^j dt &= 0 \\ \delta \int_0^T A_i^\dagger(x, u) \dot{x}^i + \mu \dot{\Theta} - \left( \frac{u^2}{2} + \mu B(x) + \phi(x) \right) dt - \\ \int_0^T \nu \frac{v^3 + v_c^3}{v^3} (ub_j \delta x^j + 2\mu \delta \Theta) dt &= 0. \end{aligned} \quad (2.110)$$

The forced phase-space Euler-Lagrange equations for the guiding center system subject to collisional drag is then given by:

$$(A_{i,j}^\dagger - A_{j,i}^\dagger) \dot{x}^i - b_j \dot{u} - \mu B_{,j} - \phi_{,j} - \nu \frac{v^3 + v_c^3}{v^3} ub_j = 0 \quad j = 1, 2, 3 \quad (2.111a)$$

$$b_i \dot{x}^i - u = 0 \quad (2.111b)$$

$$-\dot{\mu} - 2\nu \frac{v^3 + v_c^3}{v^3} \mu = 0 \quad (2.111c)$$

$$\dot{\Theta} - \mathcal{B} = 0. \quad (2.111d)$$

The applications in this dissertation will focus on the dissipative drag effects, primarily for simplicity. Pitch angle scattering may also be included in the phase-space Lagrange-d'Alembert principle, and will yield stochastic terms to the equations of motion, making the system a “stochastic differential equation”. For stochastic differential equations for the guiding center system, consult references on test particle codes [49, 52, 53].

### 2.3.5 Variational Integrators for Action Principles in Phase Space

As reviewed in Section 2.2, the conventional setting for variational integrators is regular Lagrangian systems. Variational integrators for degenerate Lagrangian systems, including the “linear in velocity” Lagrangians encountered in phase-space action principles, have received less attention in the existing literature. There do exist relevant references, however, both from the variational integrators community and from the numerical plasma physics community.

In the variational integration literature, the first application of discrete mechanics to degenerate (and phase-space) Lagrangian systems is in the context of planar point vortices [13]. Here, a canonical phase-space Lagrangian is discretized and the resulting variational integrators analyzed for accuracy and structure preservation. Following up on methods for canonical phase-space Lagrangians are the works of Refs. [90, 91], in which type-two, -three, and -four generating functions are constructed as discrete variational principles for *canonical* phase-space action principles. These methods yield one-step variational integrators that are canonically symplectic. To analyze more general degenerate Lagrangian systems, Refs. [87, 88] utilize Dirac structures [106] to develop a theory of “implicit Lagrangian systems”. This theory is put to use in electric circuit models in [89, 18]. The degeneracy of the Lagrangian for electrical circuits arises from the constraints established by Kirchoff’s laws; it is not a phase-space Lagrangian, but the results of Chapter 3 are relevant to these methods. More recently, Ref. [92] considered the standard “Veselov”-type discretizations of phase-space Lagrangians.

The structure-preserving properties of the resulting variational integrator are assessed from the perspective of whether the dynamics remain on the proper constraint sub-manifold of  $T^*M$ . It is shown that for linear one-forms, in particular, canonical Hamiltonian systems, a family of discretization methods satisfies the desired constraint.

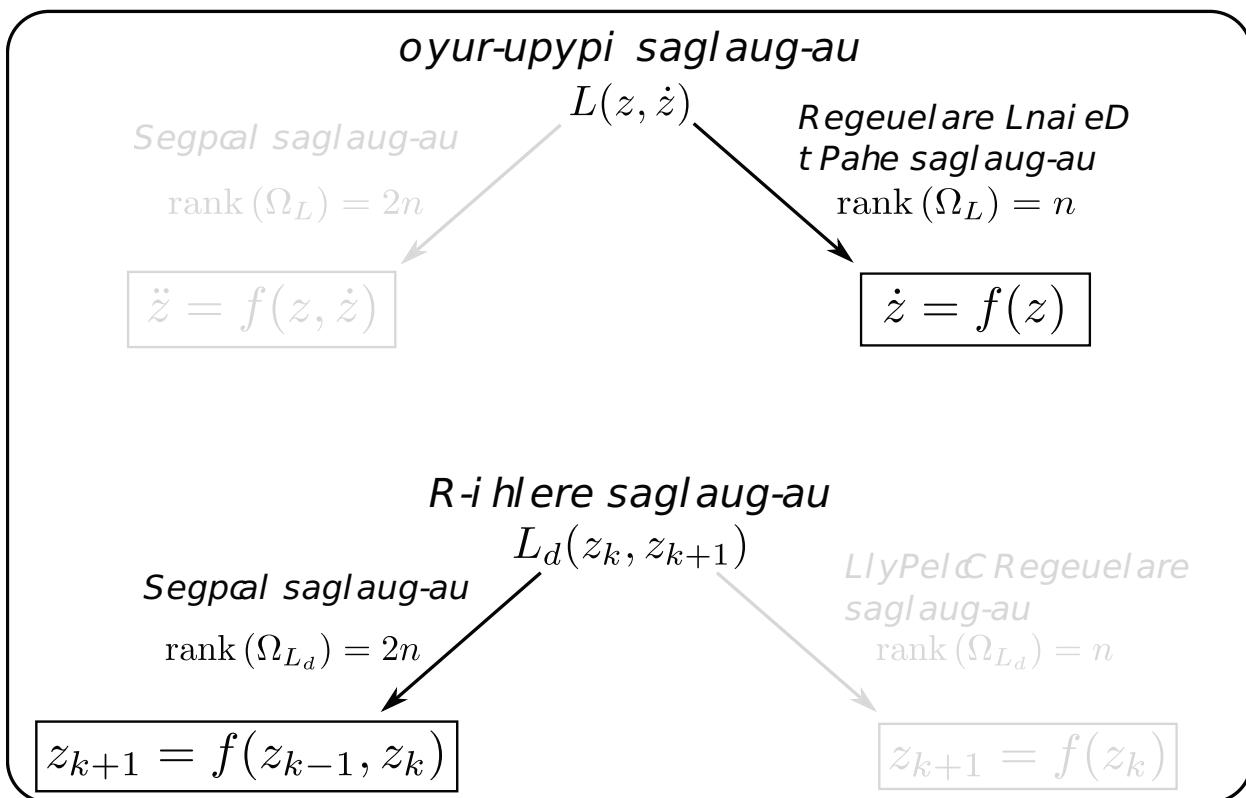
In the plasma physics literature, variational integrators were first applied to action principles in phase-space in the context of collisionless guiding center trajectories [14, 15]. A different discretization was used for a relativistic version of guiding center dynamics (intended for magnetospheric application) in Ref. [16]. The test cases encountered in these references did not exhibit the instabilities that motivated the work of Chapter 3. The linear stability properties of these algorithms was analyzed, however, from the perspective of electromagnetic gauge invariance in Ref. [96]. Several results pertaining to guiding center variational integrators are included in Ref. [17], alongside additional applications for plasma physics PDE systems. An application of variational integrators to the magnetic field line flow phase-space Lagrangian is present in Ref. [107]. Finally, several early results of this dissertation are presented in Ref. [19].

## 2.4 Dissertation Scope

The goal of this dissertation is to adapt and extend variational integration theory to discover new, structure-preserving numerical algorithms for degenerate-Lagrangian systems encountered in plasma physics. Of primary interest is the development of variational integrators for non-canonical guiding center dynamics, complete with the inclusion of collision-based dissipative effects.

# Chapter 3

## Multistep Variational Integrators



*The greatest shortcoming of the human race is our inability to understand the exponential function.*

—Prof. Albert Bartlett

## 3.1 Introduction

The first discovery of this dissertation is that all existing variational guiding center algorithms [14, 15, 16, 17, 19] and in fact most variational integrators constructed from degenerate Lagrangian systems [13, 18, 92] are *multistep methods*. Multistep methods are among the oldest methods for approximating solutions of ODEs [10, 83]. By “multistep”, it is meant that the time advance rule depends on the state at multiple previous instants in time. The implication is that the number of steps in the multistep method exceeds the order of the differential equation. For instance, if one is modeling a first-order differential equation:  $\dot{z} = X(z)$ , the “two-step” Adams-Bashforth method is given by:

$$z_{k+1} = z_k + \frac{3}{2}hX(z_k) - \frac{1}{2}hX(z_{k-1}). \quad (3.1)$$

This is in contrast to the “one-step” second-order Runge-Kutta method:

$$z_{k+1} = z_k + hX\left(z_k + \frac{1}{2}hX(z_k)\right), \quad (3.2)$$

which determines  $z_{k+1}$  as a function of  $z_k$  alone, and is therefore a multistep method only trivially (one is a number of steps, after all). One advantage of multistep methods is the ability to exceed first-order accuracy in the numerical step size  $h$  while only requiring a single evaluation of the vector field  $X$  at each new step; the previous vector field evaluations may be stored and used in later steps to improve the accuracy of the solution. One disadvantage of multistep methods is the need to supply additional initial conditions beyond those pro-

vided by the initial value problem. A more severe disadvantage is that the additional steps introduce additional modes into the numerical dynamics, and the stability of these modes is a prerequisite for the robustness of the multistep method.

The emergence of multistep methods in the context of variational integrators may be observed by comparing the number of steps in the discrete Euler-Lagrange equations to the order of the differential equation specified by the continuous Euler-Lagrange equations. Recall that the discrete Euler-Lagrange equations of a variational integrator are given by (Eq. (2.21)):

$$D_2 L_d(z_{k-1}, z_k) + D_1 L_d(z_k, z_{k+1}) = 0.$$

The discrete Euler-Lagrange equations depend on the state at two previous instants in time:  $z_k$  and  $z_{k-1}$ . If the continuous Euler-Lagrange equations form a system of second-order ODEs, then this is an appropriate feature. If the Euler-Lagrange equations instead specify only a first-order ODE system (as in Eq. (2.63)), then it is extremely important to acknowledge that the variational integrator *is a two-step method*. This discovery motivates designation of a new class of integrators: *multistep variational integrators (MVIs)*. Additionally, this discovery motivates analyzing the stability and conservation properties of multistep variational integrators using tools established over decades of numerical analysis on multistep methods [11, 93, 108, 109, 110, 111, 6, 112, 12, 95, 113, 114]. This chapter applies these tools to obtain important new results for variational integrators, multistep methods, and computational plasma physics.

Arguably, the most important aspect of multistep method theory is that multistep methods exhibit *parasitic modes*. Multistep methods evolve in a higher-dimensional space than the original dynamical system; this is evidenced by the fact that multistep methods require additional initial conditions beyond those supplied by the initial value problem. Because they evolve in a higher-dimensional space, multistep methods admit behavior not present

in the original system. The presence of this behavior may be illuminated by performing an eigenvalue analysis. A characteristic result of such an eigenvalue analysis is depicted in Fig. 3.1. Here, an eigenvalue analysis is performed on a two-dimensional two-step method, leading to four eigenvalues. Two of the eigenvalues are near positive one - these correspond to a mode near the identity map. Because the numerical step size is small, the time advance is small and expected to be near the identity map. Also present, however, are eigenvalues near *negative* one. The modes corresponding to these eigenvalues are deemed “parasitic”. They lead to rapid oscillations in the numerical solution; for parasitic modes at negative one, the oscillations have a period of two steps. Correspondingly, the parasitic oscillations may be characterized as “even-odd” or “red-black” modes. In general, parasitic modes can correspond to eigenvalues at  $i$ ,  $-i$ , or really anywhere except  $+1$  (in the  $h \rightarrow 0$  limit). If these modes are unstable, the unphysical behavior can overwhelm the numerical solution, rendering the result useless. Thus, analyzing the parasitic mode behavior of any multistep method is critical for assessing its reliability. For many commonly used multistep methods, the parasitic modes all have magnitude less than one, indicating the parasitic modes damp as the integration advances in time [93, 6]. For conservative methods, the parasitic mode behavior is not so easily managed, as we shall see.

The presence of parasitic modes in multistep methods explains the unanticipated numerical instability of the variational guiding center algorithms described in Chapter 1. Figure 3.2 replicates the study of Fig. 1.1, this time distinguishing the even-numbered steps from the odd-numbered steps using different markers. The decoupling between the even- and odd-numbered trajectories is immediately recognizable; incriminating evidence of a parasitic mode instability. Now that the instabilities are characterized, several follow-up questions naturally arise: Might a different discretization have stable parasitic modes? Can the freedom in the initial conditions be used to reduce parasitic mode behavior? Are there any other ways to mitigate parasitic modes if they grow to unacceptably large amplitude? The answer

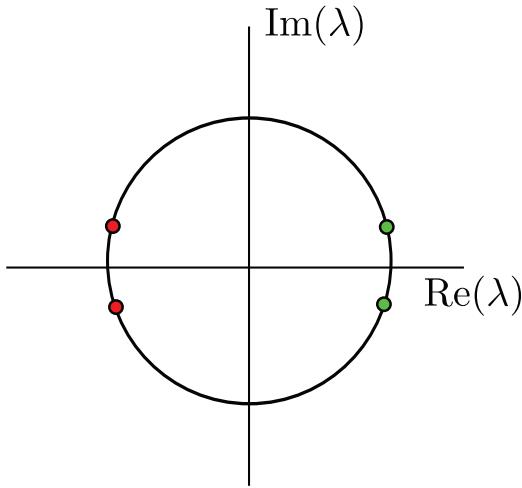


Figure 3.1: An illustration of the eigenvalues  $\lambda$  for a prototypical two-dimensional two-step method. The two roots on the unit circle near positive one correspond to the physical component of the solution, whereas the roots on the unit circle near negative one are a “parasitic mode” that oscillates rapidly.

to the first question is established by analyzing the spectral stability of multistep variational integrators in Sec. 3.3; unfortunately, because the eigenvalues of multistep variational integrators appear in reciprocal and complex conjugate pairs, the parasitic modes of MVIs are at best marginally stable. The answers to the other two questions are also explored in Sec. 3.3, with more positive results. Initial conditions may be chosen to initialize the parasitic modes to arbitrarily small amplitude using a careful initialization procedure. If the parasitic modes still grow to unacceptably large amplitude, one may re-initialize the multistep advance to reset the parasitic mode to small amplitude. Although this re-initialization step is non-conservative, it introduces a very small amount of dissipation, and can lead to the method outperforming a non-geometric integrator where it otherwise would not (see Fig. 3.12 for a demonstration of this).

Unbounded parasitic mode growth in variational guiding center algorithms directly contradicts the behavior expected of symplectic algorithms. The numerical trajectory in Fig. 3.2

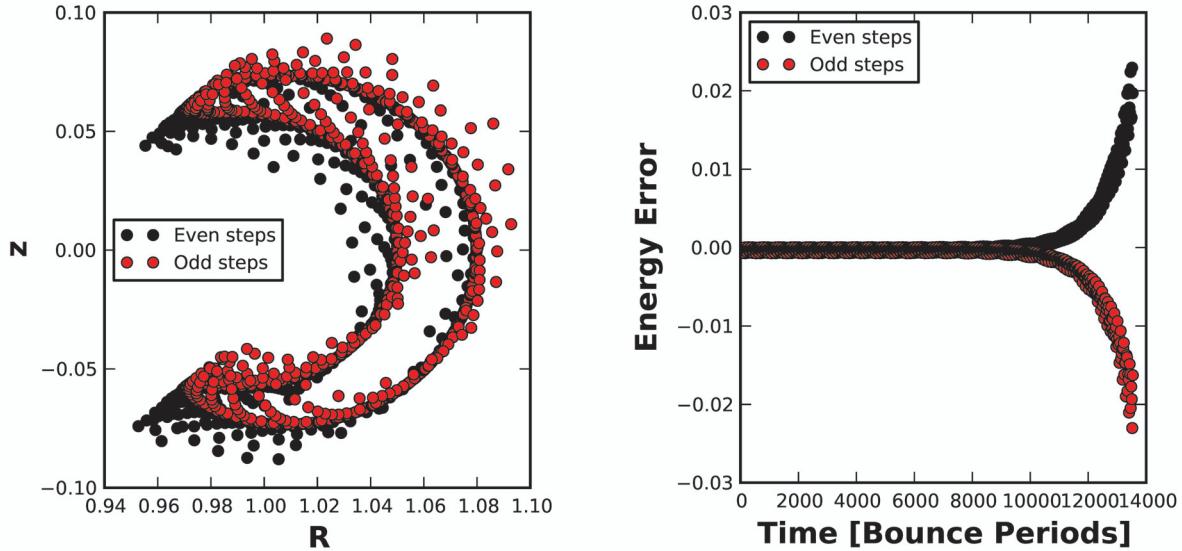


Figure 3.2: Small step even-odd

shows no evidence of KAM tori and accumulates energy error without bound. These observations must be reconciled with the fact that the multistep variational integrator preserves a two-form, per Thm. 2.2.6. The resolution of this conundrum is presented in Sec. 3.4, which analyzes the structure-preserving properties of multistep variational integrators. To summarize, because the MVI evolves on a higher-dimensional space than the continuous system, the two-form it preserves also resides on this higher-dimensional space. Informally, the continuous two-form resides on  $z$ -space, whereas the MVI two-form resides on  $(z_0, z_1)$ -space - a space twice as large. The theorems describing symplectic integrator behavior assume *the same* two-form is preserved; nothing has been established about integrators that preserve two-forms on larger spaces than that on which the original symplectic structure resides. Quite simply, the phase-space Lagrangian fails the regularity requirement of the theorem stating a variational integrator is equivalent to a symplectic integrator (Thm. 2.2.10), so the bad behavior is not contradictory to any rigorous results.

Fortunately, the story does not end here. One last question remains to be answered, and for once the answer is surprisingly positive. That question is: *If the variational integrator*

*cannot be expected to behave like a symplectic integrator, why are good results ever obtained?* For instance, the stable parameter regime shown in Fig. 1.1 demonstrated bounded energy error, which would be surprising for a non-geometric integrator that is only second-order accurate. To answer this question, one more tool from the multistep methods literature proves helpful: *the underlying one-step method* [93, 94]. “The underlying one-step method” refers to a one-step map, e.g.  $z_{k+1} = \hat{F}(z_k)$ , that generates trajectories satisfying the multistep update rule (see Eq. (3.81)). Because the underlying one-step method depends only on a single previous state, it is absent of parasitic modes. If one is interested in the long-term behavior of the non-parasitic mode of a multistep method, it is the conservation properties of the underlying one-step method that should be investigated. Section 3.4 discovers that the underlying one-step method is symplectic on the original space, albeit with a symplectic structure different from that of the original system (see Thm. 3.4.4). This result is surprising in the context of multistep methods because it has been established that - for a certain class of multistep methods - the underlying one-step method cannot be symplectic with respect to the symplectic structure of the continuous system [95]. This restriction apparently does not preclude excellent behavior of the underlying one-step method, as excellent behavior is observed from multistep variational integrators whenever the parasitic modes remain small. The good behavior may then be attributed to the underlying one-step method preserving a symplectic structure that is different from that of the continuous system.

Stepping back, this chapter draws together aspects of three separate research communities: computational plasma physics, variational integrators, and multistep methods. The existing knowledge of each of these communities is combined in a new way to obtain important results for each of the three communities. From a computational plasma physics perspective, analysis of the multistep character of variational guiding center algorithms explains both the instabilities and the good behavior observed whenever these instabilities are not present. From a variational integration perspective, discovery that variational integrators

for degenerate Lagrangian systems often constitute multistep methods reveals a dangerous class of instabilities and obtains a sharper result on symplecticity. From a multistep methods perspective, the realization that certain multistep methods may be derived from a discrete variational principle opens new tools for numerical analysis. The backward error analysis results are concisely expressed in terms of the discrete Lagrangian function and differential geometry operators (Thms. 3.3.3 and 3.3.4); because some multistep variational integrators do not belong to any other category of multistep methods, no previously presented backward error analysis expressions [12, 6] can be used to describe these methods. The new result on the symplecticity of the underlying one-step method is a surprising feature, and the abstract nature of the two-form (Eq. (3.83)) suggests such a discovery would be unlikely without the variational integration framework. Exciting avenues persist for developing the theory of multistep variational integrators, including generalization of multistep variational integrators beyond two-step methods and degenerate Lagrangian systems. This and other ideas for future work are presented in Sec. 3.7.2.

## 3.2 Definitions

Embarking on a theory of multistep variational integrators necessitates definitions of multistep methods and multistep variational integrators. It is also useful to know when one can expect to encounter multistep variational integrators. This section presents these definitions, theorems describing when multistep variational integrators emerge, and a condition for testing a given multistep method for the existence of a discrete variational formulation.

Starting in the continuous setting, suppose we seek to model the time evolution of a dynamical system evolving on a differentiable manifold  $M$  according to some first-order differential equation. The corresponding *initial value problem* identifies a *vector field*  $X \in$

$T M$  and an *initial condition*  $z_0 \in M$ , and seeks a solution  $z : [0, T] \subset \mathbb{R} \rightarrow M$  such that:

$$\frac{dz(t)}{dt} = X(z(t)) \quad (3.3)$$

$$z(0) = z_0. \quad (3.4)$$

Here, the vector field  $X$  is presumed time-independent because any time-dependent system may be re-formulated by treating time as an additional dependent variable that evolves according to  $\dot{t} = 1$ . Although the solution to the initial value problem will often be denoted  $z(t)$ , a related and useful to definition is the *flow map*  $F : M \times \mathbb{R} \rightarrow M$  of the vector field  $X$ . The flow map advances any point  $z \in M$  for some time  $t \in [0, T]$  along the “flow” of the vector field, i.e.  $\frac{d}{dt}F(z, t) = X(F(z, t))$ . Often the shorthand notation  $F^t : M \rightarrow M$  will be used, called the *time-t flow map* or the *flow map at the frozen time t*. This notation and terminology reflects the fact that  $X$  is defined on all of  $M$ , and so is a bit more general than considering a single solution for a single initial condition.

Turning to the discrete setting, consider a discretization of the time interval  $[0, T] \subset \mathbb{R}$  into  $N + 1$  equidistant instants:  $\{t_0, t_1, \dots, t_N\}$  with  $t_0 = 0$ ,  $t_N = T$  and  $t_k - t_{k-1} = h = T/N$  for  $k = 1, \dots, N$ . Corresponding to this time discretization, let  $z_k$  denote the numerical approximation to the solution  $z(t)$  at time  $t = t_k$ . This equips the following definition.

**Definition** A *multistep method* with  $p$  steps is a map  $F_d : \overbrace{M \times \dots \times M}^{p \text{ times}} \times \mathbb{R} \rightarrow M$ , so that a numerical trajectory may be generated according to:

$$z_{k+1} = F_d(z_{k+1-p}, z_{k+2-p}, \dots, z_k, h) \quad k = p - 1, \dots, N - 1. \quad (3.5)$$

Often, the  $h$  dependence of  $F_d$  will be implied and omitted from the list of arguments. Typically, the multistep method  $F_d$  is specified implicitly according to an *update rule*  $f_d :$

$\overbrace{M \times \dots \times M}^{p+1 \text{ times}} \times \mathbb{R} \rightarrow \mathbb{R}^n$ , with  $n = \dim(M)$ , such that  $z_p = F_d(z_0, \dots, z_{p-1}, h)$  satisfies:

$$f_d(z_0, z_1, \dots, z_p) = 0. \quad (3.6)$$

The implicit function theorem ensures that if  $\det(D_p f_d) \neq 0$ , the update rule specifies the numerical method map  $F_d$ , at least locally. Although one-step methods are formally encompassed within multistep methods, it shall be implied by “multistep” that  $p > 1$ . When restricting attention to first-order differential equations, whenever  $p > 1$  the multistep map requires additional initial conditions be generated using some *starting procedure*.

Familiar families of multistep methods include *linear multistep methods*, for which the update rule has the form:

$$\sum_{k=0}^p a_k z_k + b_k X(z_k) = 0, \quad (3.7)$$

with the  $a_k, b_k$  coefficients real numbers; and *partitioned linear multistep methods*, for which the update rule has the form:

$$\begin{aligned} \sum_{k=0}^p a_{k1} z_k^1 + b_{k1} X^1(z_k) &= 0 \\ \vdots \\ \sum_{k=0}^p a_{kn} z_k^n + b_{kn} X^n(z_k) &= 0, \end{aligned} \quad (3.8)$$

again with  $a_{kj}, b_{kj}$  real numbers. These categorizations include the famous Adams-Basforth [10] and Adams-Moulton [83] families of methods. Many of the primary characteristics of multistep methods may be observed in these algorithms [109, 12, 6].

A multistep variational integrator is then, concisely, a variational integrator that constitutes a multistep method for approximating the solution to the Euler-Lagrange equations.

Because the Euler-Lagrange equations can define either a first- or second-order system, the formal definition for multistep variational integrators will be:

**Definition** A *multistep variational integrator* is a variational integrator such that the dimension of the domain of the discrete Lagrangian map  $F_{L_d}$  exceeds the dimension of the domain of the time- $t$  continuous Lagrangian flow map  $F_L^t$ .

By this definition, conventional variational integrators (those with a regular discrete Lagrangian constructed by approximating a regular continuous Lagrangian) are not considered multistep methods. In this conventional setting, although the variational integrator requires two initial conditions in the configuration space, the initial value problem for the Euler-Lagrange equations specifies an initial position and velocity, so the discrete and continuous Lagrange maps have domains of the same dimension. It is fair to call the discrete Euler-Lagrange equations a “two-step method”, but this constitutes a multistep method only in the same trivial sense that one-step methods are multistep methods for first-order differential equations. The substantive delineation of “multistep variational integrators” occurs when the order of the continuous Euler-Lagrange equations is less than the number of steps in the discrete Euler-Lagrange equations; in particular, when the continuous Lagrangian is regular and the discrete Lagrangian is degenerate. This case is emphasized in the following theorem.

**Theorem 3.2.1** Consider a symplectic manifold  $(M, \Omega)$  with exact symplectic structure  $\Omega = -d\vartheta$ . Recall the phase-space Lagrangian  $L_{PS} : TM \rightarrow \mathbb{R}$  is given in terms of coordinates  $z$  on  $M$  by:

$$L(z, \dot{z}) = \vartheta_i(z)\dot{z}^i - H(z).$$

Consider a discrete Lagrangian  $L_d : M \times M \rightarrow \mathbb{R}$ ,  $(z_0, z_1) \mapsto L_d(z_0, z_1)$ . If the discrete Lagrangian is regular, i.e.  $\det(D_1 D_2 L_d) \neq 0$ , the discrete Lagrangian map  $F_{L_d}$  corresponds to a multistep variational integrator for the Euler-Lagrange equations.

**Proof** Recall that the Euler-Lagrange equations for the phase-space Lagrangian specify a first-order (Hamiltonian) differential equation, so the domain of the Lagrange map  $F_L$  is  $M$ . Also recall that the discrete Euler-Lagrange equations are given by:

$$D_2 L_d(z_{k-1}, z_k) + D_1 L_d(z_k, z_{k+1}) = 0.$$

By the regularity of the discrete Lagrangian, the discrete Lagrange map is well-defined and specifies  $z_{k+1}$  as a function of  $(z_{k-1}, z_k)$ , so  $F_{L_d} : M \times M \rightarrow M \times M$ . The dimension of the domain of the discrete Lagrange map is then twice the dimension of the initial condition provided by the initial value problem, and the discrete Euler-Lagrange equations are therefore a multistep variational integrator with  $p = 2$ . ■

This simple theorem assigns a significant new categorization to many existing variational algorithms [13, 14, 15, 16, 96, 92]. Because regularity of the discrete Lagrangian is typically expected (contributions of Chapter 4 being the exception), one is likely to encounter multistep variational integrators whenever phase-space Lagrangians have been discretized.

Now that we have identified where multistep variational integrators appear, a related question is: given a two-step method for approximating Hamilton's equations (in canonical or non-canonical coordinates), does there exist a discrete variational formulation for the multistep update rule? This question harks back to an analogous question in the continuous setting, which became known as *Helmholtz's inverse problem of the calculus of variations* [115, 116]. Namely, given some second-order differential equation, can this equation be derived as the Euler-Lagrange equations for some Lagrangian? In the discrete setting, this question has recently been answered for one-dimensional systems by identifying a *discrete Helmholtz condition*, which is a necessary and sufficient condition for a second-order finite difference scheme to be a variational integrator [117]. In the context of multistep variational

integrators, the following theorem presents a vectorized version of the condition in Ref. [117], and will be shown to be at least a necessary condition.

**Theorem 3.2.2** *Given an  $h$ -dependent two-step update rule  $f_d(z_0, z_1, z_2, t) = 0$ , the update rule possesses a discrete variational formulation only if:*

$$D_3 f_d(z_0, z_1, z_2, t) - (D_1 f_d(z_1, z_2, z_3, t+h))^T = 0, \quad (3.9)$$

for all  $(z_0, z_1, z_2, z_3, t)$ .

**Proof** Suppose that the update rule is derived from a variational integrator. Then:

$$f_d(z_0, z_1, z_2, t) = D_2 L_d(z_0, z_1, t-h) + D_1 L_d(z_1, z_2, t),$$

for some  $L_d$  that has not been presumed time independent. Calculating the terms in the condition:

$$\begin{aligned} D_3 f_d(z_0, z_1, z_2, t) &= D_2 D_1 L_d(z_1, z_2, t) \\ (D_1 f_d(z_1, z_2, z_3, t+h))^T &= (D_1 D_2 L_d(z_1, z_2, t))^T = D_2 D_1 L_d(z_1, z_2, t) \end{aligned}$$

Thus Eq. (3.9) is shown to be a necessary condition. ■

The discrete Helmholtz condition serves as a rapid test for the existence of a discrete variational formulation. As illustrated more generally in the proof, the explicit midpoint algorithm present in the introduction is sure to satisfy the discrete Helmholtz condition. For another example, consider the two-step Adams-Bashforth update given by:

$$f_{ab2}(z_0, z_1, z_2) = z_2 - z_1 - h \left( \frac{3}{2} X(z_1) - \frac{1}{2} X(z_0) \right). \quad (3.10)$$

Checking the discrete Helmholtz condition:

$$D_3 f_{ab2}(z_0, z_1, z_2) - D_1 f_{ab2}(z_1, z_2, z_3) = I - \frac{h}{2} \nabla X(z_1), \quad (3.11)$$

which is not equal to zero. Therefore, no variational formulation exists for the second-order Adams-Bashforth method.

## 3.3 Stability

### 3.3.1 Spectral Analysis

As can be seen from the definitions of multistep variational integrators and multistep methods in general, the numerical trajectories of multistep methods evolve in a higher-dimensional space than the physical dynamics. For multistep variational integrators constructed from action principles in phase space, the numerical trajectories evolve in a space with twice the dimension of the Hamiltonian phase space. A natural question is: how does the additional freedom in the numerical trajectory impact the dynamics? Major insight toward answering this question emerges from spectral analysis [12, 6].

Consider a  $p$ -step method given by an update rule of the form of Eq. (3.6). Consider the  $h = 0$  limit of the  $h$ -dependent numerical method. For a consistent method, the trajectory  $z_{k+1} = z_k = \dots = z_{k+1-p}$  must satisfy the  $h = 0$  incarnation of the numerical method; the true time-zero flow map is the identity. So,  $(\overbrace{\bar{z}, \dots, \bar{z}}^{p \text{ times}}) \in \overbrace{M \times \dots \times M}^{p \text{ times}}$  may be considered to be an equilibrium point of the potentially nonlinear numerical method (again, at zero step size) for any  $\bar{z} \in M$ . To analyze the spectral stability near the fixed point  $(\bar{z}, \dots, \bar{z})$ , construct the linearized update rule  $f_{linear}$  according to:

$$f_{linear}(z_0, z_1, \dots, z_p) = D_1 f_d(z_0 - \bar{z}) + D_2 f_d(z_1 - \bar{z}) + \dots + D_p f_d(z_p - \bar{z}), \quad (3.12)$$

where all  $D_i f_d$  are evaluated at  $(\bar{z}, \dots, \bar{z})$ . Next, extract the homogeneous portion of the linearized update rule, and posit the ansatz  $z_p = \lambda z_{p-1} = \dots = \lambda^p z_0$ . The eigenvalues are specified as the roots of the *characteristic polynomial*  $\rho : \mathbb{C} \rightarrow \mathbb{C}$ , which is given by:

$$\rho(\lambda) = \det(D_1 f_d(\bar{z}, \dots, \bar{z}) + \lambda D_2 f_d(\bar{z}, \dots, \bar{z}) + \dots + \lambda^p D_p f_d(\bar{z}, \dots, \bar{z})). \quad (3.13)$$

In general, with  $\dim(M) = n$ , the characteristic polynomial possesses  $np$  roots in the complex plane. Typically, the roots appear in sets of  $n$  repeated algebraic roots, and thus  $p$  potentially distinct values characterize the eigenspectrum of the multistep method. Consider these  $p$  values  $\lambda_1, \dots, \lambda_p$ . For consistent methods,  $n$  roots of the characteristic polynomial occur at 1, as required to converge to the identity map. So, choose  $\lambda_1 = 1$  and denote it the *principal root*. The mode of the numerical dynamics corresponding to the  $\lambda_1 = 1$  root is denoted the *principal mode*. For the remaining  $p - 1$  potentially distinct eigenvalues, the corresponding modes are denoted the *parasitic modes* and their eigenvalues *parasitic roots*. The parasitic modes indicate a component of the solution that does not converge to the identity map in the  $h \rightarrow 0$  limit, and thus this component of the solution is non-physical. The nomenclature “parasitic” serves as a warning that the non-physical modes may detract from the physically desirable solution.

*Note:* The principal and parasitic roots are *defined* by their values in the  $h \rightarrow 0$  limit. Although later analyses will study the algorithm using non-zero  $h$ , the parasitic and principle roots are *not* re-calculated, but instead represented by their  $h \rightarrow 0$  values [6]. For instance, in Fig. 3.1, there exists a “single” principle root at positive one and a “single” parasitic root at negative one. Also note: in Fig. 3.1 there are two eigenvalues near both positive and negative one, but because they have distinct eigenvectors, the convention will be to call it a single mode unless it is important to distinguish between the eigenvalues.

Eigenvalue analysis of multistep methods reveals that two classes of modes are present in the dynamics: principal and parasitic. An important aspect of the dialog on multistep methods is assessing and mitigating the contribution of the parasitic components to the numerical solution. One possibility is that the parasitic modes exponentially damp away [93, 118, 12, 6]. This occurs when all of the parasitic roots have magnitude  $|\lambda| < 1$ . In this case, the multistep method is deemed *strictly stable* or *A-stable*; popular Adams-Basforth and Adams-Moulton methods exhibit this property. Correspondingly, one need not expend effort in carefully initializing the multistep dynamics because the parasitic mode contributions will exponentially diminish in amplitude. The long term dynamics of strictly stable multistep methods are effectively equivalent to the long term dynamics of one-step methods. Another possibility is that the amplitude of the parasitic modes may be shown to remain bounded for long times [119, 111, 6], even including nonlinear effects. In this case, one may expend extra care in the initialization procedure to minimize the parasitic contributions, but one may rest assured that they will not grow exponentially. A third possibility is that the parasitic mode amplitudes do indeed grow, in which case some sort of cleaning procedure is advised if there exist compelling reasons to use the unstable numerical method [120, 121].

Having established the significance of spectral analysis for multistep methods, let us investigate the spectral properties of multistep variational integrators, beginning with the following theorem.

**Theorem 3.3.1** *For a multistep variational integrator defined by a discrete Lagrangian  $L_d(z_0, z_1, h)$ , the characteristic equation near an expansion point  $\bar{z}$  is given by:*

$$\det \left( D_1 D_2 L_d(\bar{z}, \bar{z}, 0) + \lambda (D_2 D_2 L_d(\bar{z}, \bar{z}, 0) + D_1 D_1 L_d(\bar{z}, \bar{z}, 0)) + \lambda^2 D_2 D_1 L_d(\bar{z}, \bar{z}, 0) \right) = 0. \quad (3.14)$$

*Moreover, if  $\lambda$  is a root of the characteristic equation, then the reciprocal  $\lambda^{-1}$  and the complex conjugate  $\lambda^*$  are also roots.*

**Proof** The  $h \rightarrow 0$  version of the discrete Euler-Lagrange equations is given by:

$$D_2 L_d(z_{k-1}, z_k, 0) + D_1 L_d(z_k, z_{k+1}, 0) = 0. \quad (3.15)$$

Linearizing the system about an equilibrium point  $\bar{z}$ :

$$\begin{aligned} D_1 D_2 L_d(\bar{z}, \bar{z}, 0)(z_{k-1} - \bar{z}) + (D_2 D_2 L_d(\bar{z}, \bar{z}, 0) + D_1 D_1 L_d(\bar{z}, \bar{z}, 0))(z_k - \bar{z}) + \\ D_2 D_1 L_d(\bar{z}, \bar{z}, 0)(z_{k+1} - \bar{z}) = 0. \end{aligned} \quad (3.16)$$

Next, assume  $z_{k+1} = \lambda z_k = \lambda^2 z_{k-1}$  and collect the homogeneous terms (those that depend on  $z_{k-1}$ ):

$$\begin{aligned} (D_1 D_2 L_d(\bar{z}, \bar{z}, 0) + \lambda D_2 D_2 L_d(\bar{z}, \bar{z}, 0) + \\ \lambda D_1 D_1 L_d(\bar{z}, \bar{z}, 0) + \lambda^2 D_2 D_1 L_d(\bar{z}, \bar{z}, 0)) z_{k-1} = 0. \end{aligned} \quad (3.17)$$

The condition that there exist non-trivial  $z_{k-1}$  satisfying this relation yields the characteristic equation.

Next, suppose  $\lambda$  is a root of the characteristic equation. Dividing the characteristic equation by  $\lambda^2$  yields:

$$\det \left( \lambda^{-2} D_1 D_2 L_d(\bar{z}, \bar{z}, 0) + \lambda^{-1} (D_2 D_2 L_d(\bar{z}, \bar{z}, 0) + D_1 D_1 L_d(\bar{z}, \bar{z}, 0)) + D_2 D_1 L_d(\bar{z}, \bar{z}, 0) \right) = 0. \quad (3.18)$$

To show that  $\lambda^{-1}$  is also a root of the characteristic equation, take the transpose of this relation to recover:

$$\det \left( D_1 D_2 L_d(\bar{z}, \bar{z}, 0) + \lambda^{-1} (D_2 D_2 L_d(\bar{z}, \bar{z}, 0) + D_1 D_1 L_d(\bar{z}, \bar{z}, 0)) + \lambda^{-2} D_2 D_1 L_d(\bar{z}, \bar{z}, 0) \right) = 0. \quad (3.19)$$

The final claim follows from the complex conjugate root theorem because the  $D_i D_j L_d$  tensors contain only real entries. The characteristic equation is then a polynomial equation with real coefficients in the one variable  $\lambda$ , satisfying the conditions for the complex conjugate root theorem. ■

**Corollary 3.3.2** *If a multistep variational integrator possesses an eigenvalue  $\lambda$  such that  $|\lambda| < 1$ , then there also exists an eigenvalue  $\lambda'$  such that  $|\lambda'| > 1$ . That is, multistep variational integrators cannot be strictly stable.*

**Proof** This corollary follows from the reciprocal pairs claim of the Theorem. ■

We see here that the structure-preserving property of multistep variational integrators comes at the expense of the methods being especially susceptible to parasitic mode instabilities. Unfortunately, this indicates that the poor long-term behavior of the parasitic mode observed in the opening example is more the standard than the exception. The following example illustrates the spectral analysis in practice and the numerical manifestations of parasitic modes.

**Example** The *trapezoidal rule* for approximating the integral of a function  $f(t)$  between  $t = a$  and  $t = b$  is given by:

$$\int_a^b f(t) dt \approx (b - a) \frac{f(a) + f(b)}{2}. \quad (3.20)$$

In light of this, given a phase-space Lagrangian  $L_{PS}$ , let the *trapezoidal discrete Lagrangian* be:

$$L_d(z_0, z_1) = \frac{h}{2} \left( L_{PS}(z_0, \frac{z_1 - z_0}{h}) + L_{PS}(z_1, \frac{z_1 - z_0}{h}) \right). \quad (3.21)$$

For the first case study, consider a linear oscillator system:

$$z = [q, p]^T \quad (3.22)$$

$$H(q, p) = \frac{1}{2}(q^2 + p^2) \quad (3.23)$$

$$L_{PS}(z, \dot{z}) = p\dot{q} - \frac{1}{2}(q^2 + p^2). \quad (3.24)$$

The discrete Euler-Lagrange equations for the trapezoidal discrete Lagrangian applied to the phase-space linear oscillator Lagrangian are given by:

$$-p_{k+1} + p_{k-1} - 2hH_{,q}(q_k, p_k) = 0 \quad (3.25a)$$

$$q_{k+1} - q_{k-1} - 2hH_{,p}(q_k, p_k) = 0. \quad (3.25b)$$

This is a *two-step* method, and is commonly referred to as *explicit midpoint*. The linearity of the discrete Euler-Lagrange equations for this system enables exact identification of a general solution. For a given step size  $h$ , the four eigenvalues  $(\lambda_1^+, \lambda_1^-, \lambda_2^+, \lambda_2^-)$  are given by:

$$\lambda_1^\pm = \pm ih + \sqrt{1 - h^2} \quad (3.26a)$$

$$\lambda_2^\pm = \pm ih - \sqrt{1 - h^2}. \quad (3.26b)$$

These eigenvalues reside on the unit circle and are illustrated in Fig. 3.3. The two  $\lambda_1$  eigenvalues are near positive one; they correspond to the principal mode. The two  $\lambda_2$  eigenvalues are near negative one, and correspond to the parasitic modes. Remember - the “principle roots” and “parasitic roots” are labeled by their values in the  $h \rightarrow 0$  limit, so for this system

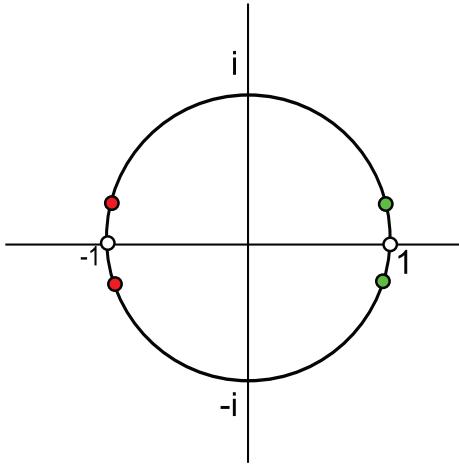


Figure 3.3: The trapezoidal multistep variational integrator for the linear oscillator problem possesses four eigenvalues. The two near positive one correspond to the principal mode (green) and the two near negative one correspond to a parasitic mode with an  $h = 0$  eigenvalue of  $-1$ .

one the principle root(s) is(are) positive one and the parasitic root(s) is(are) negative one. The “plus” eigenvalues share an eigenvector  $(1, i)^T$  and the “minus” eigenvalues share an eigenvector  $(1, -i)^T$ . So, the general solution to the discrete Euler-Lagrange equations may be written as:

$$\begin{pmatrix} q_k \\ p_k \end{pmatrix} = (c_1^+(\lambda_1^+)^k + c_2^+(\lambda_2^+)^k) \begin{pmatrix} 1 \\ i \end{pmatrix} + (c_1^-(\lambda_1^-)^k + c_2^-(\lambda_2^-)^k) \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (3.27)$$

For a specific trajectory, the two sets of initial conditions  $(q_0, p_0)$  and  $(q_1, p_1)$  dictate the values of  $c_1^\pm, c_2^\pm$ . Because the initial value problem only prescribes  $(q_0, p_0)$ , a starting procedure is required to determine  $(q_1, p_1)$ . Regardless of the initial amplitude of the parasitic modes, the marginal stability and linear dynamics maintain this amplitude throughout the numerical trajectory.

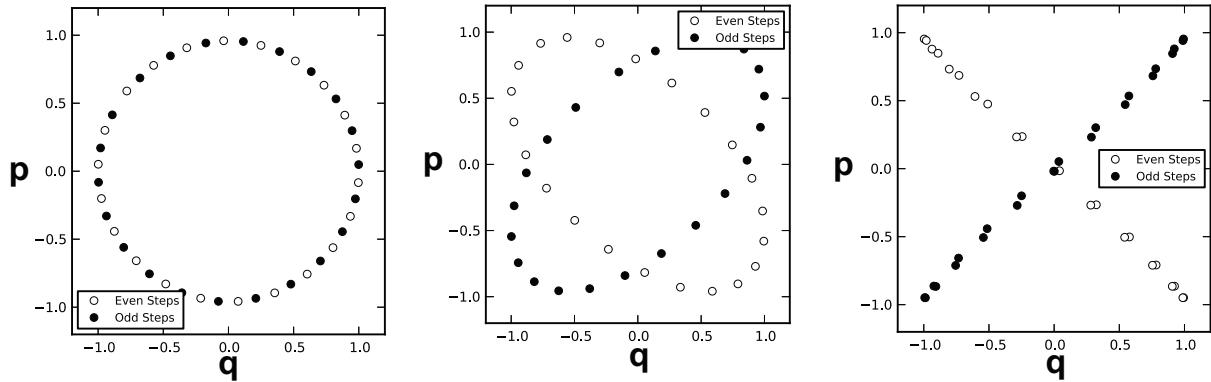


Figure 3.4: Parasitic modes are nonlinearly unstable for explicit midpoint applied to the nonlinear pendulum problem. The short-, medium-, and long-term amplitude of the parasitic even-odd oscillations is apparent in the phase portraits presented in the left, center and right panels. The full time duration shown in the portraits is  $\sim 10^5$  periods of the nonlinear pendulum.

When nonlinear effects are introduced- for instance if the linear oscillator Hamiltonian is replaced with the nonlinear pendulum Hamiltonian  $H(q, p) = \frac{1}{2}p^2 - \cos(q)$  - the parasitic modes can increase in amplitude over time. This is depicted in Fig. 3.4, in which explicit midpoint is applied to the nonlinear pendulum problem. The parasitic modes grow slowly in time, eventually overwhelming the numerical dynamics after  $\sim 10^5$  periods of the nonlinear pendulum. For this system, the eigenvalues are *the same* as those for the linear oscillator (in the  $h \rightarrow 0$  limit), so this study serves as a warning that marginal linear stability is not sufficient to ensure good long term behavior.

For an example with test particle implications, consider the trapezoidal discretization of the guiding center Lagrangian. This is the integrator presented in Ref. [15], although note that the test case here will take place with three spatial dimensions as opposed to the two spatial dimension cases of Ref. [15]. The guiding center system is described by:

$$z = [x, u]^T \quad (3.28)$$

$$L_{PS} = (A_i + ub_i)\dot{x}^i - \left(\frac{u^2}{2} + \mu B + \phi\right) \quad (3.29)$$

where  $x$  is the guiding center position,  $u$  the parallel velocity,  $A$  the magnetic vector potential,  $b$  the magnetic field unit vector,  $\mu$  the magnetic moment,  $B$  the magnetic field magnitude and  $\phi$  the scalar potential. For this system, the corresponding discrete Euler-Lagrange equations are:

$$\begin{aligned} A_{i,j}^\dagger(z_k)(x_{k+1}^i - x_{k-1}^i) - (A_j^\dagger(z_{k+1}) - A_j^\dagger(z_{k-1})) - \\ 2h(\mu B_{,j}(x_k) + \phi_{,j}(x_k)) = 0 \quad (j = 1, 2, 3) \end{aligned} \quad (3.30a)$$

$$b_i(x_k)(x_{k+1}^i - x_{k-1}^i) - (u_{k+1} + u_{k-1}) = 0, \quad (3.30b)$$

where  $A^\dagger = A + ub$ . It is interesting to note that this two-step method is not a linear, one-leg, or even general linear multistep method; the presence of the magnetic vector potential in the update rule indicates that the update rule cannot be formulated in terms of the guiding center equations of motion alone.

This multistep variational integrator also exhibits a single (four-dimensional) parasitic mode at negative one. This may be established by setting  $h = 0$  and linearizing the update rule about  $x_{k-1}$ . The impact of the parasitic mode is demonstrated in Fig. 3.5, where the growth of the parasitic mode leads to divergent even-odd decoupling during the second period of the trapped particle orbit. Clearly, this issue must be addressed before multistep variational integrators may be considered reliable.

### 3.3.2 Backward Error Analysis

Having distinguished the eigenvalue spectrum of multistep methods into principal and parasitic roots, a valuable technique for assessing the time evolution of the corresponding modes is *backward error analysis*. Although linear analysis provides preliminary insight on the dynamics, more powerful techniques are required to, for instance, assess the nonlinear behavior of marginally stable parasitic modes.

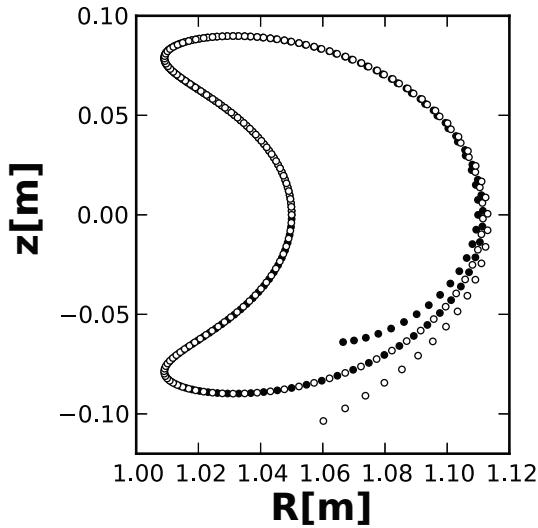


Figure 3.5: The parasitic mode at negative one leads to divergent even-odd oscillations on the second pass of a trapped particle trajectory. **Black:** Trajectory of  $z_k$  with  $k$  even. **White:** Trajectory of  $z_k$  with  $k$  odd. The integrator and field configuration may be found in Ref. [15]; cylindrical coordinates were used to perform the above study.

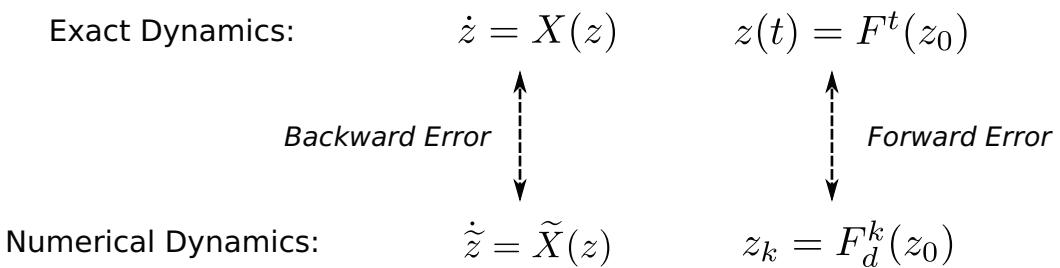


Figure 3.6: Whereas forward error analysis compares the numerical map to the true solution, backward error analysis compares a modified vector field whose solution passes through the numerical trajectory to the original vector field.

Any type of error analysis compares some aspect of the numerical dynamics to the corresponding aspect of the original system being modeled; the “error” is relative to the true system. Because the numerical method is specified as a mapping, a natural concept of error compares the numerical map to the true solution. This is referred to as *forward error analysis*. As contrasted in Fig. 3.6, backward error analysis posits the existence of some *modified vector field*  $\tilde{X}$ , different from the original, such that the flow of the modified vector field passes through the numerical trajectory at the discretized time instances  $t_k$ . Insight then emerges by comparing the modified vector field to the original. Important results established by this procedure include the Hamiltonian character of numerical trajectories generated by symplectic integrators; their modified vector fields remain Hamiltonian.

In the context of multistep methods, there exist two useful constructions for backward error analysis. The first construction ignores the prospect of parasitic modes, yielding a *smooth modified system* that describes the time evolution of the principal component of the solution. The smooth dynamics are useful for analyzing properties of the desired component of the numerical solution, and will be put to use in Section 3.6 to derive initial conditions that minimize the initial amplitude of the parasitic modes. The more general construction yields a *general modified system* by taking into account both the principal and parasitic modes. The smooth and general modified system are formally defined as follows (modeled after the definitions in Refs. [12, 6]).

**Definition** Consider a multistep method defined by an update rule  $f_d(z_0, \dots, z_p) = 0$ . The *smooth modified system* of the multistep method is determined by the *smooth modified vector field*  $\tilde{X}$  according to:

$$\dot{\tilde{z}} = \tilde{X}(\tilde{z}), \quad (3.31)$$

and the condition that every solution  $\tilde{z}(t)$  satisfies:

$$f_d(\tilde{z}(t_k), \dots, \tilde{z}(t_{k+p})) = 0. \quad (3.32)$$

This equality is in the sense of a formal power series, that is the relation holds at each power in  $h$ , and there (must) exist  $h$ -independent coefficient functions  $X_i$  such that the smooth modified vector field  $\tilde{X}$  may be expressed as the formal power series:

$$\tilde{X} = X_0 + hX_1 + h^2X_2 + \dots. \quad (3.33)$$

A few notes for clarity: For consistent multistep methods,  $X_0$  must equal the original vector field  $X$ . For multistep methods of order  $r$ , the coefficient vector fields  $X_i$  with  $0 < i < r$  are equal to zero. Finally, note that this series is in general asymptotic, so it may be necessary to truncate the series after some number of terms to obtain minimum error [12, 6]. The more general dynamics, including parasitic modes, may be modeled using the modified system defined next [6].

**Definition** Consider a  $p$ -step multistep method defined by an update rule  $f_d$  with principal root  $\lambda_1 = 1$  and  $(p-1)$  parasitic roots  $\lambda_i, i = 2, \dots, p$ . Enumerate the set of all finite products of the parasitic roots:

$$\{\lambda_i\}_{i \in \mathcal{I}} = \{\lambda = \lambda_1^{m_1} \cdot \dots \cdot \lambda_p^{m_p}; m_j \geq 0\} = \{\lambda_1, \dots, \lambda_p, \lambda_{p+1}, \dots\}, \quad (3.34)$$

and let the collection of parasitic root indices be  $\mathcal{I}^* = \mathcal{I} \setminus \{1\}$ . Supposing a solution of the form:

$$\tilde{z}(t) = z_1(t) + \sum_{i \in \mathcal{I}^*} \lambda_i^{t/h} z_i(t), \quad (3.35)$$

such that:

$$f_d(\tilde{z}(t_k), \dots, \tilde{z}(t_{k+p})) = 0, \quad (3.36)$$

the evolution of the smooth  $z_i$  functions is specified by the *general modified system*:

$$\dot{z}_1 = X_{10}(z_1, \dots, z_p) + hX_{11}(z_1, \dots, z_p) + \dots \quad (3.37a)$$

$$\dot{z}_i = X_{i0}(z_1, \dots, z_p) + hX_{i1}(z_1, \dots, z_p) + \dots \quad 2 \leq i \leq p \quad (3.37b)$$

$$z_i = hX_{i0}(z_1, \dots, z_p) + h^2X_{i1}(z_1, \dots, z_p) + \dots \quad i > p, \quad (3.37c)$$

with the  $X_{ij}$  vector fields independent of  $h$  and unique. It is assumed that the parasitic modes  $z_{i>1}$  are initialized to  $\mathcal{O}(h)$ .

The general modified system complements the spectral analysis with a more specific description of the evolution of the multistep dynamics, provided the vector fields  $X_{ij}$  of Eq. (3.37) can be found. For one, describing the numerical trajectory requires analysis of both the parasitic eigenvalues and their finite products (not equal to one), as is apparent from Eq. (3.35) and Eq. (3.37c). Should the eigenvalues of the characteristic equation form a group under multiplication (e.g. exactly two distinct eigenvalues with the principal root at positive one and a single parasitic mode at negative one), there do not exist any  $z_{i>p}$ . In connection with the smooth modified vector field, the vector fields  $X_{1i}$  of Eq. (3.37a) are equivalent to  $X_i$  of Eq. (3.33) if  $z_{i>1} = 0$  for all time [6].

The following theorems establish the smooth and general modified systems for multistep variational integrators. The first theorem describes the smooth modified system, presenting a recursive expression for defining the  $i$ -th vector field  $X_i$  in the smooth modified vector field  $\tilde{X}$ . Later, the lowest-order vector fields for the parasitic modes will be defined for multistep variational integrators using the definition of the general modified system.

**Theorem 3.3.3 (Smooth Modified System of MVIs)** Consider a consistent multistep variational integrator satisfying Theorem 3.2.1. Further suppose that the discrete Lagrangian takes the form:

$$L_d(z_0, z_1) = L_d^\vartheta(z_0, z_1) + h L_d^H(z_0, z_1), \quad (3.38)$$

where  $L_d^\vartheta$  and  $L_d^H$  are independent of  $h$ . Then there exist unique,  $h$ -independent vector fields  $X_i$  such that for every truncation index  $r$ , every solution of:

$$\dot{\tilde{z}} = X_0(\tilde{z}) + h X_1(\tilde{z}) + \dots + h^r X_r(\tilde{z}) \quad (3.39)$$

satisfies

$$D_2 L_d(\tilde{z}(t_{k-1}), \tilde{z}(t_k)) + D_1 L_d(\tilde{z}(t_k), \tilde{z}(t_{k+1})) = \mathcal{O}(h^{r+2}). \quad (3.40)$$

The vector fields  $X_{i>1}$  are recursively defined according to:

$$\sum_{\substack{m+n_1+\dots+n_m \\ n_m-1=i}} \mathcal{G}_{n_1\dots n_m}(L_d^\vartheta, z) + \sum_{\substack{m+n_1+\dots+n_m \\ n_m=i}} \mathcal{G}_{n_1\dots n_m}(L_d^H, z) = 0. \quad (3.41)$$

Here, the summations are over the positive indices  $m, n_1, \dots, n_m$  satisfying the conditions under the  $\Sigma$ s. Also,  $\mathcal{G}_{n_1\dots n_m}(L_d, z)$  is defined as:

$$\begin{aligned} \mathcal{G}_{n_1\dots n_m}(L_d, z) = & \frac{1}{m!} (\mathcal{L}_{(0, X_{n_1})} \dots \mathcal{L}_{(0, X_{n_m})} d_2 L_d(z, z) + \\ & \mathcal{L}_{(X_{n_1}, 2X_{n_1})} \dots \mathcal{L}_{(X_{n_m}, 2X_{n_m})} d_1 L_d(z, z)), \end{aligned} \quad (3.42)$$

where  $\mathcal{L}_X$  is the Lie derivative with respect to the vector field  $X$ , the slot exterior derivative  $d_i$  denotes the exterior derivative with respect to the  $i$ -th argument.

**Proof** To arrive at the defining relation Eq. (3.41) for the vector fields  $X_i$  appearing in Eq. (3.39), the procedure will be to:

- substitute the solution  $\tilde{z}(t)$  into the DEL equations,
- expand the DEL equations about  $\tilde{z}(t_{k-1})$ ,
- collect in powers of  $h$ ,
- recursively identify  $X_i$  to satisfy the update rule at each order.

In principle, the mathematics required to execute this task involve nothing more than Taylor expansions and algebraic manipulation to collect terms. However, the number of terms rapidly grows unwieldy with this straightforward approach. To maintain concise expressions, the analysis here will leverage tools from differential geometry, primarily differential forms and Lie derivatives. These geometric tools will also facilitate later calculations of explicit expressions.

As a first step, re-express the discrete Euler-Lagrange equations in terms of the exterior derivative:

$$d_2 L_d(z_0, z_1) + d_1 L_d(z_1, z_2) = 0, \quad (3.43)$$

where  $d_i$  is the  $i$ -th slot exterior derivative, so for example  $d_1 L_d(z_1, z_2) = D_1 L_d(z_1, z_2) dz_1$ . Also, the trajectory will be considered to originate at  $t_0$  as opposed to  $t_{k-1}$  because of the presumed time independence and to avoid the more tedious subscripts. Next, let  $\tilde{F}^t : M \rightarrow M$  be the time- $t$  flow of  $\tilde{X}$  (truncated at  $\mathcal{O}(h^r)$ ). Then, the task is to *choose* the  $X_i$  in Eq. (3.33) such that:

$$d_2 L_d(z_0, \tilde{F}^h(z_0)) + d_1 L_d(\tilde{F}^h(z_0), \tilde{F}^h \circ \tilde{F}^h(z_0)) = 0, \quad (3.44)$$

at each order in  $h$ . Working toward expanding about  $z_0$ , define two maps  $F_1^t : M \rightarrow M \times M, F_2^t : M \rightarrow M \times M$  according to:

$$F_1^t(z) = \left( \text{id}(z), \tilde{F}^t(z) \right) \quad (3.45)$$

$$F_2^t(z) = \left( \tilde{F}^t(z), \tilde{F}^t \circ \tilde{F}^t(z) \right), \quad (3.46)$$

where  $\text{id}$  is the identity map on  $M$ . In terms of these maps, the DEL equations become:

$$d_2 L_d(F_1^h(z_0)) + d_1 L_d(F_2^h(z_0)) = 0 \quad (3.47)$$

$$\left( F_1^{h*} d_2 L_d \right)(z_0) + \left( F_2^{h*} d_1 L_d \right)(z_0) = 0, \quad (3.48)$$

where  $F^*$  denotes the pullback under  $F$ <sup>1</sup>.

Next, to expand about  $z_0$ , the Lie derivative theorem will be applied. Given a vector field  $X$  with time- $t$  flow  $F^t$  and a one-form  $\alpha$ , the *Lie derivative theorem* states (see for instance Theorem 4.3.1 of Ref. [4]):

$$\frac{d}{dt} F^{t*} \alpha = F^{t*} \mathcal{L}_X \alpha. \quad (3.49)$$

This is useful for expanding  $F^{h*} \alpha$  for small  $h$  because a series can be constructed by iteratively integrating both sides of the Lie derivative theorem. The end result is that:

$$F^{h*} \alpha = \alpha|_{t=0} + h \mathcal{L}_X \alpha|_{t=0} + \frac{h^2}{2} \mathcal{L}_x \mathcal{L}_x \alpha|_{t=0} + \dots = \sum_{i=0}^{\infty} \frac{h^i}{i!} \mathcal{L}_x^i \alpha|_{t=0}. \quad (3.50)$$

---

<sup>1</sup>Given a map  $F : M \rightarrow N$  and a differential form  $\alpha$  on  $N$ ,  $F^* \alpha$  recovers a differential form on  $M$ . For one-forms, this may be visualized as follows. Given a tangent vector  $v$  on  $M$ , determine a real number corresponding to  $F^* \alpha(v)$  by determining a tangent vector  $w$  on  $N$  by pushing  $v$  forward to  $N$  using  $F$ , then evaluating  $\alpha(w)$ . For further details see e.g. Ref. [20]

This formula provides the Taylor expansion about time  $t = 0$ . To apply these relations to the multistep variational integrator update, note that  $F_1^t, F_2^t$  satisfy:

$$\frac{d}{dt} F_1^t(z) = (0, \tilde{X}(F_1^t(z))) \quad (3.51a)$$

$$\frac{d}{dt} F_2^t(z) = (\tilde{X}(F_2^t(z)), 2\tilde{X}(F_2^t(z))) \quad (3.51b)$$

Performing the Lie derivative expansion on the multistep variational integrator yields:

$$d_2 L_d(z_0, z_0) + d_1 L_d(z_0, z_0) + h \left( \mathcal{L}_{(0, \tilde{X})} d_2 L_d(z_0, z_0) + \mathcal{L}_{(\tilde{X}, 2\tilde{X})} d_1 L_d(z_0, z_0) \right) + \frac{h^2}{2} \left( \mathcal{L}_{(0, \tilde{X})} \mathcal{L}_{(0, \tilde{X})} d_2 L_d(z_0, z_0) + \mathcal{L}_{(\tilde{X}, 2\tilde{X})} \mathcal{L}_{(\tilde{X}, 2\tilde{X})} d_1 L_d(z_0, z_0) \right) + \dots = 0, \quad (3.52)$$

or compactly:

$$\sum_{m=0}^{\infty} \frac{h^m}{m!} \left( \mathcal{L}_{(0, \tilde{X})}^m d_2 L_d(z_0, z_0) + \mathcal{L}_{(\tilde{X}, 2\tilde{X})}^m d_1 L_d(z_0, z_0) \right) = 0. \quad (3.53)$$

The final step in the process is to collect in powers of  $h$ . Examining Eq. (3.53), recall that powers of  $h$  appear in three places: first, in the  $\frac{h^m}{m!}$  coefficients; second,  $L_d$  contains a zero-th and first-order term in  $h$ ; and third,  $\tilde{X}$  contains all powers of  $h$  up to the truncation index  $r$ . The collection of terms is facilitated by the linearity of the Lie derivative operator:

$$\mathcal{L}_X(\alpha + \beta) = \mathcal{L}_X\alpha + \mathcal{L}_X\beta \quad (3.54a)$$

$$\mathcal{L}_X(c\alpha) = c\mathcal{L}_X\alpha \quad (3.54b)$$

$$\mathcal{L}_{cX}\alpha = c\mathcal{L}_X\alpha \quad (3.54c)$$

$$\mathcal{L}_{X_1+X_2}\alpha = \mathcal{L}_{X_1}\alpha + \mathcal{L}_{X_2}\alpha, \quad (3.54d)$$

where  $c$  is taken to be a constant,  $X, X_1, X_2$  as vector fields, and  $\alpha, \beta$  are differential forms.

First, making the  $h$  dependence of  $L_d$  in Eq. (3.53) explicit:

$$\begin{aligned} & \sum_{m=0}^{\infty} \frac{h^m}{m!} \left( \mathcal{L}_{(0,\tilde{X})}^m d_2 L_d^\vartheta(z_0, z_0) + \mathcal{L}_{(\tilde{X},2\tilde{X})}^m d_1 L_d^\vartheta(z_0, z_0) \right) + \\ & \frac{h^{m+1}}{m!} \left( \mathcal{L}_{(0,\tilde{X})}^m d_2 L_d^H(z_0, z_0) + \mathcal{L}_{(\tilde{X},2\tilde{X})}^m d_1 L_d^H(z_0, z_0) \right) = 0. \end{aligned} \quad (3.55)$$

The final factors of  $h$  to extract are those introduced by  $\tilde{X}$ , which can be accomplished through the relation:

$$\begin{aligned} \mathcal{L}_{(0,\tilde{X})}^m L_d &= \left( \sum_{n=0}^{\infty} h^n \mathcal{L}_{(0,X_n)} \right)^m L_d \\ &= \sum_{n_1, \dots, n_m=0}^{\infty} h^{n_1+\dots+n_m} \mathcal{L}_{(0,X_{n_1})} \dots \mathcal{L}_{(0,X_{n_m})} L_d \end{aligned} \quad (3.56)$$

and similarly for  $\mathcal{L}_{(\tilde{X},2\tilde{X})}^m$ . The expanded discrete Euler Lagrange equations become:

$$\begin{aligned} & \sum_{m=0}^{\infty} \sum_{n_1, \dots, n_m=0}^{\infty} \frac{h^{(m+n_1+\dots+n_m)}}{m!} \left( \mathcal{L}_{(0,X_{n_1})} \dots \mathcal{L}_{(0,X_{n_m})} d_2 L_d^\vartheta(z_0, z_0) + \right. \\ & \quad \left. \mathcal{L}_{(X_{n_1},2X_{n_1})} \dots \mathcal{L}_{(X_{n_m},2X_{n_m})} d_1 L_d^\vartheta(z_0, z_0) \right) + \\ & \frac{h^{(m+n_1+\dots+n_m+1)}}{m!} \left( \mathcal{L}_{(0,X_{n_1})} \dots \mathcal{L}_{(0,X_{n_m})} d_2 L_d^H(z_0, z_0) + \right. \\ & \quad \left. \mathcal{L}_{(X_{n_1},2X_{n_1})} \dots \mathcal{L}_{(X_{n_m},2X_{n_m})} d_1 L_d^H(z_0, z_0) \right) = 0. \end{aligned} \quad (3.57)$$

It is at this point that the definition of the functional  $\mathcal{G}$  according to Eq. (3.42) is especially useful; the above equation condenses to:

$$\sum_{m=0}^{\infty} \sum_{n_1, \dots, n_m=0}^{\infty} h^{(m+n_1+\dots+n_m)} G_{n_1, \dots, n_m}(L_d^\vartheta, z_0) + h^{(m+n_1+\dots+n_m+1)} G_{n_1, \dots, n_m}(L_d^H, z_0). \quad (3.58)$$

The  $h$ -collected series is, at last, given by:

$$\sum_{i=0}^{\infty} h^i \left( \sum_{\substack{m+n_1+ \\ \dots+n_m=i}} G_{n_1\dots n_m}(L_d^\vartheta, z_0) + \sum_{\substack{m+n_1+ \\ \dots+n_m+1=i}} G_{n_1\dots n_m}(L_d^H, z_0) \right) = 0, \quad (3.59)$$

where the summations are over the non-negative indices  $m, n_1, \dots, n_m$  satisfying the conditions under the summation symbols<sup>2</sup>.

The relation defining the  $i$ -th vector field, as given in Eq. (3.41), emerges by requiring the series to be zero at *each* order in  $h$ , and noticing that the vector field  $X_i$  first appears at order  $h^{i+1}$ , namely in the term:

$$G_i(L_d^\vartheta, z_0) = \mathcal{L}_{(0, X_i)} L_d^\vartheta(z_0, z_0). \quad (3.60)$$

Equation (3.41) therefore defines  $X_i$  in terms of lower-order vector fields and their derivatives. Because  $X_0$  equals the original  $X_H$  for consistent methods,  $X_i$  may ultimately be expressed in terms of  $\vartheta, H$  and their derivatives.

The proof of the existence of  $X_i$  relies on the consistency of the method. Consider the  $i = 0$  instantiation of Eq. (3.41):

$$d_2 L_d^H + d_1 L_d^H + (\mathcal{L}_{(0, X_0)} d_2 L_d^\vartheta + \mathcal{L}_{(X_0, 2X_0)} d_1 L_d^\vartheta). \quad (3.61)$$

Because this defines  $X_0$  in terms of  $\vartheta$  and  $H$  and must yield  $X_0 = X_H$  for consistency, one may conclude:

$$(\mathcal{L}_{(0, X_0)} d_2 L_d^\vartheta(z, z) + \mathcal{L}_{(X_0, 2X_0)} d_1 L_d^\vartheta(z, z))_j = \Omega_{ij}(z) X_0^i(z) \quad (3.62a)$$

$$(d_2 L_d^H(z, z) + d_1 L_d^H(z, z))_j = H_{,j}(z). \quad (3.62b)$$

---

<sup>2</sup>For instance,  $i = 2$  allows the indices  $(m = 2, n_1 = 0, n_2 = 0)$  and  $(m = 1, n_1 = 1)$  for  $m+n_1+\dots+n_m = i$ .

Because  $\Omega$  is presumed non-degenerate by definition of the Hamiltonian vector field  $X_H$ , one may invert  $\Omega$  to determine  $X_0 = X_H$ . For  $X_{i>0}$ , Eq. (3.60) shows that  $X_i$  always appears as:

$$\mathcal{L}_{(0,X_i)} d_2 L_d^\vartheta(z, z) + \mathcal{L}_{(X_i, 2X_i)} d_1 L_d^\vartheta(z, z), \quad (3.63)$$

which by the zeroth-order vector field (the  $j$ -th component) must be equal to  $\Omega_{jk}(z) X_i^k(z)$ , so  $X_i$  is well-defined.

To conclude the proof, it remains to justify the  $\mathcal{O}(h^{r+2})$  term in Eq. (3.40). Recall that the  $\mathcal{O}(h^{r+1})$  term in the expanded DEL equations defines the  $X_r$  vector field, so truncating the smooth modified vector field at order  $r$  introduces an error term of  $\mathcal{O}(h^{r+2})$ . ■

For concreteness, the first few coefficients of  $h^i$  are given according to Eq. (3.59) as:

$$\begin{aligned} h^0 : & \quad d_2 L_d^\vartheta + d_1 L_d^\vartheta \\ h^1 : & \quad d_2 L_d^H + d_1 L_d^H + (\mathcal{L}_{(0,X_0)} d_2 L_d^\vartheta + \mathcal{L}_{(X_0, 2X_0)} d_1 L_d^\vartheta) \\ h^2 : & \quad \mathcal{L}_{(0,X_0)} d_2 L_d^H + \mathcal{L}_{(X_0, 2X_0)} d_1 L_d^H + \mathcal{L}_{(0,X_1)} d_2 L_d^\vartheta + \mathcal{L}_{(X_1, 2X_1)} d_1 L_d^\vartheta + \\ & \quad \frac{1}{2} (\mathcal{L}_{(0,X_0)} \mathcal{L}_{(0,X_0)} d_2 L_d^\vartheta + \mathcal{L}_{(X_0, 2X_0)} \mathcal{L}_{(X_0, 2X_0)} d_1 L_d^\vartheta), \end{aligned} \quad (3.64)$$

where the arguments of the Lie differentiated  $L_d^\vartheta, L_d^H$  functions are  $(z, z)$ . Setting each of these coefficients of  $h$  to zero determines terms in the modified vector field.

Turning attention now to the general modified system, the following theorem identifies the lowest-order vector fields for the parasitic modes of multistep variational integrators. Although it stops short of identifying all of the vector fields, as accomplished by the preceding theorem, one of the primary utilities of the general modified system is to assess nonlinear parasitic mode stability based on the lowest-order modified system. Thus in the interest of simplicity and practicality, the following theorem directly identifies such a lowest-order system.

**Theorem 3.3.4 (Lowest-Order Vector Field for Parasitic Modified System)**

Consider a multistep variational integrator satisfying the assumptions of Thm. 3.3.3, and suppose that the parasitic modes  $z_{2 \leq i}$  of Eq. (3.35) are initialized with amplitude  $\mathcal{O}(h)$ . The lowest-order vector field  $X_{i0}(z_1, \dots, z_p)$  for the parasitic mode  $z_i$  with  $i \geq 2$  is specified by:

$$\begin{aligned} & \left( D_2 D_1 D_2 L_d^\vartheta + \lambda_i D_2 D_2 D_2 L_d^\vartheta + \lambda_i D_1 D_1 D_1 L_d^\vartheta + 2\lambda_i D_2 D_1 D_1 L_d^\vartheta + \right. \\ & \quad \left. \lambda_i^2 D_1 D_2 D_1 L_d^\vartheta + 2\lambda_i^2 D_2 D_2 D_1 L_d^\vartheta \right)_{\alpha\beta\gamma} z_i^\beta X_H^\gamma(z_1) + \\ & \left( D_1 D_2 L_d^H + \lambda_i D_2 D_2 L_d^H + \lambda_i D_1 D_1 L_d^H + \lambda_i^2 D_2 D_1 L_d^H \right)_{\alpha\beta} z_i^\beta + \\ & \lambda_i \left( D_2 D_2 L_d^\vartheta + D_1 D_1 L_d^\vartheta + 2\lambda_i D_2 D_1 L_d^\vartheta \right)_{\alpha\beta} X_{i0}^\beta(z_1, z_i) = 0 \quad , \end{aligned} \quad (3.65)$$

where all derivatives of  $L_d^\vartheta, L_d^H$  are evaluated at the principal mode  $(z_1(t), z_1(t))$ . For  $i > p$ ,  $z_i$  has zero contribution at first order in  $h$ . If  $D_2 D_2 L_d^\vartheta(z_1, z_1) + D_1 D_1 L_d^\vartheta(z_1, z_1) + 2\lambda_i D_2 D_1 L_d^\vartheta(z_1, z_1)$  is invertible, then  $X_{i0}$  is uniquely defined.

**Proof** The proof of this Theorem proceeds similarly to that of Theorem 3.3.3, with one notable exception: the solution to the general modified system  $\tilde{z}(t) = z_1(t) + \sum_{i \in \mathcal{I}^*} \lambda_i^{t/h} z_i(t)$  is expanded about the principal mode  $z_1(t)$ . Collecting terms by common factors of  $\lambda_i$  provides the additional conditions necessary to solve for the vector fields in the modified system for the parasitic modes.

Refreshing the definitions of the general modified system and restricting attention to the lowest-order vector fields, the solution ansatz is:

$$\tilde{z}(t) = z_1(t) + \sum_{i \in \mathcal{I}^*} \lambda_i^{t/h} z_i(t),$$

with the (lowest-order) time evolution given by:

$$\begin{aligned}\dot{z}_1 &= X_H(z_1) \\ \dot{z}_i &= X_{i0}(z_1, \dots, z_p) \quad i = 2, \dots, p \\ z_i &= hX_{i0}(z_1, \dots, z_p) \quad i > p.\end{aligned}$$

The vector fields are to be chosen so that the solution  $\tilde{z}$  satisfies the discrete Euler-Lagrange equations, as specified in Eq. (3.36). In terms of the principal and parasitic modes, this condition becomes:

$$\begin{aligned}D_2 L_d \left( z_1(t) + \sum_{i \in \mathcal{I}^*} \lambda_i^{t/h} z_i(t), z_1(t+h) + \sum_{i \in \mathcal{I}^*} \lambda_i^{(t+h)/h} z_i(t+h) \right) + \\ D_1 L_d \left( z_1(t+h) + \sum_{i \in \mathcal{I}^*} \lambda_i^{(t+h)/h} z_i(t+h), z_1(t+2h) + \sum_{i \in \mathcal{I}^*} \lambda_i^{(t+2h)/h} z_i(t+2h) \right) = 0. \quad (3.66)\end{aligned}$$

Next, recall that the parasitic modes are  $\mathcal{O}(h)$  and expand about  $z_1(t)$ :

$$\begin{aligned}D_2 L_d(z_1(t), z_1(t+h)) + D_1 L_d(z_1(t+h), z_1(t+2h)) + \\ \left( \sum_{i \in \mathcal{I}^*} \lambda_i^{t/h} z_i(t) \right) \cdot D_1 D_2 L_d(z_1(t), z_1(t+h)) + \\ \left( \sum_{i \in \mathcal{I}^*} \lambda_i^{(t+h)/h} z_i(t+h) \right) \cdot (D_2 D_2 L_d(z_1(t), z_1(t+h)) + D_1 D_1 L_d(z_1(t+h), z_1(t+2h))) + \\ \left( \sum_{i \in \mathcal{I}^*} \lambda_i^{(t+2)/h} z_i(t+2h) \right) \cdot D_2 D_1 L_d(z_1(t+h), z_1(t+2h)) = 0. \quad (3.67)\end{aligned}$$

The tensor contraction is over the outermost slot derivative, where applicable. Extracting common factors of  $\lambda_i^{t/h}$  yields a term independent of  $\lambda_i$ , which is used to establish consistency (proceeding identically to the smooth modified system at lowest order), and the condition

that:

$$\begin{aligned} \lambda_i^{t/h} & \left( z_i(t) \cdot D_1 D_2 L_d(z_1(t), z_1(t+h)) + \lambda_i z_i(t+h) \cdot D_2 D_2 L_d(z_1(t), z_1(t+h)) + \right. \\ & \left. \lambda_i z_i(t+h) \cdot D_1 D_1 L_d(z_1(t+h), z_1(t+2h)) + \right. \\ & \left. \lambda_i^2 z_i(t+2h) \cdot D_2 D_1 L_d(z_1(t+h), z_1(t+2h)) \right) = 0, \end{aligned} \quad (3.68)$$

for all  $t$  and each  $i \in \mathcal{I}^*$ . Because  $\lambda_i^{t/h}$  is not zero, the expression multiplying it must be equal to zero.

To collect in powers of  $h$ , recall that  $L_d$  contributes an  $\mathcal{O}(h^0)$  and an  $\mathcal{O}(h^1)$  term and expand about  $h = 0$ . After Taylor expanding and collecting terms, a condition is identified at  $\mathcal{O}(h^1)$  and at  $\mathcal{O}(h^2)$ . The first condition requires:

$$z_i(t) \cdot \left( D_1 D_2 L_d^\vartheta(z_1(t), z_1(t)) + \lambda_i D_2 D_2 L_d^\vartheta(z_1(t), z_1(t)) + \lambda_i D_1 D_1 L_d^\vartheta(z_1(t), z_1(t)) + \right. \\ \left. \lambda_i^2 D_2 D_1 L_d^\vartheta(z_1(t), z_1(t)) \right) = 0. \quad (3.69)$$

For this condition, there are two cases to consider: (i)  $\lambda_i$  is a root of the characteristic equation (ii)  $\lambda_i$  is not a root of the characteristic equation (but instead a finite product of roots). In the event that  $\lambda_i$  is a root of the characteristic equation, this expression is identically zero (recall that  $L_d(z, z, 0)$  in Eq. (3.14) is  $L_d^\vartheta(z, z)$  for the phase-space discrete Lagrangian, so the above condition *is* the characteristic equation), yielding no information about  $z_i(t)$ . If not, i.e.  $i > p$ , then this condition requires the  $\mathcal{O}(h^1)$  contribution of  $z_i$  to be zero.

The condition at  $\mathcal{O}(h^2)$  specifies:

$$\begin{aligned}
 & \left( z_i(t), X_H(z_1(t)) \right) \cdot \left( D_2 D_1 D_2 L_d^\vartheta(z_1(t), z_1(t)) + \lambda_i D_2 D_2 D_2 L_d^\vartheta(z_1(t), z_1(t)) + \right. \\
 & \quad \lambda_i D_1 D_1 D_1 L_d^\vartheta(z_1(t), z_1(t)) + 2\lambda_i D_2 D_1 D_1 L_d^\vartheta(z_1(t), z_1(t)) + \lambda_i^2 D_1 D_2 D_1 L_d^\vartheta(z_1(t), z_1(t)) + \\
 & \quad \left. 2\lambda_i^2 D_2 D_2 D_1 L_d^\vartheta(z_1(t), z_1(t)) \right) + \\
 & z_i(t) \cdot \left( D_1 D_2 L_d^H(z_1(t), z_1(t)) + \lambda_i D_2 D_2 L_d^H(z_1(t), z_1(t)) + \lambda_i D_1 D_1 L_d^H(z_1(t), z_1(t)) + \right. \\
 & \quad \left. \lambda_i^2 D_2 D_1 L_d^H(z_1(t), z_1(t)) \right) + \\
 & \lambda_i X_{10}(z_1(t), z_i(t)) \cdot \left( D_2 D_2 L_d^\vartheta(z_1(t), z_1(t)) + D_1 D_1 L_d^\vartheta(z_1(t), z_1(t)) + \right. \\
 & \quad \left. 2\lambda_i D_2 D_1 L_d^\vartheta(z_1(t), z_1(t)) \right) = 0. \tag{3.70}
 \end{aligned}$$

This represents the defining equation of the proof, where index notation has been used to clarify the tensor contractions <sup>3</sup>. ■

Now that these two theorems have been established, the ensuing examples illustrate their utility and provide pedagogical verifications for properly applying the newly defined relations. The first two examples consider the smooth modified system, while the latter two consider the time evolution of parasitic modes. First, the canonical trapezoidal discrete Lagrangian is revisited; explicit expressions for the smooth modified vector field are produced for a linear oscillator, and the phase portrait of the original and modified system contrasted for a nonlinear pendulum. Next, the trapezoidal variational integrator for the guiding center trapped particle is studied, demonstrating the principal mode is absent of the previously observed even-odd splitting. Third, Theorem 3.3.4 is shown to recover a known result for a

---

<sup>3</sup>The invertibility condition for the uniqueness of  $X_{i0}$  is tantalizingly suggestive of an intuitive interpretation. If the system were scalar, i.e.  $\dim(z) = 1$ , then this condition is equivalent to  $\lambda_i$  being a simple root of the characteristic equation, but it's not apparent that this (or some higher-dimensional analog, like repeated eigenvalues possessing distinct eigenvectors) is true more generally. Certainly, the regularity of  $L_d$  requires  $D_2 D_1 L_d^\vartheta$  to be invertible, but in combination with the other terms, it's not clear when to expect this to hold.

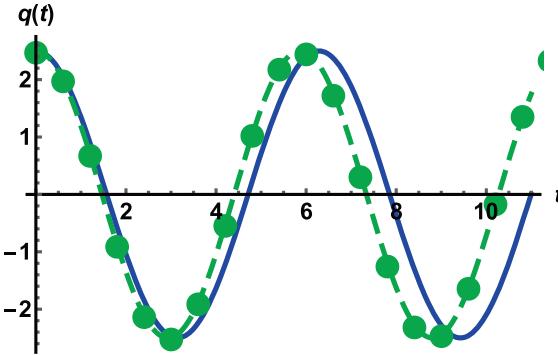


Figure 3.7: Trajectories of the linear oscillator problem modeled with a trapezoidal discretization of the phase-space Lagrangian. **Blue:** Exact solution. **Dashed Green:** Solution of smooth modified system up to  $\mathcal{O}(h^6)$ . **Green Points:** Numerical trajectory.

linear multistep method. Finally, the ability to estimate the nonlinear instability of parasitic modes is demonstrated for test particle multistep variational integrators.

**Example** Consider the trapezoidal discretization of a canonical phase-space Lagrangian, as specified in Eq. (3.21). It has already been shown that the discrete Euler-Lagrange equations for this discretization are the explicit midpoint rule (see Eq. (3.25)).

First, consider the harmonic oscillator problem, with Hamiltonian  $H(q, p) = \frac{1}{2}(q^2 + p^2)$ .

Applying the smooth modified vector field definition of Eq. (3.41) yields:

$$\tilde{X} = X_H + \frac{h^2}{6}X_H + \frac{3h^4}{40}X_H + \mathcal{O}(h^6), \quad (3.71)$$

where  $X_H = (p, -q)^T$ . All  $X_i$  with  $i$  odd are zero by the symmetry of the discrete Euler-Lagrange equations for this system. Because the coefficient functions in the smooth modified vector field simply scale the original vector field by an  $h$ -dependent constant, the numerical trajectory will generate the correct phase portrait but oscillate at an incorrect frequency. Figure 3.7 demonstrates that the flow of the smooth modified system passes through the explicit midpoint numerical trajectory.

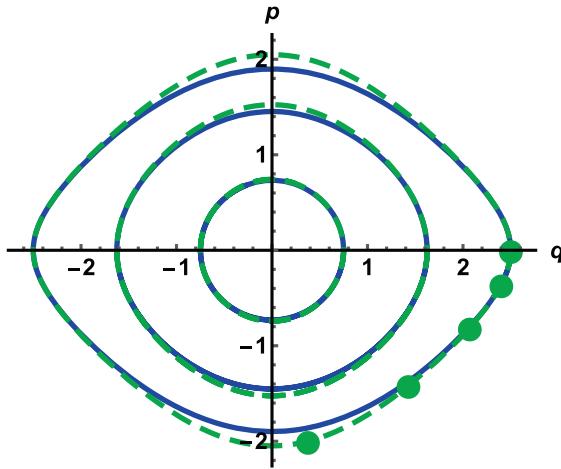


Figure 3.8: Phase portrait of the nonlinear pendulum problem modeled with a trapezoidal discretization of the phase-space Lagrangian. **Blue:** Exact solution. **Dashed Green:** Solution of smooth modified system up to  $\mathcal{O}(h^6)$ . **Green Points:** First few steps of numerical trajectory.

As a modification, consider the nonlinear pendulum Hamiltonian  $H(q, p) = \frac{p^2}{2} - \cos(q)$ .

The smooth modified vector field for this system is given by:

$$\begin{aligned} \tilde{X} = X_H + \frac{h^2}{6} & \left( \begin{array}{c} p \cos(q) \\ -\sin(q) (p^2 + \cos(q)) \end{array} \right) + \\ \frac{h^4}{360} & \left( \begin{array}{c} p (17p^2 \cos(q) + 27 \cos^2(q) - 51 \sin^2(q)) \\ -\sin(q) (7p^4 + 117p^2 \cos(q) + 27 \cos^2(q) - 21 \sin^2(q)) \end{array} \right) + \mathcal{O}(h^6), \end{aligned} \quad (3.72)$$

where  $X_H = (p, -\sin(q))$ . Figure 3.8 contrasts the phase portrait of the smooth modified system and the exact solution. The oscillatory nature of the numerical energy may be observed from this, as the smooth modified system trajectories interweave the constant energy surfaces of the exact solution.

**Example** Recovering a known result is an important means of verifying a new proposition. In the context of Theorem 3.3.4, one can analyze a discrete Lagrangian that yields a linear multistep method, then compare the result to that of the literature [6]. In particular, this

example verifies that Eq. (3.65) applied to a trapezoidal discretization of a canonical phase-space Lagrangian agrees with a known result for linear multistep methods.

For linear multistep methods, it is known that the lowest-order modified system for parasitic modes is given by [6]:

$$\dot{z}_i = \mu_i X'(z_1) z_i, \quad (3.73)$$

where  $X'$  is the derivative of the original vector field and the “growth parameter”  $\mu_i$  is given by:

$$\mu_i = \frac{\sum_{k=0}^p b_k \lambda_i^k}{\sum_{k=0}^p a_k k \lambda_i^k}. \quad (3.74)$$

As shown in the introductory example, a trapezoidal discretization of a canonical phase-space Lagrangian yields the explicit midpoint algorithm, which is a linear multistep method. One then expects Theorem 3.3.4 to specify, for the single parasitic mode  $\lambda_2 = -1$ :

$$\dot{z}_2 = -X'_H(z_1) z_2. \quad (3.75)$$

To apply Eq. (3.65) to the trapezoidal discrete Lagrangian Eq. (3.21), take note that all third-order derivatives of  $L_d^\vartheta$  are zero, and:

$$\left( D_1 D_2 L_d^H - (D_2 D_2 L_d^H + D_1 D_1 L_d^H) + D_2 D_1 L_d^H \right)(z, z) = \begin{pmatrix} H_{,qq}(z) & H_{,pq}(z) \\ H_{,qp}(z) & H_{,pp}(z) \end{pmatrix} \quad (3.76a)$$

$$\left( - (D_2 D_2 L_d^\vartheta + D_1 D_1 L_d^\vartheta) + 2 D_2 D_1 L_d^\vartheta \right)(z, z) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (3.76b)$$

Putting these together, Eq. (3.65) yields:

$$X_{20} = \begin{pmatrix} -H_{,pq}(z_1) z_2^1 - H_{,pp}(z_1) z_2^2 \\ H_{,qq}(z_1) z_2^1 + H_{,qp}(z_1) z_2^2 \end{pmatrix}, \quad (3.77)$$

which is in agreement with the expected result.

## 3.4 Structure Preservation

The presence of numerical instabilities that can increase without bound is an unexpected feature of multistep variational integrators, considering the purpose of the discrete variational principle is to constrain the admissible behavior through the retention of conservation laws. For conventional symplectic integrators, and equivalently, conventional variational integrators, the conserved symplectic structure bounds the energy error to a small value for exponentially long simulation times [6]. Although many numerical demonstrations of multistep variational integrators do exhibit excellent long-term behavior [13, 14, 15, 16, 96], the good long-term behavior is not guaranteed for all parameter regimes of a given discrete Lagrangian; alterations to the initial conditions or governing potentials may well excite the parasitic mode instabilities.

To explain the presence of large amplitude parasitic modes and refine expectations of the performance of multistep variational integrators, this Section scrutinizes the structure-preservation properties of multistep variational integrators. It is shown that the discrete symplectic structure preserved by the multistep variational integrator has *twice* the dimension of the continuous symplectic structure, so rigorous results regarding the behavior of conventional symplectic integrators cannot be expected to hold for multistep variational integrators. On a more positive note, good behavior *can* be expected of the principal mode. This follows from analyzing the symplecticity of the “underlying one-step method” of a multistep variational integrator. It is shown that the underlying one-step method preserves a discrete two-form on the original Hamiltonian phase-space, thus the non-parasitic component of the numerical solution is subject to a more restrictive constraint than that of an arbitrary trajectory satisfying the multistep update rule. The similarity between the dis-

crete and continuous two-forms suggests that good behavior can be anticipated whenever the parasitic components of the numerical trajectory remain small. More rigorous statements may be enabled by this result, for instance if it can be leveraged to demonstrate conjugate symplecticity of the underlying one-step method.

### 3.4.1 Multistep Advance

Multistep variational integrators are, by construction, structure-preserving. In fact, the multistep variational integrators emphasized in this chapter derive from regular discrete Lagrangians, so all of the standard structure-preserving proofs for variational integrators hold [9]. The admittance of poor behavior originates in the degeneracy of the *continuous* Lagrangian. The continuous dynamics evolve on a *lower* dimensional space than if the Lagrangian were regular. Because the discrete Lagrangian is ignorant of, or fails to capture, this degeneracy, the structure-preserving properties of the multistep advance are insufficiently constraining to obtain the desired behavior. This section analyzes this breakdown of the correspondence between the continuous and discrete descriptions of the dynamics.

The discrepancy between the continuous and discrete descriptions can be explained from two perspectives. In the background presented in Chapter 2, it was shown that the time advance of a regular variational integrator could be formulated as a two-step map on the “configuration space” (according to the discrete Euler-Lagrange equations), or as a one-step map on the cotangent bundle of the configuration space (according to the position-momentum formulation of the Euler-Lagrange equations). These two complementary descriptions of the time advance yield complementary analyses for the mismatch between the continuous and discrete dynamics. Analyzing the two-step discrete description first, the following theorem identifies a consequential discrepancy between the discrete and continuous symplectic structures.

**Theorem 3.4.1 (The MVI is Symplectic on  $M \times M$ )** Consider a multistep variational integrator satisfying the conditions of Theorem 3.2.1. The flow of the continuous Euler-Lagrange equations preserves a symplectic structure  $\Omega$  on  $M$  while the discrete Lagrangian map of the multistep variational integrator preserves a discrete symplectic structure  $\Omega_{L_d}$  on  $M \times M$ . That is, the discrete two-form resides on a space of twice the dimension as the continuous two-form.

**Proof** By definition of a “phase-space Lagrangian”, the continuous Euler-Lagrange equations specify a Hamiltonian vector field  $X_H$  whose flow preserves the symplectic structure  $\Omega = -d\vartheta$  on  $M$  (see Theorem 2.3.2). Recall, in coordinates:

$$\Omega = \vartheta_{i,j} dz^i \wedge dz^j.$$

By regularity of the discrete Lagrangian, the discrete Euler-Lagrange equations specify a map that preserves the symplectic structure  $\Omega_{L_d} = d_1 d_2 L_d$  on  $M \times M$  (see Theorem 2.2.6). Recall, in coordinates:

$$\Omega_{L_d} = \frac{\partial^2 L_d}{\partial z_0^i \partial z_1^j} dz_0^i \wedge dz_1^j.$$

Comparing these two previously established results verifies the claim of the theorem. ■

The mismatch between a two-step discrete description and a first-order continuous description incurs a mismatch in the dimensionality of the corresponding two-forms that are preserved. The prospect of such a mismatch introduces nuance to what should be expected of a “symplectic integrator”, or alternatively to what should be considered a “symplectic integrator”. Typically a “symplectic integrator” is understood to preserve *the same* two-form as the continuous flow map; the desirable features such as bounded energy error emerge from this property. In the numerical analysis literature, different notions of symplecticity include “discretely symplectic” [9] and “g-symplectic”, both implying a weaker sense of

structure preservation than conventional symplectic integrators. The overall consequence of this mismatch is summarized by the following Corollary.

**Corollary 3.4.2** *The rigorous expectations for symplectic integrators do not hold for multistep variational integrators.*

**Proof** Proofs of the conservation properties of symplectic integrators require that the integrator preserve the same two-form as the continuous system [6]. The Theorem explains that this is not the case for multistep variational integrators, which should be instead considered “discretely symplectic”. ■

From the perspective of the two-step discrete Lagrange map, the admittance of undesirable behavior may be attributed to a mismatch between the continuous and discrete two-forms. Even in the conventional setting wherein both the continuous and discrete Lagrangians are regular, however, the discrete and continuous two-forms are not equivalent in from the configuration-space perspective (although they do possess the same dimension, unlike above). The desirable numerical properties of regular/regular variational integrators are established by comparing the phase-space dynamics, in which the variational integrator identifies a symplectic integrator (in the conventional sense). A natural question is then: How does the one-step formulation of a multistep variational integrator compare to its continuous counterpart?

The next analysis answers this question by scrutinizing the one-step formulation of the multistep variational integrator prescribed by the position-momentum form of the discrete Euler-Lagrange equations; it follows the work of Ref. [92]. Of course, regularity of the discrete Lagrangian ensures this one-step map is canonically symplectic on  $T^*M$ . Again, the shortfall will be in the continuous dynamics being “restricted” by the degeneracy in a way that is not captured by the discrete model. Specifically, because the continuous Lagrangian is degenerate, the Legendre transform is not invertible, and the continuous description of the

dynamics on  $T^*M$  forms a differential-algebraic set of equations. The differential-algebraic equations dictate that the continuous dynamics evolve on a sub-manifold of this space; the discrete dynamics are subject to no such constraint.

**Theorem 3.4.3** *Consider a multistep variational integrator satisfying the conditions of Theorem 3.2.1 and  $\dim M = n$ . The continuous dynamics evolve on an  $n$ -dimensional submanifold of  $T^*M$  while the discrete dynamics are not restricted to this submanifold.*

**Proof** The first task is to reformulate the continuous Euler-Lagrange equations as a system on  $T^*M$ . Recall that the Euler-Lagrange equations for a phase-space Lagrangian are given by:

$$(\vartheta_{i,j}(z) - \vartheta_{j,i}(z)) \dot{z}^i - H_{,j}(z) = 0.$$

Choose coordinates  $(z^i, p_i)$ ,  $i = 1, \dots, n$  on  $T^*M$ . The Legendre transform specifies:

$$p_i = \vartheta_i(z). \quad (3.78)$$

This may be used to reformulate the Euler-Lagrange equations as the “implicit Hamiltonian system”:

$$p_j = \vartheta_j(z) \quad (3.79a)$$

$$\dot{p}_j = \vartheta_{i,j}(z) \dot{z}^i - H_{,j}(z). \quad (3.79b)$$

The equivalence with the Euler-Lagrange equations is readily verified by replacing  $\dot{p}$  in the second equation with the definition in the first equation.

Equation (3.79a) demonstrates that the dynamics do *not* evolve on all of  $T^*M$ , but instead the “momentum”  $p(t)$  at any instant is uniquely defined by the “position” at that instant  $z(t)$ . Mathematically, the dynamics are specified to evolve on the *primary constraint*

submanifold  $N = \mathbb{F}L(TM) \subsetneq T^*M$ , and the primary constraint specified by the Legendre transform shows that  $\dim(N) = \dim(M) = n$ .

Turning to the discrete dynamics, the regularity of the discrete Lagrangian indicates that for *any*  $(z_k, p_k) \in T^*M$ , a corresponding  $(z_{k+1}, p_{k+1})$  may be identified according to the position-momentum formulation of the discrete Euler-Lagrange equations:

$$\begin{aligned} p_k &= -D_1 L_d(z_k, z_{k+1}) \\ p_{k+1} &= D_2 L_d(z_k, z_{k+1}). \end{aligned}$$

The multistep variational integrator therefore accepts solution with  $(z_k, p_k) \notin N$ . ■

Even when formulating the multistep variational integrator as a one-step method, the numerical dynamics evolve in a larger space than the continuous dynamics. Instead of inquiring whether the multistep method *allows* numerical trajectories outside of the primary constraint submanifold  $N$ , one might ask whether a numerical trajectory that *starts* on the primary constraint manifold remains there (or perhaps nearby). This is similar to investigating the stability of parasitic modes as opposed to their existence. Later, in Section 3.6.3, methods that remain on the submanifold will be sought, but will only be successfully identified for a simple class of phase-space Lagrangians. For now, the preceding theorems explain why the existence of unstable parasitic modes does not contradict the structure-preservation of multistep variational integrators.

### 3.4.2 Underlying One-Step Method

In light of the excessively large range of behavior permitted by multistep variational integrators, it remains to be explained why good behavior is *ever* observed by multistep variational integrators. Between the preceding section and the present, the question has transitioned from “How can instabilities arise in integrators preserving a discrete two-form?” to “Given

the discrete two-form does not imply familiar behavior of symplectic integrators, what is the cause for the often observed long-term fidelity?”. This latter question can be addressed by adapting another concept from the multistep integration literature to multistep variational integrators: the notion of an *underlying one-step method* [93, 94, 12, 6].

Intuitively, the underlying one-step method refers to the principal mode of the dynamics, explaining that the principal mode of the multistep method essentially behaves as if it were generated by a one-step method. More formally:

**Definition** The *underlying one-step method* of a multistep method is an  $h$ -dependent map  $\hat{F} : M \rightarrow M$  given as a formal power series in the step size  $h$ :

$$\hat{F}(z) = z + h\hat{F}_1(z) + h^2\hat{F}_2(z) + \dots, \quad (3.80)$$

such that:

$$f_d(z, \hat{F}(z), \hat{F} \circ \hat{F}(z), \dots, \hat{F}^p(z)) = 0. \quad (3.81)$$

The definition in terms of a “formal power series” indicates that convergence of this series is not guaranteed. Instead, one seeks to show that the coefficient functions  $\hat{F}_i$  exist and are unique according to this definition. Although one can construct examples where the radius of convergence is zero [6], utility of the underlying one-step method can emerge from considering truncations of the expansion. A critical property of the underlying one-step method is that the  $h^0$  coefficient is  $z$ , i.e. the identity map. This is an important distinction from other maps that may satisfy Eq. (3.81).

In relation to the analysis of Section 3.3.2, the underlying one-step method bears similarity to - yet is distinct from - the flow of the smooth modified system. Both concepts ignore the prospect of parasitic modes, recover the identity map when  $h \rightarrow 0$ , and generate trajectories that satisfy the multistep update rule. The distinction between the time  $h$  flow map of the smooth modified system and the underlying one-step method is one of *convergence*.

When performing a backward error analysis, it is the modified vector field that is expanded as a formal power series in  $h$ . In contrast, the underlying one-step method directly constructs the map as a formal power series in  $h$ . Both series are in general asymptotic, and must be truncated at some optimal order to minimize error [6]. Because neither series is required to converge, one cannot simply state that the underlying one-step method *is* the flow of the smooth modified vector field, in general. The intuition remains similar, however, so in practice one chooses a description based on which will achieve minimum error or whether it is the map versus time dynamics that are of interest.

For the purposes of this section, it is of interest to identify what symplectic two-form is preserved by the underlying one-step method, if any. Because the underlying one-step method precludes parasitic mode behavior and matches the dimensionality of the true solution, one might hope that the underlying one-step method is symplectic in the conventional sense (discrete two-form matches the continuous two-form). In fact, a rather disparaging result is known about underlying one-step methods: the underlying one-step method cannot be symplectic (in the conventional sense) for linear and one-leg multistep methods [95, 122, 6]. As demonstrated in the very first example of this Chapter, it is possible to recover a consistent linear multistep method as a multistep variational integrator, so it cannot be the case that the underlying one-step method of multistep variational integrators is (conventionally) symplectic, in general. Further, it has been conjectured that a wide class of multistep methods cannot possess symplectic underlying one-step methods [6]. At first approach, this bodes poorly for what one might hope to recover regarding the symplecticity of the underlying one-step methods of multistep variational integrators.

More promisingly, the underlying one-step method need not conserve the *same* discrete two-form as the continuous method for the principal mode to exhibit excellent long-term numerical fidelity. It has been shown, for instance, that symmetric linear multistep methods possess a property known as *conjugate-symplecticity* [112]. That is, they may be shown to be

equivalent to a symplectic integrator upon coordinate transformation. Because the dynamics are equivalent whether they were generated using the canonically symplectic advance or the conjugate-symplectic advance, conjugate-symplecticity may be considered equally as good as the more conventional notion of symplecticity. This additional freedom in the discrete two-form should be kept in mind as the investigation of multistep variational integrators proceeds.

Once again, the theory of multistep variational integrators will be facilitated by the discrete variational principle. The structure-preserving properties of the underlying one-step method are specified by the following theorem:

**Theorem 3.4.4** *The underlying one-step method  $\hat{F}_{L_d}$  of a multistep variational integrator (satisfying the conditions of Theorem 3.2.1) is symplectic with respect to discrete two-form  $\hat{\Omega}_{L_d}$  on  $M$  given by:*

$$\hat{\Omega}_{L_d} = (id, \hat{F}_{L_d})^* \Omega_{L_d}, \quad (3.82)$$

or in coordinates:

$$\hat{\Omega}_{L_d}(z) = \left( D_1 D_2 L_d(z, \hat{F}_{L_d}(z)) \right)_{ij} \left( D \hat{F}_{L_d}(z) \right)_k^j dz^i \wedge dz^k. \quad (3.83)$$

**Proof** By the definition of the underlying one-step method and the multistep variational integrator,  $\hat{F}_{L_d}$  satisfies:

$$D_2 L_d(z, \hat{F}_{L_d}(z)) + D_1 L_d(\hat{F}_{L_d}(z), \hat{F}_{L_d} \circ \hat{F}_{L_d}(z)) = 0. \quad (3.84)$$

The calculation of a symplectic structure preserved by  $\hat{F}_{L_d}$  proceeds in an identical manner as the calculation of the symplectic structure preserved by any variational integrator. That is, begin by considering a discrete action restricted to act on trajectories generated by

the map of interest - this time, underlying one-step method:

$$\hat{S}_d(z_0) = \sum_{k=0}^{N-1} L_d(\hat{F}_{L_d}^k(z_0), \hat{F}_{L_d}^{k+1}(z_0)), \quad (3.85)$$

where the exponent of  $\hat{F}_{L_d}$  denotes composition of the function with itself the corresponding number of times. Next, recalling that the internal terms are zero by the discrete Euler-Lagrange equations, operating with an exterior derivative yields:

$$d\hat{S}_d(z_0) = d_1 L_d(z_0, \hat{F}_{L_d}(z_0)) + d_2 L_d(\hat{F}_{L_d}^{N-1}(z_0), \hat{F}_{L_d}^N(z_0)). \quad (3.86)$$

A second application of an exterior derivative obtains:

$$\begin{aligned} d^2 \hat{S}_d &= 0 \Rightarrow \\ d_1 d_2 L_d(z_0, \hat{F}_{L_d}(z_0)) &= d_1 d_2 L_d(\hat{F}_{L_d}^{N-1}(z_0), \hat{F}_{L_d}^N(z_0)), \end{aligned} \quad (3.87)$$

or in terms of the pullback:

$$d_1 d_2 L_d(z_0, \hat{F}_{L_d}(z_0)) = (\hat{F}_{L_d}^{N-1})^* d_1 d_2 L_d(z_0, \hat{F}_{L_d}(z_0)). \quad (3.88)$$

Therefore,  $\hat{F}_{L_d}$  is symplectic with respect to  $\hat{\Omega}_{L_d}$  as defined in Eq. (3.83). Comparison with the multistep symplectic structure  $\Omega_{L_d}$  yields Eq. (3.82). ■

This Theorem reveals that although the underlying one-step method does not preserve the *same* symplectic structure as the flow of the continuous vector field, it does preserve *a* symplectic structure on the same space. An important question - that will unfortunately remain unanswered, for now - is what can rigorously be claimed about the trajectory of the underlying one-step method based on this discrete symplecticity. One possibility is that this discrete symplecticity may be leveraged to prove conjugate symplecticity of the underlying

one-step method. Such a result would be a satisfying explanation of the observed long-term behavior whenever the parasitic modes remain small. For now, I will *conjecture* that the excellent long-term behavior of the principal component of multistep variational integrators may be attributed to this discrete symplecticity, but rigorous statements remain as future work.

## 3.5 Accuracy

A recurring theme of this Chapter is that degeneracy of the continuous Lagrangian breaks conventional results on the correspondence between variational integrators and the dynamical systems they model. One important implication of a regular continuous Lagrangian is the equality in the order of accuracy of the discrete Lagrangian and that of the discrete Hamiltonian map (see Thm. 2.2.9). Benefits of this equality include simplifying error analysis (only the truncation error of the discrete Lagrangian needs to be investigated) and guiding the construction of higher-order integrators (by approximating the “exact discrete Lagrangian” Eq. (2.55) to higher order). As will be shown in this Section (and has been observed in Ref. [92]), the same statement on accuracy *cannot* be made for multistep variational integrators; the MVI sometimes has accuracy *less* than that of the discrete Lagrangian. For example, a discretization method that yields an  $\mathcal{O}(h^6)$ -accurate variational integrator for a regular Lagrangian system might only yield an  $\mathcal{O}(h^4)$ -accurate variational integrator for a phase-space Lagrangian system. This shortfall effectively encourages the use of low-order multistep variational integrators, assuming the computational cost is fixed. Obtaining additional orders of accuracy requires *substantially* more computational effort, so it is generally preferable to use low-order multistep variational integrators with a small step sizes.

In discussing the accuracy of multistep variational integrators, a preliminary concept requiring revision to account for degenerate Lagrangians is the very *definition* of the order

of a discrete Lagrangian [92]. Section 3.2 showed how to test whether a *given* discrete Lagrangian would yield a multistep method or whether a *given* multistep method might possess a variational formulation, but it bypassed consideration of how one might *construct* a multistep variational integrator, given a continuous Lagrangian. Both the order of a discrete Lagrangian and, correspondingly, the systematic construction of a variational integrator to a specified order of accuracy refer to an “exact discrete Lagrangian”. The exact discrete Lagrangian of Eq. (2.55) requires the continuous Lagrangian to be regular in order to be well-defined, so it is at this point that modifications to the accuracy of multistep variational integrators will begin.

Recall from Eq. (2.55) that, if the continuous Lagrangian is regular, the exact discrete Lagrangian is given by:

$$L_d^E(q_0, q_1) = \int_0^h L(q_{(0,1)}(t), \dot{q}_{(0,1)}(t)) dt,$$

where  $q_{(0,1)}(t)$  is the unique trajectory satisfying the Euler-Lagrange equations with  $q_{(0,1)}(0) = q_0$  and  $q_{(0,1)}(h) = q_1$  [9]. The regularity of the Lagrangian guarantees that such a path exists for sufficiently nearby  $q_0, q_1$  and sufficiently small  $h$ . Because the Euler-Lagrange equations are second-order, each value of  $q_1$  corresponds to a unique initial condition  $q_0, \dot{q}_0$ . Variational integrators may then be constructed by choosing a discrete Lagrangian that approximates this small interval of the action; the order of the discrete Lagrangian is simply the order to which it approximates this interval of the action. If the Lagrangian is degenerate, however, there typically does *not* exist a path connecting  $q_0, q_1$  over a given time interval. For instance, if the Lagrangian is a phase-space Lagrangian, a single initial condition  $z_0$  specifies the trajectory;  $z_1$  is uniquely determined for a given time interval  $h$ . This calls into question what the “order” of a discrete Lagrangian means for degenerate

Lagrangian systems - along *what path* does a particular  $L_d(z_0, z_1)$  approximate the action integral?

Restricting attention to phase-space Lagrangians, a modified definition of the exact discrete Lagrangian has been given in Ref. [92], which will be interpreted for our purposes as:

**Definition** Given a phase-space Lagrangian  $L_{PS} : TM \rightarrow \mathbb{R}$  as defined in Eq. (2.61), the corresponding exact discrete Lagrangian  $L_d^E : M \rightarrow \mathbb{R}$  is given by:

$$L_d^E(z_0) = \int_0^h L_{PS}(z_{(0,1)}(t), \dot{z}_{(0,1)}(t)) dt, \quad (3.89)$$

where  $z_{(0,1)}(t)$  is the unique trajectory satisfying the Euler-Lagrange equations for  $L_{PS}$  originating at  $z_{(0,1)}(0) = z_0$ .

The essential distinction is that the domain of the exact discrete Lagrangian has been reduced to reflect the dimensionality of the initial value problem for the degenerate Lagrangian system.

This definition facilitates the definition of the order of a discrete Lagrangian [92], which may be applied in the context of multistep variational integrators:

**Definition** The discrete Lagrangian for a multistep variational integrator (again, satisfying Theorem 3.2.1) is of *order r* if:

$$L_d(z_0, F_L^h(z_0)) = L_d^E(z_0) + \mathcal{O}(h^r), \quad (3.90)$$

for all  $z_0 \in M$ .

The nuance addressed by this definition is: although the discrete Lagrangian for the multistep variational integrator is defined on a larger space ( $M \times M$ ) than that on which the

exact discrete Lagrangian is defined ( $M$ ), one considers the discrete Lagrangian to be a particular order if it satisfies the order condition *whenever*  $(z_0, z_1)$  are in fact connected by a solution to the phase-space Euler-Lagrange equations. This distinction from the correspondence that exists for regular Lagrangian systems has important implications for accuracy, as demonstrated by the following examples.

**Example - Accuracy Shortfall in MVIs** To demonstrate the shortfall in the order of accuracy of multistep variational integrators, this example will apply discrete Lagrangians of varying orders to the guiding center system. The outcome will be that the order of the guiding center variational integrator is sometimes *less* than the order of the discrete Lagrangian.

The discrete Lagrangians will be constructed according to a systematic procedure known as *Galerkin variational integration*, as described in Ref. [100]. The discretization procedure may be summarized as a two step process: First, one performs an approximation on the *path space*, approximating the trajectory as a finite-degree polynomial in time. For instance, the lowest-order path space approximation is that, between  $z_0$  and  $z_1$ , the trajectory is given by:  $z_0 + \left(\frac{t}{h}\right)(z_1 - z_0)$ . Extending to higher order, one might approximate the path as a quadratic polynomial in time, cubic polynomial, and so on. The second step is to perform a quadrature approximation to the interval of the action approximated by the discrete Lagrangian. That is, the action integral is approximated by a weighted sum of the integrand evaluated at different instances in time. The quadrature methods considered in Ref. [100] and used here include Gaussian quadrature - which obtains the highest order of accuracy for a given number of integrand evaluations - and Gauss-Lobatto quadrature - which obtains the highest order of accuracy for a given number of integrand evaluations *while requiring the endpoints of the interval to be used*. Not surprisingly, the overall order of accuracy of the integrator depends on both the polynomial approximation and the quadrature method. More specifically, representing the path with an  $s$ -degree polynomial in time and approximating the intervals

of the action integral with an order- $r$  quadrature method yields a discrete Lagrangian that is order  $\min(2s, r)$  accurate in the step size  $h$  [100].

To explain the Galerkin variational integrator construction procedure, two representative discrete Lagrangians will be constructed. The more general notation and trends can be extrapolated based on these representative examples. The first Galerkin variational integrator will be the *midpoint discrete Lagrangian*. Given a Lagrangian  $L : TM \rightarrow \mathbb{R}$  (this Lagrangian may be regular or degenerate), the midpoint discrete Lagrangian is given by:

$$L_{d2}(z_0, z_1) = L\left(\frac{z_0 + z_1}{2}, \frac{z_1 - z_0}{h}\right). \quad (3.91)$$

Where  $z$  represents local coordinates on  $M$ . The notation  $L_{d2}$  refers to the second-order accuracy of this discrete Lagrangian. The “Galerkin” construction of  $L_{d2}$  arises from the following choices for the path-space and quadrature approximations. Let the path joining  $(z_0, z_1)$  be approximated by a path that is linear in time:

$$\mathcal{I}_{(z_0, z_1)}(t) = z_0 + \left(\frac{t}{h}\right)(z_1 - z_0). \quad (3.92)$$

Next, choose a midpoint Gaussian quadrature method, which would prescribe:

$$\frac{1}{h} \int_0^h f(t) dt \approx f\left(t = \frac{h}{2}\right). \quad (3.93)$$

Combining the path space approximation of Eq. (3.92) with the quadrature approximation Eq. (3.93) yields:

$$\begin{aligned} \frac{1}{h} \int_0^h L(z(t_k + \tau), \dot{z}(t_k + \tau)) d\tau &\approx \frac{1}{h} \int_0^h L\left(\mathcal{I}_{(z_k, z_{k+1})}(t_k + \tau), \dot{\mathcal{I}}_{(z_k, z_{k+1})}(t_k + \tau)\right) d\tau \\ &\approx L\left(\mathcal{I}_{(z_k, z_{k+1})}(t_k + \frac{h}{2}), \dot{\mathcal{I}}_{(z_k, z_{k+1})}(t_k + \frac{h}{2})\right) \\ &= L\left(\frac{z_k + z_{k+1}}{2}, \frac{z_{k+1} - z_k}{h}\right), \end{aligned} \quad (3.94)$$

showing the Galerkin construction of Eq. (3.91). The second-order accuracy of the midpoint discrete Lagrangian  $L_{d2}$  may be justified based on the symmetry between  $z_0$  and  $z_1$ , based on the condition  $\min(2s, u)$  derived in Ref. [100], or by Taylor expanding the discrete Lagrangian and calculating the truncation error relative to the exact discrete Lagrangian (examples of this are found in Ref. [9]).

The second discrete Lagrangian constructed by the Galerkin procedure will be fourth-order accurate. To achieve greater than second-order accuracy, the degree of the polynomial path approximation needs to be increased. To do so, one must introduce “internal control points” to the discrete Lagrangian; a second-degree polynomial is uniquely specified by three points, not two, so additional information will be required beyond  $z_0, z_1$  in the discrete Lagrangian. At first, this will appear to lead to a different form of the discrete Lagrangian and different discrete Euler-Lagrange equations, however, one may ultimately express everything as a function of only  $z_0$  and  $z_1$  as shown in Ref. [9] and as will be discussed here.

To construct a fourth-order accurate discrete Lagrangian function, consider the second-degree polynomial path space approximation:

$$\mathcal{I}_{(z_0, z_{1/2}, z_1)}(t) = z_0 + \left(\frac{t}{h}\right) (4z_{1/2} - 3z_0 - 2z_1) + \left(\frac{t}{h}\right)^2 (2z_0 - 4z_{1/2} + 2z_1). \quad (3.95)$$

Notice that  $\mathcal{I}_{(z_0, z_{1/2}, z_1)}(t)$  passes through  $z_0, z_{1/2}, z_1$  at times  $t = 0, h/2, h$ , respectively. For the quadrature approximation, a two-point Gaussian quadrature method will be chosen according to:

$$\frac{1}{h} \int_0^h f(t) dt \approx \frac{1}{2} \left( f\left(\frac{h(3 - \sqrt{3})}{6}\right) + f\left(\frac{h(3 + \sqrt{3})}{6}\right) \right). \quad (3.96)$$

This Gaussian quadrature method is fourth-order accurate in  $h$ . Combined, the path space approximation in Eq. (3.95) and quadrature approximation of Eq. (3.96) yield a discrete Lagrangian:

$$\begin{aligned} L_{d4}(z_0, z_{1/2}, z_1) = & \\ & L\left(\frac{z_0(1 + \sqrt{3}) + 4z_{1/2} + z_1(1 - \sqrt{3})}{6}, -\frac{z_0(2 + \sqrt{3}) + 4z_{1/2} + z_1(2 - \sqrt{3})}{\sqrt{3}h}\right) + \\ & L\left(\frac{z_0(1 - \sqrt{3}) + 4z_{1/2} + z_1(1 + \sqrt{3})}{6}, \frac{z_0(2 - \sqrt{3}) - 4z_{1/2} + z_1(2 + \sqrt{3})}{\sqrt{3}h}\right). \end{aligned} \quad (3.97)$$

Because the discrete Lagrangian in Eq. (3.97) depends on three points,  $(z_0, z_{1/2}, z_1)$ , the discrete Euler-Lagrange equations need to be modified to include the variation with respect to  $z_{1/2}$ . The modified discrete Euler-Lagrange equations are given by:

$$D_3 L_{d4}(z_{k-1}, z_{k-1/2}, z_k) + D_1 L_{d4}(z_k, z_{k+1/2}, z_{k+1}) = 0 \quad k = 1, \dots, N-1 \quad (3.98a)$$

$$D_2 L_{d4}(z_k, z_{k+1/2}, z_{k+1}) = 0 \quad k = 0, \dots, N-1. \quad (3.98b)$$

Despite the additional variables present, this variational integrator is still a two-step method. For instance, given initial conditions  $z(t = 0), z(t = h)$ , one may apply Eq. (3.98b) to determine  $z_{1/2}$  from  $z_0, z_1$ . From there, one may iterate Eq. (3.98) to determine the rest of the numerical trajectory. Although an additional variable is present, namely the internal control points at half-integer steps, the additional equation determines these control points.

Another perspective is to consider a discrete Lagrangian  $\tilde{L}_{d4}$  of the form:

$$\tilde{L}_{d4}(z_0, z_1) = L_{d4} \left( z_0, z_{1/2}(z_0, z_1), z_1 \right), \quad (3.99)$$

where  $z_{1/2}(z_0, z_1)$  is determined according to Eq. (3.98b). The standard discrete Euler-Lagrange equations for  $\tilde{L}_{d4}$  yield the same dynamics as the modified discrete Euler-Lagrange Eqs. 3.98b for  $L_{d4}$  [9]. The takeaway is that the internal control point increases the dimensionality of the nonlinear solve at each step (increases the computational expense!) but does not alter the number of steps in the method. The fourth-order of accuracy of  $L_{d4}$  is determined from  $\min(2s, u)$ , with  $s = 2$  the degree of the polynomial and  $u = 4$  the order of the quadrature method.

Many variational integrators may be similarly constructed, varying the degree of the polynomial  $s$  and the quadrature method. A Galerkin variational integrator may be identified using three numbers [100]:  $s$  - the degree of the polynomial describing the path in time;  $r$  - the number of quadrature points; and  $u$  - the order of accuracy of the quadrature method. Discrete Lagrangians and their resulting variational integrators may then be labeled  $PsNrQu$ , with  $s, r, u$  integers (following the convention of Ref. [100]). For this study, a series of variational integrators were constructed for the guiding center system. The accuracy results are summarized in Table 3.1 Fig. 3.9. In Table 3.1, the variational integrator is labeled by the  $PsNrQu$  notation, the expected order of accuracy is shown according to  $\min(2s, u)$ , and the “actual order” of accuracy is deduced by simulating guiding center trajectories with varying step sizes and determining the rate of convergence of the results. The rates of convergence may be inspected in Fig. 3.9.

Method	Expected Order	Actual Order
P1N1Q2	2	2
P1N2Q2	2	2
P1N2Q4	2	2
P1N3Q4	2	2
P2N2Q4	4	2
P2N3Q6	4	2
P2N4Q6	4	2
P3N3Q6	6	4
P3N4Q6	6	4

Table 3.1: Multistep variational integrators can be less accurate than the order of their discrete Lagrangians. A series of Galerkin variational integrators were applied to the guiding center system; the expected order of accuracy (equal to the order of accuracy of the discrete Lagrangian) is shown in the middle column, while a numerically determined convergence rate is shown in the right column.

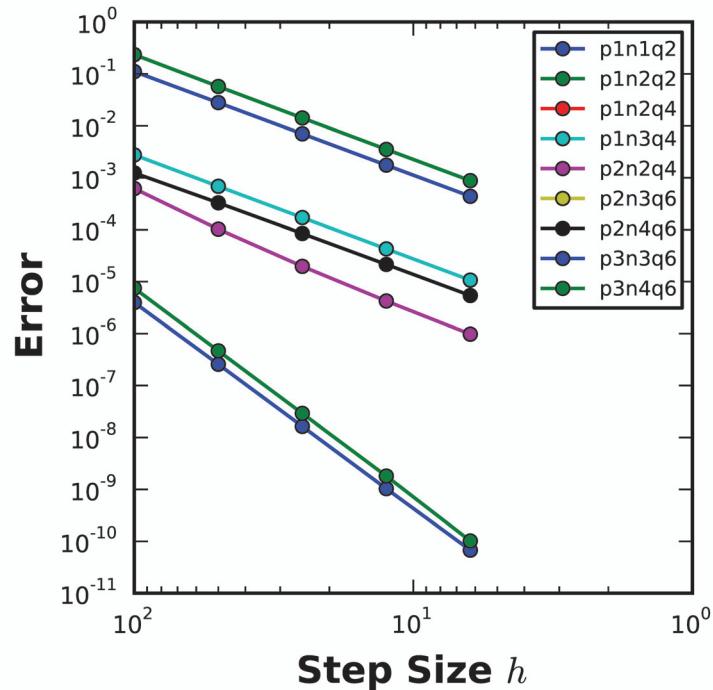


Figure 3.9: Data used to determine the rates of convergence of various Galerkin variational integrators. See also Table 3.1.

## 3.6 Parasitic Mode Mitigation

The parasitic mode instabilities must be addressed if multistep variational integrators are to achieve the desired long-term numerical fidelity. Because the modes are marginally stable, the identification of a strictly-stable multistep variational integrator is not possible. However, the marginally stable modes often grow quite slowly, if at all, so minor improvements to their stability may be sufficient to obtain desirable behavior over some desired simulation time. The first tactic for mitigating parasitic modes is then to initialize the modes with small amplitude [12, 6, 114]. This may be accomplished by a *backward error initialization* procedure, in which determination of the smooth modified system allows initializing the mode to arbitrarily small initial amplitude (for sufficiently small  $h$ ). The second tactic is to apply a smoothing procedure, either periodically or upon active identification of large parasitic mode amplitude [120, 121]. Although the smoothing step is dissipative, if applied sufficiently rarely the numerical dissipation introduced will be quite small. Finally, the third tactic is to identify the underlying one-step method in such a manner as to enable its iteration directly. So far, this has only been accomplished in the case when the phase-space Lagrangian one-form is a linear function (primarily, in canonical form). This leverages the related work of Ref. [92], but adapted to the context of multistep methods and the underlying one-step method.

### 3.6.1 Backward Error Initialization

Multistep methods exhibit additional degrees of freedom in the initial conditions that must be supplied to iterate the time advance. Although this freedom is the source of the instabilities, it may be leveraged to minimize the initial parasitic mode amplitude according to a *backward error initialization* procedure.

The standard starting procedure for multistep methods supplements the initial values provided by the initial value problem with accurate approximations of the true solution. For a two-step variational integrator, for instance, one would conventionally supply  $(z_0, z_1 = F_{LPS}^h(z_0))$ . Instead of this starting procedure, smoother trajectories result from determining the modified vector field of Thm. 3.3.3 and supplying initial conditions by approximating the flow of the *modified* vector field, e.g.  $(z_0, z_1 = \tilde{F}^h(z_0))$  [111, 6, 114]. Because the smooth modified system precludes parasitic modes, its flow map is absent of parasitic modes, and a multistep method initialized using this flow map will exhibit zero initial parasitic mode amplitude. In practice, the complexity of the vector fields in the smooth modified system implies the series will be truncated after a few terms. Correspondingly, the initial parasitic mode amplitudes will scale according to the number of terms retained in the smooth modified system. A diagram illustrating the backward error initialization procedure is shown in Fig. 3.10.

The benefits of backward error initialization are formalized in the following theorem.

**Theorem 3.6.1** *Suppose the smooth modified system Eq. (3.39) of a multistep variational integrator is truncated with truncation index  $r$ . Initializing the multistep variational integrator with the time- $h$  flow map of the smooth modified system will yield a parasitic mode with initial amplitude  $\mathcal{O}(h^{r+1})$ .*

**Proof** The proof of this follows from the backward error analysis performed in Thm. 3.3.3. To the extent which the initial condition satisfies the smooth modified system, parasitic modes cannot be present. The initial amplitude of the parasitic modes is determined by the truncation error in the smooth modified system, and thus has  $\mathcal{O}(h^{r+1})$ . ■

**Example** As a simple illustration of the backward error initialization procedure, consider the phase space Lagrangian for the linear oscillator in canonical coordinates, Eq. (3.24). Also,

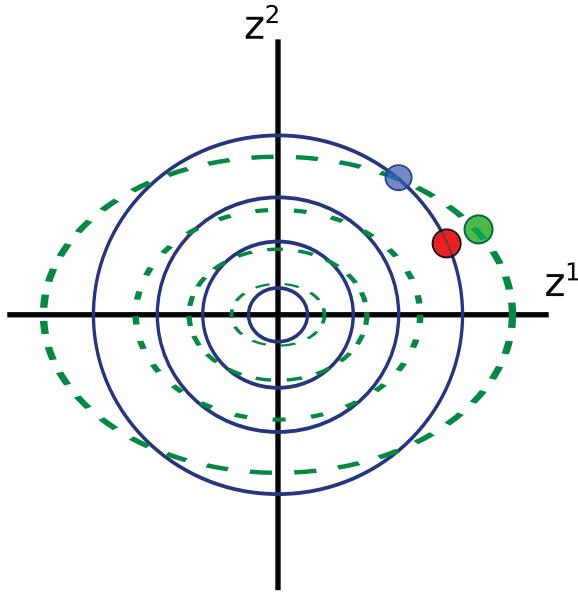


Figure 3.10: Schematic illustrating the use of backward error analysis for the generation of superior initial conditions. Integral curves of the original vector field  $X_0$  are depicted in blue while integral curves of the modified vector field  $\tilde{X}$  are shown in dashed green. Given an initial condition (blue circle), the standard procedure approximates the flow of the original vector field (red circle), while backward error initialization approximates the flow of the modified vector field (green circle).

consider the trapezoidal discrete Lagrangian Eq. (3.21), which yields the explicit midpoint algorithm Eq. (3.25) as the resulting multistep variational integrator.

Figure 3.11 illustrates the scaling of the parasitic mode amplitude as a function of the step size  $h$  when initialized using the true solution and one non-zero correction term in the smooth modified system. The results are in agreement with Thm. 3.6.1.

### 3.6.2 Re-initialization

One technique for mitigating parasitic modes is to intermittently reset the parasitic mode amplitude to small size by performing a “re-initialization” step, as used for instance in Refs. [120, 121]. Although the re-initialization step is not structure preserving (typically), one must limit the parasitic mode growth if a useful long-term result is to be obtained.

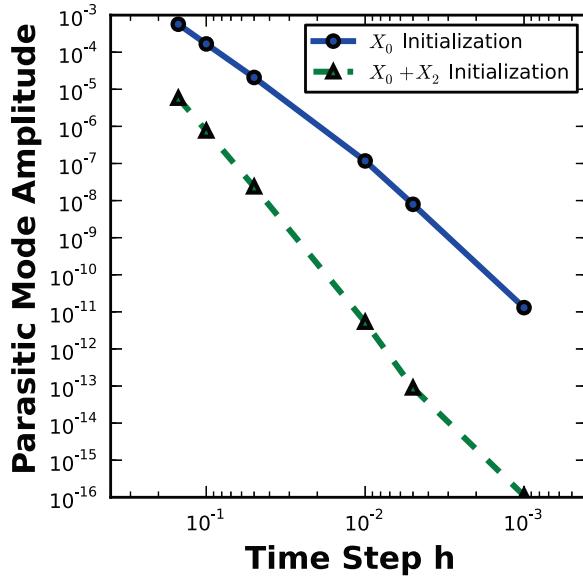


Figure 3.11: Backward error initialization reduces the initial parasitic mode amplitude in the trapezoidal multistep variational integrator for the canonical harmonic oscillator system. Because of the symmetric method, retaining a single correction term in the modified vector field leads to parasitic mode amplitudes that scale with  $\mathcal{O}(h^5)$ , shown in dashed green. In contrast, the original vector field leads to parasitic mode amplitudes that scale with  $\mathcal{O}(h^3)$ , shown in blue.

If the parasitic mode is slowly growing in time, as often is the case for the marginally stable parasitic modes of multistep variational integrators, the relative infrequency of the re-initialization step introduces very little numerical dissipation. Certainly, total elimination of the parasitic modes would be preferable, but if they cannot be avoided, re-initialization integrators safeguard against unbounded parasitic mode behavior.

Two general re-initialization integrator procedures can be envisioned. First, one can periodically perform a re-initialization step, with the re-initialization period shorter than the time for the parasitic mode to grow to unacceptably large amplitude. Alternatively, one might actively re-initialize based on some detected parasitic mode behavior. These two notions are formalized in the following definitions.

**Definition - Periodic Re-Initialization Integrator** Consider a multistep variational integrator defined by a map  $F_{MVI}^h : M \times M \rightarrow M$  and an accurate one-step method  $F_{OS}^h : M \rightarrow M$ . A *periodic re-initialization integrator* is defined by the map  $F_{PRI}^h : M \times M \rightarrow M$  according to:

$$F_{PRI}^h(z_{k-1}, z_k) = \begin{cases} F_{OS}^h(z_k) & \text{if } k \bmod (N_{RI}) = 0 \\ F_{MVI}^h(z_{k-1}, z_k) & \text{else} \end{cases}, \quad (3.100)$$

where  $N_{RI}$  is an integer representing the number of steps after which the method is to be re-initialized.

**Definition - Active Re-Initialization** Consider a multistep variational integrator for a time-independent Hamiltonian system with Hamiltonian  $H$ ; let the integrator be defined by a map  $F_{MVI}^h : M \times M \rightarrow M$ . Also consider an accurate one-step method  $F_{OS}^h : M \rightarrow M$ . Choose an energy error tolerance threshold  $\delta \in \mathbb{R}$ . An *active re-initialization integrator* is defined by the map  $F_{ARI}^h : M \times M \rightarrow M$  according to:

$$F_{ARI}^h(z_{k-1}, z_k) = \begin{cases} F_{OS}^h(z_k) & \text{if } H(z_k) - H(z_0) > \delta \\ F_{MVI}^h(z_{k-1}, z_k) & \text{else} \end{cases}. \quad (3.101)$$

The following example demonstrates that re-initialization can make the difference between a multistep variational integrator out-performing a non-geometric integrator or not.

**Example - Re-initialization of a nonlinear oscillator MVI** This example returns to the midpoint variational integrator for the nonlinear oscillator system, this time using a periodic re-initialization integrator with  $N_{RI} = 10^6$ . The parasitic mode amplitude increases slowly in time for this system. However, the accumulation of error for a high-order Runge-Kutta integration is also slow. As shown in Fig. 3.12, without re-initialization, the

accumulation of error due to the MVI parasitic mode and the numerical dissipation of the RK4 method occur at similar rates. With re-initialization, however, the MVI accumulates error much more slowly, clearly outperforming the non-geometric integration.

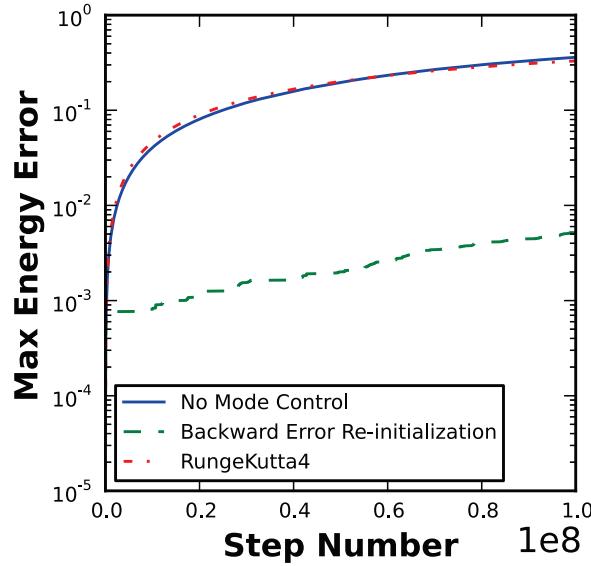


Figure 3.12: Parasitic mode mitigation using backward error initialization and re-initialization is essential for obtaining the desired long-term fidelity. In the absence of parasitic mode control (blue), the energy error of the multistep variational integrator eventually exceeds that of a non-geometric algorithm (Runge-Kutta 4, dash-dot red). With backward error initialization and re-initialization every  $\approx 10^6$  steps, the accumulation of energy error is greatly diminished.

### 3.6.3 Identification of the Underlying One-Step Method

The intent of this subsection is to elucidate a relationship between the position-momentum equations and the underlying one-step method of a multistep variational integrator. Under certain conditions, the position-momentum equations immediately identify the underlying one-step method of the discrete Euler-Lagrange equations. It is then desirable to choose a discretization procedure such that the underlying one-step method may be directly identified and iterated, bypassing the introduction of parasitic mode instabilities.

The central condition for identifying the underlying one-step method using the position-momentum equations is that the discrete momenta match the continuous momenta defined by the Legendre transform. Recall that the Legendre transform  $\mathbb{F}_{L_{PS}} : TM \rightarrow T^*M, (z, \dot{z}) \mapsto (z, p)$  according to:

$$\mathbb{F}_{L_{PS}}(z, \dot{z}) = (z, p) = (z, \frac{\partial L_{PS}}{\partial \dot{z}}) = (z, \vartheta(z)). \quad (3.102)$$

From this, we can identify the underlying one-step method of the discrete Euler-Lagrange equations according to the following theorem:

**Theorem 3.6.2** Suppose  $p_0 = \vartheta(z_0)$  implies  $p_1 = \vartheta(z_1)$  whenever  $(z_1, p_1) = \tilde{F}_{L_d}(z_0, p_0)$ , i.e.  $z_0, p_0, z_1, p_1$  satisfy the position-momentum form of the discrete Euler-Lagrange equations. Then the underlying one-step method  $\Phi$  of the discrete Euler-Lagrange equations satisfies:

$$\vartheta(z_0) = -D_1 L_d(z_0, \Phi(z_0)) \quad (3.103)$$

and

$$\vartheta(\Phi(z_0)) = D_2 L_d(z_0, \Phi(z_0)). \quad (3.104)$$

**Proof** The supposition of the proof is that:

$$\begin{aligned} \vartheta(z_0) &= -D_1 L_d(z_0, z_1) \Rightarrow \\ \vartheta(z_1) &= D_2 L_d(z_0, z_1). \end{aligned}$$

Define a map  $\phi : M \rightarrow M$  according to  $\tilde{F}_{L_d}(z_0, \vartheta(z_0)) = (\phi(z_0), \vartheta(\phi(z_0)))$ , or:

$$\begin{aligned} \vartheta(z_0) &= -D_1 L_d(z_0, \phi(z_0)) \\ \vartheta(\phi(z_0)) &= D_2 L_d(z_0, \phi(z_0)). \end{aligned}$$

Because  $z_0$  is arbitrary, we also have:

$$\vartheta(\phi(z_0)) = -D_1 L_d(\phi(z_0), \phi(\phi(z_0))).$$

From these conditions, we find that  $\phi$  satisfies the condition of the underlying one-step method:

$$D_2 L_d(z_0, \phi(z_0)) + D_1 L_d(\phi(z_0), \phi(\phi(z_0))) = 0,$$

so  $\phi = \Phi$ . ■

It's interesting to note that we could replace  $\vartheta(z_k)$  in the above theorem with any function  $f(z_k)$ . This observation may allot greater flexibility when seeking to satisfy these conditions, however, it's difficult to see a useful method of selecting  $f$  when one no longer appeals to the continuous Legendre transform, and this may have consequences on consistency of the one-step map.

Equipped with this theorem, it is desirable to iterate Eq. (3.103) to determine a trajectory for the variational integrator. The question then becomes how “difficult” it is for a discrete Lagrangian to satisfy the necessary condition of the theorem. It has been shown that certain families of Galerkin variational integrators satisfy this condition whenever  $\vartheta$  is linear in the coordinates  $z$  [92]. The example below illustrates how this works in a simple setting. From there, the hope is that one could construct a procedure for satisfying the constraint at an arbitrary polynomial order (if not, in general).

**Example- Canonical Midpoint Lagrangian** Consider a canonical phase-space Lagrangian,  $z = [q, p]^T$ ,  $\vartheta = [p, 0]$ ,

$$L_{PS}(z, \dot{z}) = p\dot{q} - H(z). \quad (3.105)$$

Suppose we choose a midpoint discrete Lagrangian:

$$L_d(z_0, z_1) = h L_{PS}\left(\frac{z_0 + z_1}{2}, \frac{z_1 - z_0}{h}\right) = \frac{p_0 + p_1}{2}(q_1 - q_0) - h H\left(\frac{z_0 + z_1}{2}\right). \quad (3.106)$$

The corresponding discrete Euler-Lagrange equations are given by:

$$z_{k+1} - z_{k-1} = h \left( X_H\left(\frac{z_{k-1} + z_k}{2}\right) + X_H\left(\frac{z_k + z_{k+1}}{2}\right) \right). \quad (3.107)$$

It is straightforward to verify the underlying one-step method of these equations is implicit midpoint, as pointed out by [123]. We will show how implicit midpoint emerges from the identification of the underlying one-step method using the position-momentum form of the discrete Euler-Lagrange equations.

First, we verify that requirement of Thm. (3.6.2) is satisfied. Note that this is assured by virtue of the midpoint discrete Lagrangian being the 1-stage Gauss method discussed in [92]. The position-momentum form of the DEL equations is given by:

$$P_k = \left( \frac{p_k + p_{k+1}}{2} + \frac{h}{2} H_{,q}\left(\frac{z_k + z_{k+1}}{2}\right), -\frac{1}{2}(q_{k+1} - q_k) + \frac{h}{2} H_{,p}\left(\frac{z_k + z_{k+1}}{2}\right) \right) \quad (3.108)$$

$$P_{k+1} = \left( \frac{p_k + p_{k+1}}{2} - \frac{h}{2} H_{,q}\left(\frac{z_k + z_{k+1}}{2}\right), \frac{1}{2}(q_{k+1} - q_k) - \frac{h}{2} H_{,p}\left(\frac{z_k + z_{k+1}}{2}\right) \right). \quad (3.109)$$

Now assume that  $P_k = \vartheta(z_k) = [p_k, 0]$ , and add the two equations to find:

$$P_{k+1} + P_k = (P_{k+1,p} + p_k, P_{k+1,q} + 0) = (p_k + p_{k+1}, 0), \quad (3.110)$$

so  $P_{k+1} = (p_{k+1}, 0)$ , and the condition for Thm. (3.6.2) is satisfied.

The underlying one-step method of the midpoint discrete Lagrangian for the canonical phase-space Lagrangian may therefore be identified from the position-momentum form of

the discrete Euler-Lagrange equations. Checking the first of the two, equivalent relations:

$$(p_k, 0) = \left( \frac{p_k + p_{k+1}}{2} + \frac{h}{2} H_{,q} \left( \frac{z_k + z_{k+1}}{2} \right), -\frac{1}{2} (q_{k+1} - q_k) + \frac{h}{2} H_{,p} \left( \frac{z_k + z_{k+1}}{2} \right) \right) \Rightarrow \quad (3.111)$$

$$(p_{k+1} - p_k, q_{k+1} - q_k) = \left( -h H_{,q} \left( \frac{z_k + z_{k+1}}{2} \right), h H_{,p} \left( \frac{z_k + z_{k+1}}{2} \right) \right), \quad (3.112)$$

recovering the expected result that implicit midpoint is the underlying one-step method. The other term of the position-momentum equations may be used to make the same identification.

This example illustrates recovery of a known underlying one-step method using Thm. (3.6.2), even though the practical impact of this example is minimal (why not simply use implicit midpoint in the first place?). The hope is that the same procedure may be applied in cases where it is otherwise unclear how to identify the underlying one-step method.

## 3.7 Discussion

### 3.7.1 Summary

By applying the analysis procedures established in the multistep numerical analysis literature to variational integrators derived from degenerate Lagrangians, significant insight has been unearthed that explains the stability and conservation properties of this interesting class of variational integrators. The presence of parasitic mode oscillations is apparent from an eigenvalue analysis of the discrete Euler-Lagrange equations when the discrete Lagrangian approximates a small interval of a phase-space action. The impact of these high-frequency modes ranges from small-scale oscillations in the conserved quantities to overwhelming degradation of the numerical trajectory. Mitigating their effects is therefore important for successful application of multistep variational integrators. The additional freedom present in supplying the initial conditions may be leveraged by sampling the initial values along the

flow of a modified vector field calculated via backward error analysis. Finally, while the symplectic structure preserved by the multistep map bears little resemblance to that of the symplectic manifold of the original Hamiltonian system, the underlying one-step method has been shown to preserve a structure on the original phase-space. All of these results emerge as immediate implications upon considering the variational integrator to be a multistep method for the integration of Hamilton's equations.

### 3.7.2 Prospects for Future Work

Several avenues for progressing the understanding and utility of multistep variational integrators persist. For one, formal results remain to be established regarding the long-term behavior of the underlying one-step method of a multistep variational integrator. The nearness of the discrete symplectic structure to the continuous one, on the original symplectic manifold, suggests bounded energy error theorems may be possible. However, the proofs require arguments beyond those provided for standard symplectic integrators, which preserve the same two-form as the flow of the original Hamiltonian system. A viable approach may be to use the discrete two-form of the underlying one-step method to prove conjugate symplecticity. Another avenue for development is in determining ways to directly iterate the underlying one-step method. If an efficient procedure can be developed for recovering the underlying one-step method analytically or numerically, one might derive a multistep variational integrator from the onset but deploy the underlying one-step method in practice. If this approach were to succeed in general, it would constitute a breakthrough in non-canonical symplectic integration.

Another exciting avenue to explore is multistep variational integrators for *regular* Lagrangian systems. Typically, the discrete Lagrangian is a function of two points, yielding a two-step method for a second-order differential equation. However, one could more generally

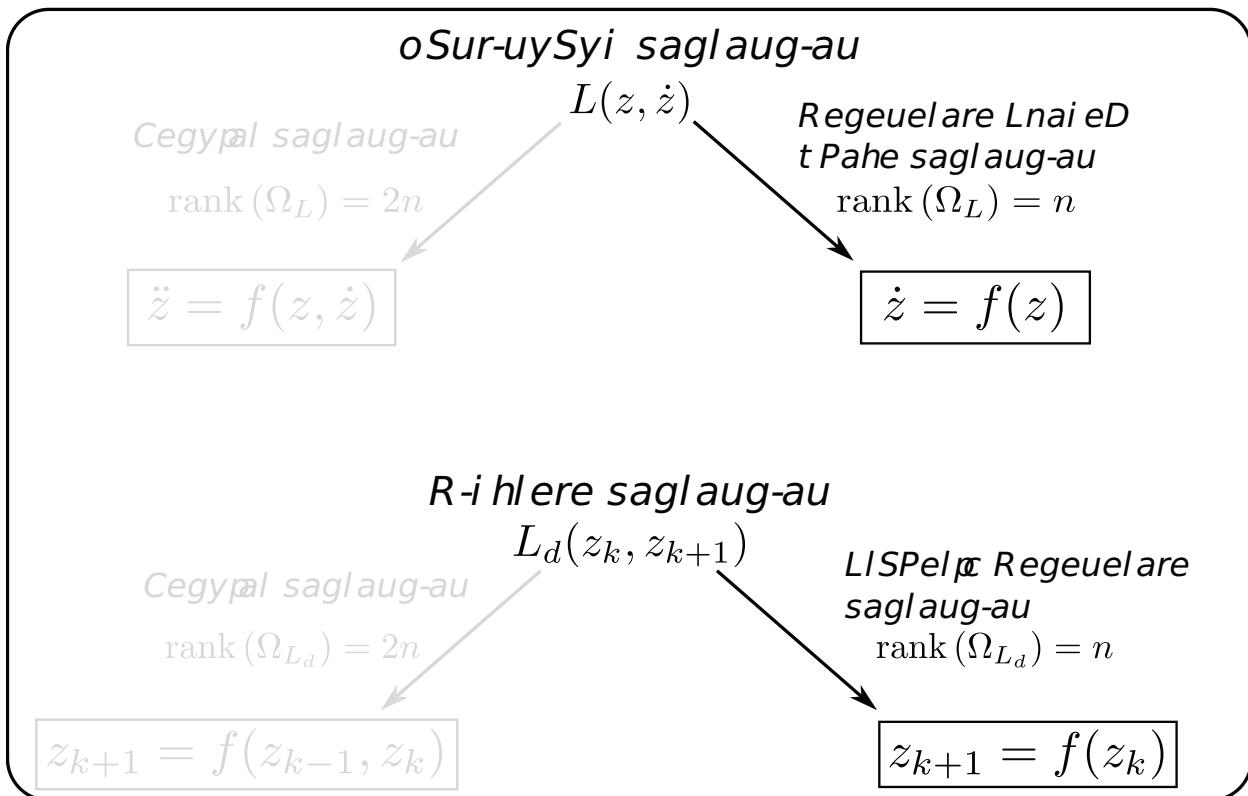
consider discrete actions of the form:

$$S_d(z_0, \dots, z_N) = \sum_{k=0}^{N-p} L_d(z_k, z_{k+1}, \dots, z_{k+p}). \quad (3.113)$$

Note that this is a *different* form of the discrete action than those used in Sec. 3.5 and Refs. [9, 100] for constructing Galerkin variational integrators and symplectic partitioned Runge-Kutta methods. The discrete Euler-Lagrange equations for this discrete action would be  $p$ -step methods, where  $p$  could be arbitrarily large. The stability results of this Chapter are not especially inspiring for such intentional construction of multistep variational integrators, but the benefits of the variational principle for numerical analysis makes this a topic worth exploring. For instance, methods that are otherwise difficult to analyze could be concisely studied in terms of their discrete Lagrangian function if they derive from a discrete action of the form in Eq. (3.113).

# Chapter 4

## Degenerate Variational Integrators



*Therefore benefit comes from what is there;*

*Usefulness from what is not there.*

—Lao Tsu, *Tao Te Ching*

## 4.1 Introduction

Chapter 3 illuminated a significant drawback to multistep variational integrators: extracting the desired long-term fidelity is contingent upon the parasitic modes remaining small. Hindering the prospect of small parasitic modes in MVIs is that they are at best marginally stable. Although judicious initial condition selection and re-initialization may be used to mitigate parasitic mode behavior, it is clearly advantageous to eliminate the unphysical modes altogether.

In this Chapter, an alternative to multistep variational integrators is developed by introducing degeneracy on a discrete level, yielding *degenerate variational integrators*. A degenerate variational integrator stems from a degenerate discrete Lagrangian, for which the discrete Legendre transforms are not invertible and the discrete two-form lacks full rank. Consequently, the discrete Euler-Lagrange equations do not uniquely define a two-step map; it remains possible, however, that a lower-dimensional map is uniquely prescribed. This discrete degeneracy is analogous to that exhibited by phase-space Lagrangians, for which the Legendre transform is also not invertible and the two-form also lacks full rank. The central result of this chapter demonstrates that when the discrete Lagrangian possesses the *same* degree of degeneracy as the continuous Lagrangian, parasitic modes are absent from the numerical dynamics. Variational integrators derived from discrete Lagrangians satisfying this property will be deemed *properly degenerate variational integrators*. Examples included in this chapter show how to obtain one-step properly-degenerate variational integrators for several non-canonical Hamiltonian systems of practical interest.

Several clues guided the realizations presented in the Chapter. The first clue was that a “staggering” of the parallel velocity coordinate  $u$  used in the discrete guiding center Lagrangians led to a dramatic reduction of the parasitic mode instabilities. This staggering was suggested in Ref. [16] for the first time; although the instabilities were not understood at that time, it was known that replacing  $(u_{k+1} + u_k)/2$  with  $u_{k+1/2}$  in the discrete Lagrangian functions led to dramatically improved (in fact, thought to be completely stable) results. For instance, consider the “unstaggered” discrete guiding center Lagrangian:

$$\begin{aligned} L_d(x_0, u_0, x_1, u_1) = & A_i^\dagger \left( \frac{x_0 + x_1}{2}, \frac{u_0 + u_1}{2} \right) (x_1^i - x_0^i) - \\ & H_{gc} \left( \frac{x_0 + x_1}{2}, \frac{u_0 + u_1}{2} \right). \end{aligned} \quad (4.1)$$

The multistep variational integrator that arises from this discrete Lagrangian requires initial conditions of the form  $(x_0, u_0, x_1, u_1)$ . In contrast, consider the “staggered” discrete guiding center Lagrangian:

$$L_d(x_0, u_{1/2}, x_1) = A_i^\dagger \left( \frac{x_0 + x_1}{2}, u_{1/2} \right) (x_1^i - x_0^i) - H_{gc} \left( \frac{x_0 + x_1}{2}, u_{1/2} \right). \quad (4.2)$$

The multistep variational integrator that arises from *this* discrete Lagrangian requires initial conditions of the form  $(x_0, u_{1/2}, x_1)$ ; there are *fewer* initial conditions for the staggered- $u$  variational integrator. Not surprisingly, this has an impact on the parasitic mode behavior. Figure 4.1 shows that, for a trapped-particle “banana” orbit, the unstaggered variational integrator degrades to the parasitic mode behavior after only a single bounce period. Meanwhile, the staggered variational integrator requires thousands of transits before the parasitic modes can be detected, even at small scale.

A second important clue was that one-step variational integrators were not difficult to derive from canonical phase-space Lagrangians. For instance, consider the canonical phase-

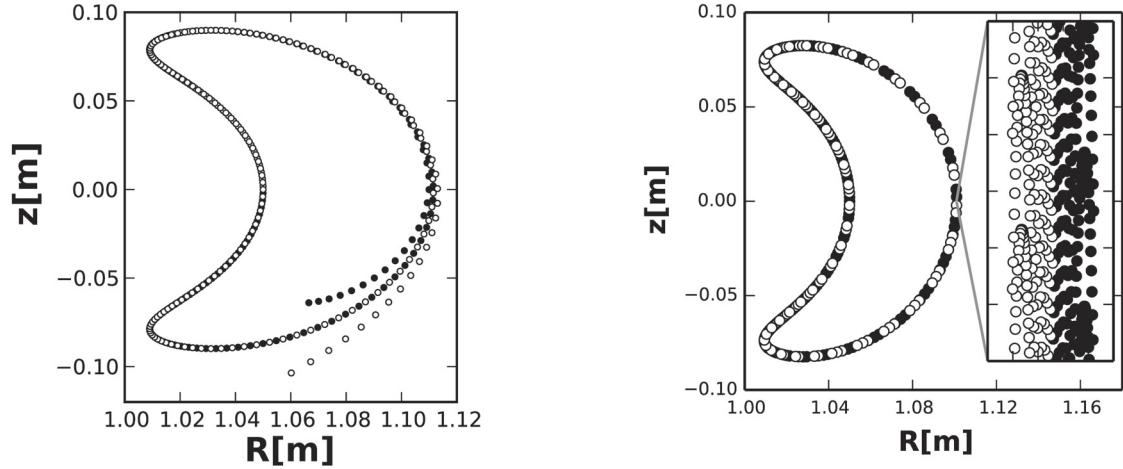


Figure 4.1: Staggering the parallel velocity  $u$  of a variational guiding center integrator reduces parasitic mode instabilities. On the left, slightly more than one period of a trapped-particle “banana orbit” is shown to be unstable for the variational integrator derived from the un-staggered discrete Lagrangian Eq. (4.1). On the right, several thousand periods are shown to be only weakly unstable for the variational integrator derived from the staggered discrete Lagrangian Eq. (4.2).

space Lagrangian:

$$L_{PS}(q, p, \dot{q}, \dot{p}) = p\dot{q} - H(q, p),$$

with Euler-Lagrange equations:

$$-\dot{p} - H_{,q}(q, p) = 0$$

$$\dot{q} - H_{,p}(q, p) = 0.$$

In contrast to the midpoint discrete Lagrangian used to derive a multistep method in Chapter 3, consider the following degenerate discrete Lagrangian:

$$L_d(q_0, p_0, q_1, p_1) = p_0(q_1 - q_0) - hH(q_1, p_0). \quad (4.3)$$

In this case, the discrete Euler-Lagrange equations are given by:

$$-p_k + p_{k-1} - hH_{,q}(q_k, p_{k-1}) = 0 \quad (4.4)$$

$$q_{k+1} - q_k - hH_{,p}(q_{k+1}, p_k) = 0. \quad (4.5)$$

As written, a unique solution for  $q_{k+1}, p_{k+1}$  does not appear to exist given  $q_{k-1}, p_{k-1}, q_k, p_k$  - this is because the discrete Lagrangian is *degenerate* per Eq. (2.31), and the discrete Euler-Lagrange equations do not define a map on  $Q \times Q$ . However, we can identify a well-defined map on a smaller dimensional space than  $Q \times Q$ , in particular,  $Q$ . Because we expect the discrete Euler-Lagrange relations to hold for all  $k$ , we may “shift” the indices of the first equation forward in time to identify a well-defined one-step map:

$$-p_{k+1} + p_k - hH_{,q}(q_{k+1}, p_k) = 0 \quad (4.6)$$

$$q_{k+1} - q_k - hH_{,p}(q_{k+1}, p_k) = 0. \quad (4.7)$$

This one-step map may be recognized as a familiar symplectic Euler scheme [6]. Note that trajectories satisfying the shifted equations also satisfy the unshifted equations, retaining the utility of a discrete variational principle for proving conservation properties despite the degeneracy of the discrete Euler-Lagrange equations.

Reflecting on these two clues, *both* examples involved discrete Lagrangians that were *themselves* degenerate. Recall, the degeneracy condition for discrete Lagrangians presented in Eq. (2.31). This observation guided the key realization of this chapter: if the variational integrators are degenerate with the *same degree* of degeneracy as the continuous system, parasitic modes are eliminated altogether. For the staggered guiding center discrete Lagrangian above, the staggered  $u$  coordinate introduces *some* degeneracy, but not enough to eliminate *all* of the parasitic modes.

Describing the elimination of parasitic modes in terms of the degeneracy of the discrete Lagrangian was instrumental for discovering one-step variational integrators for non-canonical magnetic field line flow and non-canonical guiding center trajectories. These new algorithms are derived in this chapter and tested extensively in the next. They are the first one-step variational integrators for these non-canonical systems, and do not refer to canonical coordinates for their construction. Although the guiding center degenerate variational integrator does not yet allow for completely arbitrary coordinates, a wide variety of magnetic configurations may be modeled without patching together local coordinate regions. The new method thus represents an important leap in the geometric integration of non-canonical Hamiltonian systems.

## 4.2 Definitions

Section 2.2.2 defined the regularity of a discrete Lagrangian in terms of the discrete Legendre transform. In essence, a discrete Lagrangian is degenerate whenever it is not regular. Specifically:

**Definition** A discrete Lagrangian  $L_d : Q \times Q \rightarrow \mathbb{R}$  is a *degenerate discrete Lagrangian* if:

$$\det(D_1 D_2 L_d(q_0, q_1)) = 0. \quad (4.8)$$

Correspondingly,

**Definition** A *degenerate variational integrator* (DVI) is the time advance specified by the discrete Euler-Lagrange equations (2.21) of a degenerate discrete Lagrangian.

In general, a discrete Lagrangian may be degenerate (or regular) only locally. For the concerns of this dissertation, I will be assuming a discrete Lagrangian is globally regular or globally degenerate. Consequences of a degenerate discrete Lagrangian include: the discrete

Euler-Lagrange equations do not uniquely define a map from  $Q \times Q$  into itself; the discrete Legendre transforms are not invertible; and the discrete two-form does not possess full rank.

Of course, we have already encountered similar consequences of degeneracy when discussing the phase-space Lagrangians of Sec. 2.3. In light of this, it is beneficial to characterize the degree of degeneracy of a discrete Lagrangian relative to the degree of degeneracy of a phase-space Lagrangian. The following definition - presented here for the first time - accomplishes this.

**Definition** Consider a phase-space Lagrangian  $L_{PS} : TM \rightarrow \mathbb{R}$ . A discrete Lagrangian  $L_d : M \times M \rightarrow \mathbb{R}$  is a *properly degenerate discrete Lagrangian* if:

$$\text{rank}(D_1 D_2 L_d(z_0, z_1)) = \frac{1}{2} \dim(M). \quad (4.9)$$

A definition that naturally follows is:

**Definition** The discrete Euler-Lagrange equations corresponding to a properly degenerate discrete Lagrangian define a *properly degenerate variational integrator*.

The justification for the “properly” designation is the similarity between the discrete and continuous two-forms when the condition is met. In brief,  $\text{rank}(\Omega_{L_{PS}}) = \text{rank}(\Omega_{L_d}) = \dim(M)$  when  $L_d$  is properly degenerate. In more detail, consider the tensor components of  $\Omega_L$  and  $\Omega_{L_d}$  as given by Eqs. (2.37) and (2.39) respectively. In terms of the slot derivative notation, these tensors take the form:

$$\Omega_L(q, \dot{q}) = \begin{pmatrix} D_1 D_2 L(q, \dot{q}) - D_2 D_1 L(q, \dot{q}) & -D_2 D_2 L(q, \dot{q}) \\ D_2 D_2 L(q, \dot{q}) & 0 \end{pmatrix} \quad (4.10)$$

$$\Omega_{L_d}(q_0, q_1) = \begin{pmatrix} 0 & D_1 D_2 L_d(q_0, q_1) \\ -D_2 D_1 L_d(q_0, q_1) & 0 \end{pmatrix}, \quad (4.11)$$

where the blocks are of size  $\dim(M) \times \dim(M)$ . For the special case of a phase space Lagrangian,  $D_2 D_2 L_{PS}(z, \dot{z}) = 0$  and  $D_1 D_2 L_{PS}(z, \dot{z}) - D_2 D_1 L_{PS}(z, \dot{z})$  has full rank. So,  $\text{rank}(\Omega_{L_{PS}}) = \dim(M)$ . For the discrete two-form,  $\text{rank}(\Omega_{L_d}) = 2 \times \text{rank}(D_1 D_2 L_d)$ . So, when  $\text{rank}(D_1 D_2 L_d) = \frac{1}{2} \dim(M)$ , we find  $\text{rank}(\Omega_{L_d}) = \text{rank}(\Omega_{L_{PS}})$ .

For all of the properly degenerate variational integrators encountered in this dissertation, it was possible to formulate the PDVI as a one-step method. Whether or not this is possible in general is an open question. But because it pertains to the algorithms of interest and is useful for analyzing PDVIs, it is helpful to provide a definition for discussing this formulation of the time advance.

**Definition** The *one-step formulation* of a properly degenerate variational integrator is a map  $\hat{F} : M \times \mathbb{R} \rightarrow M$ ,  $F(z_0, h) = z_1$  satisfying the discrete Euler-Lagrange equations, i.e.:

$$D_2 L_d(z_0, \hat{F}(z_0, h), h) + D_1 L_d(\hat{F}(z_0, h), \hat{F}(\hat{F}(z_0, h), h), h) = 0. \quad (4.12)$$

The  $h$ -dependence of the one-step formulated may be implied and omitted from the list of arguments.

The one-step formulation of a properly degenerate variational integrator bears resemblance to the underlying one-step method of a multistep variational integrator. The different terminology is required because properly degenerate variational integrators do not uniquely define two-step methods, so the “underlying” nomenclature is not really appropriate. The notation  $\hat{F}$  is re-used, however, as it should be apparent from context whether it is the underlying one-step method of a multistep method or the one-step formulation of a PDVI.

## 4.3 Stability Analysis

The most important consequence of proper degeneracy is the elimination of parasitic modes. The following theorem describes this in detail:

**Theorem 4.3.1** *Consider a properly degenerate discrete Lagrangian  $L_d$  constructed from a phase-space Lagrangian  $L_{PS}$ . If the properly degenerate variational integrator is consistent, then the linearized  $h \rightarrow 0$  discrete Euler-Lagrange equations possess  $n = \dim(M)$  eigenvalues at positive one. That is, properly degenerate variational integrators are absent of parasitic modes.*

**Proof** The proof requires two steps. First, it must be demonstrated that there exist at most  $n = \dim(M)$  non-zero roots of the characteristic equation from which the principal and parasitic modes are determined. Second, it must be shown that these  $n$  roots are equal to positive one. Combined, these facts establish the absence of parasitic modes in properly degenerate variational integrators.

Proceeding with the first step, recall from Thm. 3.3.1 that the characteristic equation determining the principal and parasitic modes  $\lambda_i$  for a variational integrator is given by:

$$\det(m) = 0$$

$$m = \lambda^2 D_2 D_1 L_d(\bar{z}, \bar{z}) + \lambda (D_1 D_1 L_d(\bar{z}, \bar{z}) + D_2 D_2 L_d(\bar{z}, \bar{z})) + D_1 D_2 L_d(\bar{z}, \bar{z}),$$

where  $\bar{z}$  is the point about which the  $h = 0$  DEL equations are linearized.

Next, I claim that the determinant of  $m$  may be expressed in the form:

$$\det(m) = \pm \det \begin{pmatrix} a_{ij}\lambda^2 + b_{ij}\lambda + c_{ij} & a_{ij}\lambda^2 + b_{ij}\lambda \\ b_{ij}\lambda + c_{ij} & b_{ij}\lambda \end{pmatrix}, \quad (4.13)$$

where each of the four blocks is of size  $n/2 \times n/2$ . The justification for this representation is as follows: The determinant is an  $n$ -linear function of the rows and an  $n$ -linear function of the columns. This means the determinant is unchanged upon adding a constant multiple of one row (or column) to another, and changes by a factor of negative one upon interchanging rows (or interchanging columns). Additionally, the rank of a matrix is equal to the number of linearly independent rows (and columns). So, the above representation may be constructed by first performing row operations on  $m$  to place all  $\lambda^2$  factors in the top  $n/2$  rows; the possibility of this is guaranteed by  $\text{rank}(D_2 D_1 L_d(\bar{z}, \bar{z})) = n/2$ . Afterward, perform column operations to obtain all constant factors in the left  $n/2$  columns; ensured by  $\text{rank}(D_1 D_2 L_d(\bar{z}, \bar{z})) = n/2$ . The above representation is thereby obtained, with  $a_{ij}$  some linear combination of the entries of  $D_2 D_1 L_d(\bar{z}, \bar{z})$  and  $c_{ij}$  the same linear combination of the entries of  $D_1 D_2 L_d(\bar{z}, \bar{z})$ . Because the overall factor of  $\pm 1$  does not change the roots of the characteristic equation, we may use the above representation for the eigenvalue analysis. Note that in the special case wherein the bottom two rows of  $D_2 D_1 L_d(\bar{z}, \bar{z})$  are zero, this representation is obtained immediately.

Next, consider the Leibniz formula for the determinant of  $m$ :

$$\det(m) = \frac{1}{n!} \sum_{i_1, \dots, i_n, j_1, \dots, j_n=1}^n \epsilon_{i_1 \dots i_n} \epsilon_{j_1 \dots j_n} m_{i_1 j_1} \cdot \dots \cdot m_{i_n j_n},$$

where  $\epsilon$  is the Levi-Civita symbol. The pertinent fact conveyed in this formula is that each term in the summand of  $\det(m)$  is a product of  $n$  entries of  $m$  possessing unique row and column indices. An implication of this fact is: if an entry  $m_{i^*, j^*}$  with  $i^* < n/2, j^* < n/2$  is present in a product, then an entry  $m_{i^{**}, j^{**}}$  with  $i^{**} \geq n/2, j^{**} \geq n/2$  must also be present in the product. Put simply, if an entry from the upper left block of Eq. (4.13) is present in a term of the summand, then so is an entry from the bottom right block. Arguing similarly, if an entry  $m_{i^*, j^*}$  with  $i^* < n/2, j^* \geq n/2$  (top right) is present in a product, then so is an entry

$m_{i^{**},j^{**}}$  with  $i^{**} \geq n/2, j^{**} < n/2$  (bottom left). Combining these facts and observations with the representation of Eq. (4.13), each term in the summation of the determinant is of the form:  $\lambda^{n/2} P^n(\lambda)$ , where  $P^n(\lambda)$  is an  $n$ -th degree polynomial in  $\lambda$ . Because all terms in the summation are of this form, one concludes  $\det(m) = \lambda^{n/2} P^n(\lambda)$  and therefore that the characteristic equation possesses  $n/2$  roots at  $\lambda = 0$  and  $n$  roots that are potentially non-zero.

For the (much quicker) second step, the non-zero roots are required to be positive one whenever the properly degenerate variational integrator is consistent. Consistent numerical methods must converge to the identity map in the zero step size limit, requiring  $n$  eigenvalues of the zero step size algorithm must be positive one. ■

Stepping back, the preceding theorem provides a concise condition for ensuring a variational integrator constructed from an action principle in phase space is free of parasitic modes. In short, as long as  $\text{rank}(D_1 D_2 L_d) = n/2$ , then the ranks of the continuous and discrete symplectic structures exhibit the same degree of degeneracy, and the degenerate variational integrator cannot exhibit parasitic mode oscillations.

## 4.4 Structure Preservation

Degenerate variational integrators have degenerate two-forms  $\Omega_{L_d}$ . As shown in Eq. (4.11),  $\text{rank}(\Omega_{L_d}) = 2 \text{rank}(D_1 D_2 L_d(z_0, z_1))$ , so if  $D_1 D_2 L_d(z_0, z_1)$  is degenerate, then  $\Omega_{L_d}$  is degenerate on  $M \times M$  and would therefore be categorized as a “presymplectic structure”.

The one-step formulation of a properly degenerate variational integrator, however, may be used to pull back the presymplectic structure on  $M \times M$  to identify a two-form directly on  $M$ . The mechanics of this are identical to those of the underlying one-step method of multistep variational integrator, yielding the following degenerate variational integrator analog of Thm. 3.4.4.

**Theorem 4.4.1** Consider a properly degenerate variational integrator determined by a discrete Lagrangian  $L_d$  with a one-step formulation  $\hat{F}_{L_d}$ . Then  $\hat{F}_{L_d}$  is symplectic with respect to  $\hat{\Omega}_{L_d}$ , i.e.  $(\hat{F}_{L_d})^* \hat{\Omega}_{L_d} = \hat{\Omega}_{L_d}$  where:

$$\hat{\Omega}_{L_d}(z) = (\text{id}, \hat{F}_{L_d})^* \Omega_{L_d}(z) = \left( D_1 D_2 L_d(z, \hat{F}_{L_d}(z)) \right)_{ij} \left( D \hat{F}_{L_d}(z) \right)_k^j dz^i \wedge dz^k. \quad (4.14)$$

**Proof** The proof is identical to that of Thm. 3.4.4, except interpreting  $\hat{F}_{L_d}$  as the one-step formulation of the properly degenerate variational integrator. ■

To enhance intuition regarding the rather abstract-looking two-form in Eq. (4.14), the following example shows that the physical symplectic structure can be recovered in the zero step size limit of some degenerate variational integrators.

**Example** Consider the general phase-space Lagrangian:

$$L_{PS}(z, \dot{z}) = \vartheta_i(z) \dot{z}^i - H(z),$$

and a discrete Lagrangian:

$$L_d(z_0, z_1) = \vartheta_i(z_0)(z_1^i - z_0^i) - hH(z_0).$$

Further, suppose that  $L_d$  is properly degenerate (this will only be true for certain  $\vartheta$ ).

Consider the one-step formulation  $\hat{F}$  of the properly degenerate variational integrator and let  $h = 0$ . Then  $\hat{F}(z) = \text{id}(z)$  and  $(D \hat{F}_{L_d}(z))_k^j = \delta_k^j$ . Computation of Eq. (4.14) recovers:

$$\hat{\Omega}_{L_d} = \vartheta_{i,j}(z_0) dz_0^j \wedge dz_0^i = -\Omega_{L_{PS}}(z_0).$$

The minus sign is due to the sign convention of  $\Omega_{L_{PS}}$ . Of course, if  $-\Omega_{L_{PS}}$  is preserved, so is  $\Omega_{L_{PS}}$ .

For small but finite  $h$ , the discrete symplectic structure can be expected to be non-degenerate and nearby to the physical symplectic structure. Note that this example is either identical to or works out in the same way as the degenerate variational integrators presented for field line flow and collisionless guiding center dynamics.

## 4.5 Accuracy

When attempting to construct degenerate variational integrators of high order, care must be taken to retain the desired degree of degeneracy in the discrete Lagrangian function. In this section, two remarks are made regarding the construction of higher-order accurate degenerate discrete Lagrangians. The first is that staggering techniques can be helpful for increasing the order of accuracy without changing the degree of degeneracy. The second is that one should proceed with caution when composing a degenerate variational integrator with its adjoint method. Such compositions remain as important future work, but several cautionary observations are made regarding the conservation properties of the composed method.

Staggering coordinates, particularly those whose time derivative does not appear in the phase-space Lagrangian, can be an effective means of increasing accuracy. In Sec. 4.1, the first-order accurate symplectic Euler method was constructed from the discrete Lagrangian of Eq. (4.3). In this discrete Lagrangian  $p_1$  appeared *nowhere*. This inspires the following discrete Lagrangian:

$$L_d(q_0, p_{1/2}, q_1) = p_{1/2} \cdot (q_1 - q_0) - H(q_0, p_{1/2}). \quad (4.15)$$

The DEL equations are given by:

$$-p_{k+1/2} + p_{k-1/2} - hH_{,q}(q_k, p_{k+1/2}) = 0 \quad (4.16)$$

$$q_{k+1} - q_{k-1/2} - hH_{,q}(q_k, p_{k+1/2}) = 0. \quad (4.17)$$

This second-order accurate method is referred to as the “Leapfrog” method when  $H(q, p) = p^2/2 + V(q)$  for some potential function  $V$ . The second-order accuracy may be inferred from the time-centered DEL equations. By staggering the momentum coordinate  $p$ , the integrator improved from first-order to second-order accurate in the numerical step size  $h$ .

A common method of achieving second-order accurate schemes from first-order accurate schemes is to compose the first-order accurate method with its *adjoint* method. The map defined by the composition of the original map and its adjoint is time-symmetric and therefore (at least) second-order accurate. If both the original method and its adjoint preserve *the same* symplectic structure, then the composition map will also preserve that symplectic structure. From Eq. (4.14), it is not apparent that a first-order degenerate variational integrator will preserve the same symplectic two-form as its adjoint method. One should therefore proceed with caution and look to identify what symplectic structure, if any, is preserved by the composition of a DVI and its adjoint method.

From a variational integration theory perspective, one can consider composing a discrete Lagrangian with its adjoint [9]. One might hope that doing so with degenerate variational integrators will yield a second-order accurate variational integrator with the same degree of degeneracy. However, the DEL equations resulting from the composition of two discrete Lagrangians are *not* the alternation of the DEL equations that result from each individual discrete Lagrangian alone [9, 124]. It cannot be expected in general, then, that composing discrete Lagrangians will yield degenerate variational integrators even if each of the discrete Lagrangians is itself degenerate.

Composition methods for degenerate variational integrators is an excellent topic for future study. Numerical studies could guide theoretical efforts, either by suggesting the composed degenerate variational integrators *do* preserve a two-form (if the results are good), or by discouraging effort to prove conservation properties of the composed methods (if the results are bad).

## 4.6 Applications

In a seminal development for the geometric integration of non-canonical Hamiltonian systems, this section derives the first one-step variational integrators magnetic field line flow and guiding center trajectories. The field line integrator allows arbitrary non-canonical coordinates; to obtain the discrete degeneracy, a particular gauge is chosen to eliminate one component of the vector potential. The guiding center integrator assumes coordinates such that the magnetic field unit vector has one covariant component that is zero. Systematic procedures for constructing such coordinates are known based on methods for identifying canonical guiding center coordinates [78, 46, 77, 79]. The magnetic field line and guiding center integrators are the first one-step variational integrators for these systems in terms of non-canonical coordinates.

To motivate the assumptions made below on the electromagnetic gauge and guiding center coordinates, it is helpful to review the proper degeneracy condition. For a phase-space Lagrangian  $L_{ps} = \vartheta(z)_i \dot{z}^i - H(z)$ , we seek to construct a discrete Lagrangian such that  $\text{rank}(D_1 D_2 L_d(z_0, z_1)) = \dim(M)/2$ . Typically,  $\text{rank}(D_1 D_2 L_d(z_0, z_1)) = \dim(M)$ . When choosing a discretization, the difficult term is  $\vartheta(z)_i \dot{z}^i$ . Consider a discrete Lagrangian of the form:

$$L_d(z_0, z_1) = \vartheta_i(\alpha z_0 + (1 - \alpha) z_1)(z_1^i - z_0^i) - hH(z_0), \quad (4.18)$$

where  $\alpha \in [0, 1]$  is a parameter for the discretization. The Hamiltonian term  $H(z_0)$  will contribute no components to  $D_1 D_2 L_d$ ; it does not depend on  $z_1$ . Because the velocity approximation involves *both*  $z_1^i$  and  $z_0^i$ , then non-zero components will be contributed to  $D_1 D_2 L_d$  for *any* value of the discretization parameter  $\alpha$ . The assumptions made below incorporate degeneracy by reducing the number of velocity components  $\dot{z}^i$  present in  $L_{ps}$  to  $\dim(M)/2$ . Put another way, half of the components of  $\vartheta(z)$  are set to zero. More general discretizations may relax this restriction, but it is not yet known what other techniques can be used to achieve discrete degeneracy.

### 4.6.1 Magnetic Field Line Flow

Here, a degenerate variational integrator is constructed for magnetic field line flow. To obtain the non-canonical Hamiltonian description of the field line dynamics, it is assumed that one of the coordinates  $x^3$  may be used to parameterize the motion in the other dimensions, as shown in Sec. 2.3.4. The parameterization is valid provided  $B^3 \neq 0$  along the field line(s) of interest, as explained in Sec. 2.3.4. Should  $B^3 = 0$  at some point along the trajectory, one option is to switch the independent parameter before reaching this point. To obtain the proper degeneracy condition of Eq. (4.9), it was further helpful to assume either  $A_1 = 0$  or  $A_2 = 0$ . In essence, a gauge has been chosen for the field line dynamics. Although this assumption may not be necessary to obtain a properly degenerate discrete Lagrangian, it is unknown at this time how to obtain the condition without the assumption.

Consider then a magnetic vector potential  $A$  represented in coordinates  $x = (x^1, x^2, x^3)$  in a gauge such that  $A_2 = 0$ , so:

$$A(x) = A_1(x^1, x^2, x^3)dx^1 + A_3(x^1, x^2, x^3)dx^3. \quad (4.19)$$

Recall from Eq. 2.72 that the field line flow action is given by:

$$S = \int \left( A_1(x^1, x^2, x^3) \frac{dx^1}{dx^3} + A_3(x^1, x^2, x^3) \right) dx^3,$$

where  $x^1, x^2$  are functions of  $x^3$ , and  $x^3$  is the time coordinate of the non-canonical Hamiltonian description. In this gauge, the phase-space Euler-Lagrange equations Eq. (2.75) may be written:

$$-A_{1,2}(x)\dot{x}^2 - A_{1,3}(x) + A_{3,1}(x) = 0 \quad (4.20a)$$

$$A_{1,2}(x)\dot{x}^1 + A_{3,2}(x) = 0, \quad (4.20b)$$

where  $\dot{x}^i = \frac{dx^i}{dx^3}$ . This yields a first-order system of differential equations for  $x^1, x^2$  provided the matrix

$$\begin{pmatrix} 0 & -A_{1,2}(x) \\ A_{1,2}(x) & 0 \end{pmatrix}$$

is invertible, i.e.  $-A_{1,2}(x) \sim B^3 \neq 0$ .

A properly degenerate variational integrator emerges from the discretization choice:

$$L_d(x_k, x_{k+1}) = A_1(x_{k+1}) (x_{k+1}^1 - x_k^1) + h A_3(x_{k+1}). \quad (4.21)$$

Checking the proper degeneracy condition Eq. (4.9), one calculates:

$$\text{rank}(D_1 D_2 L_d(x_k, x_{k+1})) = \text{rank} \begin{pmatrix} -A_{1,1}(x_{k+1}) & -A_{1,2}(x_{k+1}) \\ 0 & 0 \end{pmatrix} = 1, \quad (4.22)$$

satisfying the degeneracy condition for this two-dimensional (one degree of freedom) Hamiltonian system. The discrete Euler-Lagrange equations are given by:

$$A_{1,1}(x_k) (x_k^1 - x_{k-1}^1) + A_1(x_k) - A_1(x_{k+1}) + hA_{3,1}(x_k) = 0 \quad (4.23a)$$

$$A_{1,2}(x_k) (x_k^1 - x_{k-1}^1) + hA_{3,2}(x_k) = 0 \quad (4.23b)$$

Based on the proper degeneracy condition of Thm. 4.3.1, one expects a one-step method to result from the properly degenerate discrete Lagrangian. At first glance, Eqs. (4.23) appear to be *multistep*. However, we can recover a one-step method as follows. First, write Eq. (4.23b) at one time index in the future. Second, use Eq. (4.23b) at the original time to identify:

$$(x_k^1 - x_{k-1}^1) = -h \frac{A_{3,2}(x_k)}{A_{1,2}(x_k)}. \quad (4.24)$$

That is, one can *replace*  $(x_k^1 - x_{k-1}^1)$  in Eq. (4.23a) with a function of  $x_k$  *alone*, and thereby eliminate all dependence on  $x_{k-1}$ . Overall, Eq. (4.23) may be represented as a one-step method:

$$-hA_{1,1}(x_k) \left( \frac{A_{3,2}(x_k)}{A_{1,2}(x_k)} \right) + A_1(x_k) - A_1(x_{k+1}) + hA_{3,1}(x_k) = 0 \quad (4.25a)$$

$$A_{1,2}(x_{k+1}) (x_{k+1}^1 - x_k^1) + hA_{3,2}(x_{k+1}) = 0. \quad (4.25b)$$

Recall that  $A_{1,2} \neq 0$  for the field line flow equations to be well-defined, so no restrictions are present on the degenerate variational integrator that are not also present on the continuous Euler-Lagrange equations.

This method is tested in Section 5.2. For the purposes of this chapter, this example illustrates the construction of the one-step formulation and the elimination of parasitic modes by the proper degeneracy condition.

### 4.6.2 Non-Canonical Guiding Center Dynamics

The next example applies a similar technique to construct a properly degenerate variational integrator for non-canonical guiding center dynamics in axisymmetric fields. In the previous example, one component of the magnetic vector potential was assumed to be zero; this amounted to a choice for the electromagnetic gauge. Here, an additional assumption is required: the same (covariant) component of the magnetic field unit vector is assumed to be zero. Although this assumption is undesirably restrictive, the successful construction of a one-step variational integrator exhibiting no parasitic mode instabilities is a meaningful step forward. It is hoped that future developments, aided by the concise and easy-to-check condition of Thm. 4.3.1, will generalize properly degenerate variational integrators to arbitrary magnetic field configurations.

Consider then toroidal coordinates  $x^i = (r, \theta, \zeta)$ , a magnetic vector potential of the form<sup>1</sup>:

$$A(r, \theta, \zeta) = A_\theta(r, \theta)d\theta + A_\zeta(r, \theta)d\zeta, \quad (4.26)$$

with a magnetic field unit vector of the form<sup>2</sup>:

$$b(r, \theta, \zeta) = b_\theta(r, \theta)d\theta + b_\zeta(r, \theta)d\zeta. \quad (4.27)$$

The guiding center (phase-space) Lagrangian is, according to Eq. (2.103e):

$$L_{gc}(x, u, \dot{x}, \dot{u}) = (A_\theta(x) + ub_\theta(x))\dot{\theta} + (A_\zeta(x) + ub_\zeta(x))\dot{\zeta} - \left(\frac{1}{2}u^2 + \mu\mathcal{B}(x) + \phi(x)\right), \quad (4.28)$$

where  $x$  is the guiding center position,  $u$  the parallel velocity,  $\mu$  the magnetic moment, and  $\phi$  the electric potential.

---

<sup>1</sup> $A = A_\theta(r, \theta)\nabla\theta + A_\zeta(r, \theta)\nabla\zeta$

<sup>2</sup> $b = b_\theta(r, \theta)\nabla\theta + b_\zeta(r, \theta)\nabla\zeta$

Consider the discrete Lagrangian given by:

$$L_d(x_0, u_0, x_1, u_1) = h L_{gc}(x_1, u_1, \frac{x_1 - x_0}{h}, \frac{u_1 - u_0}{h}). \quad (4.29)$$

The corresponding discrete Euler-Lagrange equations are:

$$\begin{aligned} A_{\theta,r}^\dagger(x_k, u_k)(\theta_k - \theta_{k-1}) + A_{\zeta,r}^\dagger(x_k, u_k)(\zeta_k - \zeta_{k-1}) - h(\mu \mathcal{B}_{,r}(x_k) + \phi_{,r}(x_k)) &= 0 \\ A_{\theta,\theta}^\dagger(x_k, u_k)(\theta_k - \theta_{k-1}) - (A_\theta^\dagger(x_{k+1}, u_{k+1}) - A_\theta^\dagger(x_k, u_k)) - \\ A_{\zeta,\theta}^\dagger(x_k, u_k)(\zeta_k - \zeta_{k-1}) - h(\mu \mathcal{B}_{,\theta}(x_k) + \phi_{,\theta}(x_k)) &= 0 \\ -A_\zeta^\dagger(x_{k+1}, u_{k+1}) + A_\zeta^\dagger(x_k, u_k) &= 0 \\ b_\theta(x_k)(\theta_k - \theta_{k-1}) + b_\zeta(x_k)(\zeta_k - \zeta_{k-1}) - hu_k &= 0, \end{aligned} \quad (4.30)$$

where  $A^\dagger(x, u) = A(x) + ub(x)$ . Without the result of Thm. (4.3.1), one might be quick to dismiss such an algorithm as multistep based on the presence of variables at times  $t_{k-1}$ ,  $t_k$ , and  $t_{k+1}$ . However, we are assured that no parasitic modes can be present in the dynamics because  $\text{rank}(D_1 D_2 L_d(x_0, u_0, x_1, u_1)) = 2$ , and therefore the chosen discrete Lagrangian is properly degenerate.

Much like the magnetic field line DVI, the lack of parasitic modes is apparent upon construction of a one-step formulation of the time advance. This proceeds by shifting the time indices of the first and fourth equation forward, and replacing  $\theta_{k-1}$  and  $\zeta_{k-1}$  in the second equation by solving for them using the unshifted the first and fourth equations. In detail, we can determine  $\theta_{k-1}, \zeta_{k-1}$  as functions of  $(r_k, \theta_k, \zeta_k, u_k)$  according to:

$$\begin{pmatrix} A_{\theta,r}^\dagger & A_{\zeta,r}^\dagger \\ A_{\theta,u}^\dagger & A_{\zeta,u}^\dagger \end{pmatrix} \begin{pmatrix} \theta_{k-1} \\ \zeta_{k-1} \end{pmatrix} = \begin{pmatrix} A_{\theta,r}^\dagger & A_{\zeta,r}^\dagger \\ A_{\theta,u}^\dagger & A_{\zeta,u}^\dagger \end{pmatrix} \begin{pmatrix} \theta_k \\ \zeta_k \end{pmatrix} - h \begin{pmatrix} H_{gc,r} \\ H_{gc,u} \end{pmatrix}, \quad (4.31)$$

where all  $A^\dagger, H_{gc}$  functions are evaluated at  $r_k, \theta_k, u_k$ . Inverting the matrix on the left hand side:

$$\begin{pmatrix} \theta_{k-1} \\ \zeta_{k-1} \end{pmatrix} = \begin{pmatrix} \theta_k \\ \zeta_k \end{pmatrix} - h \begin{pmatrix} A_{\theta,r}^\dagger & A_{\zeta,r}^\dagger \\ A_{\theta,u}^\dagger & A_{\zeta,u}^\dagger \end{pmatrix}^{-1} \begin{pmatrix} H_{gc,r} \\ H_{gc,u} \end{pmatrix} = \begin{pmatrix} \theta_k \\ \zeta_k \end{pmatrix} - \begin{pmatrix} \Delta^\theta \\ \Delta^\zeta \end{pmatrix}. \quad (4.32)$$

Combining this relation with the index shift, the one-step formulation of the for the degenerate guiding center variational integrator is given by:

$$\begin{aligned} A_{\theta,r}^\dagger(x_{k+1}, u_{k+1})(\theta_{k+1} - \theta_k) + A_{\zeta,r}^\dagger(x_{k+1}, u_{k+1})(\zeta_{k+1} - \zeta_k) - h(\mu \mathcal{B}_{,r}(x_{k+1}) + \phi_{,r}(x_{k+1})) &= 0 \\ A_{\theta,\theta}^\dagger(x_k, u_k)\Delta^\theta(x_k, u_k) - (A_\theta^\dagger(x_{k+1}, u_{k+1}) - A_\theta^\dagger(x_k, u_k)) + \\ A_{\zeta,\theta}^\dagger(x_k, u_k)\Delta_\zeta(x_k, u_k) - h(\mu \mathcal{B}_{,\theta}(x_k) + \phi_{,\theta}(x_k)) &= 0 \\ -A_{\zeta}^\dagger(x_{k+1}, u_{k+1}) + A_{\zeta}^\dagger(x_k, u_k) &= 0 \\ A_{\theta,u}^\dagger(x_{k+1}, u_{k+1})(\theta_{k+1} - \theta_k) + A_{\zeta,u}^\dagger(x_{k+1}, u_{k+1})(\zeta_{k+1} - \zeta_k) - hH_{gc,u}(x_{k+1}, u_{k+1}) &= 0. \end{aligned} \quad (4.33)$$

Clearly, this method cannot contain parasitic modes. Additionally, note that the conservation of toroidal momentum  $p_\zeta = A_\zeta^\dagger$  directly appears as one of the equations of the degenerate variational integrator.

To illustrate the successful elimination of parasitic modes, consider the toroidal tokamak fields of Eq. (5.7) in the unperturbed  $\epsilon = 0$  configuration. Figure (4.2) compares the properly degenerate variational integrator of Eq. (4.33) to the multistep variational integrator of Ref. [16]. The later method is technically degenerate by the staggered  $u$  coordinate; this greatly improves the parasitic mode instabilities relative to the fully multistep variational integrators of Chapter 3. It is not properly degenerate, however, and the retained multistep character leads to parasitic mode instabilities that become problematic after several thousand transit periods. After this duration, the increasing separation of even- and odd-numbered

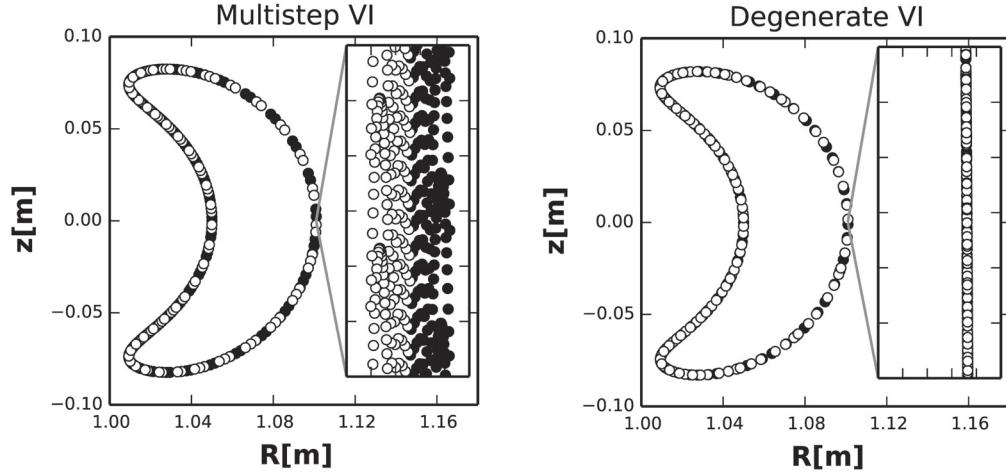


Figure 4.2: Degenerate variational integrators eliminate parasitic modes, as evidenced by this comparison of 2 keV trapped-particle simulations. Even steps are marked with filled circles and odd steps are empty. The phase portrait of the multistep integrator of Ref. [16] exhibits an increasing radial separation over the 0.6 second simulation of the 2 keV trapped particle, whereas the even and odd steps remain on a closed trajectory for the degenerate variational integrator.

steps is apparent for the multistep integrator upon close investigation of the trajectory. In contrast, the degenerate variational integrator retains the desired closed orbit along even- and odd-numbered steps. Although the multistep instability appears small in Fig. 4.2, the multistep variational integrator fails to converge shortly after the simulation duration depicted.

The elimination of parasitic modes and exactly conserved toroidal momentum are strong benefits of the degenerate variational integrator formulation of the guiding center time advance. For a comprehensive evaluation of the benefits of this technique for guiding center dynamics, refer to Section 5.3.

## 4.7 Discussion

### 4.7.1 Summary

Parasitic modes may be eliminated from variational integrators that discretize Hamilton's principle in phase space by constructing properly degenerate discrete Lagrangians. When the rank of the discrete two-form matches that of the symplectic structure, parasitic modes cannot be present. One may therefore consider the invertibility of the Legendre transform to be an important mathematical structure to be retained during the discretization procedure. Although the benefits of properly degenerate variational integrators are strong, constructing such a degenerate discrete Lagrangian is non-obvious in many cases. If two components of the phase-space one-form  $\vartheta$  are zero, a simple forward or backward discretization achieves the desired degeneracy condition, and a one-step variational integrator may be constructed. These one-step variational integrators preserve a discrete two-form on the phase-space manifold  $M$ , a property to which the good behavior observed in the applications of Chapter 5 is attributed.

### 4.7.2 Future Work

Much remains to be done regarding degenerate variational integrators, both to expand theoretical foundations and to enhance applicability to specific problems. In terms of theoretical development, high-priority tasks include: systematic construction of one-step formulations, determining the rigorous implications of the discrete symplecticity, and development of systematic methods for obtaining higher-order accuracy. On the practical impact front, high-priority tasks include: construction of a properly degenerate variational integrator for an arbitrary phase-space Lagrangian, development of higher-order methods for guiding center dynamics, and applying the methodology to altogether new problems with non-canonical Hamiltonian formulations.

One of the first theoretical advancements to be made is systematizing the construction of one-step formulations for PDVIs. In the examples presented in this chapter, the one-step formulations were determined essentially by inspection. Ongoing work is determining whether this is possible for *any* properly degenerate variational integrator or whether additional conditions are necessary for the existence of a one-step formulation. Given Theorem 4.3.1, it seems likely that a one-step formulation exists for any properly degenerate variational integrator. A systematic prescription for constructing the one-step formulation would be beneficial for cases where it is not obvious how to proceed from the DEL equations, and could play a central role in relaxing the field assumptions made on the field line and guiding center degenerate variational integrators.

Another important theoretical task is to establish the rigorous implications of the discrete symplecticity condition of Thm. 4.4.1. The good behavior of conventional symplectic integrators deploys backward error analysis to characterize the numerical dynamics, establishing for instance bounded energy error and the retention of KAM surfaces. This analysis assumes the discrete symplectic structure is *the same* as the continuous symplectic structure. For properly degenerate variational integrators, it seems that the discrete symplectic structure is only nearby to the continuous structure as determined by the step size  $h$ . Given the excellent behavior observed from PDVIs in this chapter and Chapter 5, it is likely the discrete two-form may be leveraged to establish similar rigorous expectations as those on conventional symplectic integrators. One possibility is that PDVIs are *conjugate symplectic*, meaning there exists a coordinate transformation for which they are symplectic in the conventional sense. The discrete symplecticity may be central in ensuring the existence of such a coordinate transformation.

The final noteworthy theoretical objective is establishing statements regarding the accuracy of degenerate variational integrators. For one, proper degeneracy may avoid the shortfall in accuracy observed in Chapter 3. Another open question is how compositions

methods behave for degenerate variational integrators. Because the discrete symplecticity is  $h$ -dependent, it is not clear what two-form might be preserved by the composition of two different degenerate variational integrators. Ensuring the good behavior is retained would allow construction of symmetric and Yoshida-type [36] composed degenerate variational integrators of high order.

In terms of expanding the practical impact of degenerate variational integrators, a substantial breakthrough would be to construct a properly degenerate variational integrator for an arbitrary phase-space Lagrangian. The algorithms so far rely on half of the components of the one-form  $\vartheta$  to be zero; an undesirably restrictive assumption. Clever deployment of new discretization techniques, such as those developed in [101], may obtain the proper degeneracy condition and remain consistent. Implications of this would include a one-step variational integrator for fully general non-canonical guiding center trajectories.

# Chapter 5

## Applications in Plasma Physics

*Shut up and calculate!*

—N. David Mermin, *Physics Today* [125, 126]

In this chapter, extensive numerical demonstrations are performed for multistep and degenerate variational integrators. All demonstrations were generated using the GEometric ODE Solving (GEODES) library. This C++ library was developed for this dissertation. Key features include high performance - stemming from the compiled language and GPGPU parallelization - and ease of use. Additional information about GEODES may be found in Appendix B.

### 5.1 MVI for Nearly-Integrable Non-Canonical Systems

The next example turns to variational integrators for action principles in phase-space. Specifically, phase-space Lagrangians with a special perturbative structure are studied. The integrators exploit this perturbative structure to obtain arbitrarily high order of accuracy in

the small perturbation parameter. The practical benefit of this construction is that the numerical step size is limited by the characteristic bounce time of the *perturbative* effects, rather than that of the unperturbed dynamics (which is typically much shorter). This work was performed in conjunction with Joshua Burby and Hong Qin, and may be found online in Ref. [107].

### 5.1.1 Motivation

A popular numerical technique is to leverage some known “unperturbed” dynamics to more rapidly assess the effects of some small perturbation [127, 128, 129]. For instance, given a time-dependent Hamiltonian  $\mathcal{H}(z, t) = H_t(z) + \epsilon h_t(z)$ , if one knows the flow of the Hamiltonian system when  $\epsilon = 0$ , this can be used to construct an integrator that can take steps of size  $\mathcal{O}(1)$  relative to the small perturbation parameter  $\epsilon$ . Recent work has systematized this procedure to develop  $\mathcal{O}(\epsilon^2)$ -accurate variational integrators [130], encompassing several previous symplectic integration methods for nearly integrable systems [127, 128, 129].

A natural area where this technique could be useful is the study of three-dimensional effects on magnetic field line and test particle dynamics in tokamaks. Active research topics include calculations of three-dimensional magnetic equilibria and the interactions between energetic particles and non-axisymmetric modes. An integrator exploiting the nearly integrable structure of the problem may be put to use, for instance, calculating the Poincaré sections of magnetic equilibria that are perturbations about some axisymmetric configuration for which the field line flow is known analytically. However, because the variational formulations of magnetic field line and guiding center dynamics rely on Hamilton’s principle in phase-space (Section 2.3), the existing perturbative variational integrator methods of e.g. [130] *cannot* be applied. Existing variational integrators for nearly integrable systems require two points in the configuration space to be connected by a solution of the Euler-Lagrange equations; as described in Chapter 2, this is a consequence of regularity of the Lagrangian,

and does not hold for the phase-space Lagrangians governing field line and guiding center dynamics. There therefore exists a need to develop variational integrators for nearly integrable Hamiltonian systems derived from action principles in phase space.

The work presented in this section (and more thoroughly in Ref. [107]) develops variational integrators for nearly integrable systems formulated from Hamilton's principle in phase space. The integrator for the perturbed dynamics may be constructed to arbitrarily high order in the perturbation parameter  $\epsilon$ . The practical consequence is that Poincaré sections for perturbed field line dynamics may be constructed by directly iterating from one point on the surface of section to another (step sizes of  $2\pi$  in the toroidal angle are allowed). The resulting variational integrators are multistep methods, and therefore require a starting procedure and admit parasitic modes. Unfortunately, this reduces the robustness of the method and users must check that the results are satisfactory. For the numerical demonstrations performed here, the parasitic modes remained at small amplitude for the simulated time duration.

### 5.1.2 Algorithm

A foundational component of variational integrator theory is the exact discrete Lagrangian  $L_d^E$ . As described in Section 2.2.5, the discrete Euler-Lagrange equations for the exact discrete Lagrangian are satisfied by trajectories that also satisfy the continuous Euler-Lagrange equations, that is, by true solutions. Of course, explicit formulation of the exact discrete Lagrangian often requires knowledge of the true solution, so the exact discrete Lagrangian must be approximated to obtain a useful integrator. The plan of attack for deriving variational integrators for nearly integrable non-canonical Hamiltonian systems will then be to obtain an exact discrete Lagrangian for the full, perturbative dynamics, then perform approximations that will yield practical variational integrators with a desired order of accuracy.

The exact discrete Lagrangian for nearly integrable non-canonical Hamiltonian systems is derived fully in Ref. [107]. Broadly speaking, the derivation operates by: transforming to an “interaction picture”, in which solutions of the unperturbed system are constant in the transformed path space; manipulating the action functional in the interaction picture to obtain the desired properties; then transforming back from the interaction picture to the original phase space. The properties desired of the exact discrete action  $S_d^E$  include: If  $z(t)$  is a critical point of the phase-space action, then  $z(t_k), k = 0, \dots, 1$  is a critical point of  $S_d^E$ ; The exact discrete action is of the form  $S_d^E(z_0, \dots, z_N) = \sum_{k=0}^{N-1} L_d^E(z_k, z_{k+1}, t_k)$ ; And, replacing  $S_d^E$  with its  $l$ -th order Maclaurin series yields a variational integrator of order  $l$ . The manipulations to construct such an exact discrete action are rather lengthy and technical, so only the end result is presented here.

Consider an exact symplectic manifold  $(M, \Omega = -d\vartheta)$  and a phase-space Lagrangian of the form:

$$L_{PS}(z, \dot{z}, t) = \vartheta_i(z)\dot{z}^i - (H_t(z) + \epsilon h_t(z)). \quad (5.1)$$

Let  $F_{t,s} : M \rightarrow M$  be the time-dependent flow map of the unperturbed Hamiltonian vector field  $X_{H_t}$ ;  $F_{t,s}$  advances a point  $z$  along the flow of  $X_{H_t}$  from time  $s$  to time  $t$ . The exact discrete Lagrangian  $L_d^E : M \times M \rightarrow \mathbb{R}$  for this system is given by:

$$\begin{aligned} L_d^E(z_0, z_1, t) = & f(\Phi_{t+h/2,t}^t(z_0), \Phi_{t+h/2,t+h}^t(F_{t,t+h}(z_1))) + \\ & \int_t^{t+h} \mathcal{L}_s(F_{s,t+h}(z_1)) \, ds + \epsilon \int_t^{t+h/2} \ell_s^t(\Phi_{s,t}^t(z_0)) \, ds + \\ & \epsilon \int_{t+h/2}^{t+h} \ell_s^t(\Phi_{s,t+h}^t(F_{t,t+h}(z_1))) \, ds, \end{aligned} \quad (5.2)$$

where:  $f(z_0, z_1) = \int_{I(z_0, z_1)} \vartheta$ ;  $I(z_0, z_1)$  is the unique geodesic segment directed from  $z_0$  to  $z_1$  according to an affine connection on  $M$ <sup>1</sup>;  $\Phi_{t,s}^u = F_{u,t} \circ \mathfrak{F}_{t,s} \circ F_{s,u}$  is the time-dependent

---

<sup>1</sup>For  $M = \mathbb{R}^{2n}$ ,  $I(z_0, z_1)(\lambda) = (1-\lambda)z_0 + \lambda z_1$  with  $\lambda \in [0, 1]$

flow of the Hamiltonian vector field  $X_{\epsilon F_{t,u}^* h_t}$  with Hamiltonian  $\epsilon F_{t,u}^* h_t$ ;  $\mathfrak{F}_{t,s} : M \rightarrow M$  is the time-dependent flow map of the full, perturbed system;  $\mathcal{L}_s(z) = \vartheta_i(z) X_{H_s}^i(z) - H_s(z)$ ; and  $\ell_s^u(z) = \vartheta_i(z) X_{F_{s,u}^* h_s}^i(z) - F_{s,u}^* h_s(z)$ .

As anticipated, the exact discrete Lagrangian  $L_d^E$  depends on functions that cannot be expected to be known for non-trivial problems. In particular, it presupposes knowledge of the exact solution  $\mathfrak{F}_{t,s}$  in the definition of  $\Phi_{t,s}^u$ . Fortunately, we can expand the expressions involving  $\Phi_{t,x}^u$  using the Lie derivative theorem Eq. (3.50). Truncating the expanded expressions at some order  $l$  in  $\epsilon$  yields an integrator accurate to the same order in  $\epsilon$  [107]. For instance, the zeroth-order discrete Lagrangian is given by:

$$L_d^0(z_0, z_1, t) = f(z_0, F_{t,t+h}(z_1)) + \int_t^{t+h} \mathcal{L}_s(F_{s,t+h}(z_1)) \, ds. \quad (5.3)$$

An explicit expression may be obtained for  $L_d^0$  provided the unperturbed flow map  $F_{t,t+h}$  is known. The discrete Euler-Lagrange equations corresponding to  $L_d^0$  will generate the exact solution to the unperturbed system (see examples below); that is,  $L_d^0$  is (equivalent to) the exact discrete Lagrangian for the unperturbed dynamics. If the unperturbed flow map is not known, a subsidiary expansion may be performed on  $L_d^0$  about the time step parameter  $h$ ; truncations of this system will yield an integrator with the corresponding order of accuracy in the step size  $h$ .

Continuing to one additional order, first-order perturbative effects will be captured by variational integrators obtained from the following discrete Lagrangian:

$$\begin{aligned} L_d^1(z_0, z_1, t) = & \\ & L_d^0(z_0, z_1, t) + \epsilon \int_t^{t+h/2} \ell_s^t(z) \, ds + \epsilon \int_{t+h/2}^{t+h} \ell_s^t(F_{t,t+h}(z_1)) \, ds + \\ & \epsilon \int_t^{t+h/2} \mathcal{L}_{(X_{F_{t,s}^*}, 0)} f(z_0, F_{t,t+h}(z_1)) \, ds - \epsilon \int_{t+h/2}^{t+h} \mathcal{L}_{(0, X_{F_{t,s}^*})} f(z_0, F_{t,t+h}(z_1)) \, ds. \end{aligned} \quad (5.4)$$

Again, all expressions are known provided the solution to the unperturbed dynamics is known. Continued expansion in the parameter  $\epsilon$  may be carried out to obtain a discrete Lagrangian with the desired order of accuracy.

### 5.1.3 Results

#### Perturbed Rigid Rotor

For an instructive example, consider the “perturbed rigid rotor” problem:

$$\begin{aligned} z &= \begin{pmatrix} q \\ p \end{pmatrix} \\ \vartheta(z) &= p \, dq \\ H_t(z) &= \frac{p^2}{2} \\ h_t(z) &= \frac{q^2}{2}. \end{aligned}$$

Essentially, this problem is a harmonic oscillator with frequency  $\sqrt{\epsilon}$ . By considering the quadratic potential to be a perturbation, we can deploy the machinery developed above while obtaining analytic expressions for all of the terms in Eq. (5.2).

First, the unperturbed flow map  $F_{t,s}$  is given by:

$$F_{t,s}(q, p) = (q + (t - s)p, p);$$

the rigid rotor simply translates by  $(t - s)p$ . Next, the flow map for the full system is given by the solution to the harmonic oscillator problem:

$$\mathfrak{F}_{t,s}(q, p) = \begin{pmatrix} q \cos(\sqrt{\epsilon}(t - s)) + \frac{p}{\sqrt{\epsilon}} \sin(\sqrt{\epsilon}(t - s)) \\ p \cos(\sqrt{\epsilon}(t - s)) - q\sqrt{\epsilon} \sin(\sqrt{\epsilon}(t - s)) \end{pmatrix}.$$

Combined, these yield:

$$\Phi_{t,s}^u(q, p) = \begin{pmatrix} (q + p(s - t)) \cos((t - s)\sqrt{\epsilon}) + (q(t - u) + p(\frac{1}{\epsilon} + (t - u)(s - u)))\sqrt{\epsilon} \sin((t - s)\sqrt{\epsilon}) \\ p \cos(\sqrt{\epsilon}(t - s)) - (q + (s - u)p)\sqrt{\epsilon} \sin((t - s)\sqrt{\epsilon}) \end{pmatrix}$$

Evaluating the integrals in Eq. (5.2) obtains the exact discrete Lagrangian:

$$L_d^E(z_0, z_1, t) = -\frac{1}{2} (p_1 q_0 - p_0 q_1) \cos(h\sqrt{\epsilon}) - \frac{1}{2} \left( \frac{1}{\sqrt{\epsilon}} p_0 p_1 + \sqrt{\epsilon} q_0 q_1 \right) \sin(h\sqrt{\epsilon}),$$

where terms that do not contribute to the discrete Euler-Lagrange equations have been neglected. It is straightforward to verify that the discrete Euler-Lagrange equations corresponding to this discrete Lagrangian are satisfied by trajectories generated by  $\mathfrak{F}_{t_0+h,t_0}$ .

### Non-canonical Magnetic Field Line Flow

For a non-canonical example, consider the following problem:

$$\begin{aligned} z &= \begin{pmatrix} x \\ y \end{pmatrix} \\ \vartheta &= (x^2 + y^2) (y dx - x dy) \\ H_t &= \frac{2}{9} (x^2 + y^2)^3 \\ h_t &= x \sin(t) + x^2 \sin(t). \end{aligned}$$

The analogy with magnetic field line flow considers the time  $t$  to label the toroidal coordinate and  $x, y$  to be Cartesian coordinates covering the poloidal plane. The “time-dependent” perturbation Hamiltonian represents small resonant magnetic perturbations.

In this case, the flow of the full, perturbed dynamics is unknown, so  $L_d^E$  cannot be constructed. However, one may determine  $L_d^0$  and  $L_d^1$ , verify their accuracy, and examine their phase portraits. Using the same affine connection as the rigid rotor, the following expression is obtained for  $f$ :

$$f(z_0, z_1) = \frac{1}{3} (x_1 y_0 - x_0 y_1) (x_0^2 + x_0 x_1 + x_1^2 + y_0^2 + y_0 y_1 + y_1^2).$$

Next, the flow map of the unperturbed dynamics may be verified to be:

$$\begin{aligned} F_{t,s}(x, y) = & \left( x \cos \left( \frac{1}{3} (x^2 + y^2)(t - s) \right) + y \sin \left( \frac{1}{3} (x^2 + y^2)(t - s) \right) \right. \\ & \left. - x \sin \left( \frac{1}{3} (x^2 + y^2)(t - s) \right) + y \cos \left( \frac{1}{3} (x^2 + y^2)(t - s) \right) \right). \end{aligned}$$

These maps may be used to determine  $L_d^0$  and  $L_d^1$  explicitly. The discrete Lagrangians  $L_d^0$  and  $L_d^1$  are too lengthy to reproduce here;  $L_d^0$  is manageable, extending several lines, but  $L_d^1$  involves nearly one thousand calls to sin and cos. Nonetheless, these expressions were readily produced by implementing Eq. (5.3) and Eq. (5.4) in symbolic manipulation software.

To test these expressions, the local accuracy was verified by scaling the perturbation parameter  $\epsilon$ . As shown in Fig. 5.1, the integrator generated from  $L_d^0$  successfully reproduced the unperturbed flow, exhibiting first-order in  $\epsilon$  local accuracy. The integrator generated from  $L_d^1$  obtained one degree higher local accuracy.

The phase portraits shown in Fig. 5.2 demonstrate that the lowest order perturbative effects have been captured by the  $L_d^1$  integrator. Here, the variational integrator was able to take steps of size  $2\pi$  in the time (toroidal angle) dimension. That is, the perturbative formulation of the problem allowed direct iteration of points on the Poincaré surface. The resonant perturbations are apparent in the phase portrait of the  $L_d^1$  integrator, evidenced in panel (b). Note that by retaining only the lowest order perturbative effects, higher-

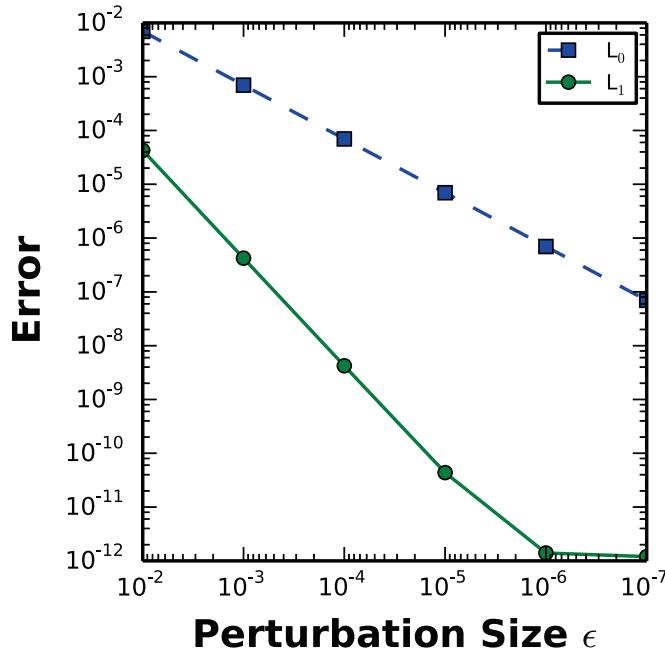


Figure 5.1: The nearly integrable variational integrators demonstrate the rate of convergence anticipated by the local order of accuracy in the perturbation size. The  $L_0$  variational integrator ignores the perturbation; the convergence depicted reflects the convergence of the perturbed dynamics to the unperturbed flow in the zero perturbation limit. Retaining the first order in the perturbation size  $\epsilon$  leads to second order convergence up to the nonlinear solve tolerance limit, as shown by the integrator labeled  $L_1$ .

order resonances are absent from the phase portrait generated by the  $L_d^1$  integrator; this is apparent upon comparison with an accurate integration of the field line flow problem, shown in panel (a). It is expected that higher-order resonances would be accurately portrayed by higher-order variational integrators, such as  $L_d^2$ .

Despite the large step size enabled for the new variational integrators, the excessive number of terms in the discrete Lagrangian functions for all but  $L_d^0$  inhibit the competitiveness of the integrators relative to straightforward integration of the field line equations. In the present implementation, it is advantageous to use accurate methods directly on the field line equations. There is likely room for optimization in the variational integrators for the per-

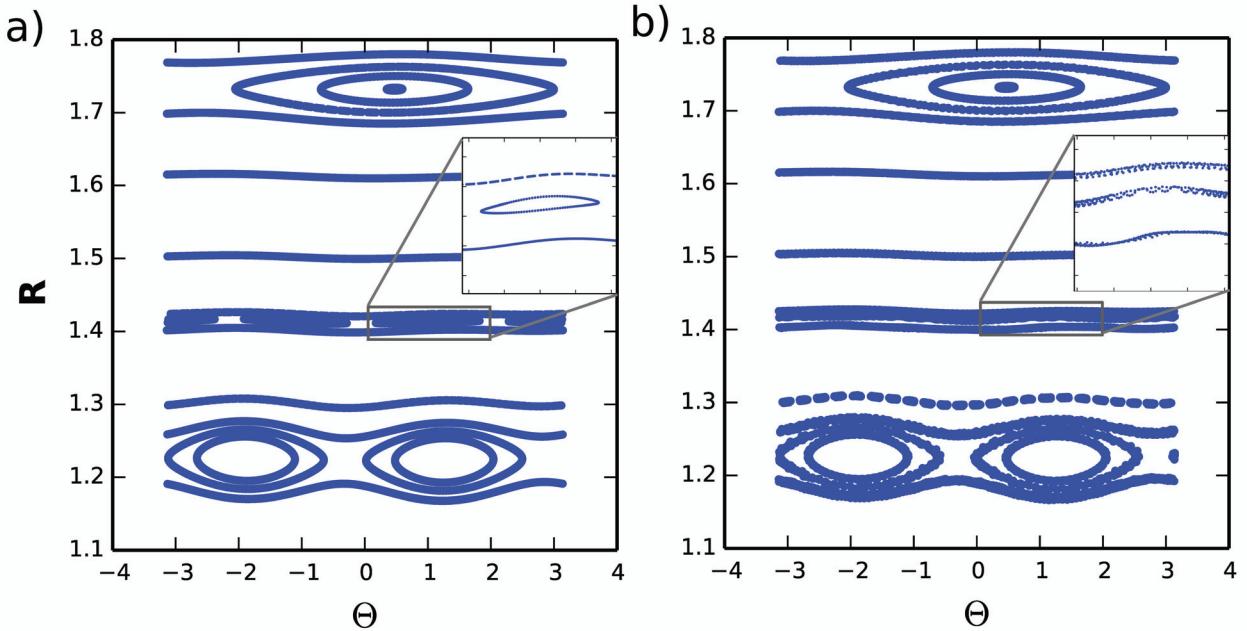


Figure 5.2: The variational integrator for nearly integrable dynamics exhibits the expected phase portrait behavior for magnetic field line flow. In this test case with two resonant magnetic perturbations, the integrator incorporating the  $\mathcal{O}(\epsilon^1)$  effects captures the lowest order resonances, as shown in panel (b). For comparison, an accurate Runge-Kutta simulation of the field line flow is shown in panel (a); the second-order resonances are apparent here.

turbed system, including analytic simplifications, caching re-used results, and approximating the discrete Lagrangians with a reduced set of basis functions.

#### 5.1.4 Discussion

In this application, the degeneracy of the phase-space Lagrangian required reformulating variational integrators for nearly integrable systems. By passing into the interaction picture and performing modifications to the action functional that would not alter its critical points, an exact discrete Lagrangian was obtained for Hamilton's principle in phase-space. Expanding the exact discrete Lagrangian using the Lie derivative theorem yields variational integrators to the desired order of accuracy in the perturbation parameter. An additional expansion in the time step parameter may be performed if the unperturbed flow map is

not known exactly. It is hoped that these methods lay the foundation for rapid assessment of weakly three-dimensional effects in the dynamical systems encountered in magnetically confined plasmas.

At present, the efficacy of the variational integrator for perturbed non-canonical Hamiltonian systems is limited by the parasitic mode instabilities and the complexity of resulting expressions in the perturbed integrator. Further development, such as construction of a degenerate variational integrator for nearly integrable systems and reduction of the computational complexity of the expressions appearing in the discrete Euler-Lagrange equations could deliver competitive methods for nearly integrable magnetic field line flow and any other non-canonical Hamiltonian system.

## 5.2 DVI for Magnetic Field Line Flow

The next example continues with variational integrators for phase-space action principles, illustrating the use of degenerate variational integrators to avoid multistep parasitic mode instabilities. The low-dimensionality and tailorabile complexity of magnetic field line flow makes this non-canonical system an excellent test case for the new methods.

### 5.2.1 Motivation

The dynamics of magnetic field line flow possess a rich Hamiltonian structure [97, 131, 98]. This two degree of freedom system exhibits many general phenomena, including Hamiltonian resonance and chaos. Due to the close relationship between magnetic field line and particle trajectories, these nonlinear phenomena can directly impact experimental objectives [132]. Accurate numerical representation of the Hamiltonian characteristics of magnetic field line flow is thus more than academic curiosity.

Correspondingly, new methods for integrating field line trajectories continue to emerge and be assessed [133, 134, 135]. Typically, these methods respect the *volume-preserving* nature of the field line dynamics - a weaker condition than symplecticity. The absence of symplectic integrators for magnetic field line flow may be attributed to its non-canonical formulation (although transforming to canonical coordinates remains a valid technique). Here, a one-step degenerate variational integrator is presented and assessed. Although the symplectic structure preserved by the DVI differs from the continuous structure (c.f. Eq. (4.14)), excellent long-term behavior is observed.

### 5.2.2 Algorithm

Consider the magnetic field line dynamical system, as defined in Eq. (2.76). Suppose the field line flow may be everywhere parameterized by the coordinate  $x^{i_3}$  (i.e.  $B^{i^3} \neq 0$ ). Further, suppose a gauge has been chosen such that one of the components  $A_{i_1}$  with  $i_1 \neq i_3$  is zero. That is, restrict attention to vector potentials of the form:

$$A = A_{i_2} dx^{i_2} + A_{i_3} dx^{i_3}.$$

A degenerate variational integrator may be constructed for this system by choosing discrete Lagrangian:

$$L_d(x_k, x_{k+1}) = A_{i_2}(x_{k+1})(x_{k+1}^{i_2} - x_k^{i_2}) + h A_{i_3}(x_{k+1}). \quad (5.5)$$

The resulting one-step formulation is given by (cf. Eq. (4.25)):

$$A_{i_2,i_1}(x_{k+1}) (x_{k+1}^{i_2} - x_k^{i_2}) + h A_{i_3,i_1}(x_{k+1}) = 0 \quad (5.6a)$$

$$-h A_{i_2,i_2}(x_k) \left( \frac{A_{i_3,i_1}(x_k)}{A_{i_2,i_1}(x_k)} \right) + A_{i_2}(x_k) - A_{i_2}(x_{k+1}) + h A_{i_3,i_2}(x_k) = 0. \quad (5.6b)$$

Recall that the “time advance” supplements the equation  $x_{k+1}^{i_3} - x_k^{i_3} = h$ .

### 5.2.3 Numerical Results

For a numerical test case, consider a “perturbed tokamak” represented in right-handed toroidal  $(r, \theta, \zeta)$  coordinates, with  $r$  the minor radius,  $\theta$  the geometric poloidal angle, and  $\zeta$  the geometric toroidal angle. The magnetic vector potential  $A = A_r dr + A_\theta d\theta + A_\zeta d\zeta$ <sup>2</sup> and magnetic field  $B = B^r \partial_r + B^\theta \partial_\theta + B^\zeta \partial_\zeta$ <sup>3</sup> are given by:

$$A_r = 0 \quad (5.7a)$$

$$A_\theta = \frac{B_0 R_0 r \cos(\theta) - B_0 R_0^2 \log(1 + \frac{r}{R_0} \cos(\theta))}{\cos^2(\theta)} \quad (5.7b)$$

$$A_\zeta = -\frac{B_0 r^2}{2q} + \sum_i \epsilon_i B_0 R_0 r \sin(m_i \theta - n_i \zeta), \quad (5.7c)$$

$$B^r = \sum_i \frac{\epsilon_i m_i B_0 R_0 \cos(m_i \theta - n_i \zeta)}{R_0 + r \cos(\theta)} \quad (5.8a)$$

$$B^\theta = \frac{B_0}{q(R_0 + r \cos \theta)} - \sum_i \frac{\epsilon_i B_0 R_0 \sin(m_i \theta - n_i \zeta)}{r(R_0 + r \cos \theta)} \quad (5.8b)$$

$$B^\zeta = -\frac{B_0 R_0}{(R_0 + r \cos \theta)^2}, \quad (5.8c)$$

where  $B_0$  is the magnetic field magnitude,  $R_0$  is the major radius,  $q$  is the on-axis safety factor, and  $\epsilon_i$  is the size of the  $i$ -th perturbation with poloidal mode number  $m_i$  and toroidal mode number  $n_i$ . A DVI for this system corresponds to Eq. (5.6) with  $(x^1, x^2, x^3) = (r, \theta, \zeta)$  and  $(i_1, i_2, i_3) = (1, 2, 3)$ .

The first demonstration is that the degenerate variational integrator is first-order accurate in the step size  $h$ . The lack of symmetry in the discretization is apparent; correspondingly,

---

<sup>2</sup>In vector calculus notation,  $A = A_r \nabla r + A_\theta \nabla \theta + A_\zeta \nabla \zeta$

<sup>3</sup> $(\partial_r, \partial_\theta, \partial_\zeta)$  are the basis vectors for contravariant vectors; they are dual to the covariant basis  $(dr, d\theta, d\zeta)$  or  $(\nabla r, \nabla \theta, \nabla \zeta)$

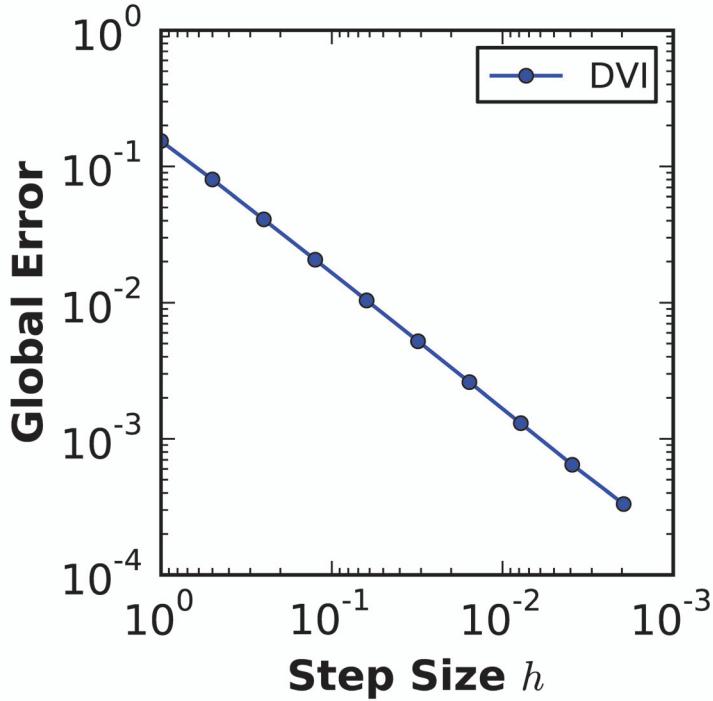


Figure 5.3: The degenerate variational integrator for magnetic field line flow is first-order accurate in the step size  $h$ .

we cannot expect second order accuracy. Indeed, Fig. 5.3 shows the method to be only first-order accurate like the forward-, backward-, and symplectic-Euler schemes.

Despite the low order of accuracy, good numerical fidelity is obtained. For instance, consider a single ( $\epsilon_1 = 10^{-4}, m_1 = 3, n_1 = 0$ ) perturbation and an on-axis safety factor value of  $q = \sqrt{2}$ . The perturbation introduces a  $B^r$  component to the magnetic field, leading to nontrivial energy conservation behavior (when  $B^r = 0$ ,  $r(t)$  is constant and therefore the Hamiltonian  $H = -A_\zeta = \frac{B_0 r^2}{2q}$  is exactly conserved by any integration method based on the vector field). By retaining axisymmetry, the system remains “time independent”, so conservation of the Hamiltonian is a useful invariant for evaluating numerical methods. As depicted in Fig. 5.4, the first-order accurate degenerate variational integrator outperforms the fourth-order Runge-Kutta method in the long term dynamics. In the figure shown, the first-order variational integrator becomes advantageous over  $10^5$  toroidal transits, when

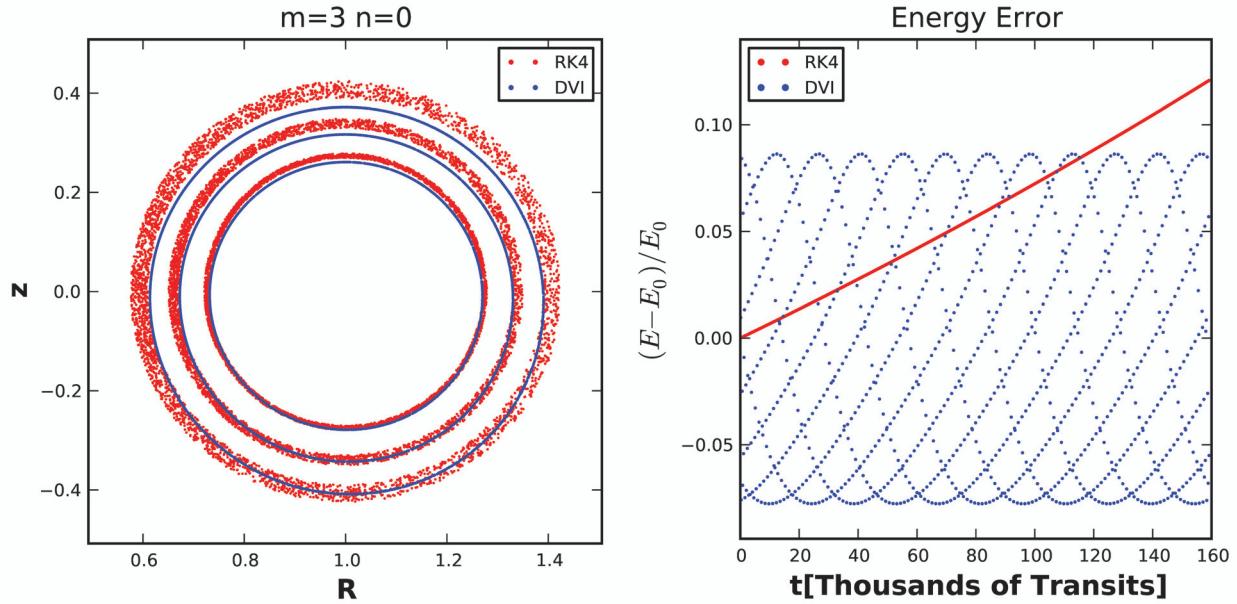


Figure 5.4: The first-order accurate DVI outperforms the fourth-order Runge-Kutta advance for the magnetic field line flow problem. Both integrators used a step size of  $\tau/6$ , where  $\tau = 2\pi$  is the transit period in radians. On the left, the phase portrait shows the DVI respects the closed orbits whereas the dissipative RK4 method allows a drift. On the right, the “energy” error is depicted, where the energy for this system is (minus) the  $\zeta$  component of the vector potential. The DVI exhibits bounded energy error; the appearance of multiple trajectories is an aliasing artifact from downsampling the data.

using equal step sizes. As the implicit integrator is more computationally expensive, similar advantages should be apparent after  $10^7$  toroidal transits using equal CPU expense (allowing three times as many steps in the same time interval for the explicit RK4 method). The Newton-Raphson iterations to solve the implicit update for the DVI Eq. (5.6) typically converge within one or two iterations, but is still more expensive than the explicit *RK4* method. Independent of the timescales over which the variational method is advantageous, the qualitatively different behavior is apparent.

Increasing the complexity of the test cases, consider now a resonant perturbation with ( $\epsilon_1 = 10^{-4}, m_1 = 3, n_1 = 2$ ). With an on-axis safety factor of  $q = \sqrt{2}$ , this resonates with the  $q = \frac{3}{2}$  rational surface located at  $r = \frac{1}{3}$ . Figure 5.5 portrays the phase portrait after

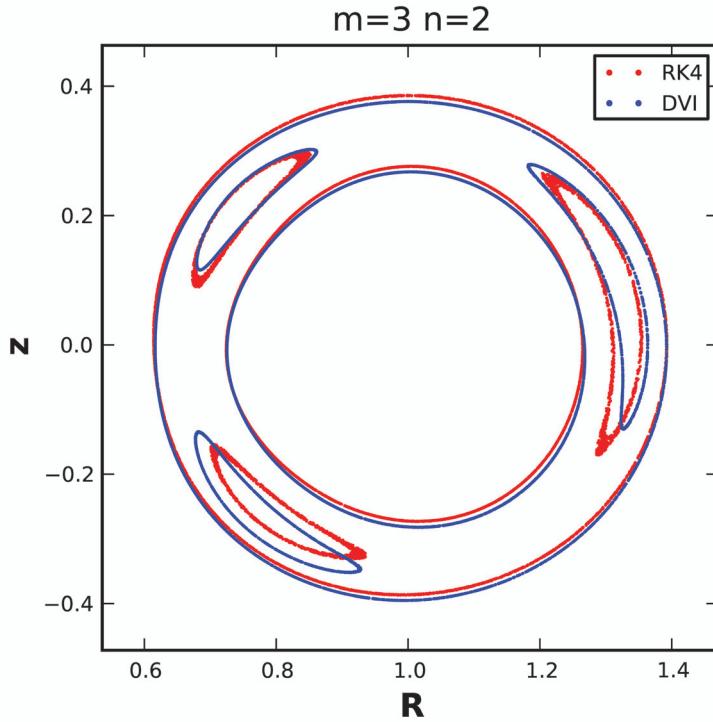


Figure 5.5: The magnetic field line DVI captures the single resonant perturbation phase portrait well over long times. Although the large step size leads to a change in the location of the islands, the qualitative (closed trajectory) behavior is well retained.

several thousand toroidal transits. Using the same parameters (except mode numbers) as the previous run, the portrait fidelity also favors the degenerate variational integrator.

The final comparison employs two resonant perturbations to introduce stochastic regions, specifically with parameters ( $\epsilon_1 = 3.5 \times 10^{-4}$ ,  $m_1 = 3$ ,  $n_1 = 2$ ,  $\epsilon_2 = 3.5 \times 10^{-4}$ ,  $m_2 = 7$ ,  $n_2 = 5$ ). The retention of - and significance of retaining - KAM surfaces is apparent in the phase portraits of Fig. 5.6. It appears no closed trajectories are present in the phase portrait generated by RK4.

#### 5.2.4 Discussion

The degenerate variational integrator is an interesting and useful new tool for calculating field line trajectories. The phase portraits reveal that the integrator captures the correct qualita-

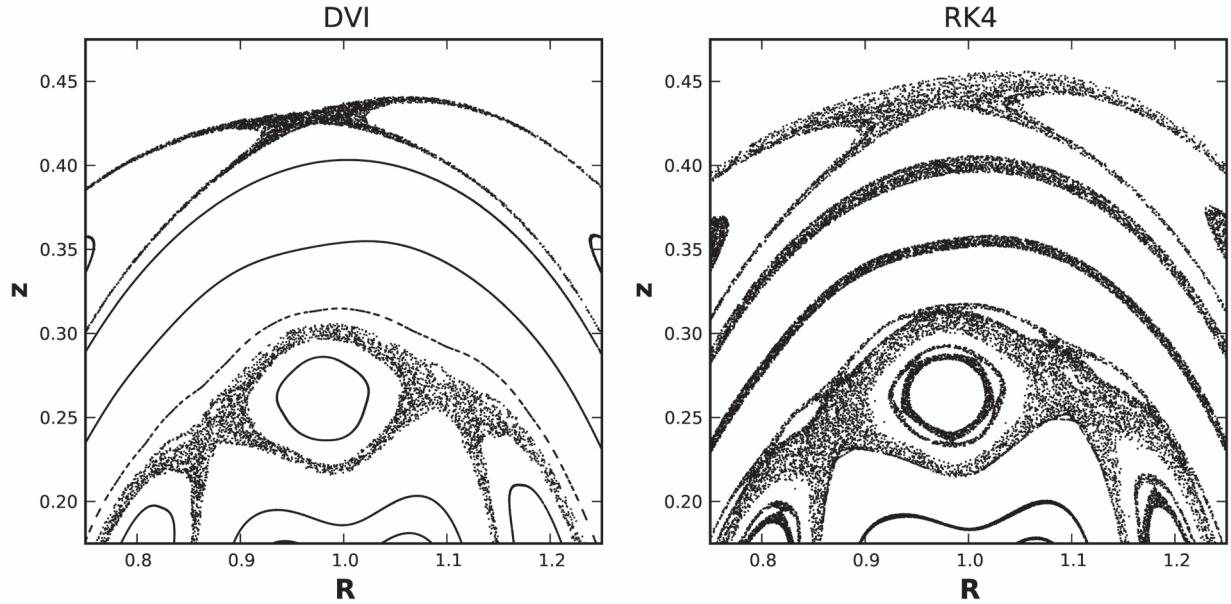


Figure 5.6: When stochastic regions are present, retention of KAM surfaces is important for accurate representation of the phase portrait. The degenerate variational integrator succeeds at retaining the invariant tori, leading to distinct regions of integrable and non-integrable trajectories. In contrast, RK4 will eventually fill all empty regions of the Poincaré section. The step size was 0.5 radians per step for both methods (not a fair computational expense comparison) and a simulation time of  $2 \times 10^5$  transits per initial condition.

tive behavior. This can be especially significant when attempting to distinguish stochastic regions from integrable regions over long timescales. Given the low order of accuracy of the DVI method, its good behavior must be attributed to its variational construction and preservation of a symplectic structure nearby to the original.

Notably, the magnetic field line DVI lacks parasitic mode instabilities. If one were to perform a naive discretization of the magnetic field line phase-space Lagrangian Eq. (2.71), the discrete Lagrangian will typically be regular and result in a multistep variational integrator. The parasitic mode instabilities would be likely to interfere with the bounded energy error and closed trajectories. With an eye toward achieving the “proper degeneracy condition”, defined in Sec. 4.2 and characterized in Thm. 4.3.1, choosing a gauge that eliminates one

component of the magnetic vector potential makes it easier to build the degeneracy into the discrete Lagrangian.

The comparisons here were performed at equal step size, using relatively large steps. The unphysical behavior will take substantially (one hundred times) longer to accumulate to similar magnitudes when performing an equal run time comparison. Depending on the application, high-order non-geometric methods are likely to be satisfactory for a given computational cost. However, further developments that increase the accuracy of the DVI or decrease the run time could make the method highly effective for practical studies even on moderate timescales (path lengths).

### 5.3 DVI for Guiding Center Trajectories

The final application turns to the original motivator for this work: the guiding center motion of magnetized particles. Because the importance of test particle and kinetic simulations has been discussed previously, no additional motivation is necessary for this section beyond that presented in Chapter 1 and elsewhere. A few comments are appropriate, however, regarding the comparisons performed in this section verses alternative algorithms.

The demonstrations performed here serve two purposes. The first purpose is to verify the degenerate variational integrators exhibit the desired long-term behavior. The second purpose is to determine parameter regimes in which it is beneficial to use the degenerate variational integrators in practice. The user is interested in how the methods perform given equal computational expense rather than equal step sizes. It is important to keep in mind, then, how adjusting the numerical step size impacts the numerical solution.

Recall that, given an integrator accurate to order  $r$  in the numerical step size  $h$ , the global error  $\delta$  scales like  $h^r$ . Consider an order- $r$  integrator and let  $\delta_0$  denote the global error for a given simulation time using step size  $h_0$ , so  $\delta_0 \sim h_0^r$ . Now consider adjusting the step

size to  $h_1$ . Then the ratio of the global errors scales as:

$$\frac{\delta_1}{\delta_0} \sim \left( \frac{h_1}{h_0} \right)^r. \quad (5.9)$$

For high-order integrators, such as  $r = 4$ , this both means: the error is rapidly reduced as the step size decreases; and, the error is rapidly increased as the step size increases. To obtain a fixed error reduction over a given time interval requires a step size reduction that is only *logarithmic* in the reduction factor, with base  $r$ . However, if the step size is increased by a fixed factor, the global error will increase by that factor raised to the  $r$ -th *power*.

When evaluating low-order conservative algorithms relative to high-order non-conservative algorithms, the fundamental question is: over what timescale does it become advantageous to use the conservative method provided the step sizes are chosen to allow equal computational expense across the methods? The strong sensitivity of the global error to the step size means that small changes in the conditions of the comparison can propagate into large changes in the final result. A factor of two in optimization or acceptable error tolerances can propagate to a factor of sixteen in the timescale over which the conservative algorithm is competitive. In favor of developing improved conservative algorithms is that small improvements in the time to solution will drastically reduce the timescale over which they are advantageous. In defense of the traditional, explicit, high-order Runge-Kutta methods, the undesirable behavior can often be deferred to acceptably long simulation times using moderate increases in computational expense. Still, it is desirable to have a robust algorithm so that the user does not need to invest time adjusting numerical step sizes to obtain satisfactory behavior. The results of this section should be evaluated according to these considerations.

### 5.3.1 Algorithms

The degenerate variational integrator for guiding center trajectories demonstrated in this section makes the following assumption about the fields: *there exists one covariant component of the magnetic field that is zero, and a gauge has been chosen such that the same (covariant) component of the magnetic vector potential is zero.* In short, this particular construction of a degenerate variational integrator requires one (covariant) component of  $A^\dagger$  to be zero. In general, it is not apparent that this is a necessary restriction for constructing a degenerate variational integrator, but it is presently unknown how to construct a DVI for arbitrary fields. Despite this restriction, the algorithm may be applied to many scientifically pertinent field configurations, including those possessing nested flux surfaces. It will be shown how to construct such an algorithm in the case when flux coordinates are known.

#### Collisionless DVI-GC Algorithm

Consider a guiding center particle described by the system Eq. (2.103) in coordinates  $z = (x^1, x^2, x^3, u)$  with vector potential  $A = A_1 dx^1 + A_2 dx^2 + A_3 dx^3$  and magnetic field unit vector  $b = b_1 dx^1 + b_2 dx^2 + b_3 dx^3$  with  $A_1 = b_1 = 0$ . The degenerate variational integrator stems from choosing a discrete Lagrangian:

$$\begin{aligned} L_d(z_k, z_{k+1}) &= h L_{GC}(z_{k+1}, \frac{z_{k+1} - z_k}{h}) \\ &= A_2^\dagger(z_{k+1})(x_{k+1}^2 - x_k^2) + A_3^\dagger(z_{k+1})(x_{k+1}^3 - x_k^3) - h H_{gc}(z_{k+1}). \end{aligned} \quad (5.10)$$

Following the procedure of Chapter 4, the one-step formulation of the degenerate variational integrator for guiding center trajectories (DVI-GC) is:

$$A_{2,1}^\dagger(z_{k+1})(x_{k+1}^2 - x_k^2) + A_{3,1}^\dagger(z_{k+1})(x_{k+1}^3 - x_k^3) - h(\mu B_{,1}(x_{k+1}) + \phi_{,1}(x_{k+1})) = 0 \quad (5.11a)$$

$$A_{2,2}^\dagger(z_k)\Delta^2 + A_{3,2}^\dagger(z_k)\Delta^3 - (A_2^\dagger(z_{k+1}) - A_2^\dagger(z_k)) - h(\mu B_{,2}(x_k) + \phi_{,2}(x_k)) = 0 \quad (5.11b)$$

$$A_{2,3}^\dagger(z_k)\Delta^2 + A_{3,3}^\dagger(z_k)\Delta^3 - (A_3^\dagger(z_{k+1}) - A_3^\dagger(z_k)) - h(\mu B_{,3}(x_k) + \phi_{,3}(x_k)) = 0 \quad (5.11c)$$

$$b_2(x_{k+1})(x_{k+1}^2 - x_k^2) + b_3(x_{k+1})(x_{k+1}^3 - x_k^3) - hu_{k+1} = 0, \quad (5.11d)$$

where  $\Delta$  is given by:

$$\begin{pmatrix} A_{2,1}^\dagger(z_k) & A_{3,1}^\dagger(z_k) \\ b_2(x_k) & b_3(x_k) \end{pmatrix} \begin{pmatrix} \Delta^2 \\ \Delta^3 \end{pmatrix} = h \begin{pmatrix} \mu B_{,1}(x_k) + \phi_{,1}(x_k) \\ u_k \end{pmatrix}.$$

For reference, this algorithm may be compared to the guiding center phase-space Euler-Lagrange equations, Eq. (2.104). Taylor expansion of the algorithm about time  $t_k$  reveals that the method is first-order accurate.

### Collisionless DVI-GC Algorithm in Flux Coordinates

Consider a magnetic configuration with nested magnetic surfaces and described by a set of flux coordinates  $x^i = (\Psi, \theta, \zeta)$ , which satisfy <sup>4</sup>:

$$A = \Psi d\theta + A_\zeta(\Psi)d\zeta, \quad (5.12)$$

and the straight field line condition:

$$\frac{d\theta}{d\zeta} = \frac{1}{q(\psi)} = \frac{1}{A_{\zeta,\Psi}(\Psi)}. \quad (5.13)$$

---

<sup>4</sup> $A = \Psi \nabla \theta + A_\zeta(\Psi) \nabla \zeta$

In these coordinates, the guiding center Lagrangian is given by:

$$L_{gc}(z, \dot{z}) = ub_\Psi(x)\dot{\Psi} + (\Psi + ub_\theta(x))\dot{\theta} + (A_\zeta(\Psi) + ub_\zeta(x))\dot{\zeta} - H_{gc}(z). \quad (5.14)$$

To obtain the desired form of the Lagrangian for constructing the degenerate variational integrator, the  $ub_\Psi\dot{\Psi}$  term needs to be eliminated. There exist two possibilities for this. The first is that one may perform a re-definition of the guiding center, which remains nearby to the original, that simply eliminates this term from the Lagrangian [57]. The second possibility is to integrate the term by parts, incorporating the  $\frac{d}{dt}(ub_\Psi)$  term into the definition of the poloidal [46] or toroidal [77] angle. With the first approach, the DVI constructed above may be directly applied. With the latter, one must map between the poloidal angle in flux coordinates and new poloidal angle each step.

Taking the first approach, neglecting the  $b_\Psi$  term in the guiding center Lagrangian constitutes a re-definition of the guiding center [57, 78] and obtains the necessary format for constructing the DVI of Eq. (5.11). In the flux coordinates, the method becomes:

$$(1 + u_{k+1}b_{\theta,\zeta}(x_{k+1}))(\theta_{k+1} - \theta_k) + (A_{\zeta,\Psi}(x_{k+1}) + u_{k+1}b_{\zeta,\Psi}(x_{k+1}))(\zeta_{k+1} - \zeta_k) - h(\mu B_{,\Psi}(x_{k+1}) + \phi_{,\Psi}(x_{k+1})) = 0 \quad (5.15a)$$

$$u_k b_{\theta,\theta}(x_k) \Delta^\theta + u_k b_{\zeta,\theta}(x_k) \Delta^\zeta - (A_\theta^\dagger(z_{k+1}) - A_\theta^\dagger(z_k)) - h(\mu B_{,\theta}(x_k) + \phi_{,\theta}(x_k)) = 0 \quad (5.15b)$$

$$u_k b_{\theta,\zeta}(x_k) \Delta^\theta + u_k b_{\zeta,\zeta}(x_k) \Delta^\zeta - (A_\zeta^\dagger(z_{k+1}) - A_\zeta^\dagger(z_k)) - h(\mu B_{,\zeta}(x_k) + \phi_{,\zeta}(x_k)) = 0 \quad (5.15c)$$

$$b_\theta(x_{k+1})(\theta_{k+1} - \theta_k) + b_\zeta(x_{k+1})(\zeta_{k+1} - \zeta_k) - hu_{k+1} = 0. \quad (5.15d)$$

where  $z = (\Psi, \theta, \zeta, u)$ ,  $A^\dagger = A + ub$ , and:

$$\begin{pmatrix} 1 + ub_{\theta,\zeta}(x_k) & A_{\zeta,\Psi}(x_k) + u_k b_{\zeta,\Psi}(x_k) \\ b_2(x_k) & b_3(x_k) \end{pmatrix} \begin{pmatrix} \Delta^2 \\ \Delta^3 \end{pmatrix} = h \begin{pmatrix} \mu B_{,\Psi}(x_k) + \phi_{,\Psi}(x_k) \\ u_k \end{pmatrix}.$$

### Collisional Drag DVI-GC Algorithm

The degenerate variational integrator for guiding center trajectories subject to collisional drag stems from a discrete forced variational principle of the form:

$$\delta \sum_{k=0}^{N-1} h L_{gc}(z_{k+1}, \frac{z_{k+1} - z_k}{h}) - \sum_{k=0}^{N-1} \nu \frac{v_k^3 + v_c^3}{v_k^3} (u_k b_2(x_k) \delta x_k^1 + u_k b_3(x_k) \delta x_k^2 + 2\mu \delta \Theta_k). \quad (5.16)$$

Let  $z = (x, u, \mu, \Theta)$  and let  $f^{Drift}(z_k, z_{k+1})$  be the update rule of the one-step formulation of the discrete Euler-Lagrange equations for the collisionless guiding center DVI, i.e. the left hand side of Eq. (5.11). The forced discrete Euler-Lagrange equations are then given by:

$$f_1^{Drift}(z_k, z_{k+1}) = 0 \quad (5.17a)$$

$$f_2^{Drift}(z_k, z_{k+1}) - h\nu \frac{v_k^3 + v_c^3}{v_k^3} u_k b_2(x_k) = 0 \quad (5.17b)$$

$$f_3^{Drift}(z_k, z_{k+1}) - h\nu \frac{v_k^3 + v_c^3}{v_k^3} u_k b_3(x_k) = 0 \quad (5.17c)$$

$$f_4^{Drift}(z_k, z_{k+1}) = 0 \quad (5.17d)$$

$$-\mu_{k+1} + \mu_k - 2h\nu \frac{v_{k+1}^3 + v_c^3}{v_{k+1}^3} \mu_{k+1} = 0. \quad (5.17e)$$

#### 5.3.2 Numerical Results

All of the comparisons of this section use an unperturbed version of the tokamak fields of Eqs. (5.7) and (5.8), that is, the perturbation parameters  $\epsilon_i$  are chosen to be zero. For ITER-like conditions,  $B_0 = 5.3$  T and  $R_0 = 6.2$  m have been chosen. Additionally, the on-axis safety factor  $q$  has been chosen to be  $\sqrt{2}$ . The particle parameters have been chosen to correspond to a 1 MeV deuteron, as will be injected via neutral beam heating in ITER. The results are qualitatively similar for a 3.5 MeV alpha particle, with minor adjustments to the timescales over which the comparisons yield analogous results.

### Collisionless Results

The benefits of retaining the non-canonical coordinates in the formulation of the time advance may be measured by comparing the time to solution for different integration methods. Figure 5.7 presents this measurement, comparing the computational expense per step for: explicit fourth-order Runge-Kutta, the degenerate variational integrator, and the canonical symplectic integration scheme of Zhang et al. [79]. The method in Ref. [79] constructs canonical coordinates for the guiding center system and uses implicit midpoint for the time advance in these canonical coordinates. Although the canonical symplectic integration scheme obtains the desired conservation properties, one must essentially invert the canonical coordinate transformation each step to evaluate the field functions at the proper location. This doubles the dimension of the nonlinear solve of the implicit time advance equations, and therefore requires double the computational expense relative to the DVI method. The reduction in time to solution obtained by the DVI will enhance its competitiveness with the RK4 method when equal CPU-expense comparisons are performed. For the comparisons in this section, a factor of 4 will be used as an equal CPU expense comparison between RK4 and the DVI.

The degenerate variational integrator for guiding center trajectories is confirmed to be first-order accurate in Fig. 5.8.

The first set of comparisons tests the qualitative behavior by comparing RK4 and the DVI near the maximum step size for which the DVI converges. Consider then the 1 MeV deuteron confined in ITER-like fields described above, with a trapped-particle pitch  $\lambda = \frac{u}{\|v\|} = 0.2$ . Fig. 5.9 shows the rapid deterioration of the RK4 banana orbit alongside the properly closed trajectory of the DVI-GC method. The orbit fidelity may be assessed from either the phase portrait or the energy and momentum invariants. The exact conservation of toroidal momentum  $A_\zeta^\dagger = A_\zeta + ub_\zeta$  of DVI-GC may be observed, in contrast to the accumulated toroidal momentum error for the RK4 method of over 7% during the 4 ms interval. This conservation results from a discrete Noether's theorem [9] and the lack of the

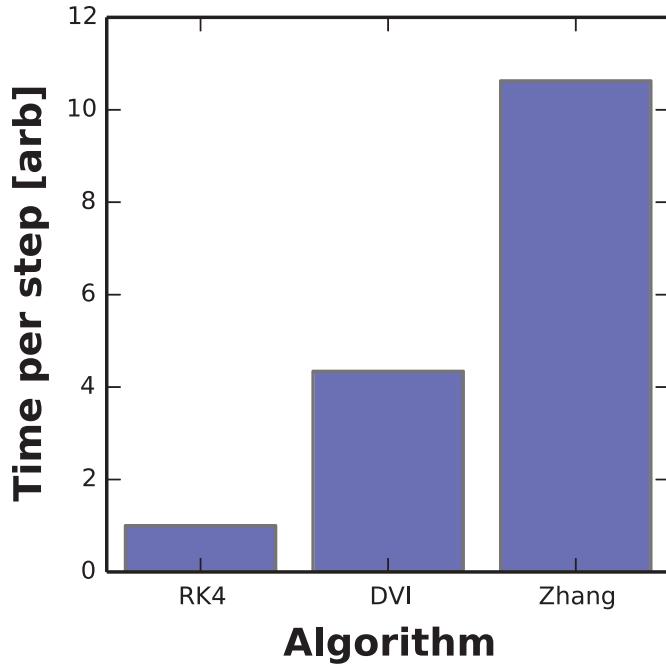


Figure 5.7: Relative computational expense per step of several guiding center integrators: fourth-order Runge-Kutta (RK4), degenerate variational integrator (DVI), and the canonical symplectic integration method of Zhang et al. [79] (Zhang). The DVI method is the fastest of the methods known to preserve a discrete two-form.

toroidal coordinate  $\zeta$  in the discrete Lagrangian. The energy error of the DVI-GC method exhibits the bounded energy error characteristic of symplectic algorithms, whereas the energy error accumulates without bound for the non-geometric RK4 algorithm, accumulating to 6% over this time interval.

Similar behavior may be observed for the 1 MeV deuteron initialized with a passing pitch of  $\lambda = 0.4$ . The phase portrait comparison is shown in Fig. 5.10. Notably, the trajectory generated by the non-variational algorithm becomes trapped over the course of the simulation as a result of the accumulated numerical error. Although the portrait is distorted for the variational integrator (the numerical trajectory is on constant toroidal

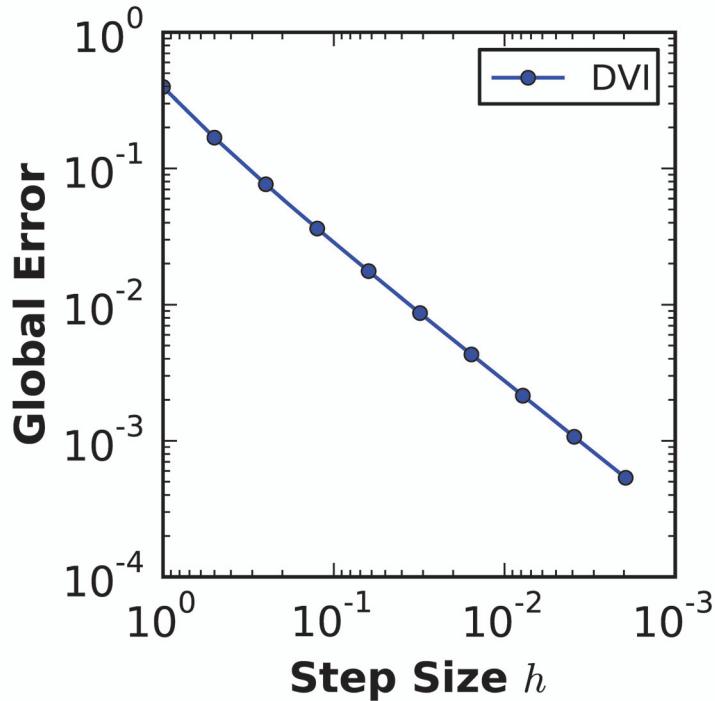


Figure 5.8: The degenerate variational integrator for guiding center dynamics is verified to be first-order accurate in the numerical step size  $h$ .

momentum  $A_\zeta^\dagger$  surfaces but not on constant energy surfaces), it retains the correct trapping status indefinitely.

Now that the expected numerical fidelity has been confirmed, a more practical comparison can be performed. Based on the time to solution study in Fig. 5.7, RK4 will use steps of size one fourth that of the DVI. The trapped-particle results are presented in Fig. 5.11 and passing particle results in Fig. 5.12. These equal expense comparisons bear strong resemblance to the equal step size comparisons; the notable distinction is the timescale over which the numerical error accumulates for RK4. The equal expense studies are approximately one thousand times longer than the equal step size studies.

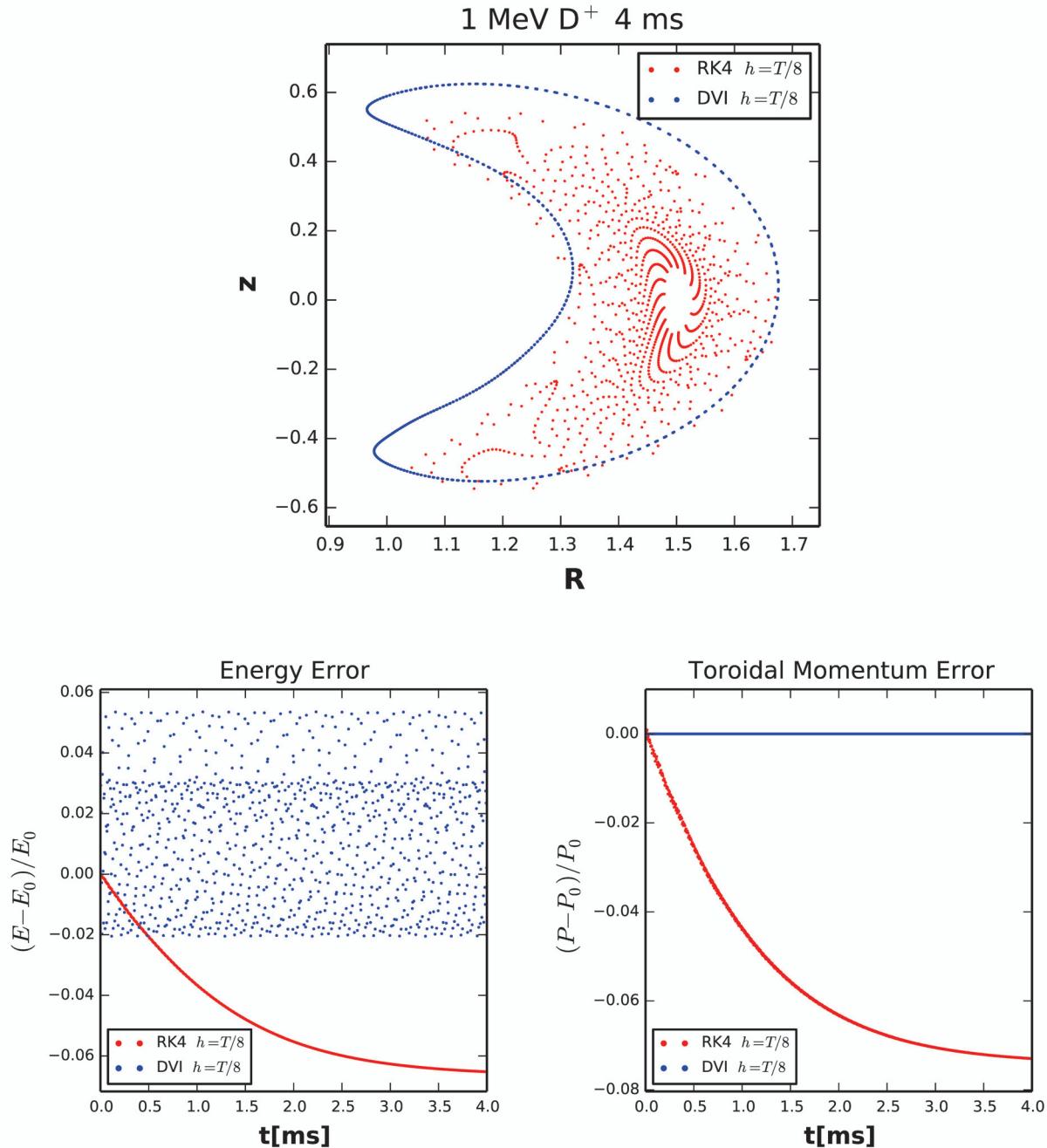


Figure 5.9: Equal step size comparison of a trapped 1 MeV deuteron. Here  $T$  is the period of the trapped-particle orbit.

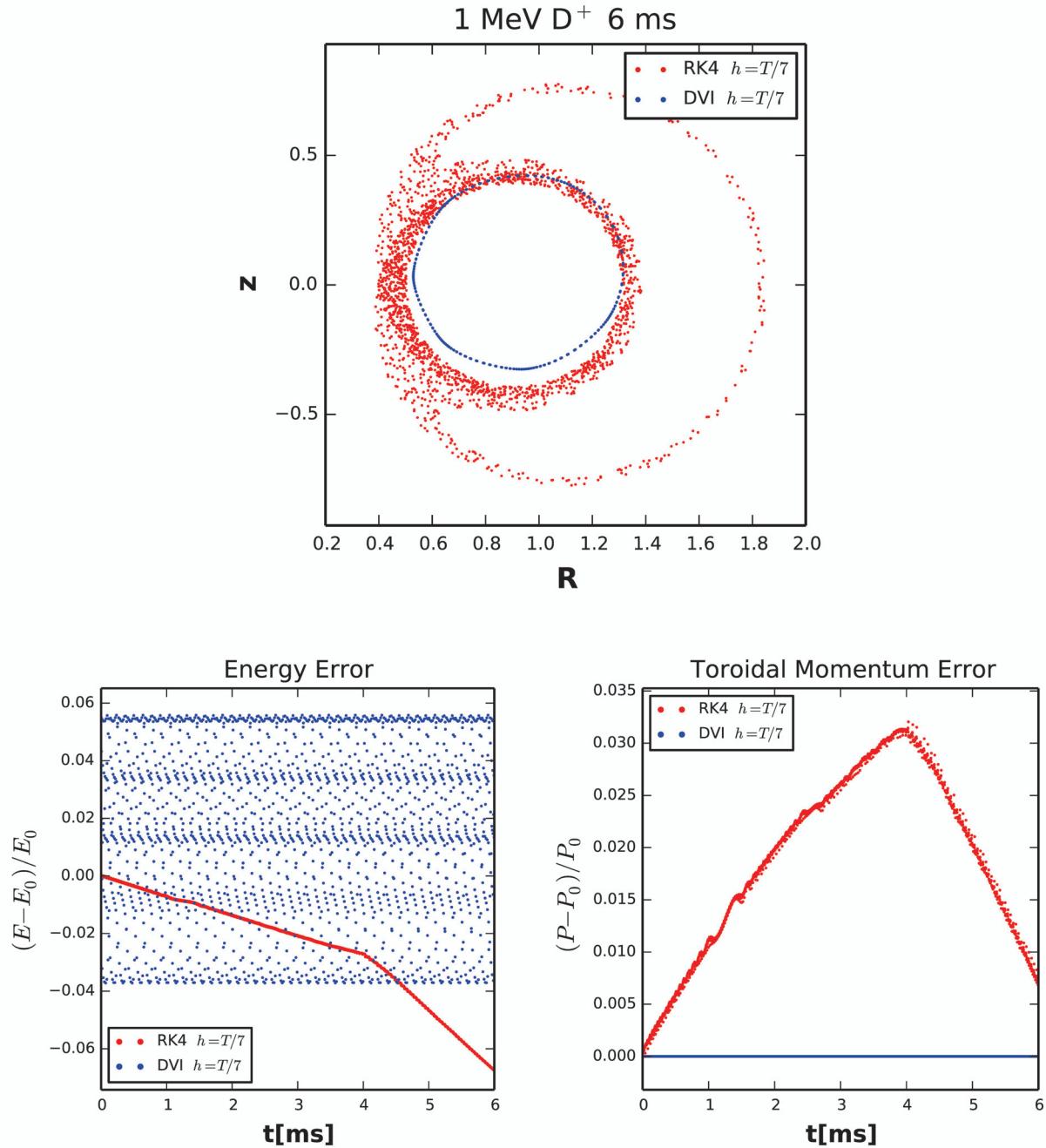


Figure 5.10: Equal step size comparison of a passing 1 MeV deuteron, where  $T$  is the period of the passing particle orbit.

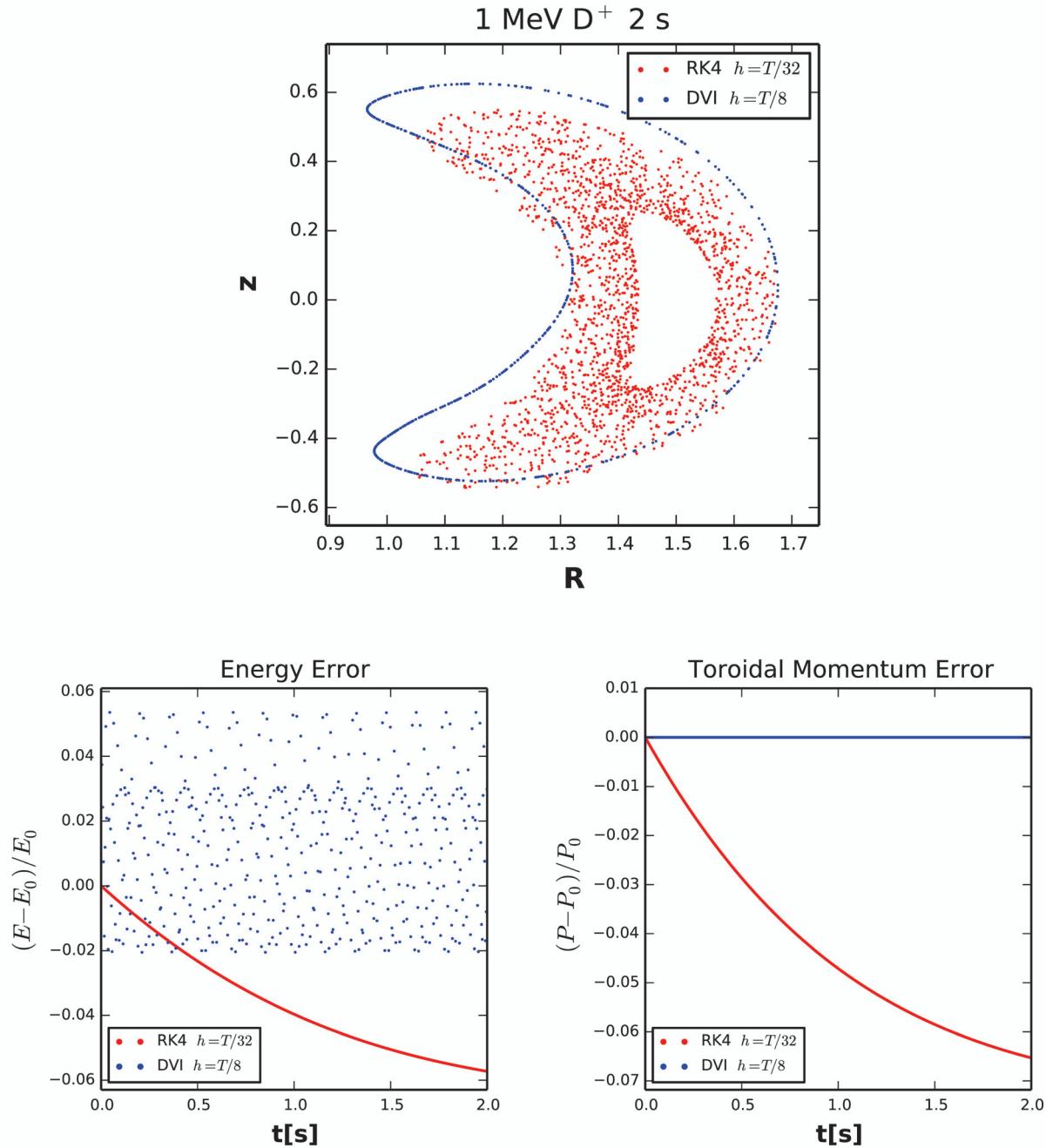


Figure 5.11: Equal computational expense comparison of RK4 and the DVI method for a trapped 1 MeV deuteron.

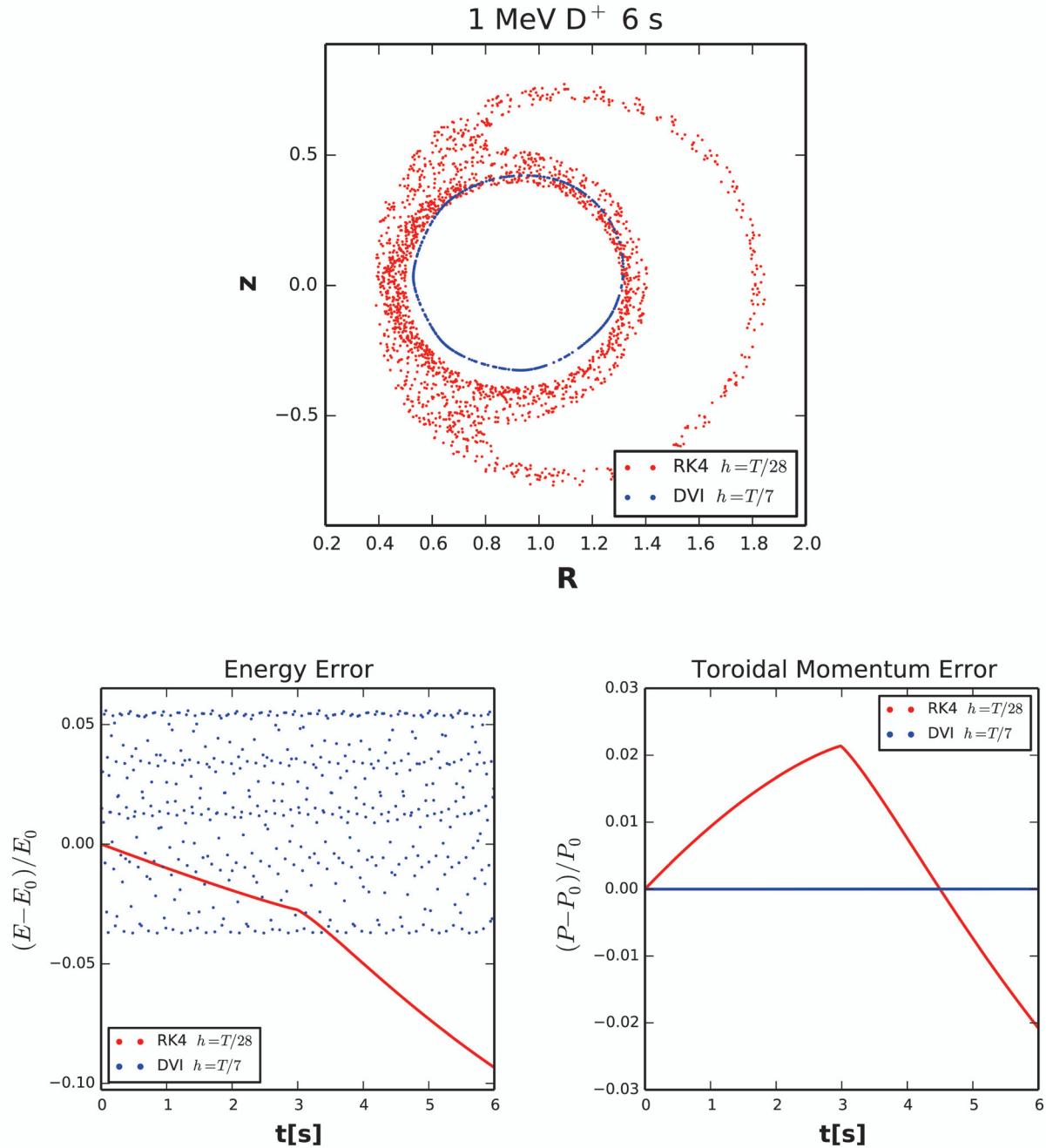


Figure 5.12: Equal computational expense comparison of RK4 and the DVI method for a passing 1 MeV deuteron.

### Collisional Drag Results

The inclusion of collisional effects is essential for realistic fast particle studies. An important consideration is then the degree to which numerical error accumulates over a fast particle thermalization time. Because deterioration of the phase portrait is expected, the collisionless invariants - such as energy and toroidal momentum - become preferred measures of performance. Figures 5.13 and 5.14 depict the toroidal momentum for a 1 MeV throughout a thermalization time. For comparison, several step sizes are used for RK4, spanning the equal step size/equal computational expense limits. Over this timescale, the equal computational expense RK4 method behaves satisfactorily, yielding a final toroidal momentum that is within a similar error margin of the true value as that obtained by the DVI. The comparisons show that, if not advantageous, the DVI method is competitive on this timescale. Additional improvements to optimization might yield a factor of 3 as the equal computational expense step size ratio, which would lead to the degenerate variational integrator method performing superior to RK4 for fast particle thermalization processes.

#### 5.3.3 Discussion

The degenerate variational integrator constructed for guiding center trajectories offers attractive benefits over many of the existing methods. Relative to pre-existing variational integrators constructed from the guiding center Lagrangian [14, 15, 16, 19], the DVI is *much* more robust. It exhibits *no* parasitic mode instabilities, and is therefore better suited to real world applications. Additionally, it is a one-step method, requiring no nuanced initialization procedure. Relative to canonical symplectic integration, e.g. that in Ref. [79], the implicit equations have fewer unknown variables and therefore are much less expensive at each step. The degerate variational integrator in this section is also the first to incorporate collisional drag effects into an otherwise symplectic or variational algorithm, and is therefore

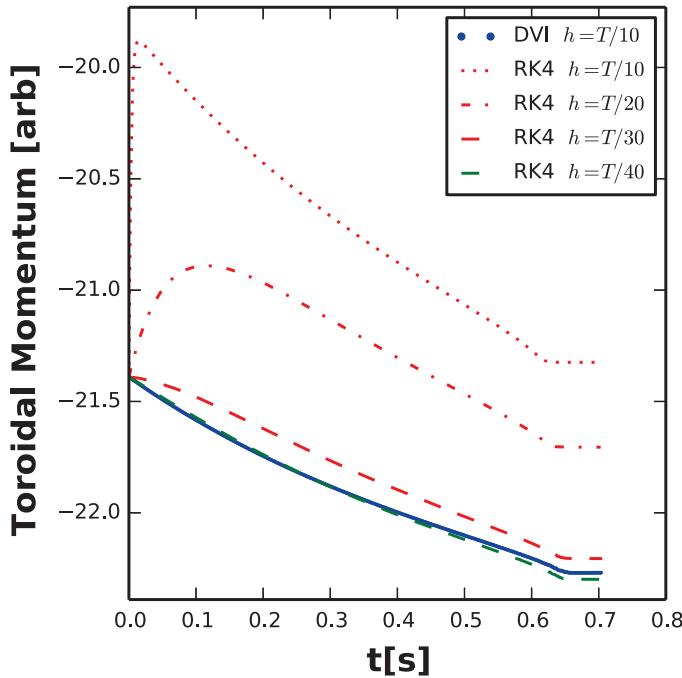


Figure 5.13: Toroidal momentum for a trapped 1 MeV deuteron subject to fast particle drag.

a more realistic model. Finally, compared to high-order explicit Runge-Kutta algorithms, the degenerate variational integrator exhibits the correct qualitative characteristics of the Hamiltonian system. Although each step of the degenerate variational integrator is more expensive, the method has been shown to be competitive during an equal CPU expense comparison of energetic particle thermalization.

The biggest drawback of the guiding center DVIs is that it assumes a particular representation of the fields. The other methods discussed above can be used for any magnetic field and vector potential. The degenerate variational integrator approach is a very new idea, and it is hoped that this restriction can be circumvented in future developments. Likely, such a generalization would also make it clear how to perform one-step variational integration of any non-canonical Hamiltonian system. Still, the present method can handle any magnetic

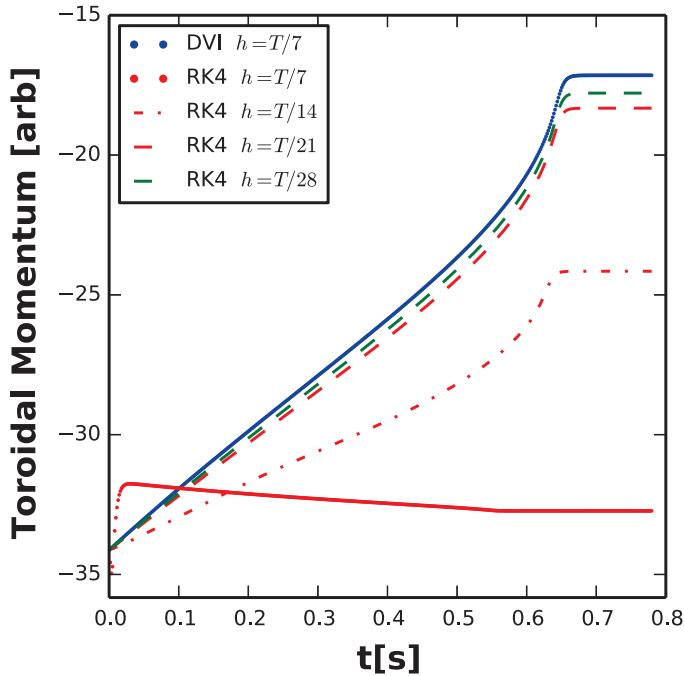


Figure 5.14: Toroidal momentum for a passing 1 MeV deuteron subject to fast particle drag.

configuration with nested flux surfaces, and is therefore a meaningful contribution as the first one-step variational integrator in non-trivial non-canonical coordinates.

Other improvements include construction of higher-order methods and adaptive stepping of the DVI. Given the first-order accuracy of the DVI, composition with the adjoint method might yield superior results. As a caution, it's not obvious that the overall scheme constructed by composition of the presented DVI and its adjoint would itself possess a variational formulation [9]. It is possible that the structure-preserving properties could be violated from such a composition; the two maps would preserve a different discrete two-form, so their composition might not be discretely symplectic. The addition of adaptive stepping (e.g. [136]) could be an excellent enhancement to the degenerate variational integrator in both the collisionless and collisional settings. The effects of adaptive stepping would be most pronounced in the *collisional* setting; presently, the numerical step size is

limited to some fraction of the *initial* particle orbit period. In the presence of collisions, the period increases as the particle slows down. Adaptive stepping could ensure the DVI uses the maximum possible step size throughout the full thermalizing trajectory, reducing the overall computational expense and strongly benefiting the performance relative to an equal cost Runge-Kutta integration.

# Chapter 6

## Conclusion

*An expert is a man who has made all the mistakes which can be made in a very narrow field.*

—Niels Bohr

### 6.1 Summary

What began as a seemingly straightforward application of variational integrators grew into an extensive detour through multistep numerical methods, degenerate Lagrangian systems, and geometric mechanics. Although the path at times diverted from the immediate goal of advancing guiding center algorithms, many lessons were learned along the way. It is my hope that the extra time spent developing and generalizing these lessons will allow others to avoid the pitfalls encountered. In the end, the one-step formulation of a variational guiding center integrator represents tangible progress toward symplectic integration of general, non-canonical Hamiltonian systems.

The developments on multistep variational integrators identify noteworthy drawbacks to the existing variational integrators for guiding center trajectories [14, 15, 16, 96, 19] and any

other nondegenerate variational integrator constructed from a degenerate Lagrangian action principle [13, 18, 92]. The demonstration that the parasitic mode instabilities are at best marginally stable indicates that effort should not be expended on obtaining a variational integrator with stable parasitic modes. The best one can hope for this class of methods is that the parasitic modes do not significantly increase over the time interval of interest or that re-initialization introduces a satisfactorily small amount of non-conservative error. On a more positive note, the discrete symplecticity of the underlying one-step method suggests the principle component of multistep variational integrators will exhibit good long term numerical fidelity. This structure preservation is especially surprising in light of the theorem that, for many multistep methods, the underlying one-step method cannot preserve the symplectic two-form of the continuous system [95]; contradiction is avoided because the discrete symplectic structure for MVIs differs from the continuous structure. The discrete symplecticity result of Thm. 3.4.4 and stability analysis of Sec. 3.3 may be put to use to rapidly analyze new multistep algorithms, including those possessing a previously unnoticed variational formulation.

Degenerate variational integrators improve upon the situation by eliminating parasitic modes. On an intuitive level, the notion that discretizations of degenerate Lagrangians should themselves be degenerate is appealing; the degeneracy of the Lagrangian may be considered another mathematical structure to be retained upon discretization. This intuitive notion is verified and formalized in Thm. 4.3.1, which proves that no parasitic modes are present whenever the ranks of the discrete and continuous Lagrangian two-forms are equal. In particular cases, it has been shown how to construct such a properly degenerate variational integrator. The magnetic field line and guiding center systems considered, although not fully general, admit no other known one-step variational or symplectic formulations without reverting to canonical coordinates. Both cases exhibit the anticipated long-term nu-

merical fidelity, suggesting that the discrete symplecticity property is effective at obtaining the desired behavior and motivating further development in this new direction.

The applications of Chapter 5 put discrete variational principles to use in several plasma-relevant settings. A new collection of multistep variational integrators is constructed for nearly integrable non-canonical Hamiltonian systems of Section 5.1. Here, the variational construction makes it apparent how to utilize the known, unperturbed flow while incorporating perturbative effects to the desired order in perturbation size. Although the final discrete Euler-Lagrange equations are horrendously complicated, the general expressions for the discrete Lagrangians are quite concise. With additional reduction of the computational cost of these methods - for instance by approximating the update rule with a reduced set of basis functions - they may prove valuable for rapid determination of three-dimensional effects in magnetically confined plasmas. The final pair of applications test the benefits of degenerate variational integrators for magnetic field line and guiding center dynamics. Although both methods are only first-order accurate, excellent long-term behavior is exhibited. Such good behavior seems unlikely unless the discrete symplecticity of the one-step formulations is valid and effective. By including collisional drag effects, it is shown both that the methods are not restricted to purely Hamiltonian systems and that the methods are advantageous over physically meaningful timescales. The degenerate guiding center integrator is an excellent tool for modeling energetic particle thermalization, and would likely excel at modeling relativistic runaway electron dynamics if appropriately adapted.

At the most practical level, the GEODES library provides a new and useful tool for an old task: solving ordinary differential equations. Although alternatives exist for rapidly estimating the solution to a differential equation, I have found these alternatives do not scale well to testing many different problems nor to performing expensive calculations. Certainly, the physics is more important than the ease of use of the tool, but the emphasis on a user-

friendly and maintainable library should free up future users to spend more of their time as studying physics.

## 6.2 Future Work

From the outset of this research, the intent was to improve algorithms for guiding center trajectories using variational integrators. The construction of a degenerate variational integrator for guiding center dynamics has made progress on this task, but falls short of the original intent due to the restrictions on the magnetic field configuration for which the degeneracy is obtained during discretization. Within the assumption that one component of the magnetic field is zero, applications that can deploy the DVI include electrostatic gyrokinetic particle-in-cell simulations [70, 71]. Still, to address many problems of interest to test particle calculations, this restriction will need to be overcome. The most important outstanding task is therefore to generalize the procedure for constructing properly degenerate Lagrangians when discretizing Hamilton's principle in phase-space.

Also with the intent of enhancing test particle calculations, opportunities are available for improving the collisional models used by geometric integrations of guiding center trajectories. Using the phase-space Lagrange-d'Alembert framework, it is possible to include stochastic pitch-angle scattering terms. By including the pitch-angle scattering as external forcing terms, arbitrary effects may be included, but the forcing terms violate the conservative properties of the integrator. An alternative is to formulate the pitch-angle scattering as a stochastic Hamiltonian system [137]; recent work has made progress toward this goal [138, 139]. Should a stochastic Hamiltonian description of pitch angle scattering be available, one could construct stochastic variational integrators [140, 141] for these dynamics. Development of stochastic variational integrators could yield improved statistical properties when simulating test particle dynamics.

A final exciting opportunity for development I'd like to mention is determining the benefits of symplectic integration in kinetic particle-in-cell simulations. Several recent developments have constructed fully symplectic time advances for particle-in-cell simulations [142, 143]. Although extensive literature exists on the benefits of symplectic integration for ODEs, there may be unexplored advantages of symplectic integration in PDE systems. These advantages could be studied in Vlasov-Maxwell particle-in-cell studies, or the degenerate variational integrators of this dissertation could be combined with symplectic field advances to build a geometric electrostatic gyrokinetic simulation. If the past successes of symplectic integration are any indication, exciting rewards may be awaiting in the realm of fully symplectic kinetic simulations.

# Appendix A

## Mathematical Preliminaries

### Notation Summary

$\mathbb{C}$	Set of complex numbers
$D_i$	Slot derivative - differentiates with respect to $i$ -th argument of a function
$d$	Exterior derivative
$d_i$	Slot exterior derivative - exterior derivative with respect to the $i$ -th argument
$dx^i$	Basis element of one-forms (covariant vectors) in the coordinates $(x^1, \dots, x^n)$
$\partial_i$	Basis element of tangent (contravariant) vectors, corresponding to $\frac{\partial}{\partial x^i}$ , where $p$ is the position
$g$	Metric tensor
$H$	Hamiltonian
$h$	Numerical time step
$I$	Identity Matrix
$i$	Component index
$\text{id}$	Identity map
$k$	Numerical time index
$\mathcal{L}$	Lie derivative
$L$	Lagrangian
$L_d$	Discrete Lagrangian
$\mathbb{F}$	Legendre transform

---

$\mathbb{F}^\pm$	Discrete Legendre transform
$F$	Flow map
$F^*$	Pull-back under map $F$
$F^t$	Time $t$ flow map
$F_d$	Discrete map (of a numerical method)
$F_H$	Hamiltonian flow map
$F_H^t$	Time $t$ Hamiltonian flow map
$F_L$	Lagrangian flow map
$F_L^t$	Time $t$ Lagrangian flow map
$F_{L_d}$	Discrete Lagrangian map
$F_{H_d}$	Discrete Hamiltonian map
$\hat{F}$	Underlying one-step method
$\hat{F}_{L_d}$	Underlying one-step method of a multistep variational integrator Also, one-step formulation of a degenerate variational integrator
$\tilde{F}$	Flow map of modified vector field
$\tilde{F}^t$	Time $t$ flow map of modified vector field
$f_d$	Numerical update rule
$f_{,i}$	Differentiation with respect to the $i$ -th component, i.e. $\frac{\partial f}{\partial x^i}$
$\mathbb{R}$	Set of real numbers
$M$	Manifold
$N$	Number of time steps
$n$	Dimension
$\pi_i$	Projection operator selecting the $i$ -th component
$Q$	Configuration space
$S$	Action functional
$S_d$	Discrete action functional
$TQ$	Tangent bundle of the configuration space
$T^*Q$	Co-tangent bundle of the configuration space
$X$	Vector field
$X_H$	Hamiltonian vector field

$\tilde{X}$	Modified vector field
$\bar{z}$	Equilibrium point
$\Omega$	Symplectic structure
$\Omega_{L_d}$	Discrete symplectic structure corresponding to a discrete Lagrangian $L_d$
$\hat{\Omega}_{L_d}$	Discrete symplectic structure preserved by the underlying one-step method of a multistep variational integrator defined by discrete Lagrangian $L_d$

## A.1 Introduction

**Disclaimer:** The intent of this chapter is *not* to surpass the presentation of this material in existing references. In particular, I highly recommend the accessible-yet-thorough presentation of Ref. [20]. For the convenience of the reader, the intent is to concisely explain notation, summarize key results used in this thesis, and aggregate pertinent results that are spread across one or more references. I hope it serves as a useful “cheat sheet” for those familiar with the material and perhaps a bare bones overview if time does not allow consulting the original sources.

This thesis utilizes *exterior calculus* as opposed to vector/tensor calculus; in place of divergence, gradient, and curl, the thesis uses the exterior derivative, wedge product, and interior product (*not* respectively). Advantages of exterior calculus include: generality (to high dimensional spaces), ease of use in curvilinear coordinates, and a decoupling of metric-dependent (geometric) quantities from metric-independent (topological) quantities. In practice, this often means more time expended on definitions and the meaning of operations and less time fussing with metric tensor factors, Cristoffel symbols, and otherwise tedious algebraic manipulations. The expense of these advantages is an initial investment of time to learn the (at times, abstract) the tools of exterior calculus.

After reading this Chapter, one should ideally: associate tangent vectors and one-forms with contravariant and covariant vectors, respectively; be familiar with the basis elements of vectors, tensors, and differential forms; understand how to construct a tensor from a differential form; operate with an exterior derivative and understand that  $d^2 = 0$ ; take an interior product; associate pull-backs and push-forwards with the calculus chain rule; and intuit the meaning of a symplectic map. Competency in the preceding list will leave one well equipped for the material encountered in the dissertation.

## A.2 Tangent Vectors and One-Forms

Physicists are often well acquainted with the notion of covariant and contravariant vectors. For instance, consider the three-dimensional Euclidean space  $\mathbb{R}^3$  with curvilinear coordinate functions  $(u^1, u^2, u^3)$ . From these coordinate functions, there exist two, position-dependent coordinate bases:  $(e_1, e_2, e_3)$  and  $(e^1, e^2, e^3)$ . A vector  $v$  might be expressed in terms of “contravariant components” according to:

$$v = v^1 e_1 + v^2 e_2 + v^3 e_3 = v^i e_i, \quad (\text{A.1})$$

where the latter expression uses the Einstein summation convention that summation over repeated indices (one subscript and one superscript) is implied. Similarly, a vector  $v$  might be expressed in terms of “covariant components” according to:

$$v = v_1 e^1 + v_2 e^2 + v_3 e^3 = v_i e^i. \quad (\text{A.2})$$

The “covariance” and “contravariance” refers to how these components transform when a coordinate transformation is performed, as will be detailed below. In the language of exterior calculus, “contravariant vectors” will be replaced by the notion of “tangent vectors”, and “covariant vectors” will be replaced by the notion of “one-forms”.

The starting point for the definitions of these entities is some point  $\rho$  in a smooth manifold  $M$  with local curvilinear coordinates  $(u^1, \dots, u^n)$ , where  $n$  is the dimension of the manifold. At this point  $\rho$ , one can identify two vector spaces. The first is the *tangent space* of  $M$  at  $\rho$ , denoted  $T_\rho M$ . A basis may be constructed for this space using the coordinates  $u^i$  and the position  $\rho$  according to:

$$e_i = \frac{\partial \rho}{\partial u^i}. \quad (\text{A.3})$$

That is, if one changes the  $i$ -th coordinate function  $u^i$  some infinitesimal amount while keeping all other coordinate functions fixed, how much does the position  $\rho$  change? The basis elements  $e_i$  span the tangent space  $T_\rho M$ , and are often denoted  $e_i = \partial_{u^i}$ . The second vector space at  $\rho$  is the *dual* of the first space; it is the *cotangent space* of  $M$  at  $\rho$ , denoted  $T_\rho^* M$ . The basis of this space is denoted:

$$e^i = du^i. \quad (\text{A.4})$$

This basis is the dual basis to the  $\partial_{u^i}$  set, meaning in particular:

$$du^i(\partial_{u^j}) = \delta_j^i, \quad (\text{A.5})$$

where  $\delta$  is the Kronecker delta function, so  $\delta_j^i = 1$  when  $i = j$  and  $\delta_j^i = 0$  when  $i \neq j$ . Using this relation, it can be determined that an element  $\alpha = a_i du^i$  in the cotangent space acts on a vector  $v = v^i \partial_i$  in the tangent space according to:

$$\alpha(v) = a_i v^i. \quad (\text{A.6})$$

The elements of the cotangent space  $T_\rho^* M$ , like  $\alpha$ , are called *one-forms*. These are linear mappings on the vectors in the tangent space, which are referred to as *tangent vectors*.

**Example - Polar Coordinates** For an example, let us construct the position-dependent basis elements for a simple curvilinear coordinate system. Consider the Euclidean space  $\mathbb{R}^2$  with Cartesian coordinates  $(x, y)$  and polar coordinates  $(r, \theta)$ . These coordinate functions

are related by:

$$x = r \cos(\theta) \quad (\text{A.7a})$$

$$y = r \sin(\theta). \quad (\text{A.7b})$$

The position  $\rho$  in  $\mathbb{R}^2$  may be denoted by the vector:

$$\rho = x\hat{x} + y\hat{y}, \quad (\text{A.8})$$

where  $\hat{x}, \hat{y}$  are the familiar unit vectors in the  $x$ - and  $y$ -directions. Calculating the basis elements of  $T_p\mathbb{R}^2$ :

$$\begin{aligned} \frac{\partial \rho}{\partial r} &= \frac{\partial x}{\partial r}\hat{x} + \frac{\partial y}{\partial r}\hat{y} \\ &= \sin(\theta)\hat{x} + \cos(\theta)\hat{y} \\ &= \frac{x\hat{x} + y\hat{y}}{\sqrt{x^2 + y^2}}, \end{aligned} \quad (\text{A.9})$$

where the fact that  $\hat{x}, \hat{y}$  are constant in space and therefore independent of  $r$  has been used. Similarly:

$$\begin{aligned} \frac{\partial \rho}{\partial \theta} &= \frac{\partial x}{\partial \theta}\hat{x} + \frac{\partial y}{\partial \theta}\hat{y} \\ &= -y\hat{x} + x\hat{y}. \end{aligned} \quad (\text{A.10})$$

To determine the dual basis  $(dr, d\theta)$ , use Eq. A.5 to deduce:

$$dr = \frac{x\hat{x} + y\hat{y}}{\sqrt{x^2 + y^2}} \quad (\text{A.11})$$

$$d\theta = \frac{-y\hat{x} + x\hat{y}}{x^2 + y^2}. \quad (\text{A.12})$$

Reflecting on this example, notice that as  $r$  increases,  $\partial_\theta$  increases in magnitude while  $d\theta$  decreases in magnitude. Correspondingly,  $\partial_\theta$  and  $d\theta$  are *not* the unit vector  $\hat{\theta}$ . As an additional observation, notice that  $(\partial_r, \partial_\theta)$  is an orthogonal (but not orthonormal) basis, as is  $(dr, d\theta)$ . This is *not* a general fact, but instead a feature of this particular curvilinear coordinate system. Still, the dual basis relation of Eq. A.5 *is* a general statement.

Let us now direct our attention to how the components of tangent vectors and one-forms transform when representing these entities by different coordinates. Consider a second set of curvilinear coordinates  $(\bar{u}^1, \bar{u}^2, \bar{u}^3)$  in a neighborhood containing  $\rho$ . Given a vector  $v$  represented in the old coordinates  $u$  as  $v^i \partial_{u^i}$ , the goal is to find some components  $\bar{v}^i$  such that:

$$\bar{v}^i \partial_{\bar{u}^i} = v^i \partial_{u^i}, \quad (\text{A.13})$$

i.e. the vector is unchanged by the coordinate transformation. To identify the new components, apply the chain rule relation:

$$\frac{\partial \rho}{\partial u^i} = \frac{\partial \rho}{\partial \bar{u}^j} \frac{\partial \bar{u}^j}{\partial u^i}, \quad (\text{A.14})$$

to obtain:

$$v^i \partial_{u^i} = v^i \frac{\partial \bar{u}^j}{\partial u^i} \partial_{\bar{u}^j}. \quad (\text{A.15})$$

Comparison of Eq. (A.15) with Eq. (A.13) identifies:

$$\bar{v}^i = v^j \frac{\partial \bar{u}^i}{\partial u^j}. \quad (\text{A.16})$$

This important relation is the *transformation rule for the contravariant components of a vector*. Thus the relationship between tangent vectors and contravariant vectors has been established.

Similar reasoning is applied to obtain the transformation rule for the components of a one-form. Consider a one-form  $\alpha$  represented in the old coordinate functions  $u^i$  as  $\alpha = a_i du^i$ . The goal is now to identify components  $\bar{a}_i$  such that:

$$\bar{a}_i d\bar{u}^i = a_i du^i. \quad (\text{A.17})$$

To do so, treat  $du^i$  as the differential of  $u^i$  considered to be a function of the  $\bar{u}^j$  (more on the  $d$  operator will be explained in Sec. A.4.2):

$$du^i = \frac{\partial u^i}{\partial \bar{u}^j} d\bar{u}^j. \quad (\text{A.18})$$

Then  $\alpha$  is given by:

$$a_i du^i = a_i \frac{\partial u^i}{\partial \bar{u}^j} d\bar{u}^j, \quad (\text{A.19})$$

identifying the new components  $\bar{a}_i$  according to:

$$\bar{a}_i = a_j \frac{\partial u^j}{\partial \bar{u}^i}. \quad (\text{A.20})$$

This relation is the *transformation rule for the covariant components of a vector*. Therefore, the components of a one-form transform like those of a covector!

When considering dynamical systems in this dissertation, it will rarely be sufficient to consider the tangent space or cotangent space at a *single* point  $\rho$  of a manifold  $M$ . The state of the system will be evolving through the manifold, with a different tangent space identifiable at each new point. A useful set is then the collection of *all* tangent spaces to a manifold  $M$ . This set is known as the *tangent bundle*, and is notated  $TM$ . Similarly, the *cotangent bundle*  $T^*M$  is the union of all cotangent spaces on  $M$ . An element of the tangent

bundle  $T^*M$  is identified by a position and a vector  $(\rho, v)$ , while an element of the cotangent bundle is identified by a position and a one-form  $(\rho, \alpha)$ .

## A.3 Tensors

The basis elements defined in Sec. A.2 can be used to construct higher dimensional tensor objects. Consider again a position  $\rho$  in a smooth manifold  $M$  with local coordinates  $u^i$ . The basis elements for an  $n$ -times covariant,  $m$ -times contravariant tensor (of rank  $n + m$ ) are given in terms of the tensor product  $\otimes$  by:

$$du^{i_1} \otimes \dots \otimes du^{i_n} \otimes \partial_{u^{j_1}} \otimes \dots \otimes \partial_{u^{j_m}}. \quad (\text{A.21})$$

This basis element appears intimidating, but all it means is that the tensor is a multilinear map that specifies a real number when provided with  $n$  tangent vectors and  $m$  one-forms as arguments. For instance, the three-times covariant once contravariant tensor:

$$T_{i_1 i_2 i_3}^j du^{i_1} \otimes du^{i_3} \otimes du^{i_3} \otimes \partial_{u^j}, \quad (\text{A.22})$$

acts on vectors  $(v, w, x)$  and a one-form  $\alpha$  according to:

$$T(v, w, x, \alpha) = T_{i_1 i_2 i_3}^j v^{i_1} w^{i_2} x^{i_3} a_j, \quad (\text{A.23})$$

with the summation implied for all four indices.

### A.3.1 The Metric Tensor

One of the most important tensors is “the” *metric tensor*,  $g$ . A metric tensor is a non-degenerate, symmetric, twice covariant tensor that allows one to take an inner product of

two tangent vectors. Note that in the preceding discussion, vectors were always paired with one-forms (or covariant tensor components) to obtain real numbers; up until now, an inner product has not been defined. Given a metric tensor, the inner product is defined as:

$$\langle v, w \rangle = g(v, w) = g_{ij}v^i w^j. \quad (\text{A.24})$$

The symmetry and non-degeneracy of  $g_{ij}$  achieves the familiar properties of inner products, including  $\langle v, w \rangle = \langle w, v \rangle$  and  $\langle v, w \rangle = 0 \quad \forall w$  implies  $v$  is the zero vector. Based on this definition, one may identify that the components of the metric tensor are:

$$g_{ij} = \langle \partial_{u^i}, \partial_{u^j} \rangle. \quad (\text{A.25})$$

**Example - Polar Coordinate Metric Tensor** For illustration, this example calculates the components of the metric tensor for polar coordinates. Using the definition of the components of the metric tensor Eq. (A.25) and the polar coordinate tangent vector basis elements Eqs. (A.9) and (A.10), one determines:

$$g_{rr} = 1 \quad (\text{A.26})$$

$$g_{\theta\theta} = r^2 \quad (\text{A.27})$$

$$g_{r\theta} = g_{\theta r} = 0. \quad (\text{A.28})$$

Given these components, one might write the polar coordinate metric tensor as:

$$g = dr \otimes dr + r^2 d\theta \otimes d\theta, \quad (\text{A.29})$$

or in matrix form as:

$$g = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}. \quad (\text{A.30})$$

Using the metric tensor, it is possible to “associate” tangent vectors with one-forms. For instance, consider a metric tensor  $g$  and a tangent vector  $v$ . The goal is then to identify some one-form  $\alpha$ . To do so, set the components  $a_j$  of  $\alpha$  according to:

$$a_j = g_{ij}v^i. \quad (\text{A.31})$$

Then given another tangent vector  $w$ ,  $\alpha(w)$  is a linear map defined by:

$$\alpha(w) = \alpha_i w^i = g_{ij}v^i w^j. \quad (\text{A.32})$$

It is in this sense that the metric tensor is said to “lower” indices. It is conventional to denote the components of the one-form associated with  $v$  through the metric tensor as  $v_j$ , i.e.

$$v_j = g_{ij}v^i. \quad (\text{A.33})$$

The inverse of the metric tensor is also symmetric and non-degenerate, but is twice contravariant. It is denoted  $g^{ij}$ , with the property that:

$$g^{ij}g_{jk} = \delta_k^i, \quad (\text{A.34})$$

as required by the definition of the inverse. The inverse metric tensor may be used to associate tangent vectors to one-forms by following a construction analogous to that above for tangent vectors. Similarly, one informally speaks of raising indices using  $g^{ij}$ .

Now is a good time to observe that metric tensors are *not* specified for every smooth manifold; the metric tensor is an *additional* structure that yields particular relations whenever

it is provided. Many of the expressions obtained using exterior calculus will be independent of this particular structure, which has the advantage that these expressions may be used in settings where no metric has been defined.

## A.4 Differential Forms

An extremely important class of tensors is *differential k-forms* or differential forms, for short. A differential  $k$ -form is a  $k$ -times covariant skew-symmetric tensor. That is, it is a multilinear operator acting on  $k$  tangent vectors to specify a real number. The “skew-symmetry” means that interchanging the order of two adjacent arguments incurs a minus sign, for instance a  $k$ -form  $\alpha$  satisfies:

$$\alpha(v_1, \dots, v_i, v_{i+1}, \dots, v_k) = -\alpha(v_1, \dots, v_{i+1}, v_i, \dots, v_k), \quad (\text{A.35})$$

where  $1 \leq i < k$ . Here, the subscripts on the  $v$ s is to denote  $k$  different tangent vectors, *not* components (the components would be labeled with superscripts, of course!). A zero-form is simply a scalar, a one-form a rank-one tensor, two-form a rank-two tensor, and so on.

### A.4.1 Wedge Product

The skew symmetry of  $k$ -forms is conveniently represented using the *wedge product*,  $\wedge$ . It is easiest to explain how the wedge product behaves for the one-form basis elements  $du^i$ , then extrapolate more general behavior based on the linearity and associativity of the wedge product. Given two one-form basis elements  $du^i$  and  $du^j$ , the wedge product of the two basis elements identifies a skew-symmetric bilinear two-form defined by:

$$du^i \wedge du^j = du^i \otimes du^j - du^j \otimes du^i. \quad (\text{A.36})$$

The skew symmetry of  $du^i \wedge du^j$  can be verified by operating on a pair of vectors  $v, w$ :

$$\begin{aligned} du^i \wedge du^j(v, w) &= du^i \otimes du^j(v, w) - du^j \otimes du^i(v, w) \\ &= v^i w^j - v^j w^i, \end{aligned}$$

which is clearly antisymmetric under interchange of  $v, w$ . An important consequence of this antisymmetry is that:

$$du^i \wedge du^i = 0, \quad (\text{A.37})$$

for all one-form basis elements  $du^i$ . Additionally, the wedge product is associative:

$$(du^{i_1} \wedge du^{i_2}) \wedge du^{i_3} = du^{i_1} \wedge (du^{i_2} \wedge du^{i_3}) \quad (\text{A.38})$$

and distributes over addition:

$$du^{i_1} \wedge (du^{i_2} + du^{i_3}) = du^{i_1} \wedge du^{i_2} + du^{i_1} \wedge du^{i_3}. \quad (\text{A.39})$$

With an understanding of how the wedge product operates on one-form basis elements, one can infer how the wedge product operates on arbitrary one-forms  $\alpha^{(1)} = a_i du^i$  and  $\beta^{(1)} = b_i du^i$ . In particular:

$$\alpha^{(1)} \wedge \beta^{(1)} = a_i b_j du^i \wedge du^j, \quad (\text{A.40})$$

so

$$\alpha^{(1)} \wedge \beta^{(1)}(v, w) = \alpha^{(1)}(v) \beta^{(1)}(w) - \beta^{(1)}(v) \alpha^{(1)}(w) = a_i v^i b_j w^j - b_i v^i a_j w^j. \quad (\text{A.41})$$

As an example, consider  $\mathbb{R}^2$  with  $\alpha = a_x dx + a_y dy$  and  $\beta = b_x dx + b_y dy$ . Then:

$$\begin{aligned}\alpha \wedge \beta &= a_x b_x dx \wedge dx + a_x b_y dx \wedge dy + a_y b_x dy \wedge dx + a_y b_y dy \wedge dy \\ &= (a_x b_y - a_y b_x) dx \wedge dy.\end{aligned}$$

Here one can recognize that the wedge product for this two-dimensional problem bears resemblance to the cross product. Three important distinctions to make are: (i) With the cross product, two vectors are combined to form a new vector; for the wedge product, two one-forms are combined to obtain a *two-form*, that is, a different geometric object than one began with. (ii) The cross product requires modification in curvilinear coordinates, while the wedge product operates *the same way* in curvilinear coordinates. (iii) The cross product is not associative, unlike the wedge product. These distinctions make the wedge product more convenient to work with in curvilinear coordinates and high dimensions.

Increasing the generality of the discussion one last time, consider now the wedge product between a  $k$ -form  $\alpha^{(k)}$  and an  $l$ -form  $\beta^{(l)}$ . The wedge product  $\alpha^{(k)} \wedge \beta^{(l)}$  yields a  $(k+l)$ -form. The resulting  $(k+l)$ -form satisfies the following property:

$$\alpha^{(k)} \wedge \beta^{(l)} = (-1)^{(k+l)} \beta^{(l)} \wedge \alpha^{(k)}. \quad (\text{A.42})$$

This property may be deduced by counting the number of times one-form basis elements must be interchanged to switch the order of  $\alpha^{(k)}$  and  $\beta^{(l)}$  in the wedge product. Another important property is that  $\alpha^{(k)} \wedge \beta^{(l)} = 0$  whenever  $(k+l) > n$ , where  $n$  is the dimension of the manifold. The reason for this is that there *must* be a repeated one-form basis element in the resulting  $(k+l)$  form. Interchanging the order of basis elements and applying associativity demonstrates that the final result must be zero.

Equipped with an understanding of the wedge product, the discussion of differential forms may be re-visited with increased specificity. Consider  $\mathbb{R}^3$  with coordinates  $(u^1, u^2, u^3)$ . A zero-form on  $\mathbb{R}^3$  is simply a scalar, and thus has the form:

$$\alpha^{(0)} = a. \quad (\text{A.43})$$

A one-form on  $\mathbb{R}^3$  has three-components, and thus in general takes the form:

$$\alpha^{(1)} = a_1 du^1 + a_2 du^2 + a_3 du^3. \quad (\text{A.44})$$

Similarly, a two-form has three components (in  $\mathbb{R}^3$ ), and has the general form:

$$\alpha^{(2)} = a_1 du^2 \wedge du^3 + a_2 du^3 \wedge du^1 + a_3 du^1 \wedge du^2. \quad (\text{A.45})$$

Finally, a three-form has only one-component, and in that sense is like a scalar in  $\mathbb{R}^3$ . It takes the form:

$$\alpha^{(3)} = adu^1 \wedge du^2 \wedge du^3. \quad (\text{A.46})$$

Reflecting on these expressions, a geometric interpretation of zero-forms associates them with points; one-forms with line segments; two-forms with area elements; and three-forms with volume elements.

Up to this point, tangent vectors and differential forms have been discussed *at a single point*  $\rho$  in the manifold. Physicists are well acquainted with the idea of a *vector field*, which assigns a different vector to *each* point of the manifold; usually this assignment is smoothly varying, specifying the vector at each point as some smooth function of the local coordinates of the manifold. A similar idea pertains to differential forms. The components of differential forms may be specified as some smooth function of the position, as represented by the local

coordinates. This gives rise to the notion of *one-form fields*, and similarly for higher degree differential forms. For instance, one might specify the one-form field on  $\mathbb{R}^3$ :

$$\alpha^{(1)} = (x + \cos(y))dx + y^2z^2dz. \quad (\text{A.47})$$

Often, the word “field” will be omitted when it can be understood from context that the one-form specified assigns a one-form at each point in the space of interest. These fields are important for our next fundamental tool: the exterior derivative.

### A.4.2 Exterior Derivative

The wedge product combines  $k$ -forms with  $l$ -forms to construct  $(k+l)$ -forms. Another essential operator for exterior calculus is the *exterior derivative*; it takes  $k$ -forms into  $(k+1)$ -forms. To illustrate how to operate with the exterior derivative, it will be shown how to operate on a representative  $k$ -form and generalized from there. Consider then a  $k$ -form field  $\alpha^{(k)} = a(u)du^{i_1} \wedge du^{i_2} \wedge \dots \wedge du^{i_k}$ . Then the exterior derivative of  $\alpha^{(k)}$  is given by:

$$d\alpha^{(k)} = \frac{\partial a}{\partial u^j} du^j \wedge du^{i_1} \wedge \dots \wedge du^{i_k}. \quad (\text{A.48})$$

In words, the procedure is to differentiate the coefficient function with respect to each coordinate  $u^j$ , then use the wedge product to prepend  $du^j$  to the pre-existing  $k$ -form basis elements. This formula justifies the procedure used in Eq. (A.18) in which  $d$  operated on a scalar function by returning the differential of that function. That is, for any zero-form field (scalar function)  $\alpha^{(0)} = a(u)$ ,

$$d\alpha^{(0)} = \frac{\partial a}{\partial u^i} du^i; \quad (\text{A.49})$$

the exterior derivative of a zero-form is the differential of that function!

The  $k$ -form  $\alpha^{(k)}$  considered above had only one non-zero component. More generally, a  $k$ -form can be represented as the *sum* of some number<sup>1</sup> of terms of the form  $a_i(u)du^{i_1} \wedge du^{i_2} \wedge \dots \wedge du^{i_k}$ . The linearity of the exterior derivative operator allows one to distribute the operation over the summation, operating on each individual term in the manner illustrated above. This practice is illustrated in the following examples.

**Example - Exterior Derivative in Cartesian  $\mathbb{R}^3$**  Consider  $\mathbb{R}^3$  space with Cartesian coordinates  $(x, y, z)$ . Given a zero-form  $\alpha^{(0)} = a(x, y, z)$ , the exterior derivative of  $\alpha^{(0)}$  is:

$$d\alpha^{(0)} = \frac{\partial a}{\partial x}dx + \frac{\partial a}{\partial y}dy + \frac{\partial a}{\partial z}dz. \quad (\text{A.50})$$

Interestingly, this one-form has components that, in Cartesian coordinates, are equal to the gradient of  $a$ . Let's continue one degree higher: Given a one-form  $\alpha^{(1)} = a_x(x, y, z)dx + a_y(x, y, z)dy + a_z(x, y, z)dz$ , the exterior derivative is given by:

$$\begin{aligned} d\alpha^{(1)} &= a_{x,y}dy \wedge dx + a_{x,z}dz \wedge dx + a_{y,x}dx \wedge dy + a_{y,z}dz \wedge dy + a_{z,x}dx \wedge dz + a_{z,y}dy \wedge dz \\ &= (a_{y,x} - a_{x,y})dx \wedge dy + (a_{z,y} - a_{y,z})dy \wedge dz + (a_{x,z} - a_{z,x})dz \wedge dx. \end{aligned} \quad (\text{A.51})$$

Here, an additional index notation convention is encountered: indices appearing after a comma denote differentiation with respect to that component, so for example  $a_{x,y} = \frac{\partial a_x}{\partial y}$ . Continuing an interesting pattern, the two-form obtained in Eq. (A.51) has components that match those of the *curl* of a vector valued function in Cartesian coordinates. To check whether this pattern continues, consider the final case of the exterior derivative of a two-form  $\alpha^{(2)} = a_x(x, y, z)dy \wedge dz + a_y(x, y, z)dz \wedge dx + a_z(x, y, z)dx \wedge dy$ . To motivate the subscripts on the  $a_i$  functions, they have been chosen according to the direction of the vector normal to the plane defined in the two-form basis element. Of course, this yields a rather fortuitous

---

<sup>1</sup> $n$  choose  $k$ , to be precise

result:

$$\begin{aligned} d\alpha^{(3)} &= a_{x,x}dx \wedge dy \wedge dz + a_{y,y}dy \wedge dz \wedge dx + a_{z,z}dz \wedge dx \wedge dy \\ &= (a_{x,x} + a_{y,y} + a_{z,z})dx \wedge dy \wedge dz, \end{aligned} \quad (\text{A.52})$$

thereby obtaining the (Cartesian) divergence! Operating according to the relatively simple properties of the wedge product and exterior derivative recovered entities whose components match those of div, grad, curl in Cartesian coordinates.

Before leaving this example, several comments are appropriate for contextualizing these results. The first is that the correspondence between the exterior derivative and div, grad, curl is *not* so straightforward in curvilinear coordinates. This is because in curvilinear coordinates div, grad, curl have additional factors originating from the metric tensor. In contrast, the exterior derivative operates *exactly* the same way in curvilinear coordinates as it does in Cartesian coordinates. For instance, with spherical coordinates  $(r, \theta, \phi)$  on  $\mathbb{R}^3$ , the exterior derivative of a zero-form  $\alpha^{(0)} = a(r, \theta, \phi)$  is given by:

$$d\alpha^{(0)} = \frac{\partial a}{\partial r}dr + \frac{\partial a}{\partial \theta}d\theta + \frac{\partial a}{\partial \phi}d\phi. \quad (\text{A.53})$$

The components of this form cannot be claimed to be the components of gradient in spherical coordinates. For the exact correspondences, which will reveal the metric tensor factors that appear in the expressions for div, grad, curl, the interested reader is referred to Ref. [20]. An inference that *is* valid, however, is that the exterior derivative and wedge product can often prove satisfactory for purposes where one would typically use vector calculus operators. Based on their simplicity in curvilinear coordinates and their validity in higher dimensions, a compelling case is made for adopting these exterior calculus formulations over more standard vector calculus operators.

Several properties are important to keep in mind regarding the exterior derivative  $d$ . The first is that, for any  $k$ -form  $\alpha$ , the exterior derivative has the following property:

$$d^2\alpha = 0. \quad (\text{A.54})$$

In the examples above, the components of the differential forms obtained using the exterior derivative matched the components of the vectors obtained by divergence, gradient, and curl. Equation (A.54) applied to the examples above will resemble the direct computation of  $\nabla \times \nabla f = 0$  and  $\nabla \cdot (\nabla \times f) = 0$  in Cartesian coordinates. This property holds not only in Cartesian coordinates in  $\mathbb{R}^3$ , but in any curvilinear coordinates in any dimensional space. The proof can be shown in a satisfying geometric manner using Stoke's theorem (amounts to the boundary of a boundary being the empty set); for details, the reader is again referred to Ref. [20].

An additional property that is helpful to keep in mind is essentially the product rule for the exterior derivative. Given a  $k$ -form  $\alpha^{(k)}$  and an  $l$ -form  $\beta^{(l)}$ , then:

$$d(\alpha^{(k)} \wedge \beta^{(l)}) = (d\alpha^{(k)}) \wedge \beta^{(l)} + (-1)^k \alpha^{(k)} \wedge (d\beta^{(l)}). \quad (\text{A.55})$$

To intuit this factor of  $(-1)^k$ , consider an example with  $k = l = 1$  and  $\alpha^{(1)} = a(x)dy$ ,  $\beta^{(1)} = b(x)dz$ . Then calculate  $d\alpha^{(1)} \wedge \beta^{(1)}$  two ways: First by representing  $\alpha^{(1)} \wedge \beta^{(1)}$  as:

$$\alpha^{(1)} \wedge \beta^{(1)} = a(x)b(x)dy \wedge dz, \quad (\text{A.56})$$

then using Eq. (A.48) and the standard product rule to determine:

$$d(\alpha^{(1)} \wedge \beta^{(1)}) = \left( \frac{\partial a}{\partial x}b + a\frac{\partial b}{\partial x} \right) dx \wedge dy \wedge dz. \quad (\text{A.57})$$

Second, apply Eq. (A.55) directly:

$$\begin{aligned}
 d(\alpha^{(1)} \wedge \beta^{(1)}) &= (d(a(x)dy)) \wedge (b(x)dz) + (-1)(a(x)dy) \wedge (d(b(x)dz)) \\
 &= \frac{\partial a}{\partial x} b dx \wedge dy \wedge dz - a \frac{\partial b}{\partial x} dy \wedge dx \wedge dz \\
 &= \left( \frac{\partial a}{\partial x} b + a \frac{\partial b}{\partial x} \right) dx \wedge dy \wedge dz,
 \end{aligned} \tag{A.58}$$

obtaining agreement with the first calculation.

### A.4.3 Symplectic Two-Forms

For the purposes of this dissertation, the most important differential forms encountered are *symplectic two-forms*. A symplectic two-form (field)  $\Omega$  is:

- A differential two-form
- Non-degenerate
- Closed (meaning  $d\Omega = 0$ )

Much like “the” metric tensor, any two-form satisfying these properties can be a symplectic two-form, however, the expectation is that one specifies a unique two-form field on a manifold  $M$  as “the” symplectic two-form. Synonymous terminology refers to “the symplectic structure” on a manifold  $M$  (see Sec. 2.3.1).

The condition that the two-form  $\Omega$  be closed is automatically satisfied if there exists a one-form  $\vartheta$  such that  $\Omega = d\vartheta^2$ . This follows from Eq. (A.54):  $d\Omega = d^2\vartheta = 0$ . If such a one-form does exist (globally), then  $\Omega$  is said to be *exact*. All exact differential forms are closed, but not all closed differential forms are exact; this is analogous to all vector fields that are the curl of another vector field having zero divergence, but zero divergence being

---

<sup>2</sup>In the context of Hamiltonian mechanics, the convention is that  $\Omega = -d\vartheta$  when such a one-form exists.

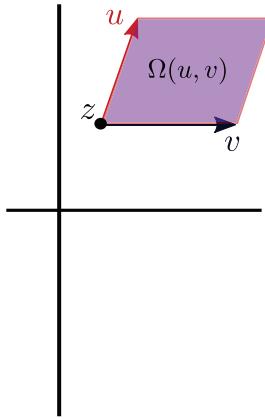


Figure A.1: A symplectic structure determines a real number  $\Omega(u, v)$  based on two tangent vectors  $u, v$  located at some point  $z$  in the symplectic manifold  $M$ .

insufficient to ensure the global existence of a vector field whose curl gives the divergence-free vector field.

As a two-form field, symplectic structures associate a two-form at each point of a manifold. Like all two-forms, one can associate a real number with the two-form and a pair of tangent vectors; see Fig. A.1. These real numbers can have intuitive geometric meanings, as illustrated in the following example.

**Example - Canonical Symplectic Structure** Suppose  $M$  is a 2 dimensional space that admits global coordinates  $(q, p)$ . One way to obtain such an  $M$  is if  $M = T^*Q$  for some one-dimensional configuration space  $Q$ , and  $p$  is identified using the Legendre transform and a regular Lagrangian on  $TQ$  (see Sec.2.2.2). On this space, “the canonical symplectic structure”  $\Omega_c$  is given by:

$$\Omega_c = dq \wedge dp. \quad (\text{A.59})$$

In terms of the tensor product, this becomes:

$$\Omega_c = dq \otimes dp - dp \otimes dq. \quad (\text{A.60})$$

One may therefore write the components of this twice covariant tensor as:

$$\Omega_c = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (\text{A.61})$$

It is instructive to operate on a pair of test vectors with the canonical symplectic structure  $\Omega_c$ . Let the first representative tangent vector be  $v = v^q \partial_q + v^p \partial_p$  and the second tangent vector be  $w = w^q \partial_q + w^p \partial_p$ . Then:

$$\begin{aligned} \Omega_c(v, w) &= dq \wedge dp(v, w) \\ &= dq(v)dp(w) - dp(v)dq(w) \\ &= v^q w^p - v^p w^q. \end{aligned} \quad (\text{A.62})$$

That is, the canonical symplectic structure determines the oriented area spanned by two tangent vectors in phase space! In terms of matrix representations, one would perform this calculation as:

$$\begin{aligned} \Omega_c(v, w) &= v^T \Omega_c w = (v^q, v^p) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} w^q \\ w^p \end{pmatrix} \\ &= v^q w^p - v^p w^q. \end{aligned} \quad (\text{A.63})$$

In  $2n$ -dimensional phase-space, the canonical symplectic structure takes the form:

$$\Omega_c = dq^i \wedge dp_i \quad (\text{A.64})$$

where  $i = 1, \dots, n$ . The geometric interpretation of the real number obtained for  $\Omega_c(v, w)$  is the sum of the oriented areas in each of the  $q^i, p_i$  planes in phase-space.

The canonical symplectic structure is relatively simple to the extent that the components of the antisymmetric tensor are not functions of the position in the phase-space manifold  $M$ . Regardless of where the tangent vectors originate in  $M$ , the canonical symplectic structure calculates their area. More generally, the symplectic structure has components that vary as functions of the position in the manifold. This is the case, for instance, for the important guiding center system, as shown next.

**Example - Guiding Center Symplectic Structure** Consider a magnetized particle drifting in a time-independent magnetic field  $B = d \times A$  with guiding center position  $x$ , parallel velocity  $u$  and magnetic moment  $\mu$ . Let the coordinates be  $x^1, x^2, x^3$ , and assume units are normalized so that  $e = m = c = 1$ . The symplectic structure  $\Omega_{GC}$  for this system (see how to construct equations of motion given a symplectic structure and a Hamiltonian in Sec. 2.3.1) is given by:

$$\Omega_{GC} = A_{i,j}^\dagger dx^i \wedge dx^j + b_i dx^i \wedge du, \quad (\text{A.65})$$

where  $A_i^\dagger(x, u) = A_i(x) + ub_i(x)$ , where  $b_i$  is the  $i$ -th covariant component of the magnetic field unit vector. Using the tensor product, this may be written:

$$\begin{aligned} \Omega_{GC} &= A_{i,j}^\dagger dx^i \otimes dx^j - A_{i,j}^\dagger dx^j \otimes dx^i + b_i dx^i \otimes du - b_i du \otimes dx^i \\ &= \left( A_{i,j}^\dagger - A_{j,i}^\dagger \right) dx^i \otimes dx^j + b_i (dx^i \otimes du - du \otimes dx^i), \end{aligned} \quad (\text{A.66})$$

so one would write the components of the rank-two twice-covariant tensor as:

$$\Omega_{GC} = \begin{pmatrix} 0 & -B_3^\dagger & B_2^\dagger & b_1 \\ B_3^\dagger & 0 & -B_1^\dagger & b_2 \\ -B_2^\dagger & B_1^\dagger & 0 & b_3 \\ -b_1 & -b_2 & -b_3 & 0 \end{pmatrix}, \quad (\text{A.67})$$

where  $B_i^\dagger = -\epsilon_i^{jk} A_{j,k}^\dagger$ , with  $\epsilon_i^{jk}$  the Levi-Civita symbol. If one were to apply  $\Omega_{GC}$  to a pair of test vectors, the result would not simply be (a summation of) the area(s) spanned by the two vectors, but instead the (summation of the) flux of the effective magnetic field  $B^\dagger$  through the area(s) spanned by the two vectors.

## A.5 Push-Forward and Pull-Back

The tools introduced in this section are essential to answering the following question: *what does it mean for a map to be symplectic?* This dissertation is heavily concerned with whether the time advance of a numerical method “preserves” a symplectic two-form. Based on familiar notions of conservation of energy and conservation of momentum, one might be tempted to guess that to preserve a symplectic two-form means that the components of the twice-covariant tensor are constant along a particle’s trajectory, for instance. Unfortunately, this guess is incorrect. The canonical symplectic structure defined in Eq. A.64 has the same tensor components everywhere in space, so it would not be noteworthy for a numerical method to generate trajectories that stay on “constant  $\Omega_c$  tensor component” paths. The story is then a bit more involved.

We did observe that  $\Omega_c$  can be used to calculate the oriented area spanned by two vectors, summed over all  $q^i$ - $p_i$  planes. The question of whether a map is symplectic (with respect to  $\Omega_c$ ) boils down to the following: Consider a map  $F : M \rightarrow M$ ; it takes some initial point  $\rho_0$  and determines a new point  $\rho_1 = F(\rho_0)$ . Consider two arbitrary tangent vectors  $v_0, w_0$  located at the initial point  $\rho_0$  (so  $v_0, w_0 \in T_{\rho_0} M$ ). Calculate the area spanned by these two vectors using  $\Omega_c(v_0, w_0)$ . Next, *advance*  $v_0, w_0$  to two new vectors  $v_1, w_1$  located at  $\rho_1$ , i.e.  $v_1, w_1 \in T_{\rho_1} M$ . Doing so amounts to using the chain rule on the map  $F$ , and is referred to as the *push-forward* of the tangent vectors. Determine the area spanned by these new vectors by evaluating  $\Omega_c(v_1, w_1)$ . The tensor components of  $\Omega_c$  are constant in space, but if they had

spatial dependence, one would evaluate them at the new position  $\rho_1$ . If the summed, oriented areas spanned by  $v_1, w_1$  is the same as the summed, oriented areas spanned by  $v_0, w_0$ ; and if this is true for *any* pair of initial tangent vectors  $v_0, w_0$ , then the map  $F$  is said to be *symplectic* or to *preserve a symplectic structure*. Because this condition is required to hold for any initial pair of vectors, whether or not this condition is satisfied depends only on  $\Omega$  and  $F$ . It can then be formulated as a relationship between only these two entities without reference to the arbitrary tangent vectors. This formulation will refer to the *pull-back* of  $\Omega$  using  $F$ .

The first unexplained step in the above procedure is how one obtains the new vectors  $v_1, w_1$  from the old vectors  $v_0, w_0$  using  $F$  and the “push-forward”. First, some notation. Given a map  $F : M \rightarrow M$  and a tangent vector  $v_0 \in T_{\rho_0}M$ , the *push-forward* of  $v_0$  is a vector  $v_1 \in T_{F(\rho_0)}M$ , written as:

$$v_1 = F_* v_0. \quad (\text{A.68})$$

Concisely,  $F_* v_0$  is the push-forward of  $v_0$  by the map  $F$ . To write the push-forward in coordinates, suppose  $\rho_1^i = F^i(\rho_0)$ , where  $\rho_1^i$  is the value of the  $i$ -th coordinate function at position  $\rho_1$ . Then the push-forward is given in coordinates by:

$$v_1^i = F_{,j}^i v_0^j. \quad (\text{A.69})$$

Sometimes this is written:

$$v_1^i = \frac{\partial \rho_1^i}{\partial \rho_0^j} v_0^j. \quad (\text{A.70})$$

That is, the tangent vectors transform by using the chain rule on the mapping  $F$ . This process is illustrated in Fig. A.2.<sup>3</sup>

---

<sup>3</sup>To better justify these expressions, a more in-depth discussion of tangent vectors is required. This more in-depth discussion considers tangent vectors as tangent to a parameterized path  $\gamma$  in  $M$  at the point  $\rho_0$ . One then calculates the push-forward by advancing the path  $\gamma$  using the map  $F$  and determining the new tangent vector at the point  $\rho_1$ .

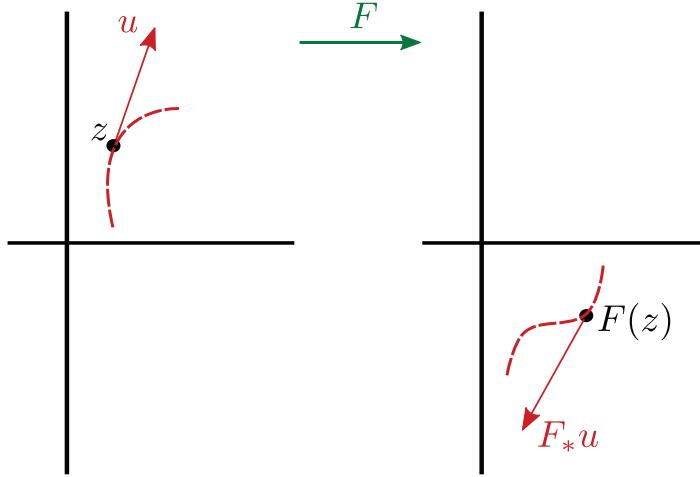


Figure A.2: A map  $F : M \rightarrow M$  advancing  $z$  to  $F(z)$  may be used to advance a tangent vector  $u$  at  $z$  to a tangent vector  $F_*u$  at  $F(z)$ .

The push-forward of a vector can be used to define the pull-back of a differential form. Consider a differential  $k$ -form field  $\alpha$  on the manifold  $M$ . The pull-back allows one to associate  $k$ -forms at the image point  $\rho_1 = F(\rho_0)$  with  $k$ -forms at the original point  $\rho_0$ . To do so, let  $\alpha_\rho$  denote the  $k$ -form specified by the  $k$ -form field  $\alpha$  evaluated at the point  $\rho$ . The pull-back  $F^*\alpha$  defines a  $k$ -form field on  $M$  whose  $k$ -form at the point  $\rho$  is defined by:

$$(F^*\alpha)_\rho(v_0, v_1, \dots, v_k) = \alpha_{F(\rho)}(F_*v_0, F_*v_1, \dots, F_*v_k). \quad (\text{A.71})$$

To describe this process in words, recall that a  $k$ -form is a multilinear operation that associates real numbers to sets of  $k$  tangent vectors at a point. To define a  $k$ -form field on  $M$ , we need to specify how it yields real numbers when given  $k$  tangent vectors at a point. The  $k$ -form field given by the pull-back of another  $k$ -form field  $\alpha$  chooses a real number by pushing forward all  $k$  tangent vectors to the new point  $F(\rho)$ , then evaluates  $\alpha_{F(\rho)}$  using the  $k$  pushed-forward vectors. It sounds a little complicated, but amounts to using the chain rule once for each vector. For instance, consider a one-form field  $\alpha^{(1)}$  given by:  $\alpha^{(1)} = a_i(u)du^i$ .

Then:

$$F^* \alpha^{(1)} = a_i(F(u)) F_{,j}^i du^j, \quad (\text{A.72})$$

where  $F_{,j}^i$  is evaluated at  $u$ . Similar reasoning applies to a two-form field. Let  $\alpha^{(2)} = a_{ij} du^i \wedge du^j$ , where the summation over  $j$  assumes  $j \geq i$  to avoid repeated two-form basis elements. Then:

$$F^* \alpha^{(2)} = a_{ij}(F(u)) F_{,k}^i F_{,l}^j du^k \wedge du^l. \quad (\text{A.73})$$

When one represents this in tensor notation, one can recover a conventional formula used in textbooks for symplectic maps, namely that  $F$  is symplectic with respect to the two-form field  $\Omega$  if:

$$\Omega = \frac{\partial \rho_1}{\partial \rho_0}^T \Omega \frac{\partial \rho_1}{\partial \rho_0}. \quad (\text{A.74})$$

This dissertation will prefer instead the notation described in the following definition.

**Definition - Symplectic Map** Given a manifold  $M$  equipped with a symplectic structure  $\Omega$ , a mapping  $F : M \rightarrow M$  is *symplectic* if:

$$F^* \Omega = \Omega. \quad (\text{A.75})$$

# Appendix B

## GEODES Library

### B.1 Overview

To test the new and existing algorithms considered in this dissertation, a “GEometric ODE Solver” (GEODES) library was developed. The primary goals of the library are: *performance*, *usability*, and *Maintainability*. Toward this end, many features have been incorporated to ensure ease of use and development. Benchmarks have been performed against existing test particle codes. Although the primary application is test particle dynamics, the library has been designed to allow the user to choose a dynamical system and integrator at will. At present, dynamical systems implemented within the GEODES library include guiding center dynamics, Lorentz force dynamics, magnetic field line flow, linear and nonlinear oscillators, and the Hénon-Heiles system [21].

## B.2 Features

### B.2.1 Language

GEODES was programmed using the C++ programming language [144]. Benefits of C++ include performance, portability, and object-oriented capabilities. Because C++ is a compiled language, the run time is typically superior to interpreted languages Mathematica and MATLAB. Further, its memory managing capabilities allow advanced parallelization techniques, such as GPGPU programming (see below). Compilers for C++ are present on all research computing facilities, as far as I am aware. Finally, the object-orienting guides a flexible and extensible design, detailed in Sec. B.2.5.

### B.2.2 Style

“Style” refers to programming conventions in: naming, comments, design, and other best practices. Adhering to a consistent style enhances readability of the code and reduces the likelihood of certain common mistakes. For GEODES, a strong effort has been made to conform to the “Google C++ Style Guide” [145]. This style was chosen because it is likely to be known by other programmers and can be found easily online.

Highlights of the Google C++ Style Guide, as they pertain to GEODES, include naming conventions, variable scoping, and function conventions. For instance, naming conventions enable classifying the entities encountered in a program by inspection of the name. For all entities, the guide advises: “Names should be descriptive; eschew abbreviation” [145]. This means GEODES has names like “RungeKutta4” and “AdamsBashforth2” rather than “rk2” or “ab2”. Variables use lower case with underscores linking words, for instance “b\_hat” and “mod\_b”, unless they are constant in which case they are preceded by a “k” and use mixed case, e.g. “kMu”. Member variables have a trailing underscore, e.g. “b\_hat\_”. Class names have mixed case, e.g. “NonlinearPendulum” and “Integrator”. Class member names also

have mixed case, e.g. “Integrator.Step” and “NonlinearPendulum.VectorField”. In terms of variable scoping, the style guide advises to “place a function’s variables in the narrowest scope possible”. For GEODES, this means there are *no* global variables. This practice greatly reduces debugging time because one does not have to search the entire code base for places where a global variable may be modified. For function conventions, the input arguments are specified first and the output arguments are specified second. It is then easy to discern the intent of a function based on its name and the intent of its arguments based on their position. Additionally, the suggestion to “prefer small and focused functions” was practiced in the development of GEODES. Many function implementations fit within thirty lines, making it easy to read the entirety of their contents on a single screen.

These style conventions should make it easy for future developers to understand existing code and contribute new code with a conformal style.

### B.2.3 Documentation

Documentation for GEODES is generated from in-source comments using Doxygen [146]. The developer uses special characters to flag comments for future parsing by the Doxygen software. Whenever a class, function, or variable is defined, the developer annotates the explanatory comments with these special characters. Later, documentation is constructed from these comments by using Doxygen to parse the source code. This workflow achieves several desirable results: systematic and thorough comments are encouraged, documentation is kept up to date without requiring additional maintenance time, and users are provided with comprehensive and well-organized documentation. Examples of the source code and the resulting documentation are shown in Figs. B.1 and B.2, respectively.

```


/*
 * @brief Destructor to delete EMFields pointer
 */
GuidingCenter::~GuidingCenter(){
    delete em_fields_;
}

/*
 * @brief Evaluates the guiding center equations of motion.
 *
 * Letting \f$ x \f$ be particle position and \f$ u \f$ the velocity:
 * \f{eqnarray}{l}
 * \dot{x} &= u B^\dagger - \frac{b \times E^\dagger}{B^\dagger} \\
 * \dot{u} &= \frac{B^\dagger \cdot \dot{E}}{B^\dagger}
 * \f}
 * @param[in] kt Time
 * @param[in] kx Position of guiding center particle [x u]
 * @param[out] fx Vector field evaluation
 * @return 0 if successful
 */
int GuidingCenter::VectorField(const double kt, const Eigen::VectorXd &kx,
Eigen::VectorXd &fx) const {

    // Declare intermediate convenience variables
    Eigen::Vector3d b_hat; // b_hat = magnetic field unit vector
    ...
}


```

Figure B.1: Sample of in-source comments that Doxygen uses to create documentation (see Fig. B.2).

#### 8.43.1.2 int GuidingCenter::VectorField( const double kt, const VectorX & kx, VectorX & fx ) const [virtual]

Evaluates the guiding center equations of motion.

Letting  $x$  be particle position and  $u$  the velocity:

$$\begin{aligned}\dot{x} &= u B^\dagger - \frac{b \times E^\dagger}{B^\dagger} \\ \dot{u} &= \frac{B^\dagger \cdot \dot{E}}{B^\dagger}\end{aligned}$$

##### Parameters

in	kt	Time
in	kx	Position of guiding center particle [x u]
out	fx	Vector field evaluation

##### Returns

0 if successful

Implements [Model](#).

The documentation for this class was generated from the following files:

- [guiding\\_center.h](#)
- [guiding\\_center.cc](#)

Figure B.2: Documentation created using Doxygen to parse in-source comments.

### B.2.4 Version Control

GEODES is maintained under version control using Git [147]. Version control software facilitates development by maintaining a recoverable history of the project throughout the development process. Git can rapidly change between different versions of the code, allowing developers to revert undesirable changes, inspect modifications made since a certain time, and experiment with new features before deploying them to the primary code base. GEODES has a history of nearly 1800 “commits”, detailing its development from the earliest of stages. New users may easily “clone” the code repository to use or develop the library.

### B.2.5 Design

To ensure flexibility and maintainability, GEODES adopts an object-oriented design. The design for the library grew out of an assignment for the Princeton University course *AST 506: Software Engineering for Scientific Computing*. Object-oriented design organizes fundamental components of a software project into *classes*. A class is a collection of functions and variables related by some unifying feature or purpose. For instance, one fundamental class for GEODES is the “Integrator” class. Any integrator should specify how to advance the state of a dynamical system by some time increment. As such, any class that “inherits” from the Integrator base class is required to implement a function called “Step”, and all integrators have a data member that specifies the step size. The object-oriented design of this class ensures a uniform interface for any integrator, decoupling the code that uses integrators (say, the driver of the ODE solver) from the actual integrators themselves. Object-oriented design also facilitates code re-use. For instance, all implicit integrators inherit from the “ImplicitIntegrator” base class, meaning they can call the nonlinear solve methods defined in ImplicitIntegrator. The nonlinear solves are therefore defined in one place alone, and changes can be made there rather than in each individual integrator that happens to be

implicit. Overall, the object-oriented approach makes GEODES well-organized and flexible for future modification.

### B.2.6 Unit testing

GEODES is systematically and automatically tested using the Google Test unit testing library [148]. Each function has “unit tests” written to verify the desired functionality. A unit test is a self-contained test of some subset of code intended to accomplish a particular task. By maintaining a library of unit tests, one may rapidly verify the behavior of the entire code

base while performing modifications. This approach

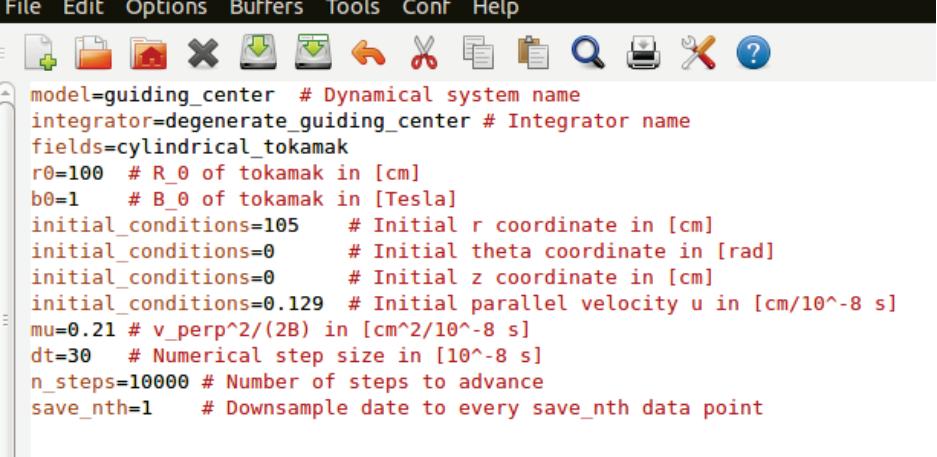
is useful both for debugging - the end result of which is often a new unit test - and for catching unintended consequences of new modifications. The tests also serve as a baseline verification that the library is working properly when being deployed on a new system. New tests may be easily contributed by using the built-in functionality of the Google Test library. GEODES presently uses over 200 unit tests, which can all be executed immediately following compilation.



```
[ OK ] ModelTest.BFieldR
[ RUN ] ModelTest.Perturbed
[ OK ] ModelTest.Perturbed
[ RUN ] ModelTest.Perturbed
[ OK ] ModelTest.Perturbed
[ RUN ] ModelTest.Perturbed
[ OK ] ModelTest.Perturbed
[ RUN ] ModelTest.Nonlinear
[ OK ] ModelTest.Nonlinear
[ RUN ] ModelTest.Nonlinear
```

### B.2.7 Input Specification

When writing scientific software, the number of user-specified options can quickly become cumbersome. To manage user-specified options - including dynamical system parameters, integrator options, initial conditions, and output settings - GEODES uses the *program\_options* library of Boost. With this library, the user may specify intuitively-named options in a configuration file or command line, and developers may add new options with a single line



```

File Edit Options Buffers Tools Conf Help
model=guiding_center # Dynamical system name
integrator=degenerate_guiding_center # Integrator name
fields=cylindrical_tokamak
r0=100 # R_0 of tokamak in [cm]
b0=1 # B_0 of tokamak in [Tesla]
initial_conditions=105 # Initial r coordinate in [cm]
initial_conditions=0 # Initial theta coordinate in [rad]
initial_conditions=0 # Initial z coordinate in [cm]
initial_conditions=0.129 # Initial parallel velocity u in [cm/10^-8 s]
mu=0.21 # v_perp^2/(2B) in [cm^2/10^-8 s]
dt=30 # Numerical step size in [10^-8 s]
n_steps=10000 # Number of steps to advance
save_nth=1 # Downsample date to every save_nth data point

```

Figure B.3: input file example

of code. All parsing of the configuration file and command line options is performed by the `program_options` library. The net result is an intuitive interface for GEODES, from which a user may access the full functionality of the software without needing to re-compile any of the code. An example of an input file is shown in Fig. B.3; the simplicity and clarity of the configuration file should be clear even for new users.

### B.2.8 Parallelization

Parallelization is essential for operating on high performance research computing platforms. Fortunately, the independence of test particles makes test particle calculations fully parallelizable. GEODES takes advantage of this parallel structure of the problem to deploy simulations on General-Purpose Graphics Processing Units (GPGPUs). In detail, the CUDA Toolkit is used to implement a driver for the GEODES library that runs on GPGPU hardware [149]. GPGPUs contain many (several thousand) cores that execute the same sets of instructions. Although any one core is much slower than a single modern CPU core, the combined capabilities of the many cores on the GPGPU exceed that of a modern CPU.

Fig. B.4 demonstrates the time-to-solution improvements obtained by performing the calculation of many guiding center trajectories using a GPGPU relative to that of a single

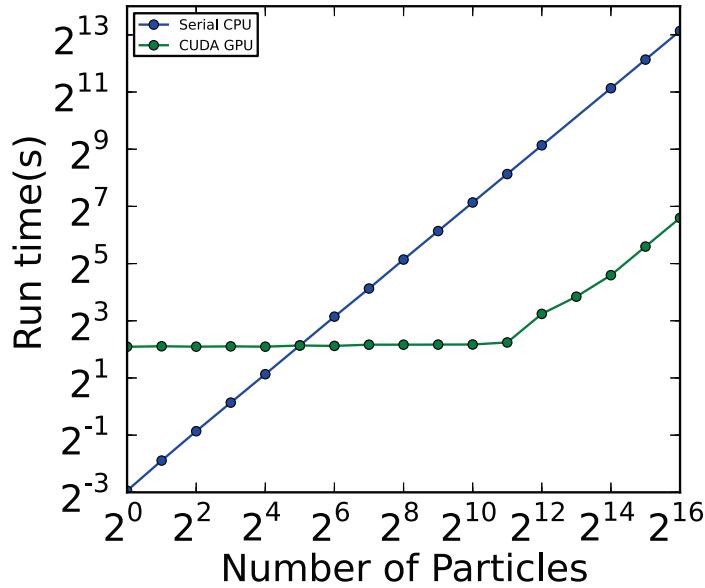


Figure B.4: Time-to-solution comparison using CPU and GPGPU hardware for guiding center trajectories. For small numbers of particles, the CPU is superior, but at large numbers of particles, the GPGPU is superior. The CPU is a 2.5 GHz AMD Opteron, while the GPGPU is a 2496 core nVidia Tesla k20m card.

core of a CPU. For small numbers of particles, the CPU achieves a faster time-to-solution, with the limiting case of a single particle being approximately thirty times faster to compute on a CPU rather than the GPU. As the number of particles is increased, however, the CPU time increases linearly while the GPGPU time increases negligibly until the number of particles exceeds the number of cores ( $2048 = 2^{11}$  cores on this GPGPU card). In the large numbers of particles limit, the GPGPU is approximately 64 times faster than the single core of the CPU. To be fair, most modern CPUs come equipped with multiple cores that can be used using MPI or OpenMP, for instance. Still, with a factor of 8-16 adjustment for this fact, the GPGPU is superior by a factor of four to eight. Additionally, many high performance computing facilities have computational nodes with both CPU and GPGPU units, so it is advantageous to take advantage of these “heterogeneous” architectures.

# Bibliography

- [1] L. D. Landau and E. M. Lifshitz, *Mechanics*. Pergamon Press, 1969.
- [2] R. Abraham and J. E. Marsden, *Foundations of Mechanics*. Addison-Wesley Publishing Company, 1987.
- [3] V. I. Arnold, *Mathematical Methods of Classical Mechanics*. Springer, 1989.
- [4] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*. Springer Science and Business Media, 1999.
- [5] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*. Addison Wesley, 2001.
- [6] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*. Springer, 2006.
- [7] R. I. McLachlan and G. R. W. Quispel, “Geometric integrators for ODEs,” *Journal of Physics A: Mathematical and General*, vol. 39, pp. 5251–5285, 2006.
- [8] R. G. Littlejohn, “Variational principles of guiding centre motion,” *Journal of Plasma Physics*, vol. 29, no. 01, pp. 111–125, 1983.
- [9] J. E. Marsden and M. West, “Discrete mechanics and variational integrators,” *Acta Numerica*, pp. 1–158, 2001.
- [10] F. Bashforth and J. C. Adams, *An attempt to test the theories of capillary action by comparing the theoretical and measured forms of drops of fluid, with an explanation of the method of integration employed in constructing the tables which give the theoretical forms of such drops*. Cambridge University Press, 1883.
- [11] G. Dahlquist, “Convergence and stability in the numerical integration of ordinary differential equations,” *Mathematica Scandinavica*, vol. 4, pp. 33–53, 1956.
- [12] E. Hairer, “Backward error analysis for multistep methods,” *Numerische Mathematik*, vol. 84, pp. 199–232, 1999.
- [13] C. W. Rowley and J. E. Marsden, “Variational integrators for degenerate Lagrangians, with application to point vortices,” *Proceedings of the 41st IEEE Conference on Decision and Control*, vol. 2, pp. 1521–1527, 2002.

- [14] H. Qin and X. Guan, “Variational symplectic integrator for long-time simulations of the guiding-center motion of charged particles in general magnetic fields,” *Physical Review Letters*, vol. 100, no. 3, p. 035006, 2008.
- [15] H. Qin, X. Guan, and W. M. Tang, “Variational symplectic algorithm for guiding center dynamics and its application in tokamak geometry,” *Physics of Plasmas*, vol. 16, no. 04, p. 042510, 2009.
- [16] J. Li, H. Qin, Z. Pu, L. Xie, and S. Fu, “Variational symplectic algorithm for guiding center dynamics in the inner magnetosphere,” *Physics of Plasmas*, vol. 18, no. 05, p. 052902, 2011.
- [17] M. Kraus, *Variational Integrators in Plasma Physics*. Doctoral Thesis, Technische Universität München, 2013.
- [18] S. Ober-Blöbaum, M. Tao, M. Cheng, H. Owhadi, and J. E. Marsden, “Variational integrators for electric circuits,” *Journal of Computational Physics*, 2013.
- [19] C. L. Ellison, J. M. Finn, H. Qin, and W. M. Tang, “Development of variational guiding center algorithms for parallel calculations in experimental magnetic equilibria,” *Plasma Physics and Controlled Fusion*, vol. 57, p. 054007, 2015.
- [20] T. Frankel, *The Geometry of Physics: An Introduction*. Cambridge University Press, third ed., 2012.
- [21] A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion*. Springer Science + Business Media, 1983.
- [22] T. Itoh and K. Abe, “Hamiltonian-conserving discrete canonical equations based on variational difference quotients,” *Journal of Computational Physics*, vol. 77, pp. 85–102, 1988.
- [23] D. Cohen and E. Hairer, “Linear energy-preserving integrators for Poisson systems,” *BIT Numer Math*, vol. 51, pp. 91–101, 2011.
- [24] S. Reich, “Momentum conserving symplectic integrators,” *Physica D: Nonlinear Phenomena*, vol. 76, pp. 375–383, 1994.
- [25] J. C. Simo and O. Gonzalez, “Assessment of energy-momentum and symplectic schemes for stiff differential systems,” *ASME Winter Annual Meeting*, 2011.
- [26] C. Kane, J. E. Marsden, and M. Ortiz, “Symplectic-energy-momentum preserving variational integrators,” *Journal of Mathematical Physics*, vol. 40, p. 3353, 1999.
- [27] S. Zaijiu, “On the construction of the volume-preserving difference schemes for source-free systems via generating functions,” *Journal of Computational Mathematics*, vol. 12, pp. 265–272, 1994.

- [28] G. R. W. Quispel, “Volume-preserving integrators,” *Physics Letters A*, vol. 206, pp. 26–20, 1995.
- [29] G. R. W. Quispel and C. P. Dyt, “Volume-preserving integrators have linear error growth,” *Physics Letters A*, vol. 242, pp. 25–30, 1998.
- [30] Y. He, Y. Sun, J. Liu, and H. Qin, “Volume-preserving algorithms for charged particle dynamics,” *Journal of Computational Physics*, vol. 281, pp. 135–147, 2015.
- [31] R. D. Ruth, “A canonical integration technique,” *IEEE Transactions on Nuclear Science*, vol. NS-30, pp. 2669–2671, 1983.
- [32] J. M. Sanz-Serna, “Runge-Kutta schemes for Hamiltonian systems,” *BIT*, vol. 28, pp. 877–883, 1988.
- [33] J. M. Sanz-Serna, “Symplectic integrators for hamiltonian problems: an overview,” *Acta Numerica*, vol. 1, pp. 243–286, 1992.
- [34] P. J. Channell and C. Scovel, “Symplectic integration of Hamiltonian systems,” *Nonlinearity*, vol. 3, pp. 231–259, 1990.
- [35] E. Forest and R. D. Ruth, “Fourth-order symplectic integration,” *Physica D*, vol. 43, pp. 105–117, 1990.
- [36] H. Yoshida, “Construction of higher order symplectic integrators,” *Physics Letters A*, vol. 150, no. 5, 6, 7, pp. 262–268, 1990.
- [37] J. Candy and W. Rozmus, “A symplectic integration algorithm for separable Hamiltonian functions,” *Journal of Computational Physics*, vol. 92, pp. 230–256, 1991.
- [38] V. I. Arnold, “Instability of dynamical systems with several degrees of freedom,” *Sov. Math. Doklady*, vol. 5, pp. 581–585, 1964.
- [39] B. V. Chirikov, “A universal instability of many-dimensional oscillator systems,” *Physics Reports*, vol. 52, no. 5, pp. 263 – 379, 1979.
- [40] H. Wobig and D. Pfirsch, “On guiding centre orbits of particles in toroidal systems,” *Plasma Physics and Controlled Fusion*, vol. 43, pp. 695–716, 2001.
- [41] C. Froeschlé, M. Guzzo, and E. Lega, “Graphical evolution of the arnold web: From order to chaos,” *Science*, vol. 289, pp. 2108 – 2110, 2000.
- [42] A. Seibert, S. Denisov, A. V. Ponomarev, and P. Hägggi, “Mapping the Arnold web with a graphic processing unit,” *Chaos*, vol. 21, p. 043123, 2011.
- [43] J. Wisdom and M. Holman, “Symplectic maps for the  $n$ -body problem: stability analysis,” *The Astronomical Journal*, vol. 104, no. 5, p. 2022, 1992.

- [44] H. Yoshida, “Recent progress in the theory and application of symplectic integrators,” *Celestial Mechanics and Dynamical Astronomy*, vol. 56, pp. 27–43, 1993.
- [45] É. Forest, “Geometric integration for particle accelerators,” *Journal of Physics A: Mathematical and General*, vol. 39, pp. 5321–5377, 2006.
- [46] R. B. White and M. S. Chance, “Hamiltonian guiding center drift orbit calculation for plasmas of arbitrary cross section,” *Physics of Fluids*, vol. 27, no. 10, p. 2455, 1984.
- [47] R. J. Akers, E. Verwichte, T. J. Martin, S. D. Pinches, and R. Lake, “GPGPU Monte Carlo calculation of gyro-phase resolved fast ion and n-state resolved neutral deuterium distributions,” in *Proceedings of the 39th EPS Conference on Plasma Physics*, p. P5.088, 2012.
- [48] J. L. Velasco, A. Bustos, F. Castejon, L. A. Fernandez, V. Martin-Mayor, and A. Tarancón, “ISDEP: Integrator of stochastic differential equations for plasmas,” *Computer Physics Communications*, vol. 183, no. 9, p. 1877, 2012.
- [49] G. J. Kramer, R. V. Budny, A. Bortolon, E. D. Fredrickson, G. Y. Fu, W. W. Heidbrink, R. Nazikian, E. Valeo, and M. A. V. Zeeland, “A description of the full-particle-orbit-following spiral code for simulating fast-ion experiments in tokamaks,” *Plasma Physics and Controlled Fusion*, vol. 55, p. 025013, 2013.
- [50] M. McMillan and S. A. Lazerson, “BEAMS3D neutral beam injection model,” *Plasma Physics and Controlled Fusion*, vol. 56, p. 095019, 2014.
- [51] D. Pfefferl  , J. P. Graves, W. A. Cooper, C. Misev, I. T. Chapman, and M. T. ans S Sangaroon, “NBI fast ion confinement in the helical core of MAST hybrid-like plasmas,” *Nuclear Fusion*, vol. 54, p. 064020, 2014.
- [52] E. Hirvijoki, O. Asunta, T. Koskela, T. Kurki-Suonio, J. Meittunen, S. Sipil  , A. Snicker, and S.   k  slompolo, “Ascot: Solving the kinetic equation of minority particle species in tokamak plasmas,” *Computational Physics Communications*, vol. 185, pp. 1310–1321, 2014.
- [53] E. Hirvijoki, T. Kurki-Suonio, S.   k  slompolo, J. Varje, T. Koskela, and J. Meittunen, “Monte carlo method and high performance computing for solving fokker-plank equation of minority plasma particles,” *Journal of Plasma Physics*, vol. 81, p. 435810301, 2015.
- [54] R. J. Goldston and H. H. Towner, “Effects of toroidal field ripple on suprathermal ions in tokamak plasmas,” *Journal of Plasma Physics*, vol. 26, pp. 283–307, 10 1981.
- [55] N. N. Gorelenkov and R. B. White, “Perturbative study of energetic particle redistribution by Alfv  n eigenmodes in ITER,” *Plasma Physics and Controlled Fusion*, vol. 49, no. 09, p. 095021, 2013.

- [56] K. G. McClements, R. J. Akers, W. U. Boeglin, M. Cecconello, D. Keeling, O. M. Jones, A. Kirk, I. Klimek, R. V. Perez, K. Shinohara, and K. Tani, “The effects of resonant magnetic perturbations on fast ion confinement in the Mega Amp Spherical Tokamak,” *Plasma Physics and Controlled Fusion*, vol. 57, p. 075003, 2015.
- [57] R. B. White, “High energy particles in tokamaks,” in *AIP Conference Proceedings*, p. 59, 2008.
- [58] A. Bustos, F. Castejon, L. A. Fernandez, J. Garcia, V. Martin-Mayor, J. M. Reynolds, R. Seki, and J. L. Velasco, “Impact of 3d features on ion collisional transport in ITER,” *Nuclear Fusion*, vol. 50, no. 12, p. 125007, 2010.
- [59] R. B. White, N. Gorelenkov, W. W. Heidbrink, and M. A. Van Zeeland, “Beam distribution modification by alfvén modes),” *Physics of Plasmas*, vol. 17, no. 5, pp. –, 2010.
- [60] R. B. White, N. Gorelenkov, W. W. Heidbrink, and M. A. V. Zeeland, “Particle distribution modification by low amplitude modes,” *Plasma Physics and Controlled Fusion*, vol. 52, no. 4, p. 045012, 2010.
- [61] I. T. Chapman, “Controlling sawtooth oscillations in tokamak plasmas,” *Plasma Physics and Controlled Fusion*, vol. 53, no. 1, p. 013001, 2011.
- [62] N. Gorelenkov, S. Pinches, and K. Toi, “Energetic particle physics in fusion research in preparation for burning plasma experiments,” *Nuclear Fusion*, vol. 54, no. 12, p. 125001, 2014.
- [63] S. D. Pinches, I. T. Chapman, P. W. Lauber, H. J. C. Oliver, S. E. Sharapov, K. Shinohara, and K. Tani, “Energetic ions in ITER plasmas,” *Physics of Plasmas*, vol. 22, p. 021807, 2015.
- [64] A. Fasoli, C. Gormenzano, H. L. Berk, B. Breizman, S. Briguglio, D. S. Darrow, N. Gorelenkov, W. W. Heidbrink, A. Juan, S. V. Konovalov, R. Nazikian, J.-M. Norterdaeme, S. Sharapov, K. Shinohara, D. Testa, K. Tobita, Y. Todo, G. Vlad, and F. Zonca, “Chapter 5: Physics of energetic ions,” *Nuclear Fusion*, vol. 47, pp. S246–S284, 2007.
- [65] H. ITER Physics Expert Group on Energetic Particles, C. Drive, and I. P. B. Editors, “Chapter 5: Physics of energetic ions,” *Nuclear Fusion*, vol. 39, no. 12, p. 2471, 1999.
- [66] X. Guan, H. Qin, and N. J. Fisch, “Phase-space dynamics of runaway electrons in tokamaks,” *Physics of Plasmas*, vol. 17, p. 092502, 2010.
- [67] A. H. Boozer, “Theory of runaway electrons in ITER: Equations, important parameters, and implications for mitigation,” *Physics of Plasmas*, vol. 22, p. 032504, 2015.

- [68] A. J. Brizard and T. S. Hahm, “Foundations of nonlinear gyrokinetic theory,” *Reviews of Modern Physics*, vol. 79, p. 421, 2007.
- [69] J. A. Krommes, “The gyrokinetic description of microturbulence in magnetized plasmas,” *Annual Review of Fluid Mechanics*, vol. 44, pp. 175–201, 2012.
- [70] C. Chang, S. Ku, and H. Weitzner, “Numerical study of neoclassical plasma pedestal in tokamak geometry,” *Physics of Plasmas*, vol. 11, no. 5, p. 2649, 2004.
- [71] S. Ethier, W. M. Tang, and Z. Lin, “Gyrokinetic particle-in-cell simulations of plasma microturbulence on advanced computing platforms,” *Journal of Physics: Conference Series*, vol. 16, pp. 1–15, 2005.
- [72] R. A. Kolesnikov, W. Wang, F. L. Hinton, G. Rewoldt, and W. M. Tang, “Drift-kinetic simulation of neoclassical transport with impurities in tokamaks,” *Physics of Plasmas*, vol. 17, no. 02, p. 2506, 2010.
- [73] J. W. Burby, A. J. Brizard, P. J. Morrison, and H. Qin, “Hamiltonian gyrokinetic Vlasov-Maxwell system,” *Physics Letters A*, vol. 379, no. 36, pp. 2073–2077, 2015.
- [74] B. Obama, “Executive Order – Creating a National Strategic Computing Initiative.” Executive Order, 2015.
- [75] B. Karasözen, “Poisson integrators,” *Mathematical and Computer Modelling*, vol. 40, pp. 1225–1244, 2004.
- [76] G. Darboux, “Sur le problème de Pfaff,” *Bulletin des Sciences Mathématiques*, vol. 6, pp. 14–36, 49–68, 1882.
- [77] R. White and L. E. Zakharov, “Hamiltonian guiding center equations in toroidal magnetic configurations,” *Physics of Plasmas*, vol. 10, no. 3, p. 573, 2003.
- [78] J. R. Cary and A. J. Brizard, “Hamiltonian theory of guiding-center motion,” *Reviews of Modern Physics*, vol. 81, no. 02, p. 693, 2009.
- [79] R. Zhang, J. Liu, Y. Tang, H. Qin, J. Xiao, and B. Zhu, “Canonicalization and symplectic simulation of the gyrocenter dynamics in time-independent magnetic fields,” *Physics of Plasmas*, vol. 21, no. 03, p. 0325047, 2014.
- [80] J. C. Butcher, “Numerical methods for ordinary differential equations in the 20th century,” *Journal of Computational and Applied Mathematics*, vol. 125, pp. 1–29, 2000.
- [81] C. Runge, “Über die numerische Auflösung von Differentialgleichungen,” *Math. Ann.*, vol. 46, pp. 167–178, 1895.
- [82] W. Kutta, “Beitrag zur näherungsweisen Integration totaler Differentialgleichungen,” *Z. Math. Phys.*, vol. 46, pp. 435–453, 1901.

- [83] F. R. Moulton, *New methods in exterior ballistics*. University of Chicago, 1926.
- [84] H. F. Trotter, “On the product of semi-groups of operators,” *Proceedings of the American Mathematical Society*, vol. 10, pp. 545–551, 1959.
- [85] G. Strang, “On the construction and comparison of difference schemes,” *SIAM Journal on Numerical Analysis*, vol. 5, no. 3, pp. 506–517, 1968.
- [86] M. West, *Variational Integrators*. Ph.d. thesis, California Institute of Technology, 2004.
- [87] H. Yoshimura and J. E. Marsden, “Dirac structures and Lagrangian mechanics. Part I: Implicit Lagrangian systems,” *Journal of Geometry and Physics*, vol. 57, no. 1, pp. 133–156, 2006.
- [88] H. Yoshimura and J. E. Marsden, “Dirac structures and Lagrangian mechanics. Part II: Variational structures,” *Journal of Geometry and Physics*, vol. 57, no. 1, pp. 209–250, 2006.
- [89] H. Yoshimura and J. E. Marsden, “Dirac structures and implicit Lagrangian systems in electric networks,” *17th Internatoinal Symposium on Mathematical Theory of Networks and Systems*, pp. 1444–1449, 2006.
- [90] S. Lall and M. West, “Discrete variational Hamiltonian mechanics,” *Journal of Physics A*, vol. 39, pp. 5509–5519, 2006.
- [91] M. Leok and J. Zhang, “Discrete Hamiltonian variational integrators,” *IMA Journal of Numerical Analysis*, vol. 31, no. 4, pp. 1497–1532, 2011.
- [92] T. Tyranowski and M. Desbrun, “Variational partitioned Runge-Kutta methods for Lagrangians linear in velocities,” *arXiv:1401.7904*, 2014.
- [93] U. Kirchgraber, “Multi-step methods are essentially one-step methods,” *Numerische Methematik*, vol. 48, pp. 85–90, 1986.
- [94] K. Feng, “The step-transition operators for multi-step methods of ODEs,” *Journal of Computational Mathematics*, vol. 16, no. 3, pp. 193–202, 1998.
- [95] Y.-F. Tang, “The symplecticity of multi-step methods,” *Computers Math. Applic.*, vol. 25, no. 3, pp. 83–90, 1993.
- [96] J. Squire, H. Qin, and W. M. Tang, “Gauge properties of the guiding center variational symplectic integrator,” *Physics of Plasmas*, vol. 19, no. 05, p. 052501, 2012.
- [97] J. R. Cary and R. G. Littlejohn, “Noncanonical Hamiltonian mechanics and its application to magnetic field line flow,” *Annals of Physics*, vol. 151, pp. 1–34, 1983.
- [98] P. J. Morrison, “Magnetic field lines, Hamiltonian dynamics, and nontwist systems,” *Physics of Plasmas*, vol. 7, p. 2279, 2000.

- [99] T. H. Stix, “Heating of toroidal plasmas by neutral injection,” *Plasma Physics*, vol. 14, pp. 367–384, 1972.
- [100] S. Ober-Blöbaum and N. Saake, “Construction and analysis of higher order Galerkin variational integrators,” *arXiv*, vol. arXiv:1304.1398, 2013.
- [101] M. Leok and T. Shingel, “Prolongation-collocation variational integrators,” *IMA Journal of Numerical Analysis*, vol. 32, pp. 1194–1216, 2011.
- [102] M. Leok and T. Shingel, “General techniques for constructing variational integrators,” *Frontiers of Mathematics in China*, vol. 7, pp. 273–303, 2012.
- [103] J. Hall and M. Leok, “Spectral variational integrators,” *Numerische Mathematik*, pp. 1–60, 2014.
- [104] J. W. Burby, J. Squire, and H. Qin, “Automation of the guiding center expansion,” *Physics of Plasmas*, vol. 20, p. 072105, 2013.
- [105] A. J. Dragt and J. M. Finn, “Insolubility of trapped particle motion in a magnetic dipole field,” *Journal of Geophysical Research*, vol. 81, no. 13, p. 2327, 1976.
- [106] P. A. M. Dirac, “Generalized Hamiltonian dynamics,” *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 246, pp. 326–332, 1958.
- [107] J. W. Burby, C. L. Ellison, and H. Qin, “Variational integrators for perturbed non-canonical Hamiltonian systems,” *Unpublished*, 2014.
- [108] G. D. Quinlan and S. Tremaine, “Symmetric multistep methods for the numerical integration of planetary orbits,” *Astronomical Journal*, vol. 100, no. 5, pp. 1694–1700, 1990.
- [109] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer Series in Computational Mathematics, Vol 8, 1993.
- [110] G. D. Quinlan, “Resonances and instabilities in symmetric multistep methods,” *Unpublished*, 1999.
- [111] E. Hairer and C. Lubich, “Symmetric multistep methods over long times,” *Numerische Mathematik*, vol. 97, pp. 699–723, 2004.
- [112] E. Hairer, “Symmetric linear multistep methods,” *BIT Numerical Mathematics*, vol. 46, no. 3, pp. 515–524, 2006.
- [113] P. Console, E. Hairer, and C. Lubich, “Symmetric multistep methods for constrained Hamiltonian systems,” *Numerische Mathematik*, vol. 124, pp. 517–539, 2012.
- [114] R. D’Ambrosio and E. Hairer, “Long-term stability of multi-value methods for ordinary differential equations,” *Journal of Scientific Computing*, vol. 60, pp. 627–640, 2014.

- [115] H. Helmholtz, “Über die physikalische Bedeutung des Prinzips der kleinsten Wirkung,” *J. Reine Angew. Math.*, vol. 100, pp. 137–166, 1887.
- [116] J. Douglas, “Solution of the inverse problem of the calculus of variations,” *Trans. Amer. Math. Soc.*, vol. 50, pp. 137–166, 1941.
- [117] L. Bourdin and J. Cresson, “Helmholtz’s inverse problem of the discrete calculus of variations,” *Journal of Difference Equations and Applications*, vol. 19, no. 9, pp. 1417–1436, 2013.
- [118] D. Stoffer, “General linear methods: connection to one step methods and invariant curves,” *Numerische Mathematik*, vol. 64, pp. 395–407, 1993.
- [119] B. Cano and J. M. Sanz-Serna, “Error growth in the numerical integration of periodic orbits by multistep methods, with application to reversible systems,” *IMA Journal of Numerical Analysis*, vol. 18, pp. 57–75, 1998.
- [120] A. Aoyagi and K. Abe, “Runge-Kutta smoother for suppression of computational-mode instability of leapfrog scheme,” *Journal of Computational Physics*, vol. 93, no. 2, pp. 287–296, 1991.
- [121] K. C. B. New, K. Watt, C. W. Misner, and J. M. Centrella, “Stable 3-level leapfrog integration in numerical relativity,” *Physical Review D*, vol. 58, no. 6, p. 064022, 1998.
- [122] E. Hairer and P. Leone, “Order barriers for symplectic multi-value methods,” in *Proc. of the 17th Dundee Biennial Conference*, pp. 133–149, Pitman Research, 1997.
- [123] J. Vankerschaver and M. Leok, “A novel formulation of point vortex dynamics on the sphere: Geometrical and numerical aspects,” *Journal of Nonlinear Science*, pp. 01–37, 2013.
- [124] C. L. Ellison, J. W. Burby, and H. Qin, “Comment on ”symplectic integration of magnetic systems”: a proof that the boris algorithm is not variational,” *Journal of Computational Physics*, vol. 301, pp. 489–493, 2015.
- [125] N. D. Mermin, “What’s wrong with this pillow?,” *Physics Today*, vol. 42, pp. 9–11, 1989.
- [126] N. D. Mermin, “Could Feynman have said this?,” *Physics Today*, vol. 57, pp. 9–11, 2004.
- [127] R. I. McLachlan, “Composition methods in the presence of small parameters,” *BIT Numerical Mathematics*, vol. 35, no. 2, pp. 258–268, 1995.
- [128] J. E. Chambers and M. A. Murison, “Pseudo-high-order symplectic integrators,” *The Astronomical Journal*, vol. 119, pp. 425–433, 2000.

- [129] J. Laskar and P. Robutel, “High order symplectic integrators for perturbed Hamiltonian systems,” *Celestial Mechanics and Dynamical Astronomy*, vol. 80, no. 1, pp. 39–62, 2001.
- [130] W. M. Farr, “Variational integrators for almost-integrable systems,” *Celestial Mechanics and Dynamical Astronomy*, vol. 103, no. 2, pp. 105–118, 2009.
- [131] Y.-T. Lau and J. M. Finn, “Dynamics of a three-dimensional incompressible flow with stagnation points,” *Physica D*, vol. 57, pp. 283–310, 1992.
- [132] E. C. da Silva, I. L. Caldas, and R. L. Viana, “Field line diffusion and loss in a tokamak with an ergodic magnetic limiter,” *Physics of Plasmas*, vol. 8, pp. 2855–2865, 2001.
- [133] J. M. Finn and L. Chacón, “Volume preserving integrators for solenoidal fields on a grid,” *Physics of Plasmas*, vol. 12, p. 054503, 2005.
- [134] J. M. Finn, “Issues in measure-preserving three dimensional flow integrators: Self-adjointness, reversibility, and non-uniform time stepping,” *Physics of Plasmas*, vol. 22, p. 032508, 2015.
- [135] R. Albanese, M. D. Magistris, R. Fresa, F. Maviglia, and S. Minucci, “Accuracy assessment of numerical tracing of three-dimensional magnetic field lines in tokamaks with analytical invariants,” *Fusion Science and Technology*, vol. 68, p. 741, 2015.
- [136] A. S. Richardson and J. M. Finn, “Symplectic integrators with adaptive time steps,” *Plasma Physics and Controlled Fusion*, vol. 54, no. 01, p. 014004, 2012.
- [137] J. A. L. Cami, *Stochastic Geometric Mechanics*. Doctoral dissertation, Universidad de Zaragoza, 2008.
- [138] J. W. Burby, A. I. Zhmoginov, and H. Qin, “Hamiltonian mechanics of stochastic acceleration,” *Physical Review Letters*, vol. 111, no. 19, p. 195001, 2013.
- [139] J. W. Burby, *Chasing Hamiltonian Structure in Gyrokinetic Theory*. Doctoral dissertation, Princeton University, 2015.
- [140] N. Bou-Rabee and H. Owhadi, “Stochastic variational integrators,” *IMA Journal of Numerical Analysis*, vol. 29, no. 2, p. 421, 2008.
- [141] N. Bou-Rabee and H. Owhadi, “Long-run accuracy of variational integrators in the stochastic context,” *SIAM Journal of Numerical Analysis*, vol. 48, no. 2, pp. 278–297, 2010.
- [142] J. Squire, H. Qin, and W. M. Tang, “Geometric integration of the Vlasov-Maxwell system with a variational particle-in-cell scheme,” *Physics of Plasmas*, vol. 19, no. 08, p. 084501, 2012.

- [143] H. Qin, J. Liu, J. Xiao, R. Zhang, Y. He, Y. Sun, J. Burby, C. Ellison, and Y. Zhou, “Canonical symplectic particle-in-cell method for long-term large-scale simulations of the Vlasov-Maxwell equations,” *Nuclear Fusion*, vol. 56, p. 014001, 2016.
- [144] Standard C++ Foundation, “ISO C++.” <https://isocpp.org>. Accessed: 2016-02-23.
- [145] Google Inc., “Google C++ Style Guide.” <https://google.github.io/styleguide/cppguide.html>. Accessed: 2016-02-23.
- [146] D. van Heesch, “Doxygen.” [www.stack.nl/~dimitri/doxygen](http://www.stack.nl/~dimitri/doxygen). Accessed: 2016-02-23.
- [147] L. Torvalds *et al.*, “Git.” <https://git-scm.com>. Accessed: 2016-02-23.
- [148] Google Inc., “Google Test.” <https://github.com/google/googletest>. Accessed: 2016-02-23.
- [149] NVIDIA Corporation, “CUDA Toolkit.” <https://developer.nvidia.com/cuda-toolkit>. Accessed: 2016-02-23.