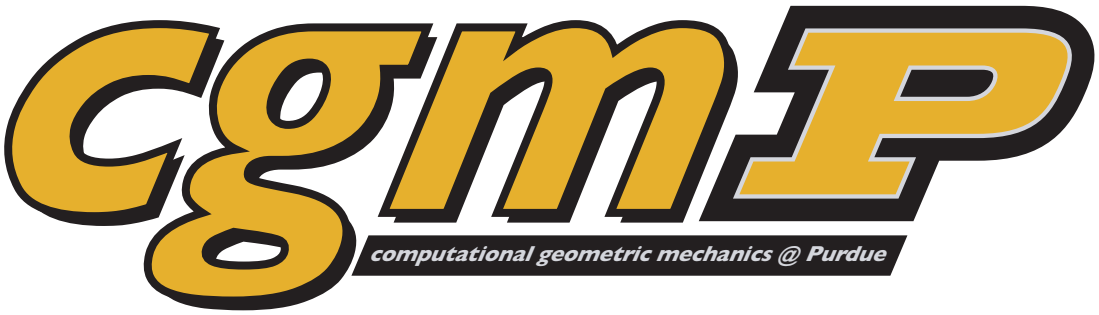
Melvin Leok, Ph.D.
Assistant Professor of Mathematics
Purdue University, West Lafayette

Papers and Preprints



2008

Melvin Leok: Curriculum Vitae

Department of Mathematics	*phone*:	+1(765)496-3578
Purdue University	*fax*:	+1(765)494-0548
Mathematical Sciences Building, 150 North University Street	*e-mail*:	mleok@math.purdue.edu
West Lafayette, IN 47907-2067, USA.	*homepage*:	http://www.math.purdue.edu/~mleok/

Education

California Institute of Technology

Ph.D. Control & Dynamical Systems, Applied & Computational Mathematics (minor) Oct 2000–Jun 2004
Thesis: *Foundations of Computational Geometric Mechanics*
Committee: *Jerrold E. Marsden (advisor), Thomas Y. Hou, Richard M. Murray, Michael Ortiz, and Alan D. Weinstein (Mathematics, UC Berkeley).*

M.S. Mathematics Oct 1999–Jun 2000

B.S. Mathematics (*with honor*) Oct 1996–Jun 2000

Professional Experience

Assistant Professor (Tenure-Track), Mathematics, Purdue University. Aug 2006–present

T.H. Hildebrandt Research Assistant Professor, Mathematics, University of Michigan. Sep 2004–Aug 2006

Postdoctoral Scholar, Control and Dynamical Systems, California Institute of Technology. Jul 2004–Aug 2004

Research Interests

Computational Geometric Mechanics: Applying discrete differential geometry and discrete Lagrangian and Hamiltonian mechanics to the construction of geometric structure-preserving numerical algorithms.

Computational Geometric Control Theory: Construction of real-time digital feedback control of mechanical systems using techniques from geometric control theory and computational geometric mechanics.

Numerical Analysis: Derivation of accurate and efficient numerical schemes with good long-time geometric stability properties by combining geometric integration with adaptive, spectral, and multiscale techniques.

Research Prizes

SciCADE New Talent Prize, International Conference on Scientific Computation and Differential Equations. 2007

SIAM Student Paper Prize, Society for Industrial and Applied Mathematics. 2003

Leslie Fox Prize in Numerical Analysis (second prize), Institute of Mathematics and its Applications, UK. 2003

Research Grants

(Total of $746,298 in single PI awards)

NSF Faculty Early Career Development (CAREER) Award, DMS-0747659. $455,188, single PI. 2008–2013

NSF Computational Mathematics LTB Grant, DMS-0714223. $163,743, single PI. 2007–2010

NSF Applied Mathematics Grant, DMS-0504747 (transferred to DMS-0726263). $108,067, single PI. 2005–2009

Margaret and Herman Sokol Spring/Summer Research Grant, University of Michigan. $4,000, single PI. 2006

Horace H. Rackham Faculty Fellowship, University of Michigan. $7,000, single PI. 2005

Horace H. Rackham Faculty Grant, University of Michigan. $7,300, single PI. 2004–2005

Grant-in-Aid of Research, Sigma Xi, The Scientific Research Society. $1,000, single PI. 2001

Travel Awards

USACM Young Investigator Fellowship, World Congress on Computational Mechanics. 2008

International Congress of Mathematicians Travel Grant, NSF DMS-0514413, American Mathematical Society. 2006

ICIAM Travel Grant, NSF DMS-0218051, Society for Industrial and Applied Mathematics. 2003

London Mathematical Society Bursary, Short Course on Computational Differential Equations. 2002

SIAM Student Travel Award, SIAM Conference on Applications of Dynamical Systems. 2001

Graduate Fellowships

Josephine de Kármán Fellowship (established by Theodore von Kármán). 2003–2004

International Fellowship, Agency for Science, Technology and Research, Singapore. 2002–2004

Poincaré Fellowship (Betty and Gordon Moore Fellowship), Caltech. 2000–2004

Tau Beta Pi Fellowship, Tau Beta Pi, National Engineering Honor Society. 2000–2001

Selected Honors and Awards

Herbert J. Ryser Scholarship, Caltech Mathematics Department. 1999

E. T. Bell Undergraduate Mathematics Research Prize, Caltech Mathematics Department. 1999

Jack E. Froehlich Memorial Award, Caltech. 1999

Sigma Xi, The Scientific Research Society (Associate Member / Full Member). 1999, 2005

Tau Beta Pi, National Engineering Honor Society. 1999

Loke Cheng-Kim Foundation Scholarship, Singapore. 1996–2000

Melvin Leok: Curriculum Vitae

Publications Available for download at `http://www.math.purdue.edu/~mleok/`

Refereed Journal Papers

1. *Discrete Poincaré Lemma* (with M. Desbrun, J.E. Marsden), Appl. Numer. Math. **53** (2–4), 231–248, 2005.
2. *Discrete Routh Reduction* (with S.M. Jalnapurkar, J.E. Marsden, M. West), J. Phys. A: Math. Gen. **39**, 5521-5544 (Geometric Integration Special Issue, invited paper), 2006.
3. *Lie Group Variational Integrators for the Full Body Problem* (with T. Lee, N.H. McClamroch), Comput. Methods Appl. Mech. Engrg. **196** (29-30), 2907–2924, 2007.
4. *Lie Group Variational Integrators for the Full Body Problem in Orbital Mechanics* (with T. Lee, N.H. McClamroch), Celestial Mechanics and Dynamical Astronomy **98** (2), 121–144, 2007.
5. *Global Optimal Attitude Estimation using Uncertainty Ellipsoids* (with T. Lee, A.K. Sanyal, N.H. McClamroch), Systems and Control Letters, **57** (3), 236–245, 2008.
6. *Optimal Attitude Control of a Rigid Body using Geometrically Exact Computations on SO*(3) (with T. Lee, N.H. McClamroch), Journal of Dynamical and Control Systems, accepted, 2007.
7. *Discrete Control Systems* (with T. Lee, N. H. McClamroch), invited article for the Springer Encyclopedia of Complexity and Systems Science, accepted, 2008.
8. *Estimating the Attractor Dimension of the Equatorial Weather System*, Acta Phys. Pol. A **85**, S27–S35, 1994.

Refereed Conference Papers

9. *A Lie Group Variational Integrator for the Attitude Dynamics of a Rigid Body with Applications to the 3D Pendulum* (with T. Lee, N.H. McClamroch), Proc. IEEE Conf. on Control Applications, 962–967, 2005.
10. *Controlled Lagrangians and Stabilization of the Discrete Cart-Pendulum System* (with A.M. Bloch, J.E. Marsden, D.V. Zenkov), Proc. IEEE Conf. on Decision and Control, 6579–6584, 2005.
11. *Attitude Maneuvers of a Rigid Spacecraft in a Circular Orbit* (with T. Lee, N.H McClamroch), Proc. American Control Conf., 1742–1747, 2006.
12. *Polyhedral Potential and Lie Group Variational Integrator Computations for the Full Two Body Problem* (with E. Fahnestock, T. Lee, N.H. McClamroch, and D. Scheeres), Proc. AIAA/AAS Astrodynamics Specialist Conf., AIAA-2006-6289, 2006.
13. *Optimal Control of a Rigid Body using Geometrically Exact Computations on SE(3)* (with T. Lee, N.H. McClamroch), Proc. IEEE Conf. on Decision and Control, 2710–2715, 2006.
14. *Deterministic Global Attitude Estimation* (with T. Lee, A.K. Sanyal, N.H. McClamroch), Proc. IEEE Conf. on Decision and Control, 3174–3179, 2006.
15. *Controlled Lagrangians and Potential Shaping for Stabilization of Discrete Mechanical Systems* (with A.M. Bloch, J.E. Marsden, D.V. Zenkov), Proc. IEEE Conf. on Decision and Control, 3333–3338, 2006.
16. *A Discrete Variational Integrator for Optimal Control Problems on SO(3)* (with A.M. Bloch, I.I. Hussein, A.K. Sanyal), Proc. IEEE Conf. on Decision and Control, 6636–6641, 2006.
17. *Global Attitude Estimation using Single Direction Measurements* (with T. Lee, N.H. McClamroch, A.K. Sanyal), Proc. American Control Conf., 3659–3664, 2007.
18. *Optimal Attitude Control for a Rigid Body with Symmetry* (with T. Lee, N.H. McClamroch), Proc. American Control Conf., 1073–1078, 2007.
19. *Propagation of Uncertainty Sets for Rigid Body Attitude Flows* (with N.A. Chaturvedi, T. Lee, N.H. McClamroch, A.K. Sanyal), Proc. IEEE Conf. on Decision and Control, 2689–2694, 2007.
20. *A Combinatorial Optimization Problem for Spacecraft Formation Reconfiguration* (with T. Lee, N.H. McClamroch), Proc. IEEE Conf. on Decision and Control, 5370–5375, 2007.
21. *Time Optimal Attitude Control for a Rigid Body* (with T. Lee, N.H. McClamroch), Proc. American Control Conf., 5210–5215, 2008.
22. *Global Symplectic Uncertainty Propagation on SO(3)* (with T. Lee, N.H. McClamroch), Proc. IEEE Conf. on Decision and Control, accepted, 2008.

Papers under Review

23. *Geometric Structure-Preserving Optimal Control of the Rigid Body* (with A.M. Bloch, I.I. Hussein, A.K. Sanyal), Journal of Dynamical and Control Systems, submitted, 2007.
24. *Computational Geometric Optimal Control of Rigid Bodies* (with T. Lee, N.H. McClamroch), Brockett Legacy Special Issue, Communications in Information and Systems, submitted, 2008.
25. *Lagrangian Mechanics and Variational Integrators on Two-Spheres* (with T. Lee, N.H. McClamroch), International Journal for Numerical Methods in Engineering, submitted, 2008.
26. *Nonlinear Dynamics of the 3D Pendulum* (with N.A. Chaturvedi, T. Lee, N.H. McClamroch), Journal of Nonlinear Science, submitted, 2008.

Publications (continued) Available for download at `http://www.math.purdue.edu/~mleok/`

Papers under Revision

27. *Controlled Lagrangians and Stabilization of Discrete Mechanical Systems I* (with A.M. Bloch, J.E. Marsden, D.V. Zenkov), 2008.

Preprints

28. *Generalized Galerkin Variational Integrators*, 2004.
29. *A Discrete Theory of Connections on Principal Bundles* (with J.E. Marsden, A.D. Weinstein), 2004.
30. *Discrete Exterior Calculus* (with M. Desbrun, A.N. Hirani, J.E. Marsden), 2003.

Papers in Preparation

31. *Global Attitude Tracking with Optimal Attitude State Estimation* (with A.K. Sanyal, S.B. Singh, and N. Nordkvist)
32. *The Discrete Hamilton-Pontryagin Principle: Unifying Discrete Lagrangian and Hamiltonian Mechanics.*

Invited Seminar Talks

Scientific Computing Seminar, Department of Mathematics, Tsinghua University, Beijing, China.	Jun 2008
Institute of Computational Mathematics Seminar, Chinese Academy of Sciences, Beijing, China.	Jun 2008
Applied and Computational Mathematics Seminar, University of Auckland, New Zealand.	Dec 2007
Applied Mathematics and PDE Seminar, University of Wisconsin, Madison, WI.	Oct 2007
Applied Mathematics Seminar, Imperial College, London, UK.	May 2007
Highly Oscillatory Problems Seminar, Newton Institute, University of Cambridge, UK.	May 2007
Center for Applied Mathematics Colloquium, University of Notre Dame, Notre Dame, IN.	Apr 2007
Mathematics Colloquium, University of Iowa, Iowa City, IA.	Mar 2007
CDS/CIMMS Lunchtime Seminar, Caltech, Pasadena, CA.	Dec 2006
Applied Mathematics Colloquium, University of Maryland, Baltimore County, MD.	Nov 2006
Mathematical Sciences Colloquium, Rensselaer Polytechnic Institute, Troy, NY.	Oct 2006
Geometry and Dynamical Systems with Applications Seminar, Arizona State University, Tempe, AZ.	Apr 2006
Dynamics Seminar, University of Colorado, Boulder, CO.	Feb 2006
Mathematics Colloquium, Colorado School of Mines, CO.	Feb 2006
Applied and Interdisciplinary Mathematics Seminar, University of Michigan, Ann Arbor, MI.	Feb 2006
Mathematics Colloquium, University of South Carolina, SC.	Feb 2006
Mathematics Colloquium, Texas A&M University, College Station, TX.	Feb 2006
Mathematics Colloquium, Colorado State University, Fort Collins, CO.	Feb 2006
Mathematics Special Colloquium, Purdue University, West Lafayette, IN.	Feb 2006
Numerical Analysis and Differential Equations Seminar, North Carolina State University, Raleigh, NC.	Jan 2006
Applied Mathematics Seminar, Mathematics, University of Waterloo, Canada.	Dec 2005
Control Seminar, College of Engineering, University of Michigan, Ann Arbor, MI.	Oct 2005
Differential Geometry and Analysis Seminar, University of Toledo, Toledo, OH.	Sep 2005
Computer and Computational Sciences, CCS-2, LANL, Los Alamos, NM.	Aug 2005
Special Seminar, Mathematics, University of California, Berkeley, CA.	Apr 2005
Mathematical Physics Seminar, University of Minnesota, Twin Cities, MN.	Mar 2005
Numerical Analysis and Differential Equations Seminar, North Carolina State University, Raleigh, NC.	Jan 2005
Mathematics Colloquium, University of California, San Diego, CA.	Dec 2004
Flight Dynamics and Control Seminar, Aeronautics, University of Michigan, Ann Arbor, MI.	Oct 2004
Geometry Seminar, University of Michigan, Ann Arbor, MI.	Oct 2004
Applied Mathematics Seminar, University of California, San Diego, CA.	Aug 2004
Temasek Laboratories, National University of Singapore.	Jul 2004
Mathematics Department, National University of Singapore. (4 talks)	Jul 2004
Institute for High Performance Computing, National University of Singapore.	Jul 2004
SIAM Student Chapter, Caltech, Pasadena, CA.	Jan 2004
Applied Math and Numerical Analysis Seminar, University of Minnesota, Twin Cities, MN.	Jan 2004
School of Engineering and Science, International University of Bremen, Germany.	Sep 2003
Paderborn Institute for Scientific Computation, University of Paderborn, Germany.	Sep 2003
Caltech/JPL Nonlinear Astrodynamics Group, Pasadena, CA.	Aug 2003
Institute for High Performance Computing, National University of Singapore.	Feb 2003
Center for Integrative Multiscale Modeling and Simulation, Caltech, Pasadena, CA.	May 2002

Invited Seminar Talks (Continued)

Culham Electromagnetics and Lightning Ltd, Abingdon, Oxfordshire, UK.	Apr 2002
Department of Informatics, University of Bergen, Bergen, Norway.	Oct 2001
Department of Mathematical Sciences, NTNU, Trondheim, Norway.	Oct 2001
Mathematics Department, National University of Singapore.	Aug 2000
Center for Remote Imaging, Sensing and Processing, National University of Singapore.	Sep 1999

Invited Conference Talks

Fourth Annual Structured Integrators Workshop, Stanford, CA.	Apr 2008
Geometric Mechanics Workshop, Banff International Research Station, Banff, Canada.	Aug 2007
Discrete Differential Geometry, Berlin, Germany.	Jul 2007
New Talent Plenary Lecture, SciCADE 2007, Saint Malo, France.	Jul 2007
Effective Computational Methods for Highly Oscillatory Problems: The Interplay between Mathematical Theory and Applications, Newton Institute, University of Cambridge, UK.	Jul 2007
First International Summer School on Geometry, Mechanics, and Control, Castro Urdiales, Spain. (One of three principal lecturers, 7 hours)	Jun 2007
Turbulence Working Group Workshop, T-7, LANL, Santa Fe, NM.	Dec 2003
Full Body Problem Workshop, Caltech, Pasadena, CA.	Nov 2003
Leslie Fox Prize Meeting, University of Cambridge, UK.	Jun 2003
Student Paper Prize Presentation, SIAM Annual Meeting, Montréal, Canada.	Jun 2003
Geometrical Mechanics and Turbulence Modeling, CNLS, LANL, Santa Fe, NM.	Nov 2002
Geometry, Symmetry and Mechanics II, University of Warwick, UK.	Jul 2002
Invariant and Symmetry-Preserving Integrators for N-Body Simulation, University of Leicester, UK.	Apr 2002
Reduced Dimensional Modeling Workshop, CNLS, LANL, Los Alamos, NM.	Nov 2001

Invited Minisymposium Talks

Advances in Time-Integration, World Congress on Computational Mechanics, Venice, Italy.	Jul 2008
Geometric Integration and Computational Mechanics, Foundations of Computational Mathematics, City University of Hong Kong, Hong Kong, China.	Jun 2008
Geometric Numerical Integration, Joint Meeting of the AMS - NZMS, Wellington, New Zealand.	Dec 2007
Geometric and Symplectic Integration, SciCADE 2007, Saint-Malo, France.	Jul 2007
Geometric Methods in Dynamical Systems, SIAM Dynamical Systems, Snowbird, UT.	May 2007
Applications of the Geometric Phase in Classical Mechanics, SIAM Annual Meeting, Boston, MA.	Jul 2006
Contemporary Dynamical Systems, AMS Annual Meeting, San Antonio, TX.	Jan 2006
Geometric Dynamics and its Applications, SIAM Dynamical Systems, Snowbird, UT.	May 2005
Geometric Dynamics and Applications, AIMS Fifth International Conference on Dynamical Systems and Differential Equations, Pomona, CA.	Jun 2004
Geometric Methods for PDEs, NUMDIFF, University of Halle, Germany.	Sep 2003
Non-Grid based Methods for Geophysical and Astrophysical Flows, ICIAM, Sydney, Australia.	Jul 2003
Structure-Preserving Algorithms, SciCADE 2003, Trondheim, Norway.	Jul 2003
Discrete Geometry and Geometric Integration, SIAM Dynamical Systems, Snowbird, UT.	May 2003
Geometric Integration and Computational Dynamics, Foundations of Computational Mathematics, Minneapolis, MN.	Aug 2002
Geometric Integration, SIAM Dynamical Systems, Snowbird, UT. *SIAM Student Travel Award.*	May 2001

Contributed Talks

MSRI Workshop on Application of Topology in Science and Engineering, Berkeley, CA.	Sep 2006
Geometric Numerical Integration Workshop, Oberwolfach, Germany.	Mar 2006
Frontiers of Applied Analysis, Pittsburgh, PA.	Sep 2005
IPAM Relativistic Astrophysics Workshop, Los Angeles, CA.	May 2005
Southern California Applied Mathematics Symposium, Claremont, CA.	Apr 2004
Auckland Numerical Ordinary Differential Equations, Auckland, New Zealand.	Jul 2003
International Congress on Industrial and Applied Mathematics, Sydney, Australia.	Jul 2003
Biennial Conference on Numerical Analysis, Dundee, Scotland.	Jun 2003
Mechanics and Symmetry European Summer School, Peyresq, France.	Sep 2001
Southwest Regional Workshop on New Directions in Dynamical Systems, USC, Los Angeles, CA.	Nov 2000
Caltech SURF Seminar Day, Pasadena, CA.	Oct 1999
Caltech SURF Seminar Day, Pasadena, CA. *Semi-finalist, Perpall Speaking Competition.*	Oct 1998

Contributed Posters

Advanced Computational Electromagnetics Workshop, Boston, MA. May 2006
New Paradigms in Computation, IMA Tutorial/Workshop, Minneapolis, MN. Mar 2005
Compatible Spatial Discretizations for Partial Differential Equations,
Institute for Mathematics and its Applications "Hot Topics" Workshop, Minneapolis, MN. May 2004
ARO-Institute for Collaborative Biotechnologies Review, UCSB, Santa Barbara, CA. Jul 2003, Feb 2004
International Congress on Industrial and Applied Mathematics, Sydney, Australia. Jul 2003
CIMMS-IPAM Workshop on Molecular Modeling and Computation, Pasadena, CA. Nov 2002
DARPA/NSF OPAAL Workshop, Seattle, WA. May 2001
Southern California Applied Mathematics Symposium, Caltech, Pasadena, CA. May 2001
Dynamics Days 2000, Santa Fe, NM. Jan 2000
SIAM Conference on Applications of Dynamical Systems, Snowbird, UT. May 1999, 2003
NSF-KDI/IGPP Workshop on accurate simulation and modeling of physical systems,
San Diego, CA. *Student Poster Silver Medal Award in 1998.* Nov 1998, 1999, 2000

Research Visits, Conferences and Summer Schools

Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics,
Chinese Academy of Sciences, Beijing, China. Host: Prof. Jialin Hong.[‡,♯] Jun 2008
Visiting Fellow, Highly Oscillatory Problems: Computation, Theory and Application,
Newton Institute, University of Cambridge.[†,‡] May 2007
Geometry of Mechanism Science, Notre Dame, IN.[†] Mar 2007
MSRI Workshop on Application of Topology in Science and Engineering, Berkeley, CA.[†] Sep 2006
IMA Tutorial: Algebraic Geometric Methods in Engineering, Minneapolis, MN.[†] Sep 2006
International Congress of Mathematicians, Madrid, Spain.[‡] Aug 2006
Geometric Numerical Integration Workshop, Oberwolfach, Germany.[†,‡] Mar 2006
Multiscale Modeling and Computation - Basic Theory and the Geosciences, Caltech, Pasadena, CA.[†] Nov 2005
IPAM Bridging Time and Length Scales in Materials Science and Bio-Physics,
Multiscale Analysis and Computation, Los Angeles, CA.[†] Nov 2005
IMA New Directions Short Course: Quantum Computation, Minneapolis, MN.[†] Aug 2005
International Forum on Multiscale Methods and Partial Differential Equations, Los Angeles, CA.[†] Aug 2005
Quantum Control Summer School, Caltech, Pasadena, CA.[†] Aug 2005
IPAM Grand Challenge Problems in Computational Astrophysics, Los Angeles, CA.
Relativistic Astrophysics.[†] May 2005
N-Body Problems in Astrophysics.[†] Apr 2005
Department of Mathematics, National University of Singapore, Singapore.[‡,♯] Jul 2004
2004 CNA Summer School, Advances in Nonlinear Analysis,
Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA.[†] May–Jun 2004
DARPA Workshop on Design of Robust Dynamical Systems, UTRC, East Hartford, CT.[†] Jan 2004
Southern California Applied Mathematics Symposium, Irvine, CA. May 2003
Advances and Mathematical Issues in Large Scale Simulation,
Institute for Mathematical Sciences, National University of Singapore.[‡] Feb 2003
Mathematical Challenges in Scientific and Engineering Computation, Newton Institute,
University of Cambridge.[‡] Jan 2003
Workshop on Geometry, Dynamics, and Mechanics, Fields Institute, Toronto, Canada.[†] Aug 2002
Workshop on Astrodynamics, University of Surrey.[†] Apr 2002
Workshop on Classical N-Body Systems and Applications, University of Warwick.[†] Apr 2002
LMS/EPSRC Short Course on Computational Differential Equations.[†] Mar 2002
Groupoidfest, University of California, Berkeley, CA.[‡] Nov 2001
Geometric Integration Group of Prof. Hans Munthe-Kaas, Department of Informatics,
University of Bergen, Bergen, Norway.[♯] Oct 2001
Numerics Group of Prof. Brynjulf Owren, Department of Mathematical Sciences,
Norwegian University of Science and Technology, Trondheim, Norway.[♯] Oct 2001
Mechanics and Symmetry European Summer School, Peyresq, France.[‡] Sep 2001
Numerical Analysis Group of Prof. Arieh Iserles, Department of Applied Mathematics
and Theoretical Physics, University of Cambridge, United Kingdom.[‡] Aug–Sep 2001
Surface Water Waves, Newton Institute EuroConference, University of Cambridge.[†] Aug 2001

Research Visits, Conferences and Summer Schools (Continued)

5th PIMS Industrial Problem Solving Workshop, University of Washington, Seattle, WA.† Jun 2001
4th PIMS Graduate Mathematics Modelling Camp, University of Victoria, BC, Canada.† Jun 2001
2001 CNA Summer School, Multiscale Problems in Nonlinear Analysis,
Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA.† Jun 2001

†*Funded by conference organizers.* ‡*Funded by research grants.* ♯*Funded by host institution.*

Postdoctoral Scholars Advised

Diana Sosa Martín, Visiting Assistant Professor of Mathematics, Purdue University, 2008–2009.

Graduate Students Advised

Tomoki Ohsawa (2008 summer research advisor)
Ph.D. Student, Applied and Interdisciplinary Mathematics, University of Michigan.
Daniel J. Grebow (graduate advisory committee member, advisor: Kathleen C. Howell)
Ph.D. Student, Aeronautics and Astronautics, Purdue University.
Taeyoung Lee (co-advised with N. Harris McClamroch)
(Distinguished Achievement Award, Ivor K. Mclvor Award, BGCE Student Paper Prize finalist,
Rackham International Student Fellow, Rackham Predoctoral Fellow)
Ph.D., Aerospace Engineering, University of Michigan, Ann Arbor, Spring 2008.
Nalin A. Chaturvedi (dissertation committee member, advisors: N. Harris McClamroch and Dennis S. Bernstein)
Ph.D., Aerospace Engineering, University of Michigan, Ann Arbor, Spring 2007.
Masako Kishida (co-advised with Dennis S. Bernstein)
M.S., Applied and Interdisciplinary Mathematics, University of Michigan, Ann Arbor, Spring 2006.

Undergraduate Students Advised

Wooi-Chen Ng, Summer Research Fellow, Computer Science and minor in Mathematics, Purdue University, 2008.
Charles Roldan, Summer REU Fellow, Mathematics and minor in Physics, Purdue University, 2007, 2008.
Nathan Orlow, Summer REU Fellow, Mathematics and Statistics, Purdue University, 2007.

Professional Service

Editor, *Journal of Nonlinear Science* (Springer).
Editor, *Journal of Geometric Mechanics* (American Institute of Mathematical Sciences, inaugural issue in 2009).
Referee, *Aerospace Science and Technology*, *Applied Numerical Mathematics*,
Celestial Mechanics and Dynamical Astronomy, *Communications in Contemporary Mathematics*,
Communications in Numerical Methods in Engineering, *Computational Science & Discovery*,
Discrete and Continuous Dynamical Systems, *Foundations of Computational Mathematics*,
IEEE Transactions on Automatic Control, *IET Control Theory & Applications*,
IMA Journal of Numerical Analysis, *Journal of Computational Physics*, *Journal of Mathematical Physics*,
Journal of Nonlinear Science, *Journal of Physics A*, *Journal of Symplectic Geometry*, *Nonlinearity*,
Numerical Algorithms, *Physica D*, *Physics Letters A*, *Proceedings of the Royal Society A*,
SIAM Applied Dynamical Systems, *SIAM Multiscale Modeling and Simulation*,
SIAM Journal on Numerical Analysis, *SIAM Journal on Scientific Computing*,
SIGMA (Symmetry, Integrability and Geometry: Methods and Applications),
Soft Computing and Automation Journal, *Transport in Porous Media*.
Reviewer, *IEEE Conference on Decision and Control 2005, 2006, 2007, 2008*, *IEEE Multi-conference on Systems and Control 2007*, *American Control Conference 2008*, *Mathematical Reviews*, *Springer Books*, *Swiss National Science Foundation*.
Panel member, Computational Mathematics, National Science Foundation. Mar 2008
Organizer, computational mechanics minisymposium, SciCADE 2009, Beijing, China. May 2009
Co-organizer, nonlinear dynamics and control of mechanical systems invited session,
IEEE Conference on Decision and Control. Dec 2006
Co-organizer, contemporary dynamical systems special session, AMS Annual Meeting. Jan 2006
Co-organizer, geometric dynamics and its applications minisymposium (MS 59, 70), SIAM DS05. May 2005
Co-organizer, CIMMS Workshop on Discrete Geometry for Mechanics, Pasadena, CA. Oct 2003
Co-organizer, discrete geometry and geometric integration minisymposium (MS 38, 62), SIAM DS03. May 2003
Co-organizer, CIMMS Workshop on Networks, Optimization and Duality, Pasadena, CA. July 2002
Co-organizer, geometric integration minisymposium (MS 51, 69), SIAM DS01. May 2001

University Service

Organizer, Computational and Applied Mathematics Seminar, Purdue University.	Fall 2008
Member, Engagement Pillar Group Committee, College of Science, Purdue University.	2008
Member, Computer Committee, Mathematics, Purdue University.	2007–2008
Judge, Undergraduate Research and Poster Symposium, Purdue University.	Apr 2007
Faculty advisor, Purdue Singaporean Students Association.	Oct 2006–present
Panelist, Proposal Writing, Faculty Professional Development Program, University of Michigan.	Sep 2005
Panelist, Graduate School Discussion Panel, Undergraduate Math Club, University of Michigan.	Oct 2004
Officer, SIAM Student Chapter, Caltech.	2004
Member, Caltech Project for Effective Teaching, Caltech.	2002–2004
Judge, Semifinals, Doris S. Perpall Speaking Competition, Caltech.	Nov 2003
Judge, Mathematics/Applied Mathematics Session, Caltech SURF Seminar Day	Oct 2002, 2003
Member, Academics Committee, Graduate Student Council, Caltech.	2002
Graduate Student Representative, Committee on Institute Programs, Caltech.	2001–2003
Director, CDS Option Representative, Graduate Student Council, Caltech.	2001–2002
Member, Committee on Teaching Assistant Training, Graduate Dean's Office, Caltech.	2001–2002
Student Representative, Feynman Teaching Prize Selection Committee, Caltech.	2000
Director for Academic Affairs, Associated Students of Caltech (ASCIT).	1999–2000
Chairman, Academics and Research Committee, Caltech.	1999–2000
Student Representative, Core Curriculum Steering Committee, Caltech.	1999–2000
House Representative, Academics and Research Committee, Caltech.	1998–1999
Student Representative, Academic Policies and Curriculum Committee, Caltech.	1997–1999

Teaching Experience

Purdue University

Instructor, Numerical Analysis (Math/CS 514)	Fall 2008
Instructor, Geometric Numerical Integration (Math 692A)	Spring 2008
Instructor, Ordinary Differential Equations (Math 366)	Spring 2008
Instructor, Introduction to Differential Geometry and Topology (Math 562)	Fall 2007
Instructor, Ordinary Differential Equations (Math 266)	Spring 2007

University of Michigan, Ann Arbor

Instructor, Numerical Methods for Engineers and Scientists (Math 371/Engr 371)	Winter 2005/06, Fall 2005
Instructor, Applied Honors Calculus II (Math 156)	Fall 2004

California Institute of Technology

Teaching Assistant, Advanced Topics in Geometric Mechanics (CDS 280)	Spring 2004
Teaching Assistant, Geometric Mechanics (CDS 205)	Spring 2003
Teaching Assistant, Applied Operator Theory (CDS 201)	Fall 2002
Teaching Assistant, Graduate Abstract Algebra (Math 120)	Fall 1999, Winter 2000, Spring 2000

Professional Memberships

American Mathematical Society, Foundations of Computational Mathematics (Geometric Integration Interest Group), London Mathematical Society, Mathematical Association of America, Singapore Mathematical Society, Society for Industrial and Applied Mathematics (Dynamical Systems Activity Group).

Miscellaneous

Nationality: Permanent resident of the United States of America (Green Card). Singapore citizen.

References

Anthony M. Bloch
Alexander Ziwet Collegiate Professor of Mathematics;
Chair, Department of Mathematics, University of Michigan, Ann Arbor, MI.
Department of Mathematics, The University of Michigan, Ann Arbor, MI 48109, USA.
(734)647-4980, abloch@umich.edu

Mathieu Desbrun
Associate Professor of Computer Science
and Computational Science & Engineering, California Institute of Technology.
158-79, Caltech, Pasadena, CA 91125, USA.
(626)395-6230, mathieu@cs.caltech.edu

Darryl D. Holm
Professor of Applied Mathematics, Imperial College of Science, Technology, and Medicine, UK.
Department of Mathematics, South Kensington Campus, Imperial College, London SW7 2AZ, UK.
+44 20 7594 8531, d.holm@imperial.ac.uk

Laboratory Fellow, Los Alamos National Laboratory.
T-7, MS B284, Los Alamos, NM 87545, USA.
(505) 667-6398, dholm@lanl.gov

Thomas Y. Hou
Charles Lee Powell Professor of Applied & Computational Mathematics, California Institute of Technology.
217-50, Caltech, Pasadena, CA 91125, USA.
(626)395-4546, hou@acm.caltech.edu

Arieh Iserles
Professor in Numerical Analysis of Differential Equations, University of Cambridge, UK.
Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK.
+44 1223 337891, A.Iserles@damtp.cam.ac.uk

Jerrold E. Marsden
Carl F. Braun Professor of Engineering and Control & Dynamical Systems, California Institute of Technology.
107-81, Caltech, Pasadena, CA 91125, USA.
(626)395-4176, marsden@cds.caltech.edu

N. Harris McClamroch
Professor of Aerospace Engineering, University of Michigan, Ann Arbor.
Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109, USA.
(734)763-2355, nhm@engin.umich.edu

Alan D. Weinstein
Professor of Mathematics, University of California, Berkeley.
Chair, Department of Mathematics, University of California, Berkeley.
Department of Mathematics, University of California, Berkeley, CA 94720, USA.
(510)642-3518, alanw@math.berkeley.edu

Melvin Leok: Bibliography

Department of Mathematics	*phone*:	+1(765)496-3578
Purdue University	*fax*:	+1(765)494-0548
150 N. University Street	*e-mail*:	mleok@math.purdue.edu
West Lafayette, IN 47907-2067, USA.	*homepage*:	http://www.math.purdue.edu/~mleok/

Publications Available for download at http://www.math.purdue.edu/~mleok/

Refereed Journal Papers

1. Desbrun, M., M. Leok, J.E. Marsden [2005], *Discrete Poincaré Lemma* Appl. Numer. Math. **53** (2–4), 231–248.
2. Jalnapurkar, S.M., M. Leok, J.E. Marsden, M. West [2006], *Discrete Routh Reduction*, J. Phys. A: Math. Gen. **39**, 5521–5544 (Geometric Integration Special Issue, invited paper).
3. Lee, T., M. Leok, N.H. McClamroch [2007], *Lie Group Variational Integrators for the Full Body Problem*, Comput. Methods Appl. Mech. Engrg. **196** (29-30), 2907–2924.
4. Lee, T. M. Leok, N.H. McClamroch [2007], *Lie Group Variational Integrators for the Full Body Problem in Orbital Mechanics*, Celestial Mechanics and Dynamical Astronomy **98** (2), 121–144.
5. Sanyal, A.K., T. Lee, M. Leok, N.H. McClamroch [2008], *Global Optimal Attitude Estimation using Uncertainty Ellipsoids*, Systems and Control Letters, **57** (3), 236–245.
6. Lee, T., M. Leok, N.H. McClamroch [2007], *Optimal Attitude Control of a Rigid Body using Geometrically Exact Computations on SO*(3), Journal of Dynamical and Control Systems, accepted.
7. Lee, T., M. Leok, N.H. McClamroch [2008], *Discrete Control Systems*, invited article for the Springer Encyclopedia of Complexity and Systems Science, accepted.
8. Leok, M. [1994], *Estimating the Attractor Dimension of the Equatorial Weather System*, Acta Phys. Pol. A **85**, S27–S35.

Refereed Conference Papers

9. Lee, T., M. Leok, N.H. McClamroch [2005], *A Lie Group Variational Integrator for the Attitude Dynamics of a Rigid Body with Applications to the 3D Pendulum*, Proc. IEEE Conf. on Control Applications, 962–967.
10. Bloch, A.M., M. Leok, J.E. Marsden, D.V. Zenkov [2005], *Controlled Lagrangians and Stabilization of the Discrete Cart-Pendulum System*, Proc. IEEE Conf. on Decision and Control, 6579–6584.
11. Lee, T., M. Leok, N.H. McClamroch [2006], *Attitude Maneuvers of a Rigid Spacecraft in a Circular Orbit*, Proc. American Control Conf., 1742–1747.
12. Fahnestock, E., T. Lee, M. Leok, N.H. McClamroch, D. Scheeres [2006], *Polyhedral Potential and Lie Group Variational Integrator Computations for the Full Two Body Problem*, Proc. AIAA/AAS Astrodynamics Specialist Conf., AIAA-2006-6289.
13. Lee, T., N.H. McClamroch, M. Leok [2006], *Optimal Control of a Rigid Body using Geometrically Exact Computations on SE(3)*, Proc. IEEE Conf. on Decision and Control, 2710–2715.
14. Lee, T., A.K. Sanyal, M. Leok, N.H. McClamroch, *Deterministic Global Attitude Estimation*, Proc. IEEE Conf. on Decision and Control, 3174–3179.
15. Bloch, A.M., M. Leok, J.E. Marsden, D.V. Zenkov [2006], *Controlled Lagrangians and Potential Shaping for Stabilization of Discrete Mechanical Systems*, Proc. IEEE Conf. on Decision and Control, 3333–3338.
16. Hussein, I.I., M. Leok, A.K. Sanyal, A.M. Bloch [2006], *A Discrete Variational Integrator for Optimal Control Problems on SO(3)*, Proc. IEEE Conf. on Decision and Control, 6636–6641.
17. Lee, T., M. Leok, N.H. McClamroch, A.K. Sanyal [2007], *Global Attitude Estimation using Single Direction Measurements*, Proc. American Control Conf., 3659–3664.
18. Lee, T., N.H. McClamroch, M. Leok [2007], *Optimal Attitude Control for a Rigid Body with Symmetry*, Proc. American Control Conf., 1073–1078.
19. Lee, T., N.A. Chaturvedi, A.K. Sanyal, M. Leok, N.H. McClamroch [2007], *Propagation of Uncertainty in Rigid Body Attitude Flows*, Proc. IEEE Conf. on Decision and Control, 2689–2694.
20. Lee, T., M. Leok, N.H. McClamroch [2007], *A Combinatorial Optimization Problem for Spacecraft Formation Reconfiguration*, Proc. IEEE Conf. on Decision and Control, 5370–5375.
21. Lee, T., M. Leok, N.H. McClamroch [2008], *Time Optimal Attitude Control for a Rigid Body*, Proc. American Control Conf., 5210–5215.
22. Lee, T., M. Leok, N.H. McClamroch [2008], *Global Symplectic Uncertainty Propagation on SO(3)*, Proc. IEEE Conf. on Decision and Control, accepted.

Melvin Leok: Bibliography

Publications (continued) Available for download at `http://www.math.purdue.edu/~mleok/`

Papers under Review

23. Bloch, A.M., I.I. Hussein, M. Leok, A.K. Sanyal [2007], *Geometric Structure-Preserving Optimal Control of the Rigid Body*, Journal of Dynamical and Control Systems, submitted.
24. Lee, T., M. Leok, N.H. McClamroch [2008], *Computational Geometric Optimal Control of Rigid Bodies*, Brockett Legacy Special Issue, Communications in Information and Systems, submitted.
25. Lee, T., M. Leok, N.H. McClamroch [2008], *Lagrangian Mechanics and Variational Integrators on Two-Spheres*, International Journal for Numerical Methods in Engineering, submitted.
26. Chaturvedi, N.A., T. Lee, M. Leok, N.H. McClamroch [2008], *Nonlinear Dynamics of the 3D Pendulum*, Journal of Nonlinear Science, submitted.

Preprints

27. Leok, M. [2004], *Generalized Galerkin Variational Integrators*, preprint.
28. Leok, M., J.E. Marsden, A.D. Weinstein [2004], *A Discrete Theory of Connections on Principal Bundles*, preprint.
29. Desbrun, M., A.N. Hirani, M. Leok, J.E. Marsden [2003], *Discrete Exterior Calculus*, preprint.

Thesis

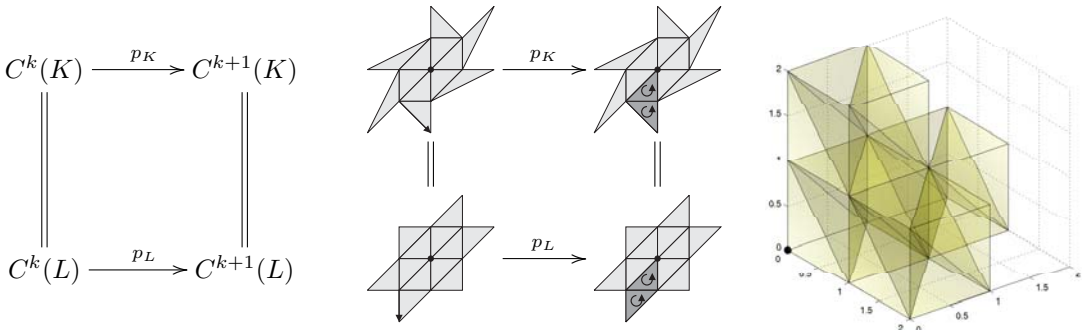
30. Leok, M. [2004], *Foundations of Computational Geometric Mechanics*, Ph.D. thesis, California Institute of Technology. A preliminary version of this thesis received the *SIAM Student Paper Prize*, and the *Leslie Fox Prize in Numerical Analysis* (second prize) in 2003.

Abstracts

1. **Discrete Poincaré Lemma** *(with M. Desbrun, and J.E. Marsden)*
   Appl. Numer. Math. **53** (2–4), 231–248, 2005.

   This paper proves a discrete analogue of the Poincaré lemma in the context of a discrete exterior calculus based on simplicial cochains. The proof requires the construction of a generalized cone operator, $p : C_k(K) \to C_{k+1}(K)$, as the geometric cone of a simplex cannot, in general, be interpreted as a chain in the simplicial complex. The corresponding cocone operator $H : C^k(K) \to C^{k-1}(K)$ can be shown to be a homotopy operator, and this yields the discrete Poincaré lemma.

   The generalized cone operator is a combinatorial operator that can be constructed for any simplicial complex that can be grown by a process of local augmentation. In particular, regular triangulations and tetrahedralizations of $\mathbb{R}^2$ and $\mathbb{R}^3$ are presented, for which the discrete Poincaré lemma is globally valid.

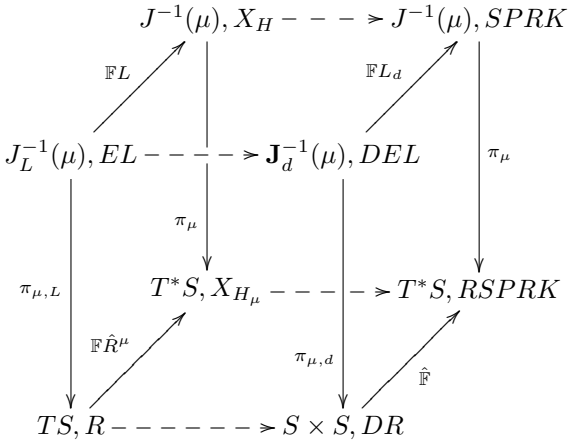


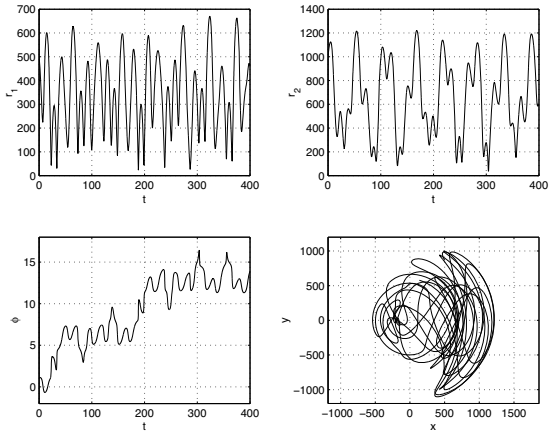
2. **Discrete Routh Reduction** *(with S.M. Jalnapurkar, J.E. Marsden, and M. West)*
J. Phys. A: Math. Gen. **39**, 5521–5544 (Geometric Integration Special Issue, invited paper), 2006.

This paper develops the theory of abelian Routh reduction for discrete mechanical systems and applies it to the variational integration of mechanical systems with abelian symmetry. The reduction of variational Runge-Kutta discretizations is considered, as well as the extent to which symmetry reduction and discretization commute. These reduced methods allow the direct simulation of dynamical features such as relative equilibria and relative periodic orbits that can be obscured or difficult to identify in the unreduced dynamics.

The methods are demonstrated for the dynamics of an Earth orbiting satellite with a non-spherical J_2 correction, as well as the double spherical pendulum. The J_2 problem is interesting because in the unreduced picture, geometric phases inherent in the model and those due to numerical discretization can be hard to distinguish, but this issue does not appear in the reduced algorithm, where one can directly observe interesting dynamical structures in the reduced phase space (the cotangent bundle of shape space), in which the geometric phases have been removed.

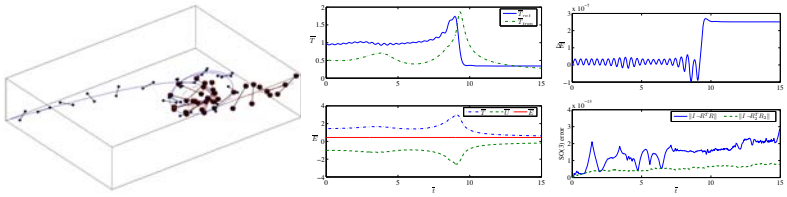
The main feature of the double spherical pendulum example is that it has a nontrivial magnetic term in its reduced symplectic form. Our method is still efficient as it can directly handle the essential non-canonical nature of the symplectic structure. In contrast, a traditional symplectic method for canonical systems could require repeated coordinate changes if one is evoking Darboux' theorem to transform the symplectic structure into canonical form, thereby incurring additional computational cost. Our method allows one to design reduced symplectic integrators in a natural way, despite the noncanonical nature of the symplectic structure.





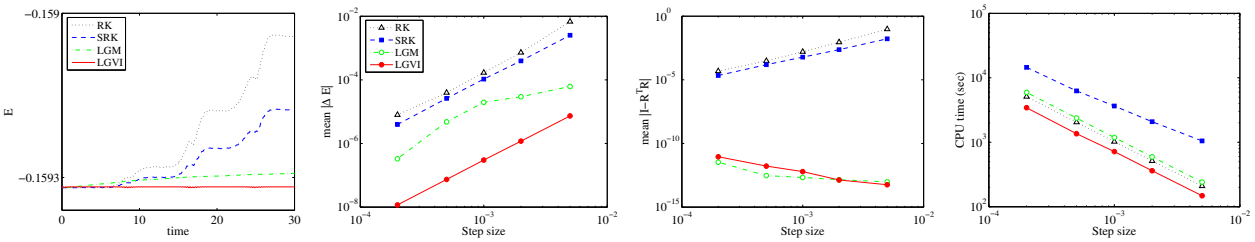
3. **Lie Group Variational Integrators for the Full Body Problem** *(with T. Lee, and N.H. McClamroch)*
Comput. Methods Appl. Mech. Engrg., **196** (29-30), 2907–2924, 2007. doi:10.1016/j.cma.2007.01.017

We develop the equations of motion for full body models that describe the dynamics of rigid bodies, acting under their mutual gravity. The equations are derived using a variational approach where variations are defined on the Lie group of rigid body configurations. Both continuous equations of motion and variational integrators are developed in Lagrangian and Hamiltonian forms, and the reduction from the inertial frame to a relative coordinate system is also carried out. The Lie group variational integrators are shown to be symplectic, to preserve conserved quantities, and to guarantee exact evolution on the configuration space. One of these variational integrators is used to simulate the dynamics of two rigid dumbbell bodies.



4. **Lie Group Variational Integrators for the Full Body Problem in Orbital Mechanics**
(with T. Lee, and N.H. McClamroch)
Celestial Mechanics and Dynamical Astronomy, **98** (2), 121–144, 2007. doi:10.1007/s10569-007-9073-x

   Equations of motion, referred to as full body models, are developed to describe the dynamics of rigid bodies acting under their mutual gravitational potential. Continuous equations of motion and discrete equations of motion are derived using Hamilton's principle. These equations are expressed in an inertial frame and in relative coordinates. The discrete equations of motion, referred to as a Lie group variational integrator, provide a geometrically exact and numerically efficient computational method for simulating full body dynamics in orbital mechanics; they are symplectic and momentum preserving, and they exhibit good energy behavior for exponentially long time periods. They are also efficient in only requiring a single evaluation of the gravity forces and moments per time step. The Lie group variational integrator also preserves the group structure without the use of local charts, reprojection, or constraints. Computational results are given for the dynamics of two rigid dumbbell bodies acting under their mutual gravity; these computational results demonstrate the superiority of the Lie group variational integrator compared with integrators that are not symplectic or do not preserve the Lie group structure.



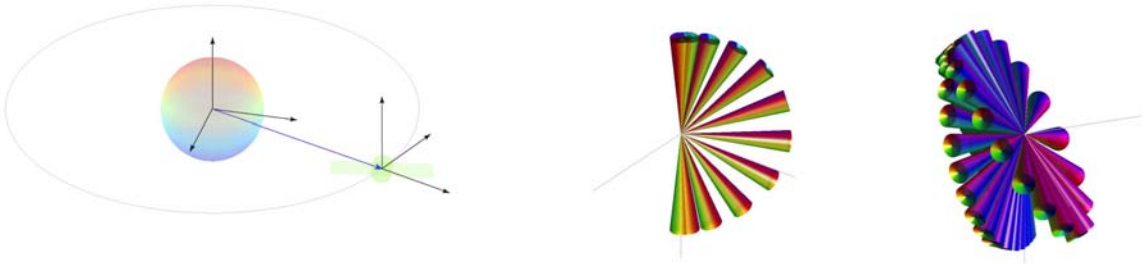
Computational properties of explicit Runge-Kutta (RK), symplectic Runge-Kutta (SRK), Lie group method (LGM), and Lie group variational integrator (LGVI).

5. **Global Optimal Attitude Estimation using Uncertainty Ellipsoids**
(with T. Lee, A.K. Sanyal, N.H. McClamroch)
Systems and Control Letters, **57** (3), 236–245, 2007.

   A deterministic attitude estimation problem for a rigid body in a potential field, with bounded attitude and angular velocity measurement errors is considered. An attitude estimation algorithm that globally minimizes the attitude estimation error is obtained. Assuming that the initial attitude, the initial angular velocity and measurement noise lie within given ellipsoidal bounds, an uncertainty ellipsoid that bounds the attitude and the angular velocity of the rigid body is obtained. The center of the uncertainty ellipsoid provides point estimates, and the size of the uncertainty ellipsoid measures the accuracy of the estimates. The point estimates and the uncertainty ellipsoids are propagated using a Lie group variational integrator and its linearization, respectively. The attitude and angular velocity estimates are optimal in the sense that the sizes of the uncertainty ellipsoids are minimized.

6. **Optimal Attitude Control of a Rigid Body using Geometrically Exact Computations on $SO(3)$** *(with T. Lee, and N.H. McClamroch)*
Journal of Dynamical and Control Systems, accepted, 2007.

An efficient and accurate computational approach is proposed for optimal attitude control of a rigid body. The problem is formulated directly as a discrete time optimization problem using a Lie group variational integrator. Discrete necessary conditions for optimality are derived, and an efficient computational approach is proposed to solve the resulting two point boundary value problem. The use of geometrically exact computations on $SO(3)$ guarantees that this optimal control approach has excellent convergence properties even for highly nonlinear large angle attitude maneuvers. Numerical results are presented for attitude maneuvers of a 3D pendulum and a spacecraft in a circular orbit.



7. **Discrete Control Systems** *(with T. Lee and N.H. McClamroch)*
Invited article for the Springer Encyclopedia of Complexity and Systems Science, accepted, 2008.

Discrete control systems, as considered here, refer to the control theory of discrete-time Lagrangian or Hamiltonian systems. These discrete-time models are based on a discrete variational principle, and are part of the broader field of geometric integration. Geometric integrators are numerical integration methods that preserve geometric properties of continuous systems, such as conservation of the symplectic form, momentum, and energy. They also guarantee that the discrete flow remains on the manifold on which the continuous system evolves, an important property in the case of rigid-body dynamics.

In nonlinear control, one typically relies on differential geometric and dynamical systems techniques to prove properties such as stability, controllability, and optimality. More generally, the geometric structure of such systems plays a critical role in the nonlinear analysis of the corresponding control problems. Despite the critical role of geometry and mechanics in the analysis of nonlinear control systems, nonlinear control algorithms have typically been implemented using numerical schemes that ignore the underlying geometry.

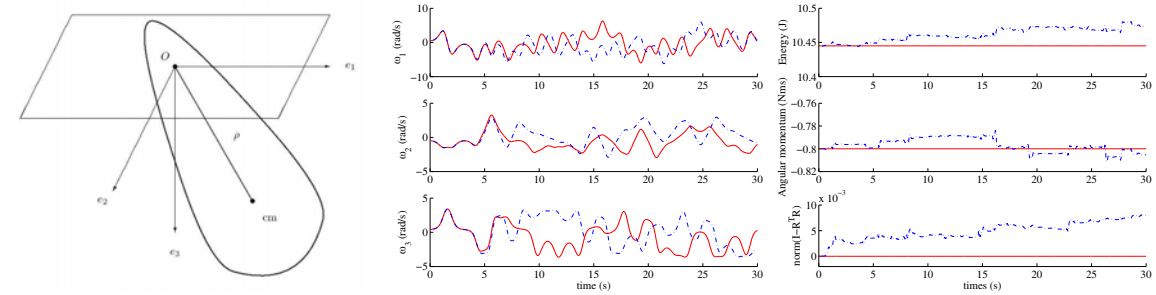
The field of discrete control system aims to address this deficiency by restricting the approximation to choice of a discrete-time model, and developing an associated control theory that does not introduce any additional approximation. In particular, this involves the construction of a control theory for discrete-time models based on geometric integrators that yields numerical implementations of nonlinear and geometric control algorithms that preserve the crucial underlying geometric structure.

8. **Estimating the Attractor Dimension of the Equatorial Weather System**
Acta Phys. Pol. A **85**, 1994.

The correlation dimension and limit capacity serve theoretically as lower and upper bounds, respectively, of the fractal dimension of attractors of dynamic systems. In this paper we show that estimates of the correlation dimension grow rapidly with increasing noise level in the time-series, while estimates of the limit capacity remain relatively unaffected. It is therefore proposed that the limit capacity can be used in studies of noisy data, despite its heavier computational requirements. An analysis of Singapore wind data with the limit capacity estimate revealed a surprisingly low dimension (≈ 2.5). It is suggested that further studies can be made with comprehensive equatorial weather data.

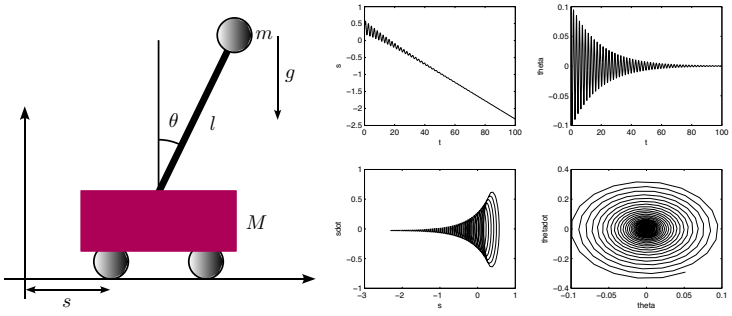
9. **A Lie Group Variational Integrator for the Attitude Dynamics of a Rigid Body with Applications to the 3D Pendulum** *(with T. Lee, and N.H. McClamroch)*
Proc. IEEE Conf. on Control Applications, 962–967, 2005.

   A numerical integrator is derived for a class of models that describe the attitude dynamics of a rigid body in the presence of an attitude dependent potential. The numerical integrator is obtained from a discrete variational principle, and exhibits excellent geometric conservation properties. In particular, by performing computations at the level of the Lie algebra, and updating the solution using the matrix exponential, the attitude automatically evolves on the rotation group embedded in the space of matrices. The geometric conservation properties of the numerical integrator imply long time numerical stability. We apply this variational integrator to the uncontrolled 3D pendulum, that is a rigid asymmetric body supported at a frictionless pivot acting under the influence of uniform gravity. Interesting dynamics of the 3D pendulum are exposed.



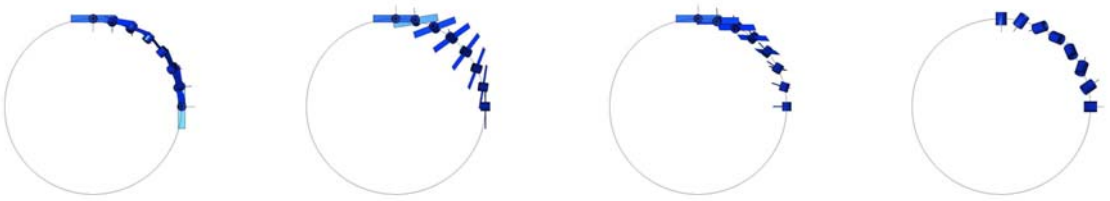
10. **Controlled Lagrangians and Stabilization of the Discrete Cart-Pendulum System**
(with A.M. Bloch, J.E. Marsden, and D.V. Zenkov)
Proc. IEEE Conf. on Decision and Control, 6579–6584, 2005.

    Matching techniques are developed for discrete mechanical systems with symmetry. We describe new phenomena that arise in the controlled Lagrangian approach for mechanical systems in the discrete context. In particular, one needs to either make an appropriate selection of momentum levels or introduce a new parameter into the controlled Lagrangian to complete the matching procedure. We also discuss digital and model predictive control.



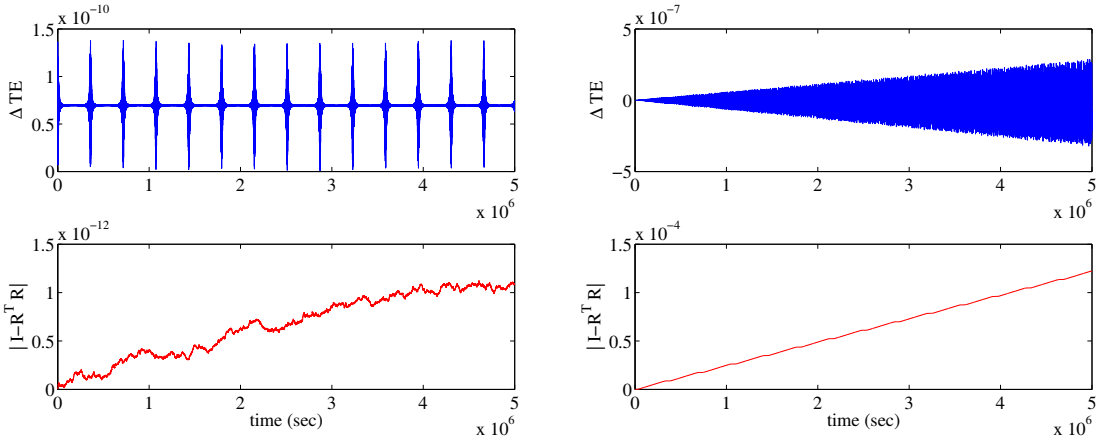
11. **Attitude Maneuvers of a Rigid Spacecraft in a Circular Orbit** *(with T. Lee, and N.H. McClamroch)*
Proc. American Control Conf., 1742–1747, 2006.

A global model is presented that can be used to study attitude maneuvers of a rigid spacecraft in a circular orbit about a large central body. The model includes gravity gradient effects that arise from the non-uniform gravity field and characterizes the spacecraft attitude with respect to the uniformly rotating local vertical local horizontal coordinate frame. The model is globally defined and provides the basis for development and analysis of various attitude maneuvers. An accurate computational approach for solving a nonlinear boundary value problem, with given initial attitude, given terminal attitude and given maneuver time, is proposed, assuming that control torque impulses can be applied at initiation and at termination of the maneuver. If the terminal attitude condition is relaxed, then an accurate computational approach for solving the minimal impulse optimal control problem is presented.



12. **Polyhedral Potential and Lie Group Variational Integrator Computations for the Full Two Body Problem**, *(with E. Fahnestock, T. Lee, N.H. McClamroch, and D. Scheeres)*
Proc. AIAA/AAS Astrodynamics Specialist Conf., AIAA-2006-6289, 2006.

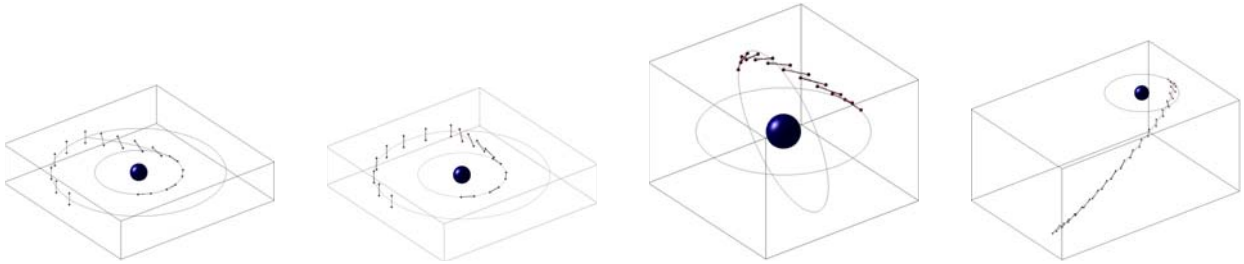
We present a combination of tools which allows for investigation of the coupled orbital and rotational dynamics of two rigid bodies with nearly arbitrary shape and mass distribution, under the influence of their mutual gravitational potential. Methods for calculating that mutual potential and resulting forces and moments for a polyhedral body representation are simple and efficient. Discrete equations of motion, referred to as the Lie Group Variational Integrator (LGVI), preserve the structure of the configuration space, SE(3), as well as the geometric features represented by the total energy and the total angular momentum. The synthesis of these approaches allows us to simulate the full two body problem accurately and efficiently. Simulation results are given for two octahedral rigid bodies for comparison with other integration methods and to show the qualities of the results thus obtained. A significant improvement is seen over other integration methods while correctly capturing the interesting effects of strong orbit and attitude dynamics coupling, in multiple scenarios.



Qualitative comparisons of behavior of deviation in total energy and rotation matrix error over time with similar computational load for LGVI and RKF7(8).

13. **Optimal Control of a Rigid Body using Geometrically Exact Computations on SE(3)** *(with T. Lee, and N.H. McClamroch)*
Proc. IEEE Conf. on Decision and Control, 2710–2715, 2006.

Optimal control problems are formulated and efficient computational procedures are proposed for combined orbital and rotational maneuvers of a rigid body in three dimensions. The rigid body is assumed to act under the influence of forces and moments that arise from a potential and from control forces and moments. The key features of this paper are its use of computational procedures that are guaranteed to preserve the geometry of the optimal solutions. The theoretical basis for the computational procedures is summarized, and examples of optimal spacecraft maneuvers are presented.

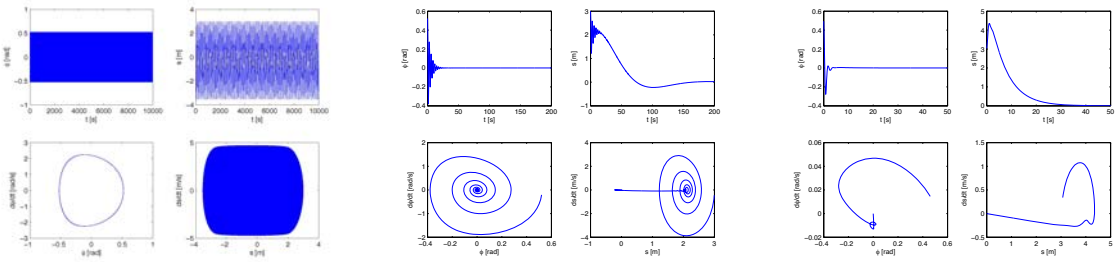


14. **Deterministic Global Attitude Estimation** *(with T. Lee, A.K. Sanyal, N.H. McClamroch)*
Proc. IEEE Conf. on Decision and Control, 3174–3179, 2006.

A deterministic attitude estimation problem for a rigid body in an attitude dependent potential field with bounded measurement errors is studied. An attitude estimation scheme that does not use generalized coordinate representations of the attitude is presented here. Assuming that the initial attitude, angular velocity and measurement noise lie within given ellipsoidal bounds, an uncertainty ellipsoid that bounds the attitude and the angular velocity of the rigid body is obtained. The center of the uncertainty ellipsoid provides point estimates, and its size gives the accuracy of the estimates. The point estimates and the uncertainty ellipsoids are propagated using a Lie group variational integrator and its linearization, respectively. The estimation scheme is optimal in the sense that the attitude estimation error and the size of the uncertainty ellipsoid is minimized at each measurement instant, and it is global since the attitude is represented by a rotation matrix.

15. **Controlled Lagrangians and Potential Shaping for Stabilization of Discrete Mechanical Systems** *(with A.M. Bloch, J.E. Marsden, D.V. Zenkov)*
Proc. IEEE Conf. on Decision and Control, 3333–3338, 2006.

The method of controlled Lagrangians for discrete mechanical systems is extended to include potential shaping in order to achieve complete state-space asymptotic stabilization. New terms in the controlled shape equation that are necessary for matching in the discrete context are introduced. The theory is illustrated with the problem of stabilization of the cartpendulum system on an incline. We also discuss digital and model predictive control.

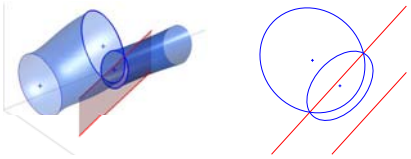


16. **A Discrete Variational Integrator for Optimal Control Problems on SO(3)** *(with A.M. Bloch, I.I. Hussein, A.K. Sanyal)*
Proc. IEEE Conf. on Decision and Control, 6636–6641, 2006.

    In this paper we study a discrete variational optimal control problem for the rigid body. The cost to be minimized is the external torque applied to move the rigid body from an initial condition to a pre-specified terminal condition. Instead of discretizing the equations of motion, we use the discrete equations obtained from the discrete Lagrange–d'Alembert principle, a process that better approximates the equations of motion. Within the discrete-time setting, these two approaches are not equivalent in general. The kinematics are discretized using a natural Lie-algebraic formulation that guarantees that the flow remains on the Lie group $SO(3)$ and its algebra $\mathfrak{so}(3)$. We use Lagrange's method for constrained problems in the calculus of variations to derive the discretetime necessary conditions. We give a numerical example for a three-dimensional rigid body maneuver.

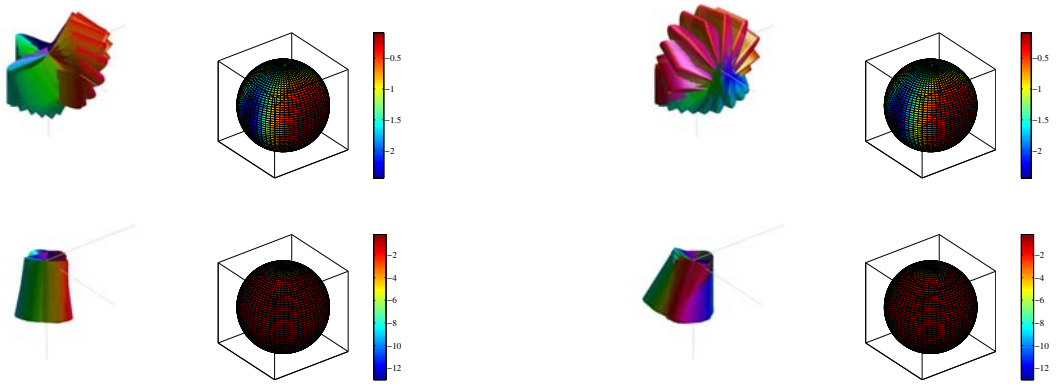
17. **Global Attitude Estimation using Single Direction Measurements**
(with T. Lee, N.H. McClamroch, A.K. Sanyal)
Proc. American Control Conf., 3659–3664, 2007.

    A deterministic attitude estimator for a rigid body under an attitude dependent potential is studied. This estimator requires only a single direction measurement to a known reference point at each measurement instant. The measurement cannot completely determine the attitude, but an attitude estimation scheme based on this measurement is developed; a feasible set compatible with the measurement is described and it is combined with an attitude dynamics model to obtain an attitude estimate. The attitude is globally represented by a rotation matrix, and the uncertainties are described by ellipsoidal sets. A numerical example for a spacecraft in a circular orbit is presented.



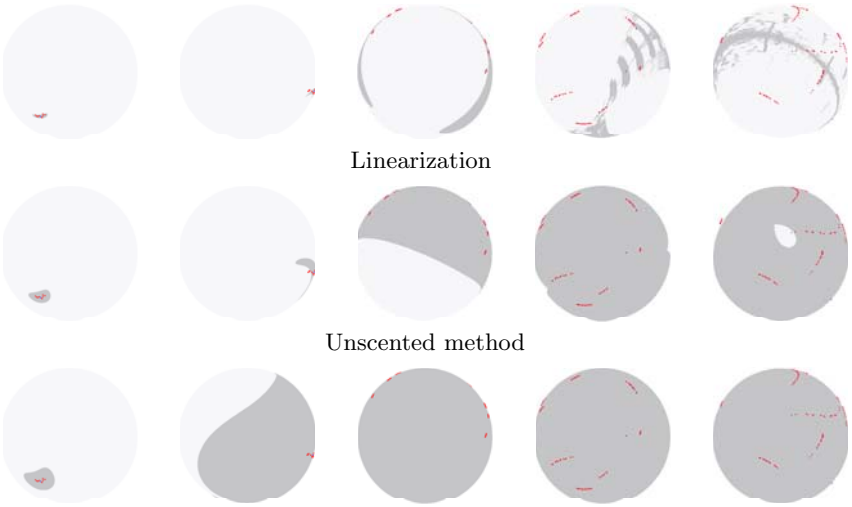
18. **Optimal Attitude Control for a Rigid Body with Symmetry**
(with T. Lee, N.H. McClamroch)
Proc. American Control Conf., 1073–1078, 2007.

    Optimal control problems are formulated and efficient computational procedures are proposed for attitude dynamics of a rigid body with symmetry. The rigid body is assumed to act under a gravitational potential and under a structured control moment that respects the symmetry. The symmetry in the attitude dynamics system yields a conserved quantity, and it causes a fundamental singularity in the optimal control problem. The key feature of this paper is its use of computational procedures that are guaranteed to avoid the numerical ill-conditioning that originates from this symmetry. It also preserves the geometry of the attitude dynamics. The theoretical basis for the computational procedures is summarized, and examples of optimal attitude maneuvers for a 3D pendulum are presented.



19. **Propagation of Uncertainty in Rigid Body Attitude Flows**
(with N.A. Chaturvedi, T. Lee, A.K. Sanyal and N.H. McClamroch)
Proc. IEEE Conf. on Decision and Control, 2689–2694, 2007.

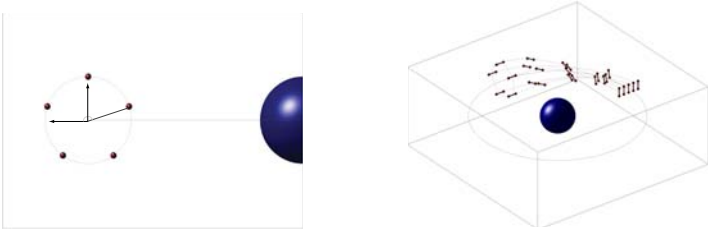
Motivated by attitude control and attitude estimation problems for a rigid body, computational methods are proposed to propagate uncertainties in the angular velocity and the attitude. The nonlinear attitude flow is determined by Euler-Poincaré equations that describe the rotational dynamics of the rigid body acting under the influence of an attitude dependent potential and by a reconstruction equation that describes the kinematics expressed in terms of an orthogonal matrix representing the rigid body attitude. Uncertainties in the angular velocity and attitude are described in terms of ellipsoidal sets that are propagated through this highly nonlinear attitude flow. Computational methods are proposed, one method based on a local linearization of the attitude flow and two methods based on propagation of a small (unscented) sample selected from the initial uncertainty ellipsoid. Each of these computational methods is constructed using the Lie group variational integrator algorithm, viewed as a discretization of the attitude flow dynamics. Computational results are obtained that indicate (1) the strongly nonlinear attitude flow characteristics and (2) the limitations of each of these methods, and indeed any method, in providing effective global bounds on the nonlinear attitude flow.



Uncertainty projected onto the reduced attitude on S^2 for irregular attitude flow.

20. **A Combinatorial Optimization Problem for Spacecraft Formation Reconfiguration**
(with T. Lee, and N.H. McClamroch)
Proc. IEEE Conf. on Decision and Control, 5370–5375, 2007.

We consider a spacecraft formation reconfiguration problem in the case of identical spacecraft. This introduces in the optimal reconfiguration problem a permutational degree of freedom, in addition to the choice of individual spacecraft trajectories. We approach this using a coupled combinatorial and continuous optimization framework, in which the inner loop consists of computing the costs associated with a particular assignment by using a geometrically exact and numerically efficient discrete optimal control method based on Lie group variational integrators. In the outer optimization loop, combinatorial techniques are used to determine the optimal assignments based on the costs computed in the inner loop. The proposed method is demonstrated on the optimal reconfiguration problem for 5 identical spacecraft to go from an inline configuration to one equally spaced on a circle.

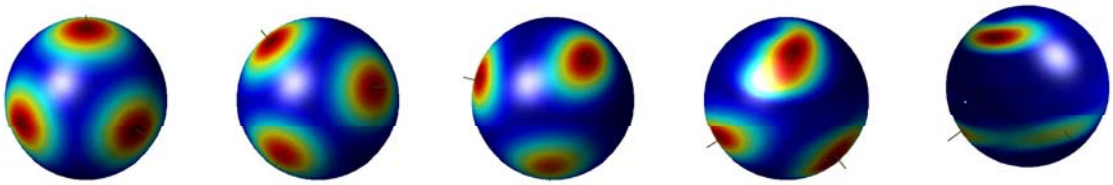


21. **Time Optimal Attitude Control for a Rigid Body** *(with T. Lee and N.H. McClamroch)*
Proc. American Control Conf., 5210–5215, 2008.

A time optimal attitude control problem is studied for the dynamics of a rigid body. The objective is to minimize the time to rotate the rigid body to a desired attitude and angular velocity while subject to constraints on the control input. Necessary conditions for optimality are developed directly on the special orthogonal group using rotation matrices. They completely avoid singularities associated with local parameterizations such as Euler angles, and they are expressed as compact vector equations. In addition, a discrete control method based on a geometric numerical integrator, referred to as a Lie group variational integrator, is proposed to compute the optimal control input. The computational approach is geometrically exact and numerically efficient. The proposed method is demonstrated by a large-angle maneuver for an elliptic cylinder rigid body.

22. **Global Symplectic Uncertainty Propagation on** $SO(3)$ *(with T. Lee and N.H. McClamroch*
Prof. IEEE Conf. on Decision and Control, accepted, 2008.

This paper introduces a global uncertainty propagation scheme for rigid body dynamics, through a combination of numerical parametric uncertainty techniques, noncommutative harmonic analysis, and geometric numerical integration. This method is distinguished from prior approaches, as it allows one to consider probability densities that are global, and are not supported on only a single coordinate chart on the manifold. The use of Lie group variational integrators, that are symplectic and stay on the Lie group, as the underlying numerical propagator ensures that the advected probability densities respect the geometric properties of uncertainty propagation in Hamiltonian systems, which arise as consequence of the Gromov nonsqueezing theorem from symplectic geometry. We also describe how the global uncertainty propagation scheme can be applied to the problem of global attitude estimation.

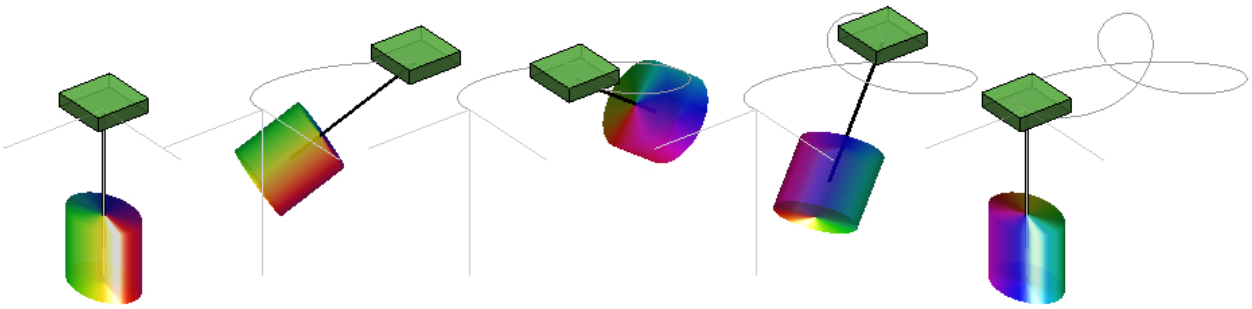


23. **Geometric Structure-Preserving Optimal Control of the Rigid Body**
(with A.M. Bloch, I.I. Hussein, A.K. Sanyal)
Journal of Dynamical and Control Systems, submitted, 2007.

In this paper we study a discrete variational optimal control problem for the rigid body. The cost to be minimized is the external torque applied to move the rigid body from an initial condition to a pre-specified terminal condition. Instead of discretizing the equations of motion, we use the discrete equations obtained from the discrete Lagrange–d'Alembert principle, a process that better approximates the equations of motion. Within the discrete-time setting, these two approaches are not equivalent in general. The kinematics are discretized using a natural Lie-algebraic formulation that guarantees that the flow remains on the Lie group $SO(3)$ and its algebra $\mathfrak{so}(3)$. We use Lagranges method for constrained problems in the calculus of variations to derive the discrete-time necessary conditions. We give a numerical example for a three-dimensional rigid body maneuver.

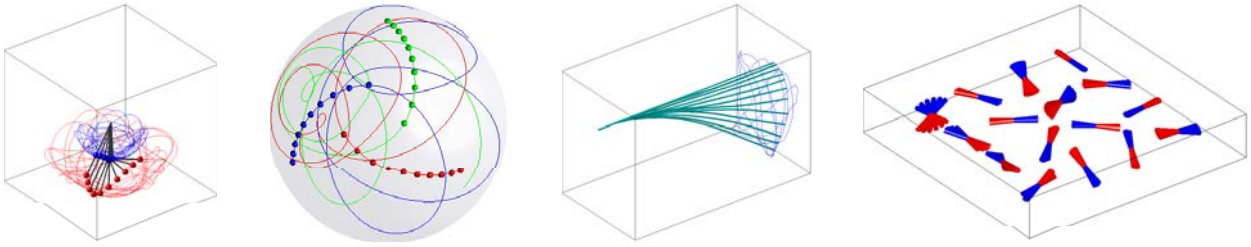
24. **Computational Geometric Optimal Control of Rigid Bodies** *(with T. Lee, N.H. McClamroch)*
Brockett Legacy Special Issue, Communications in Information and Systems, submitted, 2008.

This paper formulates optimal control problems for rigid bodies in a geometric manner and it presents computational procedures based on this geometric formulation for numerically solving these optimal control problems. The dynamics of each rigid body is viewed as evolving on a configuration manifold that is a Lie group. Discrete-time dynamics of each rigid body are developed that evolve on the configuration manifold according to a discrete version of Hamilton's principle so that the computations preserve geometric features of the dynamics and guarantee evolution on the configuration manifold; these discrete-time dynamics are referred to as Lie group variational integrators. Rigid body optimal control problems are formulated as discrete-time optimization problems for discrete Lagrangian/Hamiltonian dynamics, to which standard numerical optimization algorithms can be applied. This general approach is illustrated by presenting results for several different optimal control problems for a single rigid body and for multiple interacting rigid bodies. The computational advantages of the approach, that arise from correctly modeling the geometry, are discussed.



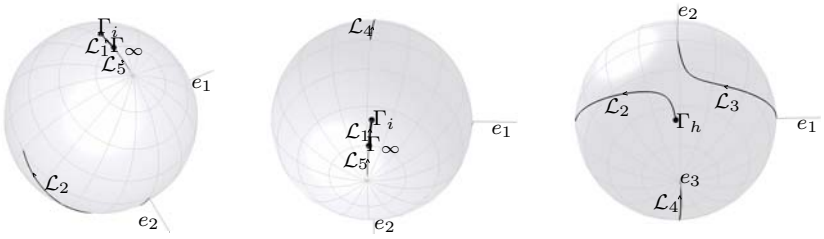
25. **Lagrangian Mechanics and Variational Integrators on Two-Spheres**
(with T. Lee and N.H. McClamroch)
International Journal for Numerical Methods in Engineering, submitted, 2008.

Euler-Lagrange equations and variational integrators are developed for Lagrangian mechanical systems evolving on a product of two-spheres. The geometric structure of a product of two-spheres is carefully considered in order to obtain global equations of motion. Both continuous equations of motion and variational integrators completely avoid the singularities and complexities introduced by local parameterizations or explicit constraints. We derive global expressions for the Euler-Lagrange equations on two-spheres which are more compact than existing equations written in terms of angles. Since the variational integrators are derived from Hamilton's principle, they preserve the geometric features of the dynamics such as symplecticity, momentum maps, or total energy, as well as the structure of the configuration manifold. Computational properties of the variational integrators are illustrated for several mechanical systems. In addition, we describe how Lie group variational integrators can be used to integrate Lagrangian flows on more general homogeneous spaces by lifting the discrete Hamilton's principle on homogeneous spaces to a constrained discrete variational principle on the Lie group by choosing a discrete connection.



26. **Nonlinear Dynamics of the 3D Pendulum** *(with N.A. Chaturvedi, T. Lee, N.H. McClamroch)*
Journal of Nonlinear Science, submitted, 2008

A 3D pendulum consists of a rigid body, supported at a fixed pivot, with three rotational degrees of freedom. The pendulum is acted on by a gravitational force. 3D pendulum dynamics have been much studied in integrable cases that arise when certain physical symmetry assumptions are made. This paper treats the non-integrable case of the 3D pendulum dynamics when the rigid body is asymmetric and the center of mass is distinct from the pivot location. Full and reduced models of the 3D pendulum are introduced and used to study important features of the nonlinear dynamics: conserved quantities, equilibria, relative equilibria, invariant manifolds, local dynamics, and presence of chaotic motions. The paper provides a unified treatment of the 3D pendulum dynamics that includes prior results and new results expressed in the framework of geometric mechanics. These results demonstrate the rich and complex dynamics of the 3D pendulum.



(a) 3D view (b) Top view (c) Bottom view

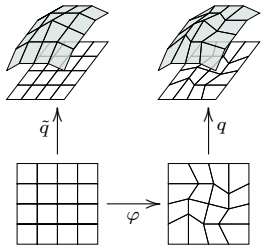
Relative equilibria attitudes for an elliptic cylinder 3D pendulum model

(d) $E = -2.65$ (e) $E = 0$ (f) $E = 2.03$ (g) $E = 8.83$ (h) $E = 11.95$

Poincaré maps for 3D pendulum with varying total energy

27. **Generalized Galerkin Variational Integrators**
Preprint, 2004.

We introduce generalized Galerkin variational integrators, which are a natural generalization of discrete variational mechanics, whereby the discrete action, as opposed to the discrete Lagrangian, is the fundamental object. This is achieved by approximating the action integral with appropriate choices of a finite-dimensional function space that approximate sections of the configuration bundle and numerical quadrature to approximate the integral. We discuss how this general framework allows us to recover higher-order Galerkin variational integrators, asynchronous variational integrators, and symplectic-energy-momentum integrators. In addition, we will consider function spaces that are not parameterized by field values evaluated at nodal points, which allows the construction of Lie group, multiscale, and pseudospectral variational integrators. The construction of pseudospectral variational integrators is illustrated by applying it to the (linear) Schrödinger equation. G-invariant discrete Lagrangians are constructed in the context of Lie group methods through the use of natural charts and interpolation at the level of the Lie algebra. The reduction of these G-invariant Lagrangians yield a higher-order analogue of discrete Euler–Poincaré reduction. By considering nonlinear approximation spaces, spatio-temporally adaptive variational integrators can be introduced as well.



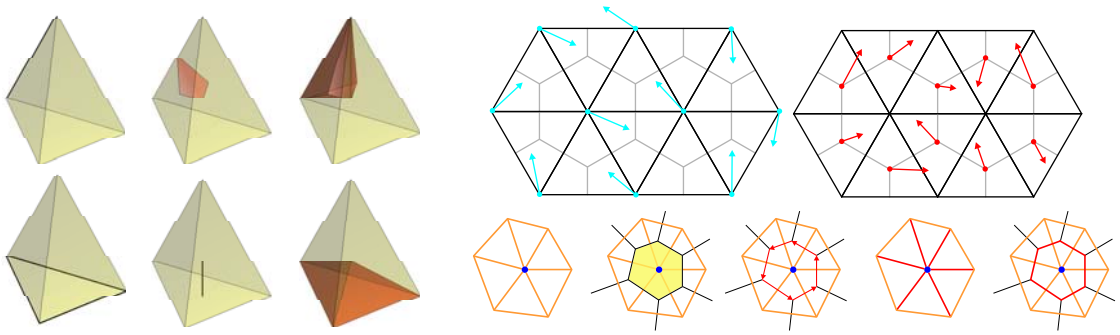
28. **A Discrete Theory of Connections on Principal Bundles** *(with J.E. Marsden, and A.D. Weinstein)*
Preprint, 2004.

Connections on principal bundles play a fundamental role in expressing the equations of motion for mechanical systems with symmetry in an intrinsic fashion. A discrete theory of connections on principal bundles is constructed by introducing the discrete analogue of the Atiyah sequence, with a connection corresponding to the choice of a splitting of the short exact sequence. Equivalent representations of a discrete connection are considered, and an extension of the pair groupoid composition, that takes into account the principal bundle structure, is introduced. Computational issues, such as the order of approximation, are also addressed. Discrete connections provide an intrinsic method for introducing coordinates on the reduced space for discrete mechanics, and provide the necessary discrete geometry to introduce more general discrete symmetry reduction. In addition, discrete analogues of the Levi-Civita connection, and its curvature, are introduced by using the machinery of discrete exterior calculus, and discrete connections.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \tilde{G} & \underset{(\pi_1,\mathcal{A}_d)}{\overset{(q,gq)}{\rightleftarrows}} & (Q\times Q)/G & \underset{(\cdot,\cdot)^h}{\overset{(\pi,\pi)}{\rightleftarrows}} & S\times S & \longrightarrow & 0 \\ & & \| 1_{\tilde{G}} & & \downarrow \alpha_{\mathcal{A}_d} & & \| 1_{S\times S} & & \\ 0 & \longrightarrow & \tilde{G} & \underset{\pi_1}{\overset{i_1}{\rightleftarrows}} & \tilde{G}\oplus(S\times S) & \underset{i_2}{\overset{\pi_2}{\rightleftarrows}} & S\times S & \longrightarrow & 0 \end{array}$$

29. **Discrete Exterior Calculus** *(with M. Desbrun, A.N. Hirani, J.E. Marsden)*
Preprint, 2003.

We present a theory and applications of discrete exterior calculus on simplicial complexes of arbitrary finite dimension. This can be thought of as calculus on a discrete space. Our theory includes not only discrete differential forms but also discrete vector fields and the operators acting on these objects. This allows us to address the various interactions between forms and vector fields (such as Lie derivatives) which are important in applications. Previous attempts at discrete exterior calculus have addressed only differential forms. We also introduce the notion of a circumcentric dual of a simplicial complex. The importance of dual complexes in this field has been well understood, but previous researchers have used barycentric subdivision or barycentric duals. We show that the use of circumcentric duals is crucial in arriving at a theory of discrete exterior calculus that admits both vector fields and forms.



Thesis Abstract

Foundations of Computational Geometric Mechanics

Ph.D., Control and Dynamical Systems, California Institute of Technology, 2004.

Thesis Advisor: Jerrold E. Marsden

A preliminary version of this thesis received the ***SIAM Student Paper Prize*** *and the* ***Leslie Fox Prize in Numerical Analysis*** *(second prize).*

Geometric mechanics involves the study of Lagrangian and Hamiltonian mechanics using geometric and symmetry techniques. Computational algorithms obtained from a discrete Hamilton's principle yield a discrete analogue of Lagrangian mechanics, and they exhibit excellent structure-preserving properties that can be ascribed to their variational derivation.

We construct discrete analogues of the geometric and symmetry methods underlying geometric mechanics to enable the systematic development of computational geometric mechanics. In particular, we develop discrete theories of reduction by symmetry, exterior calculus, connections on principal bundles, as well as generalizations of variational integrators.

Discrete Routh reduction is developed for abelian symmetries, and extended to systems with constraints and forcing. Variational Runge–Kutta discretizations are considered in detail, including the extent to which symmetry reduction and discretization commute. In addition, we obtain the Reduced Symplectic Runge–Kutta algorithm, which is a discrete analogue of cotangent bundle reduction.

Discrete exterior calculus is modeled on a primal simplicial complex, and a dual circumcentric cell complex. Discrete notions of differential forms, exterior derivatives, Hodge stars, codifferentials, sharps, flats, wedge products, contraction, Lie derivative, and the Poincaré lemma are introduced, and their discrete properties are analyzed. In examples such as harmonic maps and electromagnetism, discretizations arising from discrete exterior calculus commute with taking variations in Hamilton's principle, which implies that directly discretizing these equations yield numerical schemes that have the structure-preserving properties associated with variational schemes.

Discrete connections on principal bundles are obtained by introducing the discrete Atiyah sequence, and considering splittings of the sequence. Equivalent representations of a discrete connection are considered, and an extension of the pair groupoid composition that takes into account the principal bundle structure is introduced. Discrete connections provide an intrinsic coordinatization of the reduced discrete space, and the necessary discrete geometry to develop more general discrete symmetry reduction techniques.

Generalized Galerkin variational integrators are obtained by discretizing the action integral through appropriate choices of finite-dimensional function space and numerical quadrature. Explicit expressions for Lie group, higher-order Euler–Poincaré, higher-order symplectic-energy-momentum, and pseudospectral variational integrators are presented, and extensions such as spatio-temporally adaptive and multiscale variational integrators are briefly described.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Numerical Mathematics 53 (2005) 231–248
APPLIED NUMERICAL MATHEMATICS
www.elsevier.com/locate/apnum

Discrete Poincaré lemma

Mathieu Desbrun [a], Melvin Leok [b,*], Jerrold E. Marsden [c]

[a] *256-80, Department of Computer Science, Caltech, Pasadena, CA 91125, USA*
[b] *Department of Mathematics, University of Michigan, 525 East University Ave., Ann Arbor, MI 48109, USA*
[c] *107-81, Control and Dynamical Systems, Caltech, Pasadena, CA 91125, USA*

Available online 10 December 2004

Abstract

This paper proves a discrete analogue of the Poincaré lemma in the context of a discrete exterior calculus based on simplicial cochains. The proof requires the construction of a generalized cone operator, $p : C_k(K) \to C_{k+1}(K)$, as the geometric cone of a simplex cannot, in general, be interpreted as a chain in the simplicial complex. The corresponding cocone operator $H : C^k(K) \to C^{k-1}(K)$ can be shown to be a homotopy operator, and this yields the discrete Poincaré lemma.

The generalized cone operator is a combinatorial operator that can be constructed for any simplicial complex that can be grown by a process of local augmentation. In particular, regular triangulations and tetrahedralizations of $\mathbb{R}^2$ and $\mathbb{R}^3$ are presented, for which the discrete Poincaré lemma is globally valid.

Keywords: Discrete geometry; Discrete exterior calculus; Compatible discretizations

1. Introduction

The exactness properties of differential complexes such as the de Rham complex have recently been shown to play an important role in the design and stability of numerical methods for partial differential equations [3]. In computational electromagnetism, numerical schemes that discretize the de Rham complex, using interpolation by Whitney forms, have become increasingly prevalent [4,9], and yield generalizations of the Yee staggered-mesh algorithm [13].

* Corresponding author.
E-mail addresses: desbrun@cs.caltech.edu (M. Desbrun), mleok@umich.edu (M. Leok), marsden@cds.caltech.edu (J.E. Marsden).

0168-9274/$30.00 © 2004 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.apnum.2004.09.035

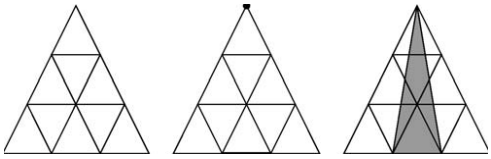
The geometrical structure underlying these numerical schemes is that of exterior calculus, and in Desbrun et al. [5], we introduced a discrete exterior calculus based on simplicial cochains. The cochain representation is attractive as it is particularly simple, and avoids the need for interpolation of forms.

The cochain representation of discrete differential forms uses the formalism of simplicial cohomology (see Munkres [11]) and identifies a cochain with a discrete differential form, and the coboundary with the exterior derivative. Consequently, a discrete differential form can only be evaluated on a chain.

The standard proof of the Poincaré lemma involves the construction of a homotopy operator though the cocone construction. This construction is unsatisfactory for developing a combinatorial proof of the discrete Poincaré lemma, since there is no canonical way to express the combinatorial cone of a k-simplex as a chain consisting of existing $(k+1)$-simplices.

Only by choosing a geometric realization of the abstract simplicial complex does it make sense to ask whether the cone of a simplex is expressible as a chain in the original simplicial complex. Even if we chose a geometric realization of the abstract simplicial complex, we find that this does not, in general, yield a representation of the cone operator as a map from chains to chains.

As an example, consider the figure below. Given the simplicial complex on the left, consisting of triangles, edges and nodes, we wish, in the center figure, to consider the cone of the bold edge with respect to the top most node. Clearly, the resulting cone in the right figure, which is shaded grey, cannot be expressed as a combination of the triangles in the original complex. As such, we cannot express the geometric cone as a combinatorial map in an obvious way.



In this paper, a generalized cone operator that is valid for chains is developed with the essential homotopy properties to yield the discrete Poincaré lemma. In the rest of this section, we will review some basic results from simplicial algebraic topology, and relate them to discrete exterior calculus.

In the second section, we will first consider simplicial complexes on Euclidean space and introduce trivially star-shaped complexes, where the geometric cone operator is expressible as a map from k-chains to $(k+1)$-chains. Next, we will consider logically star-shaped complexes, which are simplicially isomorphic to trivially star-shaped complex. Here, the logical cone operator can be constructed by conjugating the standard cone operator with the isomorphism that relates the vertex scheme of the logically-star shaped simplicial complex to the vertex scheme of the trivially star-shaped simplicial complex.

Finally, drawing upon intuition developed in studying trivially star-shaped complexes, we will construct a generalized cone operator that is valid for any simplicial complex that can be grown by a process of local augmentation. The discrete Poincaré lemma is valid for these generalized star-shaped complexes, which include regular triangulations and tetrahedralizations of $\mathbb{R}^2$ and $\mathbb{R}^3$.

For discrete analogues of the Poincaré lemma which are obtained through the use of interpolated forms, the reader is referred to the work of Bossavit [4] and Hiptmair [8]. Discrete homotopy operators in the context of logically rectangular meshes were addressed in Mansfield and Hydon [10].

1.1. Primal simplicial complex

To discretize a continuous problem using discrete exterior calculus, we first discretize the manifold as a simplicial complex. This is typically a simplicial complex in Euclidean space, but it could also be an abstract simplicial complex. This is sufficient in this paper as differential forms are metric-independent, but the full theory of discrete exterior calculus [5] requires a local metric.

We will now recall some basic definitions of simplices, simplicial complexes and abstract simplicial complexes, which are standard from algebraic topology. A more extensive treatment can be found in Munkres [11, Chapter 1, §1–§3].

Definition 1. A *k-simplex* is the convex span of $k+1$ geometrically independent points,

$$\sigma^k = [v_0, v_1, \ldots, v_k] = \left\{ \sum_{i=0}^{k} \alpha^i v_i \mid \alpha^i \geqslant 0,\ \sum_{i=0}^{n} \alpha^i = 1 \right\}.$$

The points $v_0, \ldots, v_k$ are the *vertices* of the simplex, and k is the *dimension* of the simplex. Any simplex spanned by a (proper) subset of $\{v_0, \ldots, v_k\}$ is a (*proper*) *face* of σ^k. If σ^l is a proper face of σ^k then we write $\sigma^l \prec \sigma^k$.

Example 2. Consider 3 noncolinear points v_0, v_1 and v_2 in $\mathbb{R}^3$. Then these three points individually are examples of 0-simplices which are assumed to have no orientation. Examples of 1-simplices are the oriented line segments $[v_0, v_1]$, $[v_1, v_2]$ and $[v_0, v_2]$. By writing the vertices in that order we have given orientations to these 1-simplices, i.e., $[v_0, v_1]$ is oriented from v_0 to v_1. The triangle $[v_0, v_1, v_2]$ is a 2-simplex oriented in counter clockwise direction. Note that the orientation of $[v_0, v_2]$ does not agree with that of the triangle.

Definition 3. A *simplicial complex* K in $\mathbb{R}^N$ is a collection of simplices in $\mathbb{R}^N$ such that:

(1) Every face of a simplex of K is in K.
(2) The intersection of any two simplices of K is a face of each of them.

Definition 4. The *polytope* of K, denoted $|K|$, is the geometric union of the simplices of K. A *simplicial triangulation* of a polytope $|K|$ is a simplicial complex K such that the union of the simplices of K recover the polytope $|K|$.

Definition 5. If L is a subcollection of K that contains all faces of its elements, then L is a simplicial complex in its own right, and it is called a *subcomplex* of K. The collection of all simplices of K of dimension at most k, is a subcomplex which is called the *k-skeleton* of K and is denoted $K^{(k)}$.

We now introduce the notion of an abstract simplicial complex, which captures the topology of a simplicial complex by encoding its connectivity.

Definition 6. An *abstract simplicial complex* is a collection $\mathcal{S}$ of finite nonempty sets, such that if A is in $\mathcal{S}$, so is every nonempty subset of A.

An element A of $\mathcal{S}$ is a *simplex* of $\mathcal{S}$, and its *dimension* is one less than the number of elements. Every nonempty subset of A is a *face* of A. The *dimension* of $\mathcal{S}$ is the largest dimension of its simplices. The *vertex set* V of $\mathcal{S}$ is the union of the one-point elements of $\mathcal{S}$, and the vertices $v \in V$ are identified with the 0-simplices $\{v\} \in \mathcal{S}$. A subcollection of $\mathcal{S}$ that is itself a complex is called a *subcomplex* of $\mathcal{S}$.

Two abstract simplicial complexes $\mathcal{S}$ and $\mathcal{T}$ are *isomorphic* if there is a bijection f mapping the vertex set of $\mathcal{S}$ to the vertex set of $\mathcal{T}$ such that $\{v_0, \ldots, v_k\} \in \mathcal{S}$ if and only if $\{f(v_0), \ldots, f(v_k)\} \in \mathcal{T}$.

Definition 7. If K is a simplicial complex, and V is its vertex set, let $\mathcal{K}$ be the collection of all subsets $\{v_0, \ldots, v_k\}$ of V such that $[v_0, \ldots, v_k] \in K$. Then, $\mathcal{K}$ is an abstract simplicial complex called the *vertex scheme* of K.

Definition 8. K is a *geometric realization* of $\mathcal{S}$ if the abstract simplicial complex $\mathcal{S}$ is isomorphic to the vertex scheme of the simplicial complex K.

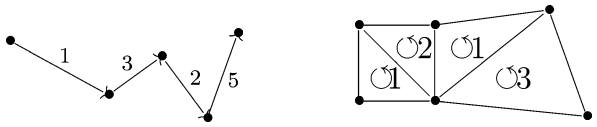
As with simplicial complexes, we will denote the orientation of a simplex in an abstract simplicial complex using the ordered representation $[v_0, \ldots, v_k]$.

1.2. Differential forms and exterior derivative

We will now define discrete differential forms. We will use some terms (which we will define) from algebraic topology but it will become clear by looking at the examples that one can gain a clear and working notion of what a discrete form is without any algebraic topology. We start with a few definitions for which more details can be found in [11, pp. 26–27].

Definition 9. Let K be a simplicial complex. We denote the free Abelian group generated by a basis consisting of oriented k-simplices by, $C_k(K;\mathbb{Z})$. This is the space of finite formal sums of the k-simplices, with coefficients in $\mathbb{Z}$. Elements of $C_k(K;\mathbb{Z})$ are called *k-chains*.

Example 10.



1-chain 2-chain

We view discrete k-forms as maps from the space of k-chains to $\mathbb{R}$. Recalling that the space of k-chains is a group, we require the maps that define the forms to be homomorphisms into the additive group $\mathbb{R}$. Thus, discrete forms are cochains in algebraic topology. We will define cochains below in the definition of forms, but for more context and more details, readers can refer to any algebraic topology text, for example, Munkres [11, p. 251].

This point of view of forms as cochains is not new. The idea of defining forms as cochains appears, for example, in the works of Adams [2], Dezin [6], Hiptmair [7], Sen et al. [12]. Our point of departure is that the other authors go on to develop a theory of discrete exterior calculus of forms by using interpolation which we will be able to avoid. The formal definition of discrete forms follows.

Definition 11. A *primal discrete k-form* α is a homomorphism from the chain group $C_k(K;\mathbb{Z})$ to the additive group $\mathbb{R}$. Thus, a discrete k-form is an element of $\mathrm{Hom}(C_k(K),\mathbb{R})$, the space of *cochains*. This space becomes an Abelian group if we add two homomorphisms by adding their values in $\mathbb{R}$. The standard notation for $\mathrm{Hom}(C_k(K),\mathbb{R})$ in algebraic topology is $C^k(K;\mathbb{R})$. But we will often use the notation $\Omega_d^k(K)$ for this space as a reminder that this is the space of discrete (hence the d subscript) k-forms on the simplicial complex K. Thus

$$\Omega_d^k(K) := C^k(K;\mathbb{R}) = \mathrm{Hom}\big(C_k(K),\mathbb{R}\big).$$

Note that by the above definition for k-chain $\sum_i a_i c_i^k$ (where $a_i \in \mathbb{Z}$) and a discrete k-form α

$$\alpha\left(\sum_i a_i c_i^k\right) = \sum_i a_i \alpha\big(c_i^k\big),$$

and for two discrete k-forms $\alpha, \beta \in \Omega_d^k(K)$ and k-chain $c \in C_k(K;\mathbb{Z})$

$$(\alpha+\beta)(c) = \alpha(c) + \beta(c).$$

In exterior calculus on smooth manifolds, integration of k-forms on a k-manifold is defined in terms of integration in $\mathbb{R}^k$. This is done by doing the integration in local coordinates, which is independent of the choice of charts by the change of variables theorem. For details, see the first few pages of Chapter 7 of Abraham et al. [1]. We will not try to introduce the notion of integration of discrete forms on a simplicial complex. Instead, we will work with the natural bilinear pairing of cochains and chains defined by evaluation. More formally we have the following definition.

Definition 12. The *natural pairing* of a k-form α and a k-chain c is defined as the bilinear pairing

$$\langle \alpha, c\rangle = \alpha(c).$$

As mentioned above, in discrete exterior calculus this natural pairing plays the role that integration of forms on chains plays in the usual exterior calculus. The two are related by a procedure done at the time of discretization. Indeed consider a simplicial triangulation K of a polyhedron in $\mathbb{R}^n$, i.e., consider a "flat" discrete manifold. Consider a continuous problem with some smooth forms defined in the space $|K| \subset \mathbb{R}^n$. To define the discrete form α_d^k corresponding to α^k, one integrates α^k on all the p-simplices in K. Then the evaluation of α_d^p on a k-simplex σ^p is defined by $\alpha_d^k(\sigma^k) := \int_{\sigma^k} \alpha^k$. Thus, discretization is the only place where integration plays a role in our discrete exterior calculus.

Now we can define the discrete exterior derivative, which we will call **d** as in the usual exterior calculus. The discrete exterior derivative will be defined as the dual with respect to the natural pairing defined above, of the boundary operator which is defined below.

Definition 13. The *boundary* operator $\partial_k : C_k(K;\mathbb{Z}) \to C_{k-1}(K;\mathbb{Z})$ is a homomorphism defined by defining it on a simplex $\sigma^k = [v_0,\ldots,v_k]$,

$$\partial_k \sigma^p = \partial_k\big([v_0, v_1, \ldots, v_k]\big) = \sum_{i=0}^{k} (-1)^i [v_0,\ldots,\hat{v}_i,\ldots,v_k],$$

where $[v_0,\ldots,\hat{v}_i,\ldots,v_k]$ is the $(k-1)$-simplex obtained by omitting the vertex v_i. Note that $\partial_k \circ \partial_{k+1} = 0$.

Example 14. Given an oriented triangle $[v_0, v_1, v_2]$, the boundary, by the above, is $[v_1, v_2] - [v_0, v_2] + [v_0, v_1]$ which are the boundary edges of the triangle.

Definition 15. On a simplicial complex of dimension n, a *chain complex* is a collection of chain groups and homomorphisms ∂_k such that

$$0 \to C_n(K) \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_{k+1}} C_k(K) \xrightarrow{\partial_k} \cdots \xrightarrow{\partial_1} C_0(K) \to 0,$$

and $\partial_k \circ \partial_{k+1} = 0$.

Definition 16. The *coboundary operator* $\delta^k : C^k(K) \to C^{k+1}(K)$ defined by duality to the boundary operator using the natural bilinear pairing between discrete forms and chains. Specifically, for a discrete form $\alpha^k \in \Omega_d^k(K)$ and a chain $c_{k+1} \in C_{k+1}(K; \mathbb{Z})$ we define δ^k by

$$\langle \delta^k \alpha^k, c_{k+1} \rangle = \langle \alpha^k, \partial_{k+1} c_{k+1} \rangle \tag{1}$$

that is

$$\delta^k(\alpha^k) = \alpha^k \circ \partial_{k+1}.$$

This definition of the coboundary operator induces the *cochain complex*,

$$0 \leftarrow C^n(K) \xleftarrow{\delta^{n-1}} \cdots \xleftarrow{\delta^k} C^k(K) \xleftarrow{\delta^{k-1}} \cdots \xleftarrow{\delta^0} C^0(K) \leftarrow 0,$$

where it is easy to see that $\delta^{k+1} \circ \delta^k = 0$.

Definition 17. The *discrete exterior derivative* denoted by $\mathbf{d} : \Omega_d^k(K) \to \Omega_d^{k+1}(K)$ is defined to be the coboundary operator δ^k.

Remark 18. With the above definition of the exterior derivative $\mathbf{d} : \Omega_d^k(K) \to \Omega_d^{k+1}(K)$ and the relationship between the natural pairing and integration one can regard equation (1) as a discrete generalized Stokes' theorem. Thus, given a k-chain c and a discrete k-form α, the discrete Stokes' theorem, which is true by definition, states that

$$\langle \mathbf{d}\alpha, c \rangle = \langle \alpha, \partial c \rangle.$$

Furthermore, it also follows immediately that $\mathbf{d}^2 = 0$.

2. Discrete Poincaré lemma

In this section, we will prove the discrete Poincaré lemma using the cocone construction. We will first consider the case of simplicial complexes that are trivially star-shaped, followed by logically star-shaped abstract complexes, before generalizing the result to generalized star-shaped abstract complexes.

As we have shown in the introduction, the cone is not a well-defined map from chains to chains for arbitrary simplicial complexes, and we will first consider trivially star-shaped complexes for which the cone is a well-defined map from chains to chains, before extending the construction to more general complexes.

Definition 19. Given a k-simplex $\sigma^k = [v_0, \ldots, v_k]$ we construct the *cone* with vertex w and base σ^k as follows,

$$w \diamond \sigma^k = [w, v_0, \ldots, v_k].$$

Lemma 20. *The geometric cone operator satisfies the following property,*

$$\partial\big(w \diamond \sigma^k\big) + w \diamond \big(\partial\sigma^k\big) = \sigma^k.$$

Proof. This is a standard result from simplicial algebraic topology. □

2.1. Trivially star-shaped complexes

We first introduce the notion of a trivially star-shaped complex, for which the standard cone construction yields a well-defined map from chains to chains, and for which the standard proof of the Poincaré lemma extends to the discrete case. We will also construct an arbitrarily dense tetrahedralization of an open neighborhood about a point using trivially star-shaped complexes.

It should be noted that the notion of a trivially star-shaped complex depends on the geometric realization, and is therefore not an intrinsic property of an abstract simplicial complex. Intrinsic generalizations will be considered in the next two subsections.

Definition 21. A complex K is called *trivially star-shaped* if there exists a vertex $w \in K^{(0)}$ such that for all $\sigma^k \in K$, the cone with vertex w and base σ^k is (geometrically) expressible as a chain in K. That is to say,

$$\exists w \in K^{(0)} \mid \forall \sigma^k \in K,\ w \diamond \sigma^k \in C_{k+1}(K).$$

We denote the cone operation with respect to w as $p : C_k(K) \to C_{k+1}(K)$.

Lemma 22. *In trivially star-shaped complexes, the cone operator $p : C_k(K) \to C_{k+1}(K)$ satisfies the following identity,*

$$p\partial + \partial p = I,$$

at the level of chains.

Proof. Follows from the identity for cones, and the fact that the cone is well-defined at the level of chains on trivially star-shaped complexes. □

Definition 23. The *cocone* operator $H : C^k(K) \to C^{k-1}(K)$ is defined by,

$$\big\langle H\alpha^k, \sigma^{k-1}\big\rangle = \big\langle \alpha^k, p\big(\sigma^{k-1}\big)\big\rangle.$$

This operator is well-defined on trivially star-shaped simplicial complexes.

Lemma 24. *The cocone operator $H : C^k(K) \to C^{k-1}(K)$ satisfies the following identity,*

$$H\mathbf{d} + \mathbf{d}H = I,$$

at the level of cochains.

Proof. A simple duality argument applied to the cone identity,

$$p\partial + \partial p = I,$$

yields the following,

$$\begin{aligned}\langle \alpha^k, \sigma^k \rangle &= \langle \alpha^k, (p\partial + \partial p)\sigma^k \rangle = \langle \alpha^k, p\partial\sigma^k \rangle + \langle \alpha^k, \partial p\sigma^k \rangle \\ &= \langle H\alpha^k, \partial\sigma^k \rangle + \langle \mathbf{d}\alpha^k, p\sigma^k \rangle = \langle \mathbf{d}H\alpha^k, \sigma^k \rangle + \langle H\mathbf{d}\alpha^k, \sigma^k \rangle \\ &= \langle (\mathbf{d}H + H\mathbf{d})\alpha^k, \sigma^k \rangle.\end{aligned}$$

Therefore,

$$H\mathbf{d} + \mathbf{d}H = I,$$

at the level of cochains. □

Corollary 25 (Discrete Poincaré lemma for trivially star-shaped complexes). *Given a closed cochain α^k, that is to say, $\mathbf{d}\alpha^k = 0$, there exists a cochain β^{k-1} such that $\mathbf{d}\beta^{k-1} = \alpha^k$.*

Proof. Applying the identity for cochains,

$$H\mathbf{d} + \mathbf{d}H = I,$$

we have,

$$\langle \alpha^k, \sigma^k \rangle = \langle (H\mathbf{d} + \mathbf{d}H)\alpha^k, \sigma^k \rangle$$

but $\mathbf{d}\alpha^k = 0$, so,

$$= \langle \mathbf{d}\big(H\alpha^k\big), \sigma^k \rangle.$$

Therefore, $\beta^{k-1} = H\alpha^k$ is such that $\mathbf{d}\beta^{k-1} = \alpha^k$ at the level of cochains. □

Example 26. We construct an arbitrarily dense tetrahedralization of the cone of a $(n-1)$-simplex over the origin.

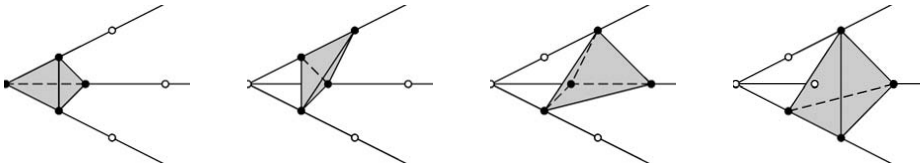
If we denote by v_i^k, the projection of the v_i vertex to the kth concentric sphere, where the 0th concentric sphere is simply the central point, then we fill up the cone $[c, v_1, \ldots, v_n]$ with simplices as follows:

$$\big[v_1^0, v_1^1, \ldots, v_n^1\big], \qquad \big[v_1^2, v_1^1, \ldots, v_n^1\big], \qquad \big[v_1^2, v_2^2, v_2^1, \ldots, v_n^1\big], \ldots, \qquad \big[v_1^2, \ldots, v_n^2, v_n^1\big].$$

Since S^{n-1} is orientable, we use a consistent triangulation of S^{n-1} and the n-cones to triangulate B^n so that the resulting triangulation is star-shaped.

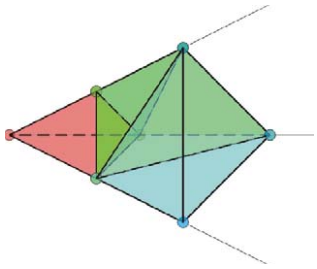
This fills up the region to the 1st concentric sphere, and we repeat the process by leapfrogging at the last vertex to add $[v_1^2, \ldots, v_n^2, v_n^3]$, and continuing the construction, to fill up the annulus between the 1st and 2nd concentric sphere. Thus, we can keep adding concentric shells to create an arbitrarily dense triangulation of a n-ball about the origin.

In three dimensions, these simplices are given by,



$[c, v_1^1, v_2^1, v_3^1]$, $[v_1^2, v_1^1, v_2^1, v_3^1]$, $[v_1^2, v_2^2, v_2^1, v_3^1]$, $[v_1^2, v_2^2, v_3^2, v_3^1]$.

Putting them together, we obtain,



Triangulation of a 3-dimensional cone

This example is significant, since we can construct an arbitrarily dense trivially star-shaped triangulation of a n-ball, and recover the continuous Poincaré lemma from the discrete Poincaré lemma for trivially star-shaped complexes.

2.2. *Logically star-shaped complexes*

We now consider logically star-shaped complexes which are simplicially isomorphic to trivially star-shaped complexes, and as such inhert a cone operator with the desired homotopy properties from the trivially star-shaped complex. This property is intrinsic as it only depends on the vertex scheme of a complex.

Definition 27. A simplicial complex is *logically star-shaped* if its vertex scheme is isomorphic to the vertex scheme of a trivially star-shaped complex.

Example 28. Two simplicial complexes whose vertex schemes are isomorphic.

Trivially star-shaped complex $\cong$ Logically star-shaped complex

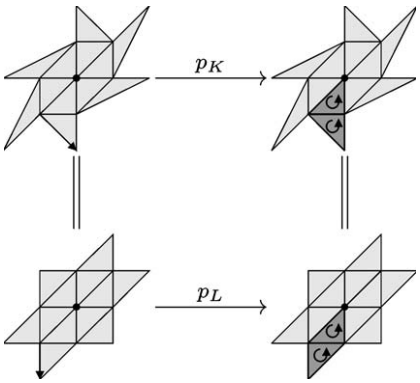
Definition 29. The *logical cone* operator $p : C^k(L) \to C^{k+1}(L)$ satisfies,

$$\begin{array}{ccc} C^k(K) & \xrightarrow{p_K} & C^{k+1}(K) \\ \| & & \| \\ C^k(L) & \xrightarrow{p_L} & C^{k+1}(L) \end{array}$$

Which is to say that given the isomorphism $\varphi : K \to L$, we define,

$$p_L = \varphi \circ p_K \circ \varphi^{-1}.$$

Example 30. An example of the construction of the logical cone operator.



This definition of the logical cone operator results in the identities for the cone and cocone operator to follow from the trivially star-shaped case, and we record the results as follows.

Lemma 31. *In logically star-shaped complexes, the logical cone operator satisfies the following identity,*

$$p\partial + \partial p = I,$$

at the level of chains.

Proof. Follows immediately by pushing forward the result for trivially star-shaped complexes using the isomorphism. □

Lemma 32. *In logically star-shaped complexes, the logical cocone operator satisfies the following identity,*

$$H\mathbf{d} + \mathbf{d}H = I,$$

at the level of cochains.

Proof. Follows immediately by pushing forward the result for trivially star-shaped complexes using the isomorphism. □

Thus, we have a discrete Poincaré lemma for logically star-shaped complexes.

Corollary 33 (Discrete Poincaré lemma for logically star-shaped complexes). *Given a closed cochain* α^k*, that is to say,* $\mathbf{d}\alpha^k = 0$*, there exists a cochain* β^{k-1} *such that* $\mathbf{d}\beta^{k-1} = \alpha^k$.

Proof. Follows from the above lemma using the proof for the trivially star-shaped case. □

2.3. Generalized star-shaped complexes

We now introduce generalized star-shaped complexes, which are constructed using a process of local augmentation. We can recursively construct a generalized cone operator such that it satisfies the homotopy identity,

$$p\partial + \partial p = I,$$

which is the crucial property of the cone operator, from the point of view of proving the discrete Poincaré lemma.

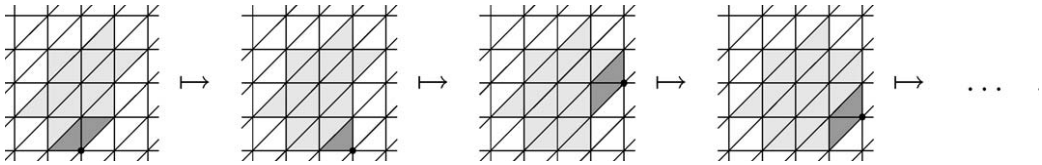
Definition 34. Given a complex K, and a vertex v, the *one-ring* of v is the set of all simplices of K which contain v as a vertex.

Definition 35. Given a n-complex K, consider a vertex w that is not already contained in the complex, and a $(n-1)$-chain c_{n-1} that is contained on the boundary of K, and is included in the one-ring of some vertex on ∂K. Then, a *one-ring cone augmentation* of K is the complex obtained by adding the n-cone $w \diamond c_{n-1}$, and all its faces to the complex.

Definition 36. A complex is *generalized star-shaped* if it can be constructed by repeated one-ring cone augmentation of an initially logically star-shaped complex.

Note that the simplest example of a logically star-shaped complex is the complex consisting of an n-simplex and all its faces. We will now introduce an example of a regular triangulation of the plane that is generalized star-shaped.

Example 37. The regular 2-dimensional triangulation can be obtained by the successive application of the one-ring cone augmentation procedure, as the following sequence illustrates,



Remark 38. Since each simplex in the one-ring of v contains v, and each simplex is connected, a chain consisting of simplices of a one-ring is connected. As such, a noncontractible complex cannot be constructed by inductive one-ring cone augmentation, as it will involve adding a cone $w \diamond c_{n-1}$, where c_{n-1} is disjoint and therefore cannot be included in the one-ring of a vertex.

We construct our generalized cone operator $p : C^k(K) \to C^{k+1}(K)$ recursively. Since we are restricting ourselves to generalized star-shaped complexes, which are obtained by repeated one-ring augmentation of an initially logically star-shaped complex, it suffices to show that we can define a generalized cone operator on the newly added simplices that satisfies the homotopy property.

We proceed by induction. The base case is a generalized star-shaped complex that is logically star-shaped, and has a logical cone operator which satisfies the homotopy property. The generalized cone operator is defined to coincide with the logical cone operator, and satisfies the homotopy property as well.

Given a generalized star-shaped complex K_i, we augment using one-ring cone augmentation by adding the n-cone $w \diamond c_{n-1}$, and all its faces to obtain K_{i+1}. We need to define p for simplices in $K_{i+1} \setminus K_i$.

To define $p(\sigma^k)$ for $\sigma^k \in K_{i+1} \setminus K_i$, we choose a $\sigma^{k+1} \in K_{i+1} \setminus K_i$, to include in the cone $p(\sigma^k)$, such that $\sigma^{k+1} \succ \sigma^k$, and σ^{k+1} and σ^k are consistently oriented. Let us require that the homotopy property of p holds on σ^{k+1},

$$\begin{aligned}\sigma^{k+1} &= (p\partial + \partial p)\big(\sigma^{k+1}\big) = p\big(\partial\sigma^{k+1}\big) + \partial p\big(\sigma^{k+1}\big) \\ &= p\big(\partial\sigma^{k+1} - \sigma^k + \sigma^k\big) + \partial p\big(\sigma^{k+1}\big) = p\big(\sigma^k\big) + p\big(\partial\sigma^{k+1} - \sigma^k\big) + \partial p\big(\sigma^{k+1}\big).\end{aligned}$$

The geometric cone operator suggests how to proceed. If a $(k+1)$-simplex σ^k is part of the geometric cone of a k-simplex σ^k with respect to the point x, then σ^{k+1} is contained in a $(k+1)$-hyperplane containing x. Therefore, the geometric cone of σ^{k+1} is $(k+1)$-dimensional. When the geometric cone is viewed as a map from $(k+1)$-chains to $(k+2)$-chains, it follows that $p(\sigma^{k+1}) = \emptyset$. Thus, if σ^{k+1} is a term in $p(\sigma^k)$ for some σ^k, then $p(\sigma^{k+1}) = \emptyset$.

Since we chose to include σ^{k+1} as one of the terms in $p(\sigma^k)$, let us follow the intuition suggested by the geometric cone and define $p(\sigma^{k+1}) = \emptyset$. Then,

$$\sigma^{k+1} = p\big(\sigma^k\big) + p\big(\partial\sigma^{k+1} - \sigma^k\big) + \partial(\emptyset) = p\big(\sigma^k\big) + p\big(\partial\sigma^{k+1} - \sigma^k\big).$$

Rearranging, we obtain an expression for what $p(\sigma^k)$ needs to be,

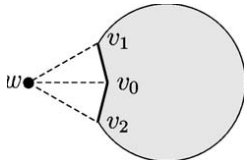
$$p\big(\sigma^k\big) = \sigma^{k+1} - p\big(\partial\sigma^{k+1} - \sigma^k\big).$$

To recap, we define $p(\sigma^k)$ by choosing $\sigma^{k+1} \succ \sigma^k$, such that σ^{k+1} and σ^k are consistently oriented. Then, we define,

$$p\big(\sigma^k\big) = \sigma^{k+1} - p\big(\partial\sigma^{k+1} - \sigma^k\big), \qquad p\big(\sigma^{k+1}\big) = \emptyset.$$

It remains to show that in defining $p(\sigma^k)$, we can order the definition of p on the simplices of the cone so that the simplices in the chain $\partial\sigma^{k+1} - \sigma^k$ already have p defined on it. If we can construct p in the above fashion so that it is well-defined, the homotopy property will automatically hold by construction. We will now define the generalized cone operator for 2- and 3-dimensions.

Definition 39. In 2-dimensions, the 1-ring condition implies that the base of the cone consists of either one or two 1-simplices. To aid in visualization, consider the following diagram,



One-ring cone augmentation of a complex in 2-dimensions

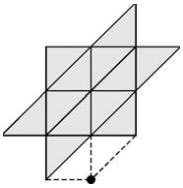
In the case of one 1-simplex, $[v_0, v_1]$, when we augment using the cone construction with the new vertex w, we define,

$$p\big([w]\big) = [v_0, w] + p\big([v_0]\big), \qquad p\big([v_0, w]\big) = \emptyset,$$
$$p\big([v_1, w]\big) = [v_0, v_1, w] - p\big([v_0, v_1]\big), \qquad p\big([v_0, v_1, w]\big) = \emptyset.$$

In the case of two 1-simplices, $[v_0, v_1]$, $[v_0, v_2]$, we have,

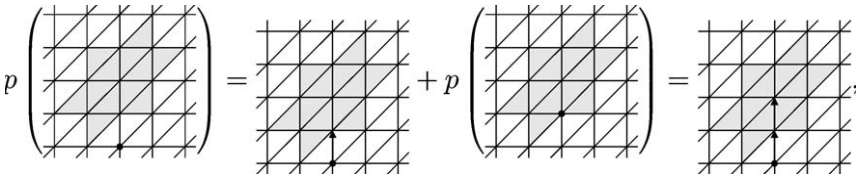
$$p\big([w]\big) = [v_0, w] + p\big([v_0]\big), \qquad p\big([v_0, w]\big) = \emptyset,$$
$$p\big([v_1, w]\big) = [v_0, v_1, w] - p\big([v_0, v_1]\big), \qquad p\big([v_0, v_1, w]\big) = \emptyset,$$
$$p\big([v_2, w]\big) = [v_0, v_2, w] - p\big([v_0, v_2]\big), \qquad p\big([v_0, v_2, w]\big) = \emptyset.$$

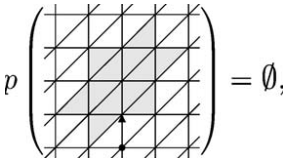
Example 40. We will now compute the generalized cone operator for part of a regular 2-dimensional triangulation that is not logically star-shaped. Consider a logically star-shaped complex, and augment with a new vertex.

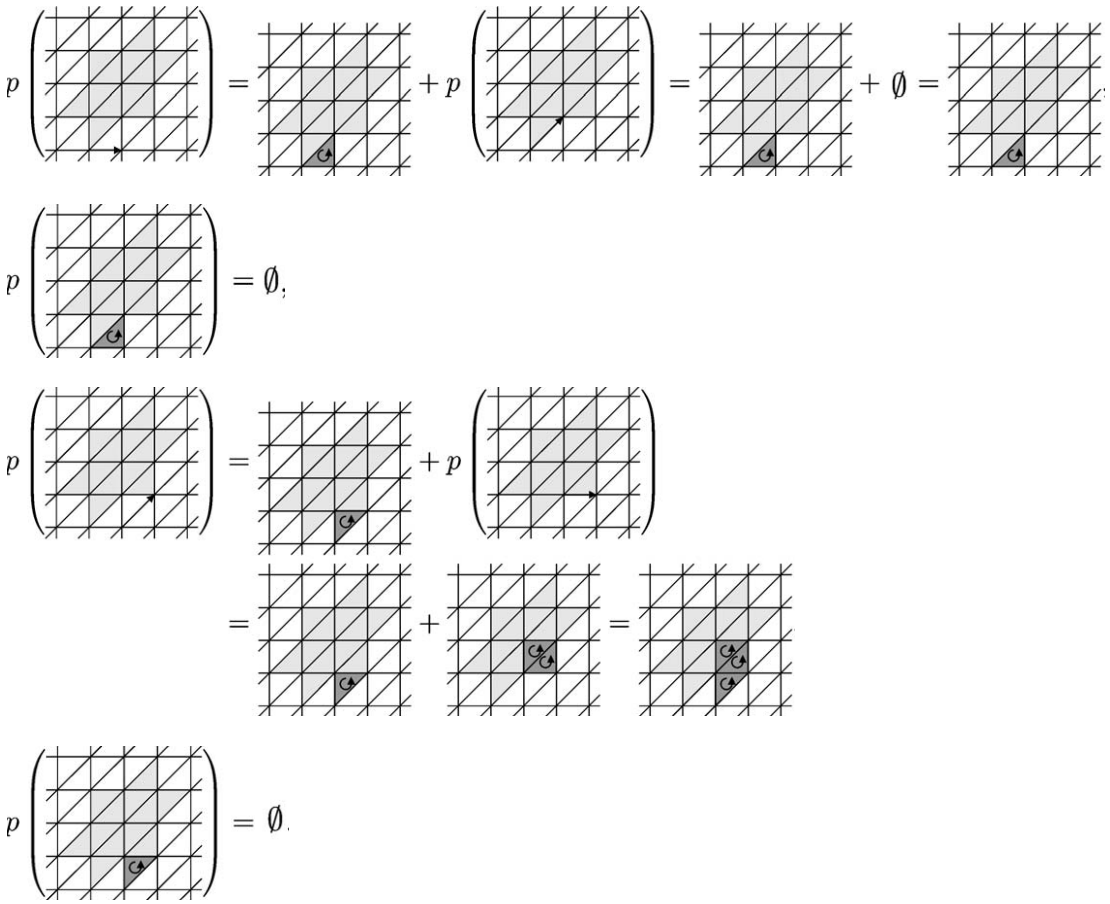


Logically star-shaped complex augmented by cone

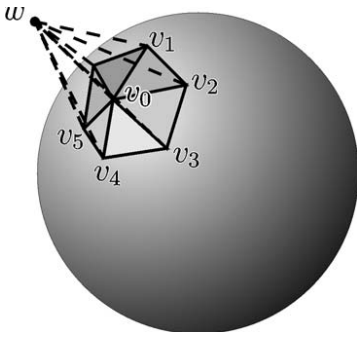
We use the logical cone operator for the subcomplex that is logically star-shaped, and the definition above for the newly introduced simplices. This yields,







Definition 41. We now define the generalized cone operator in 3-dimensions. Denote by v_0 the center of the 1-ring on the 2-surface, to which we are augmenting the new vertex w. The other vertices of the 1-ring are enumerated in order $v_1, \ldots, v_m$. To aid in visualization, consider the following diagram,



One-ring cone augmentation of a complex in 3-dimensions

If the 1-ring does not surround v_0, we denote the missing term by $[v_0, v_1, v_m]$.

$$
\begin{aligned}
k=0:\quad & p\big([w]\big) = [v_0, w] + p\big([v_0]\big), && p\big([v_0, w]\big) = \emptyset,\\
k=1:\quad & p\big([v_1, w]\big) = [v_0, v_1, w] - p\big([v_0, v_1]\big), && p\big([v_0, v_1, w]\big) = \emptyset,\\
& p\big([v_m, w]\big) = [v_0, v_m, w] - p\big([v_0, v_m]\big), && p\big([v_0, v_m, w]\big) = \emptyset,\\
k=2:\quad & p\big([v_1, v_2, w]\big) = [v_0, v_1, v_2, w] + p\big([v_0, v_1, v_2]\big), && p\big([v_0, v_1, v_2, w]\big) = \emptyset,\\
& p\big([v_{m-1}, v_m, w]\big) = [v_0, v_{m-1}, v_m, w] + p\big([v_0, v_{m-1}, v_m]\big), && p\big([v_0, v_{m-1}, v_m, w]\big) = \emptyset.
\end{aligned}
$$

If it does go around completely,

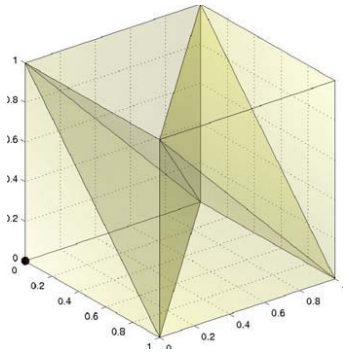
$$
p\big([v_m, v_1, w]\big) = [v_0, v_m, v_1, w] + p\big([v_0, v_m, v_1]\big), \qquad p\big([v_0, v_m, v_1, w]\big) = \emptyset.
$$

Example 42. We provide a tetrahedralization of the unit cube that can be tiled to yield a regular tetrahedralization of $\mathbb{R}^3$. The 3-simplices are as follows,

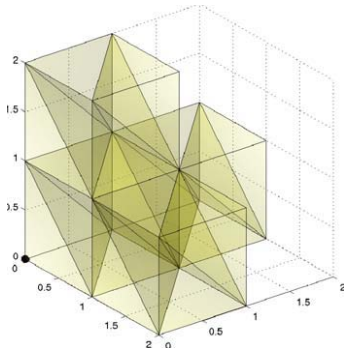
$$
[v_{000}, v_{001}, v_{010}, v_{10}], \qquad [v_{001}, v_{010}, v_{100}, v_{101}], \qquad [v_{001}, v_{010}, v_{011}, v_{101}],
$$

$$
[v_{010}, v_{100}, v_{101}, v_{110}], \qquad [v_{010}, v_{011}, v_{101}, v_{110}], \qquad [v_{011}, v_{101}, v_{110}, v_{111}].
$$

The tetrahedralization of the unit cube can be visualized as follows,



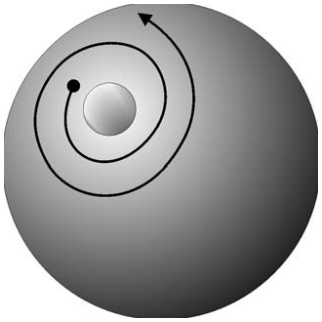
Tileable tetrahedralization of the unit cube



Partial tiling

Since this regular tetrahedralization can be constructed by the successive application of the one-ring cone augmentation procedure, the discrete Poincaré lemma can be extended to the entire regular tetrahedralization of $\mathbb{R}^3$.

In higher dimensions, we extend the construction of the generalized cone operator by choosing an appropriate enumeration of the base chain. The base chain is topologically the cone of S^{n-2} (with possibly an open $n-2$ ball removed) with respect to the central point.



Spiral enumeration of S^{n-2}, $n = 4$

By spiraling around S^{n-2}, starting from the boundary of the $n-2$ ball, and covering the rest of S^{n-2}, we obtain higher dimensional generalizations of Definitions 39, 41. Since $S^{2-2} = S^0$ is disjoint, $n = 2$ is distinguished, and we were unable to use spiral enumeration of the simplices in 2-dimensions.

The generalized cone operator is constructed so that the homotopy property holds automatically.

Lemma 43. *In generalized star-shaped complexes, the generalized cone operator satisfies the following identity,*

$$p\partial + \partial p = I,$$

at the level of chains.

Proof. By construction of the generalized cone operator. □

Lemma 44. *In generalized star-shaped complexes, the generalized cocone operator satisfies the following identity,*

$$H\mathbf{d} + \mathbf{d}H = I,$$

at the level of cochains.

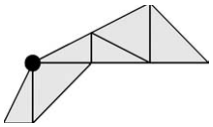
Proof. Follows immediately from applying the proof in the trivially star-shaped case, and using the identity in the previous lemma. □

We have a discrete Poincaré lemma for generalized star-shaped complexes.

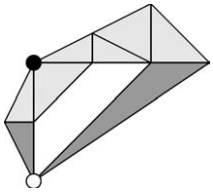
Corollary 45 (Discrete Poincaré lemma for generalized star-shaped complexes). *Given a closed cochain α^k, that is to say, $\mathbf{d}\alpha^k = 0$, there exists a cochain β^{k-1} such that $\mathbf{d}\beta^{k-1} = \alpha^k$.*

Proof. Follows from the above lemma using the proof for the trivially star-shaped case. □

Example 46. We will show how the Poincaré lemma fails when the complex is not contractible. We consider a trivially star-shaped complex, and augment by one vertex so as to make it non-contractible.



Trivially star-shaped complex

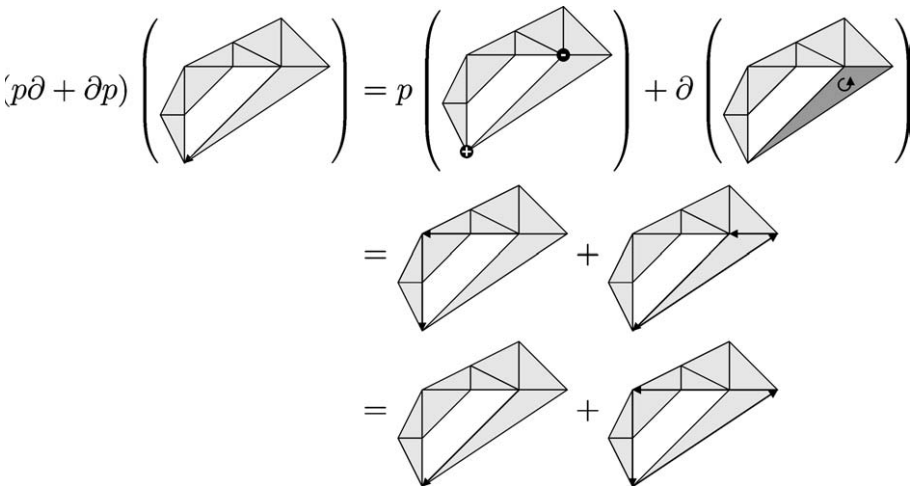


Non-contractible complex

When we attempt to verify the identity,

$$p\partial + \partial p = I,$$

we see that it is only true up to a chain that is homotopic to the inner boundary.



Since the second term is not the boundary of a 2-chain, it contributes a non-trivial term, even on closed discrete forms, and the Poincaré lemma breaks.

3. Conclusions

In summary, we have presented a constructive method of obtaining a local neighborhood in an unstructured mesh for which the discrete Poincaré lemma holds in the context of discrete exterior calculus. Furthermore, we introduced examples of regular space-filling triangulations and tetrahedralizations for which the exactness properties hold globally.

In the future, higher order analogues of the discrete theory of exterior calculus are desirable, but the cochain representation which assigns numerical quantities to a discrete set of geometric objects remains attractive due to its conceptual simplicity and the elegance of representing discrete operators as combinatorial operations on the mesh.

It is desirable to reconcile the two, by ensuring that higher-order interpolation and the combinatorial operations are consistent. This would yield more direct proofs of the exactness properties using the

standard cocone construction on the interpolated differential form, while giving a discrete homotopy operator that could be efficiently realized as a combinatorial operation on the mesh.

References

[1] R. Abraham, J.E. Marsden, T.S. Ratiu, Manifolds, Tensor Analysis and Applications, Springer, Berlin, 1988.
[2] D.H. Adams, R-torsion and linking numbers from simplicial Abelian gauge theories, hep-th/9612009.
[3] D.N. Arnold, Differential complexes and numerical stability, in: Proceedings of the International Congress of Mathematicians, Plenary Lectures, vol. I, Beijing, 2002.
[4] A. Bossavit, Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements, Academic Press Series in Electromagnetism, Academic Press, New York, 1998.
[5] M. Desbrun, A.N. Hirani, M. Leok, J.E. Marsden, Discrete exterior calculus, Preprint, 2003.
[6] A.A. Dezin, Multidimensional Analysis and Discrete Models, CRC Press, Boca Roton, FL, 1995.
[7] R. Hiptmair, Canonical construction of finite elements, Math. Comp. 68 (1999) 1325–1346.
[8] R. Hiptmair, Higher order Whitney forms, Progr. Electromagnetics Res. 32 (2001) 271–299.
[9] R. Hiptmair, Finite elements in computational electromagnetism, Acta Numer. (2002) 237–339.
[10] E.L. Mansfield, P.E. Hydon, A variational complex for difference equations, J. FoCM 4 (2) (2004) 187–217.
[11] J.R. Munkres, Elements of Algebraic Topology, Addison-Wesley, Reading, MA, 1984.
[12] S. Sen, S. Sen, J.C. Sexton, D.H. Adams, Geometric discretization scheme applied to the Abelian Chern–Simons theory, Phys. Rev. E 61 (2000) 3174–3185.
[13] K. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, IEEE Trans. Antennas Propagation 14 (1966) 302–307.

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 5521–5544 doi:10.1088/0305-4470/39/19/S12

Discrete Routh reduction

Sameer M Jalnapurkar[1], Melvin Leok[2], Jerrold E Marsden[3] and Matthew West[4]

[1] Department of Mathematics, Indian Institute of Science, Bangalore, India
[2] Department of Mathematics, University of Michigan, East Hall, 530 Church Street, Ann Arbor, MI 48109-1043, USA
[3] Control and Dynamical Systems 107-81, California Institute of Technology, Pasadena, CA 91125-8100, USA
[4] Department of Aeronautics and Astronautics, Stanford University, Stanford, CA 94305-4035, USA

E-mail: mleok@umich.edu

Received 24 August 2005, in final form 5 January 2006
Published 24 April 2006
Online at stacks.iop.org/JPhysA/39/5521

Abstract

This paper develops the theory of Abelian Routh reduction for discrete mechanical systems and applies it to the variational integration of mechanical systems with Abelian symmetry. The reduction of variational Runge–Kutta discretizations is considered, as well as the extent to which symmetry reduction and discretization commute. These reduced methods allow the direct simulation of dynamical features such as relative equilibria and relative periodic orbits that can be obscured or difficult to identify in the unreduced dynamics. The methods are demonstrated for the dynamics of an Earth orbiting satellite with a non-spherical J_2 correction, as well as the double spherical pendulum. The J_2 problem is interesting because in the unreduced picture, geometric phases inherent in the model and those due to numerical discretization can be hard to distinguish, but this issue does not appear in the reduced algorithm, where one can directly observe interesting dynamical structures in the reduced phase space (the cotangent bundle of shape space), in which the geometric phases have been removed. The main feature of the double spherical pendulum example is that it has a non-trivial magnetic term in its reduced symplectic form. Our method is still efficient as it can directly handle the essential non-canonical nature of the symplectic structure. In contrast, a traditional symplectic method for canonical systems could require repeated coordinate changes if one is evoking Darboux' theorem to transform the symplectic structure into canonical form, thereby incurring additional computational cost. Our method allows one to design reduced symplectic integrators in a natural way, despite the non-canonical nature of the symplectic structure.

PACS numbers: 02.40.Yy, 02.60.Cb, 45.10.−b
Mathematics Subject Classification: 37J15, 65L05, 70F15

1. Introduction

This paper addresses reduction theory for discrete mechanical systems with Abelian symmetry groups and its relation to variational integration. To establish the setting of the problem, a few aspects of the continuous theory are first recalled (see [29] for general background).

1.1. Continuous reduction theory

Consider a mechanical system with configuration manifold Q and a symmetry group G (with Lie algebra $\mathfrak{g}$) acting freely and properly on Q and hence, by cotangent lift on T^*Q, with the corresponding (standard, equivariant) momentum map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$. Recall from reduction theory (see [25, 36], and references therein) that, under appropriate regularity and nonsingularity conditions, the flow of a G-invariant Hamiltonian $H : T^*Q \to \mathbb{R}$ naturally induces a Hamiltonian flow on the reduced space $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$, where G_μ is the isotropy subgroup of a chosen point $\mu \in \mathfrak{g}^*$. In the Abelian case, if one chooses a connection $\mathfrak{A}$ on the principal bundle $Q \to Q/G$, then P_μ is symplectically isomorphic to $T^*(Q/G)$ carrying the canonical symplectic structure modified by magnetic terms, that is terms induced from the μ-component of the curvature of $\mathfrak{A}$.

The Lagrangian version of this theory is also well developed. In the Abelian case, it goes by the name of *Routh reduction* (see, for instance, [29], section 8.9). The reduced equations are again equations on $T(Q/G)$ and are obtained by dropping the variational principle, expressed in terms of the Routhian, from Q to Q/G. The non-Abelian version of this theory was originally developed in [32, 33], with important contributions and improvements given in [15, 30].

Of course, reduction has been enormously important for many topics in mechanics, such as stability and bifurcation of relative equilibria, integrable systems, etc. We need not review the importance of this process here as it is extensively documented in the literature.

1.2. Purpose, main results and examples

This paper presents the theory and illustrative numerical implementation for the reduction of discrete mechanical systems with Abelian symmetry groups. The discrete reduced space has a similar structure as in the continuous theory, but the curvature will be taken in a discrete sense. The paper studies two examples in detail, namely, satellite dynamics in the presence of the bulge of the Earth (the J_2 effect) and the double spherical pendulum (which has a non-trivial magnetic term). In each case the benefit of studying the numerics of the reduced problem is shown. Roughly, the reduced computations reveal dynamical structures that are hard to pick out in the unreduced dynamics in a way that is reminiscent of the phenomena of pattern evocation, as in [34, 35]. Another interesting application of the theory is that of orbiting multibody systems, studied in [41, 42].

We refer to [37] for a review of discrete mechanics, its numerical implementation, some history, as well as references to the literature. The value of geometric integrators has been documented in a number of references, such as [9]. In the present paper, we shall focus, to be specific, on discrete Euler–Lagrange and variational symplectic Runge–Kutta schemes and their reductions. One could, of course, use other schemes as well, such as Newmark, Störmer-Verlet or Shake schemes. However, we wish to emphasize that *without theoretical guidelines, coding algorithms for the reduced dynamics need not be a routine procedure since the reduced equations are not in canonical form because of non-trivial magnetic terms*. For example, using Darboux' theorem to put the structure into canonical form so that standard algorithms

can be used is not practical. We also remind the reader that there are real advantages to taking the variational approach to the construction of symplectic integrators. For example, as in [23], the variational approach provides the design flexibility to take different time steps at spatially different points in an asynchronous way and still retain all the advantages of symplecticity even though the algorithms are not strictly symplectic in the naive sense; such an approach is well known to be useful in molecular systems, for instance.

1.3. Motivation for discrete reduction

Besides its considerable theoretical interest, there are several practical reasons for carrying out discrete Routh reduction. These are as follows.

(i) Features that are clear in the reduced dynamics, such as relative equilibria and relative periodic orbits, can be obscured in the unreduced dynamics, and appear more complicated through the process of reconstruction and associated geometric phases. This is related to the phenomenon of *pattern evocation* that is an important practical feature of many examples, such as the double spherical pendulum [34, 35] and the stepping pendulum [12]. Going to a suitable (but non-obvious) rotating frame can 'evoke' such phenomena (see the movie at http://www.cds.caltech.edu/~marsden/research/demos/movies/Wendlandt/pattern.mpg). This is essentially a window to the reduced dynamics, which the theory in the present paper allows one to compute directly.

(ii) While directly studying the reduced dynamics can yield some benefits, it can be difficult to code using traditional methods. In particular, the presence of magnetic terms in the reduced symplectic form, as is the case with the double spherical pendulum, means that traditional symplectic methods for canonical systems do not directly apply; if one attempts to do so, it may result (and has in the literature) in many inefficient coordinate changes when evoking Darboux' theorem to put things into canonical form.

(iii) Although simulating the reduced dynamics involves an initial investment of time in computing geometric quantities symbolically, these additional terms do not appreciably affect the sparsity of the system of equations to be solved. As such, direct coding of the reduced algorithms can be quite efficient, due to its reduced dimensionality.

1.4. Two obvious generalizations

The free and proper assumption that we make on the group action means that we are dealing with *nonsingular*, that is, *regular* reduction (see [38] for the general theory of singular reduction and references to the literature). It would be interesting to extend the work here to the case of singular reduction but already the regular case is non-trivial and interesting. While our examples have singular points and the dynamics near these points is interesting, there is no attempt to study this aspect in the present paper.

Secondly, it would be interesting to generalize the present work to the case of non-Abelian groups and to develop a discrete version of non-Abelian Routh reduction (as in [15, 30]). We believe that such a generalization will require the further development of the theory of discrete connections, which is currently part of the research effort on discrete differential geometry (see [20], and references therein). Other future directions are discussed in the conclusions.

1.5. Other discrete reduction results

We briefly summarize some related results that have been obtained in the area of reduction for discrete mechanics. First of all, there is the important case of *discrete Euler–Poincaré and*

Lie–Poisson reduction that were obtained in [2, 27, 28]. This theory is appropriate for rigid body mechanics, for instance.

Another important case is that of *discrete semidirect product reduction* that was obtained in [3, 4] and applied to the case of the heavy top, with interesting links to discrete elastica. This case is of interest in the present study since with the heavy top, as with the general theory of semidirect product reduction (see [13, 31]), one can view the S^1 reduction of this problem as Routh reduction. Linking these two approaches is an interesting topic for future research.

1.6. Outline

After recalling the notation from continuous reduction theory, section 2 develops discrete reduction theory, derives a reduced variational principle and proves the symplecticity of the reduced flow. The relationship between continuous- and discrete-time reduction is also discussed. How the variational (and hence symplectic) Runge–Kutta algorithm induces a reduced algorithm in a natural way is shown in section 3. In section 4, we put together in a coherent way the main theoretical results of the paper up to that point. In section 5, the numerical example of satellite dynamics about an oblate Earth is given, and in section 6, the example of the double spherical pendulum, which has a non-trivial magnetic term, is given. Lastly, in section 7, we address some computational and efficiency issues.

2. Discrete reduction

In this section, it is assumed that the reader is familiar with continuous reduction theory as well as the theory of discrete mechanics; reference is made to the relevant parts of the literature as needed. It will be useful to first recall some facts about discrete mechanical systems with symmetry (see, for instance, [37] for proofs).

2.1. Discrete mechanical systems with symmetry

Let G be a Lie group (which shortly will be assumed to be Abelian) that acts freely and properly (on the left) on a configuration manifold Q. Given a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$ that is invariant under the diagonal action of G on $Q \times Q$, the corresponding *discrete momentum map* $\mathbf{J}_d : Q \times Q \to \mathfrak{g}^*$ is defined by

$$\mathbf{J}_d(q_0, q_1) \cdot \xi = D_2 L_d(q_0, q_1) \cdot \xi_Q(q_1), \tag{1}$$

where D_2 denotes the derivative in the second slot and where ξ_Q is the infinitesimal generator associated with $\xi \in \mathfrak{g}$. The map $\mathbf{J}_d$ is equivariant with respect to the diagonal action of G on $Q \times Q$ and the coadjoint action on $\mathfrak{g}^*$. The *discrete Noether theorem* states that the discrete momentum is conserved along solutions of the DEL (discrete Euler–Lagrange) equations,

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0. \tag{2}$$

Note that

$$\mathbf{J}_d(q_0, q_1) \cdot \xi = \mathbf{J}(D_2 L_d(q_0, q_1)) \cdot \xi,$$

where $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ is the momentum map on T^*Q; i.e., $\mathbf{J}_d = \mathbf{J} \circ \mathbb{F}L_d$, where $\mathbb{F}L_d = D_2 L_d : Q \times Q \to T^*Q$ is the discrete Legendre transform. Thus, for $\mu \in \mathfrak{g}^*$, we have $\mathbb{F}L_d\left(\mathbf{J}_d^{-1}(\mu)\right) \subset \mathbf{J}^{-1}(\mu)$. The symplectic algorithm (usually called the *position-momentum form of the algorithm*) obtained on T^*Q from that on $Q \times Q$ via the discrete Legendre transform thus preserves the standard momentum map $\mathbf{J}$.

There will be a standing assumption in this paper, namely that the given discrete Lagrangian L_d is *regular*; that is, for a point $(q, q) \in Q \times Q$ on the diagonal, the iterated derivative $D_2D_1L_d(q, q) : T_qQ \times T_qQ \to \mathbb{R}$ is a non-degenerate bilinear form. By the implicit function theorem, this implies that a point (q_{k-1}, q_k) near the diagonal and the DEL equations (2) uniquely determine the subsequent point q_{k+1} in a neighbourhood of the diagonal in $Q \times Q$ (or, if one prefers, for small time steps); in other words, the DEL algorithm is well defined by the DEL equations. Regularity also implies that the discrete Legendre transformation $\mathbb{F}L_d = D_2L_d : Q \times Q \to T^*Q$ is a local diffeomorphism from a neighbourhood of the diagonal in $Q \times Q$ to a neighbourhood in T^*Q. For a detailed discussion, see [37].

2.2. Reconstruction

The following lemma gives a basic result on the reconstruction of discrete curves in the configuration manifold Q from those in *shape space*, defined to be $S = Q/G$. The lemma is similar to its continuous counterpart, as in, for example, [15], lemma 2.2. The natural projection to the quotient will be denoted by $\pi_{Q,G} : Q \to Q/G$; $q \mapsto x = [q]_G$ (the equivalence class of $q \in Q$). Let $\mathrm{Ver}(q)$ denote the *vertical space* at q, namely the space of all vectors at the point q that are infinitesimal generators $\xi_Q(q) \in T_qQ$ or, in other words, the tangent space to the group orbit through q. We say that the discrete Lagrangian L_d is *group-regular* if the bilinear map $D_2D_1L_d(q, q) : T_qQ \times T_qQ \to \mathbb{R}$ restricted to the subspace $\mathrm{Ver}(q) \times \mathrm{Ver}(q)$ is nondegenerate. In addition to regularity, we shall make group-regularity a standing assumption in the paper. The following result is fundamental for what follows.

Lemma 2.1 (Reconstruction lemma)**.** *Fix $\mu \in \mathfrak{g}^*$ and let $x_0, x_1, \ldots, x_n$ be a sufficiently closely spaced discrete curve in S. Let $q_0, q_1 \in Q$ be such that $[q_0]_G = x_0$, $[q_1]_G = x_1$ and $\mathbf{J}_d(q_0, q_1) = \mu$. Then there is a unique closely spaced discrete curve $q_1, q_2, \ldots, q_n$ such that $[q_k]_G = x_k$ and $\mathbf{J}_d(q_{k-1}, q_k) = \mu$, for $k = 1, 2, \ldots, n$.*

Proof. We must construct a point q_2 close to q_1 such that $[q_2]_G = x_2$ and $\mathbf{J}_d(q_1, q_2) = \mu$; the construction of the subsequent points $q_3, \ldots, q_n$ then proceeds in a similar fashion.

To do this, pick a local trivialization of the bundle $\pi_{Q,G} : Q \to Q/G$, where locally $Q = S \times G$, and write points in this trivialization as $q_0 = (x_0, g_0)$ and $q_1 = (x_1, g_1)$, etc. Given the points $q_0 = (x_0, g_0), q_1 = (x_1, g_1)$ with $\mathbf{J}_d(q_0, q_1) = \mu$, we seek a near identity group element $k \in G$ such that $q_2 := (x_2, kg_1)$ satisfies $\mathbf{J}_d(q_1, q_2) = \mu$. According to equation (1), this means that we must satisfy the condition $D_2L_d(q_1, q_2) \cdot \xi_Q(q_2) = \langle \mu, \xi \rangle$ for all $\xi \in \mathfrak{g}$. In the local trivialization, this reads

$$D_2L_d((x_1, g_1), (x_2, kg_1)) \cdot \big(0, TR_{kg_1}\xi\big) = \langle \mu, \xi \rangle, \tag{3}$$

where R_g is right translation on G by g. Consider solving the equation

$$D_2L_d((\bar{x}_1, \bar{g}_1), (\bar{x}_2, k\bar{g}_1)) \cdot \big(0, TR_{k\bar{g}_1}\xi\big) = \langle \mu, \xi \rangle, \tag{4}$$

for k as a function of the variables $\bar{g}_1, \bar{x}_1, \bar{x}_2$ with μ fixed. By assumption, there is a solution for the case $\bar{x}_1 = x_0, \bar{x}_2 = x_1$ and $\bar{g}_1 = g_0$, namely $k = k_0 = g_1g_0^{-1}$ (a near identity group element). The implicit function theorem shows that when the point g_0, x_0, x_1 is replaced by the nearby point g_1, x_1, x_2, there will be a unique solution for k near k_0 provided that the derivative of the defining relation (3) with respect to k at the identity is invertible, which is true by group regularity. Since group regularity is G-invariant, the above argument remains valid as k_i drifts from the identity. □

Note that the above lemma makes no hypotheses about the sequences x_n or q_n satisfying any discrete evolution equations.

To carry out reconstruction in the continuous case, in addition to the requirements that the lifted curve in TQ lie on the momentum surface, and that it projects to the reduced curve $x(t) \in S$ under $\pi_{Q,G}$, one also requires that it be second order, which is to say that it is of the form $(q(t), \dot{q}(t))$. If a connection is given, then the lifted curve is obtained by integrating the *reconstruction equation*—again, see [15] for details. The discrete analogue of the second-order curve condition is explained as follows. Consider a given discrete curve as a sequence of points, $(x_0, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n)$, in $S \times S$. Lift each of the points in $S \times S$ to the momentum surface $\mathbf{J}_d^{-1}(\mu) \subset Q \times Q$. This yields the sequence, $\left(q_0^0, q_1^0\right), \left(q_0^1, q_1^1\right), \ldots, \left(q_0^{n-1}, q_1^{n-1}\right)$, which is unique up to an overall diagonal group action. The discrete analogue of the second-order curve condition is that this sequence in $Q \times Q$ defines a discrete curve in Q, which corresponds to requiring that $q_1^k = q_0^{k+1}$, for $k = 0, \ldots, n-1$, which is clearly possible in the context of the reconstruction lemma.

Discrete reconstruction naturally leads to issues of discrete geometric phases, and it would be interesting to express the discrete geometric phase in terms of the discrete curvature on shape space; this will surely involve some ideas from the currently evolving subject of discrete differential geometry and so we do not attempt to push this idea further at this point.

While many of the computations we present in this paper are in the setting of local trivializations, the results are valid globally through the construction given below.

2.3. Identification of the quotient space

Now assume that G is Abelian so that $G = G_\mu$ acts on $\mathbf{J}_d^{-1}(\mu)$ and that the quotient space $\mathbf{J}_d^{-1}(\mu)/G$ makes sense. We assume the above regularity hypotheses and freeness and properness of the action of G so that this quotient is a smooth manifold. It is clear that the map $\varphi_\mu : \mathbf{J}_d^{-1}(\mu)/G \to S \times S$ given by $[(q, q')]_G \to (x, x')$ is well defined, where the square brackets denote the equivalence class with respect to the given G action, and where $x = [q]_G$. The argument given in the reconstruction lemma shows that for a point $(q_0, q_1) \in \mathbf{J}_d^{-1}(\mu)$, φ_μ is a local diffeomorphism in a neighbourhood of the point $[(q_0, q_1)]_G$. In fact, the uniqueness part of that argument shows that for two nearby points (q_1, q_2) and (q_1', q_2') in $\mathbf{J}_d^{-1}(\mu)$, if $q_1 = g_1 q_1'$ and $q_2 = g_2 q_2'$, then $g_1 = g_2$. Thus, there is a neighbourhood U of a given a chain of closely spaced points lying in $\mathbf{J}_d^{-1}(\mu)$ with this property. Saturating this neighbourhood with the group action, we can assume that U is G-invariant. Restricted to U, φ_μ becomes a diffeomorphism to a neighbourhood of the diagonal of $S \times S$.

Assume, as above, that $L_d : Q \times Q \to \mathbb{R}$ is a discrete Lagrangian that is invariant under the action of an Abelian Lie group G on $Q \times Q$. In view of the preceding discussion, L_d restricted to $\mathbf{J}_d^{-1}(\mu)$ (and in the neighbourhood of a given chain of points in this set) induces a well-defined function $\hat{L}_d(x_0, x_1)$ of pairs of points (x_0, x_1) in $S \times S$. This *discrete reduced Lagrangian* will play an important role in what follows.

2.4. Discrete reduction

Let $\mathbf{q} := \{q_0, \ldots, q_n\}$ be a solution of the discrete Euler–Lagrange (DEL) equations. Let the value of the discrete momentum along this trajectory be μ. Let $x_i = [q_i]_G$, so that $\mathbf{x} := \{x_0, \ldots, x_n\}$ is a discrete shape space trajectory. Since $\mathbf{q}$ satisfies the discrete variational principle, it is appropriate to ask if there is a reduced variational principle satisfied by $\mathbf{x}$.

An important issue in dropping the discrete variational principle to the shape space is whether we require that the varied curves are constrained to lie on the level set of the momentum

map. The constrained approach is adopted in [15], and the unconstrained approach is used in [30]. In the rest of this section, we will adopt the unconstrained approach of [30] and show that the variations in the discrete action sum, when evaluated on a solution of the discrete Euler-Lagrange equations without assuming that the variations at the endpoints vanish, depends only on the quotient variations, and therefore drop to the shape space without constraints on the variations.

If $\mathbf{q}$ is a solution of the DEL equations, the interior terms in the variation of the discrete action sum vanish, leaving only the boundary terms; that is,

$$\delta \sum_{k=0}^{n-1} L_d(q_k, q_{k+1}) = D_1 L_d(q_0, q_1) \cdot \delta q_0 + D_2 L_d(q_{n-1}, q_n) \cdot \delta q_n. \tag{5}$$

Given a principal connection $\mathfrak{A}$ on Q, there is a horizontal–vertical split of each tangent space to Q denoted by $v_q = \operatorname{hor} v_q + \operatorname{ver} v_q$ for $v_q \in T_q Q$. Thus,

$$D_2 L_d(q_{n-1}, q_n) \cdot \delta q_n = D_2 L_d(q_{n-1}, q_n) \cdot \operatorname{hor} \delta q_n + D_2 L_d(q_{n-1}, q_n) \cdot \operatorname{ver} \delta q_n.$$

As in continuous Routh reduction, we will rewrite the terms involving vertical variations using the fact that we are on a level set of $\mathbf{J}_d$. Namely, write the vertical variation as $\operatorname{ver} \delta q_n = \xi_Q(q_n)$, where $\xi = \mathfrak{A}(\delta q_n)$ and use definition (1) of $\mathbf{J}_d$ to give

$$\begin{aligned} D_2 L_d(q_{n-1}, q_n) \cdot \operatorname{ver} \delta q_n &= D_2 L_d(q_{n-1}, q_n) \cdot \xi_Q(q_n) = \mathbf{J}_d(q_{n-1}, q_n) \cdot \xi \\ &= \langle \mu, \xi \rangle = \langle \mu, \mathfrak{A}(\delta q_n) \rangle = \mathfrak{A}_\mu(q_n) \cdot \delta q_n. \end{aligned} \tag{6}$$

Thus, the boundary terms can be expressed as

$$D_2 L_d(q_{n-1}, q_n) \cdot \delta q_n = D_2 L_d(q_{n-1}, q_n) \cdot \operatorname{hor} \delta q_n + \mathfrak{A}_\mu(q_n) \cdot \delta q_n, \tag{7}$$

$$D_1 L_d(q_0, q_1) \cdot \delta q_0 = D_1 L_d(q_0, q_1) \cdot \operatorname{hor} \delta q_0 - \mathfrak{A}_\mu(q_0) \cdot \delta q_0, \tag{8}$$

and so (5), the variation of the discrete action sum, becomes

$$\begin{aligned} \delta \sum_{k=0}^{n-1} L_d(q_k, q_{k+1}) &= D_1 L_d(q_0, q_1) \cdot \operatorname{hor} \delta q_0 + D_2 L_d(q_{n-1}, q_n) \cdot \operatorname{hor} \delta q_n \\ &\quad + \mathfrak{A}_\mu(q_n) \cdot \delta q_n - \mathfrak{A}_\mu(q_0) \cdot \delta q_0, \end{aligned} \tag{9}$$

when restricted to solutions of the discrete Euler–Lagrange equations.

Motivated by the preceding equation, introduce the 1-form $\mathcal{A}$ on $Q \times Q$ defined by

$$\mathcal{A} = \pi_2^* \mathfrak{A}_\mu - \pi_1^* \mathfrak{A}_\mu, \tag{10}$$

where $\pi_1, \pi_2 : Q \times Q \to Q$ are projections onto the first and the second components respectively. This allows us to expand the boundary terms involving $\mathfrak{A}_\mu$ into a telescoping sum, and rewrite (9) in terms of the 1-form $\mathcal{A}$ as

$$\sum_{k=0}^{n-1} (DL_d - \mathcal{A})(q_k, q_{k+1}) \cdot (\delta q_k, \delta q_{k+1}) = D_1 L_d(q_0, q_1) \cdot \operatorname{hor} \delta q_0 + D_2 L_d(q_{n-1}, q_n) \cdot \operatorname{hor} \delta q_n. \tag{11}$$

We now drop (11) to the reduced space $S \times S$. Consider the projection maps $\pi : Q \times Q \to (Q \times Q)/G$ and $\pi_{\mu,d} : \mathbf{J}_d^{-1}(\mu) \to \mathbf{J}_d^{-1}(\mu)/G$, and the inclusion maps $\iota_{\mu,d} : \mathbf{J}_d^{-1}(\mu) \hookrightarrow Q \times Q$ and $\tilde{\iota}_{\mu,d} : \mathbf{J}_d^{-1}(\mu)/G \hookrightarrow (Q \times Q)/G$. Then clearly, the following diagram commutes:

$$\begin{array}{ccc} \mathbf{J}_d^{-1}(\mu) & \xhookrightarrow{\iota_{\mu,d}} & Q \times Q \\ \downarrow{\scriptstyle \pi_{\mu,d}} & & \downarrow{\scriptstyle \pi} \\ \mathbf{J}_d^{-1}(\mu)/G & \xhookrightarrow{\tilde{\iota}_{\mu,d}} & (Q \times Q)/G \end{array}$$

By the G-invariance, L_d drops to a function $\tilde{L}_d$ on the quotient $(Q\times Q)/G$ so that $L_d = \tilde{L}_d \circ \pi$. The pullback (in this case, the restriction) of $\tilde{L}_d$ to $\mathbf{J}_d^{-1}(\mu)/G$ is called the *discrete reduced Lagrangian* and is denoted by as $\hat{L}_d$. Thus, $\hat{L}_d = \tilde{L}_d \circ \tilde{\imath}_{\mu,d}$; identifying $\mathbf{J}_d^{-1}(\mu)/G$ with $S \times S$ this definition agrees with $\hat{L}_d$ as defined earlier.

Lemma 2.2. *The 1-form $\mathcal{A}$ on $Q \times Q$ restricted to $\mathbf{J}_d^{-1}(\mu)$ drops to a 1-form $\hat{\mathcal{A}}$ on $\mathbf{J}_d^{-1}(\mu)/G$ and induces (for closely spaced points), via the map φ_μ, a 1-form that we denote by the same letter, on $S \times S$. Similarly, the 1-form $DL_d - \mathcal{A}$ on $Q \times Q$ restricted to the momentum level set $\mathbf{J}_d^{-1}(\mu)$ then drops to the 1-form $D\hat{L}_d - \hat{\mathcal{A}}$ on on $\mathbf{J}_d^{-1}(\mu)/G$ and induces, via the map φ_μ, a 1-form that we denote by the same letter, on $S \times S$.*

Proof. The equivariance of the projections π_i with respect to the diagonal action on $Q \times Q$ and the given action on Q, together with the invariance of the 1-form $\mathfrak{A}_\mu$ on Q, implies that $\mathcal{A}$ is invariant. Since $\mathfrak{A}_\mu$ vanishes on vertical vectors for the bundle $Q \to Q/G$, it follows that $\mathcal{A}$ vanishes on vertical vectors for the bundle $Q \times Q \to (Q \times Q)/G$. Therefore, there is a 1-form $\tilde{\mathcal{A}}$ on $(Q \times Q)/G$ such that $\mathcal{A} = \pi^*\tilde{\mathcal{A}}$.

Since $L_d = \tilde{L}_d \circ \pi$, and the exterior derivative commutes with pullback, it follows that $dL_d = \pi^* d\tilde{L}_d$. From $\pi \circ \iota_{\mu,d} = \tilde{\imath}_{\mu,d} \circ \pi_{\mu,d}$, we get $\iota_{\mu,d}^*\mathcal{A} = \iota_{\mu,d}^*\pi^*\tilde{A} = \pi_{\mu,d}^*\tilde{\imath}_{\mu,d}^*\tilde{\mathcal{A}}$. Thus, the 1-form $\mathcal{A}$ restricted to $\mathbf{J}_d^{-1}(\mu)$ drops to the 1-form $\hat{\mathcal{A}} = \tilde{\imath}_{\mu,d}^*\tilde{\mathcal{A}}$ on $\mathbf{J}_d^{-1}(\mu)/G$. Similarly, $\iota_{\mu,d}^* dL_d = \pi_{\mu,d}^*\tilde{\imath}_{\mu,d}^* d\tilde{L}_d$ and so dL_d restricted to $\mathbf{J}_d^{-1}(\mu)$ drops to the 1-form $\tilde{\imath}_{\mu,d}^* d\tilde{L}_d = d\hat{L}_d$ on $\mathbf{J}_d^{-1}(\mu)/G$. These 1-forms push forward under the map $\varphi_\mu : \mathbf{J}_d^{-1}(\mu)/G \to S \times S$ in the manner that was explained earlier. □

With the preceding lemma, and equation (11), we conclude that

$$\sum_{k=0}^{n-1}(D\hat{L}_d - \hat{\mathcal{A}})(x_k, x_{k+1}) \cdot (\delta x_k, \delta x_{k+1}) = D_1 L_d(q_0, q_1) \cdot \text{hor}\, \delta q_0 + D_2 L_d(q_{n-1}, q_n) \cdot \text{hor}\, \delta q_n. \tag{12}$$

Assuming that $\delta\mathbf{x}$ vanishes at the endpoints, hor $\delta q_0 = 0$, and hor $\delta q_1 = 0$ and consequently, the boundary terms vanish and we obtain the *reduced discrete variational principle*

$$\delta \sum_{k=0}^{n-1} \hat{L}_d(x_k, x_{k+1}) = \sum_{k=0}^{n-1} \hat{\mathcal{A}}(x_k, x_{k+1}) \cdot (\delta x_k, \delta x_{k+1}). \tag{13}$$

In an analogous fashion to rewriting $D\hat{L}_d(x_k, x_{k+1}) \cdot (\delta x_k, \delta x_{k+1})$ as $D_1\hat{L}_d(x_k, x_{k+1}) \cdot \delta x_k + D_2\hat{L}_d(x_k, x_{k+1}) \cdot \delta x_{k+1}$, we do the same for the $\hat{\mathcal{A}}$ term by defining

$$\hat{\mathcal{A}}(x_0, x_1) \cdot (\delta x_0, \delta x_1) = \hat{\mathcal{A}}_1(x_0, x_1) \cdot \delta x_0 + \hat{\mathcal{A}}_2(x_0, x_1) \cdot \delta x_1.$$

Then, equating terms involving δx_k on the left-hand side of (13) to the corresponding terms on the right-hand side, we get the *discrete Routh (DR) equations* giving dynamics on $S \times S$:

$$D_2\hat{L}_d(x_{k-1}, x_k) + D_1\hat{L}_d(x_k, x_{k+1}) = \hat{\mathcal{A}}_2(x_{k-1}, x_k) + \hat{\mathcal{A}}_1(x_k, x_{k+1}). \tag{14}$$

Note that these equations depend on the value of momentum μ. Thus, if $\mathbf{q}$ is a discrete curve satisfying the discrete Euler–Lagrange equations, the curve $\mathbf{x}$ obtained by projecting $\mathbf{q}$ down to S satisfies the DR equations (14).

Now consider the converse, the discrete reconstruction procedure: given a discrete curve $\mathbf{x}$ on S that satisfies the DR equations, is $\mathbf{x}$ the projection of a discrete curve $\mathbf{q}$ on Q that satisfies the DEL equations?

Let the pair (q_0, q_1) be a lift of (x_0, x_1) such that $\mathbf{J}_d(q_0, q_1) = \mu$. Let $\mathbf{q} = \{q_0, \ldots, q_n\}$ be the solution of the DEL equations with initial condition (q_0, q_1). Note that $\mathbf{q}$ has momentum

μ. Let $\mathbf{x}' = \{x'_0, \ldots, x'_n\}$ be the curve on S obtained by projecting $\mathbf{q}$. By our arguments above, $\mathbf{x}'$ solves the DR equations. However, $\mathbf{x}'$ has the initial condition (x_0, x_1), which is the same as the initial condition of $\mathbf{x}$. By uniqueness of the solutions of the DR equations, $\mathbf{x}' = \mathbf{x}$. Thus, $\mathbf{x}$ is the projection of a solution $\mathbf{q}$ of the DEL equations with momentum μ. Also, for a given initial condition q_0, there is a unique lift of $\mathbf{x}$ to a curve with momentum μ. Such a lift can be constructed using the method described in lemma 2.1. Thus, lifting $\mathbf{x}$ to a curve with momentum μ yields a solution of the discrete Euler–Lagrange equations, which projects down to $\mathbf{x}$.

We summarize the results of this section in the following theorem.

Theorem 2.3. *Let* $\mathbf{x}$ *be a discrete curve on* S*, and let* $\mathbf{q}$ *be a discrete curve on* Q *with momentum* μ *that is obtained by lifting* $\mathbf{x}$*. Then the following are equivalent.*

1. $\mathbf{q}$ *solves the DEL equations.*
2. $\mathbf{q}$ *is a solution of the discrete Hamilton's variational principle*

$$\delta \sum_{k=0}^{n-1} L_d(q_k, q_{k+1}) = 0$$

for all variations $\delta\mathbf{q}$ *of* $\mathbf{q}$ *that vanish at the endpoints.*

3. $\mathbf{x}$ *solves the DR equations*

$$D_2\hat{L}_d(x_{k-1}, x_k) + D_1\hat{L}_d(x_k, x_{k+1}) = \hat{\mathcal{A}}_2(x_{k-1}, x_k) + \hat{\mathcal{A}}_1(x_k, x_{k+1}).$$

4. $\mathbf{x}$ *is a solution of the reduced variational principle*

$$\delta \sum_{k=0}^{n-1} \hat{L}_d(x_k, x_{k+1}) = \sum_{k=0}^{n-1} \hat{\mathcal{A}}(x_k, x_{k+1}) \cdot (\delta x_k, \delta x_{k+1})$$

for all variations $\delta\mathbf{x}$ *of* $\mathbf{x}$ *that vanish at the endpoints.*

Note that for smooth group actions, the order of accuracy will be equal for the reduced and unreduced algorithms.

2.5. *Preservation of the reduced discrete symplectic form*

The DR equations define a discrete flow map $\hat{F}_k : S \times S \to S \times S$. We already know that the flow of the DEL equations preserves the symplectic form Ω_{L_d} on $Q \times Q$. In this section, we show that the reduced flow $\hat{F}_k$ preserves a *reduced* symplectic form $\Omega_{\mu,d}$ on $S \times S$, and that this reduced symplectic form is obtained by restricting Ω_{L_d} to $\mathbf{J}_d^{-1}(\mu)$ and then dropping to $S \times S$. In other words, $\pi_{\mu,d}^*\Omega_{\mu,d} = \iota_{\mu,d}^*\Omega_{L_d}$. The continuous analogue of this equation is $\pi_\mu^*\Omega_\mu = i_\mu^*\Omega_Q$.

Recall from continuous reduction theory on the Hamiltonian side that in the case of cotangent bundles, the projection $\pi_\mu : J^{-1}(\mu) \to T^*S$ can be defined as follows: if $\alpha_q \in J^{-1}(\mu)$, then the *momentum shift* $\alpha_q - \mathfrak{A}_\mu(q)$ annihilates all vertical tangent vectors at $q \in Q$, as shown by the following calculation:

$$\langle \alpha_q - \mathfrak{A}_\mu(q), \xi_Q(q) \rangle = J(\alpha_q) \cdot \xi - \langle \mu, \xi \rangle = \langle \mu, \xi \rangle - \langle \mu, \xi \rangle = 0.$$

Thus, $\alpha_q - \mathfrak{A}_\mu(q)$ induces an element of T_x^*S and $\pi_\mu(\alpha_q)$ is defined to be this element.

Let $\mathbb{F}' : \mathbf{J}_d^{-1}(\mu) \to J^{-1}(\mu)$ be the restriction of $\mathbb{F}L_d$ to $\mathbf{J}_d^{-1}(\mu)$. Thus $\mathbb{F}' \circ \iota_\mu = \iota_{\mu,d} \circ \mathbb{F}L_d$, where $\iota_\mu : J^{-1}(\mu) \to T^*Q$ and $\iota_{\mu,d} : \mathbf{J}_d^{-1}(\mu) \to Q \times Q$ are inclusions. Define the map $\hat{\mathbb{F}} : S \times S \to T^*S$ by $\hat{\mathbb{F}}(x_0, x_1) = D_2\hat{L}_d(x_0, x_1) - \hat{\mathcal{A}}_2(x_0, x_1)$. By equation (6) and lemma 2.2, this map is well defined. The map $\hat{\mathbb{F}}$ will play the role of a reduced discrete Legendre transform,

and in contrast to the continuous theory, the momentum shift appears in the map $\hat{\mathbb{F}}$, as opposed to the projection $\pi_{\mu,d}$. As in the continuous theory, $\Omega_{\mu,d}$ does involve magnetic terms.

Lemma 2.4. *The following diagram commutes:*

$$\begin{array}{ccc} \mathbf{J}_d^{-1}(\mu) & \xrightarrow{\mathbb{F}'} & J^{-1}(\mu) \\ \Big\downarrow{\scriptstyle \pi_{\mu,d}} & & \Big\downarrow{\scriptstyle \pi_\mu} \\ S \times S & \xrightarrow{\hat{\mathbb{F}}} & T^*S \end{array}$$

Proof. Let $(q_0, q_1) \in \mathbf{J}_d^{-1}(\mu)$. Thus $D_2L_d(q_0, q_1) \in J^{-1}(\mu)$, and

$$\pi_\mu(\mathbb{F}'(q_0, q_1)) = \pi_\mu(D_2L_d(q_0, q_1)).$$

As noted above, $\pi_\mu(D_2L_d(q_0, q_1))$ is the element of $T^*_{x_1}S$ determined by $(D_2L_d(q_0, q_1) - \mathfrak{A}_\mu(q_1))$. For $\delta q_1 \in T_{q_1}Q$, we have

$$\begin{aligned} \langle D_2L_d(q_0, q_1) - \mathfrak{A}_\mu(q_1), \delta q_1\rangle &= (DL_d - \mathcal{A})(q_0, q_1) \cdot (0, \delta q_1) \\ &= (D\hat{L}_d - \hat{\mathcal{A}})(x_0, x_1) \cdot (0, \delta x_1) \\ &= D_2\hat{L}_d(x_0, x_1) \cdot \delta x_1 - \hat{\mathcal{A}}_2(x_0, x_1) \cdot \delta x_1, \end{aligned}$$

where in the second equality, we used lemma 2.2. Thus,

$$\pi_\mu(D_2L_d(q_0, q_1)) = D_2\hat{L}_d(x_0, x_1) - \hat{\mathcal{A}}_2(x_0, x_1),$$

which means $\hat{\mathbb{F}} \circ \pi_{\mu,d} = \pi_\mu \circ \mathbb{F}'$. □

Theorem 2.5. *The flow of the DR equations preserves the symplectic form*

$$\Omega_{\mu,d} = \hat{\mathbb{F}}^*(\Omega_S - \pi^*_{T^*S,S}\beta_\mu),$$

where β_μ is the 2-form on S obtained by dropping $\mathbf{d}\mathfrak{A}_\mu$.

Furthermore, $\Omega_{\mu,d}$ can be obtained by dropping to $S \times S$ the restriction of Ω_{L_d} to $\mathbf{J}_d^{-1}(\mu)$. In other words, $\pi^*_{\mu,d}\Omega_{\mu,d} = \iota^*_{\mu,d}\Omega_{L_d}$.

Proof. We give an outline of the steps involved; the details are routine to fill in. The strategy is to first show that the restriction to $\mathbf{J}_d^{-1}(\mu)$ of the symplectic form Ω_{L_d} drops to a 2-form $\Omega_{\mu,d}$ on $S \times S$. The fact that the discrete flow on $Q \times Q$ preserves the symplectic form Ω_{L_d} is then used to show that the reduced flow preserves $\Omega_{\mu,d}$. The steps involved are as follows.

(i) Consider the 1-form Θ_{L_d} on $Q \times Q$ defined by $\Theta_{L_d}(q_0, q_1) \cdot (\delta q_0, \delta q_1) = D_2L_d(q_0, q_1) \cdot \delta q_1$. The 1-form Θ_{L_d} is G-invariant, and thus the Lie derivative $\mathcal{L}_{\xi_{Q\times Q}}\Theta_{L_d}$ is zero.

(ii) Since $\Omega_{L_d} = -\mathbf{d}\Theta_{L_d}$, Ω_{L_d} is G-invariant. If $\iota_{\mu,d} : \mathbf{J}_d^{-1}(\mu) \to Q \times Q$ is the inclusion, $\Theta'_{L_d} = \iota^*_{\mu,d}\Theta_{L_d}$ and $\Omega'_{L_d} = \iota^*_{\mu,d}\Omega_{L_d}$ are the restrictions of Θ_{L_d} and Ω_{L_d} respectively to $\mathbf{J}_d^{-1}(\mu)$. One checks that Θ'_{L_d} and Ω'_{L_d} are invariant under the action of G on $\mathbf{J}_d^{-1}(\mu)$.

(iii) If $\xi_{\mathbf{J}_d^{-1}(\mu)}$ is an infinitesimal generator on $\mathbf{J}_d^{-1}(\mu)$, then

$$\xi_{\mathbf{J}_d^{-1}(\mu)} \lrcorner \Omega'_{L_d} = -\xi_{\mathbf{J}_d^{-1}(\mu)} \lrcorner \mathbf{d}\Theta'_{L_d} = -\mathcal{L}_{\xi_{\mathbf{J}_d^{-1}(\mu)}}\Theta'_{L_d} + \mathbf{d}\xi_{\mathbf{J}_d^{-1}(\mu)} \lrcorner \Theta'_{L_d} = 0.$$

This follows from the G-invariance of Θ'_{L_d}, and the fact that $\Theta'_{L_d} \cdot \xi_{\mathbf{J}_d^{-1}(\mu)} = \langle \mu, \xi\rangle$.

(iv) By steps 2 and 3, the form Ω'_{L_d} drops to a reduced form $\Omega_{\mu,d}$ on $\mathbf{J}_d^{-1}(\mu)/G \approx S \times S$. Thus, if $\pi_{\mu,d} : \mathbf{J}_d^{-1}(\mu) \to S \times S$ is the projection, then $\pi^*_{\mu,d}\Omega_{\mu,d} = \Omega'_{L_d}$. Note that the closure of $\Omega_{\mu,d}$ follows from the fact that Ω'_{L_d} is closed, which in turn follows from the closure of Ω_{L_d} and the relation $\Omega'_{L_d} = \iota^*_{\mu,d}\Omega_{L_d}$.

(v) If $F_k : Q \times Q \to Q \times Q$ is the flow of the DEL equations, let F'_k be the restriction of this flow to $\mathbf{J}_d^{-1}(\mu)$. We know that F'_k drops to the flow $\hat{F}_k$ of the DR equations on $S \times S$. Since F_k preserves Ω_{L_d}, F'_k preserves Ω'_{L_d}. Using this, it can be shown that $\hat{F}_k$ preserves $\Omega_{\mu,d}$. Note that it is sufficient to show that $\pi^*_{\mu,d}(\hat{F}_k^*\Omega_{\mu,d}) = \pi^*_{\mu,d}\Omega_{\mu,d}$.

(vi) It now remains to compute a formula for the reduced form $\Omega_{\mu,d}$. Using lemma 2.4, it follows that

$$\begin{aligned}\pi^*_{\mu,d}\Omega_{\mu,d} &= \iota^*_{\mu,d}\Omega_{L_d} = \iota^*_{\mu,d}\mathbb{F}L_d^*\Omega_Q = (\mathbb{F}')^* i_\mu^*\Omega_Q \\ &= (\mathbb{F}')^*\pi_\mu^*(\Omega_S - \pi^*_{T^*S,S}\beta_\mu) = \pi^*_{\mu,d}\hat{\mathbb{F}}^*(\Omega_S - \pi^*_{T^*S,S}\beta_\mu).\end{aligned}$$

Thus $\pi^*_{\mu,d}\Omega_{\mu,d} = \pi^*_{\mu,d}\hat{\mathbb{F}}^*(\Omega_S - \pi^*_{T^*S,S}\beta_\mu)$, from which it follows that $\Omega_{\mu,d} = \hat{\mathbb{F}}^*(\Omega_S - \pi^*_{T^*S,S}\beta_\mu)$. Incidentally, this expression shows that $\Omega_{\mu,d}$ is nondegenerate provided the map $\hat{\mathbb{F}} = D_2\hat{L}_d - \hat{\mathcal{A}}_2$ is a local diffeomorphism. □

One can alternatively prove symplecticity of the reduced flow directly from the reduced variational principle—see section 2.3.4 of [21].

2.6. Relating discrete and continuous reduction

As shown in [37], if the discrete Lagrangian L_d approximates the Jacobi solution of the Hamilton–Jacobi equation, then the DEL equations give us an integration scheme for the EL equations. In our commutative diagrams we will denote the relationship between the EL and DEL equations by a dashed arrow as follows:

$$(TQ, EL) \dashrightarrow (Q \times Q, DEL).$$

This arrow can thus be read as 'the corresponding discretization'. By the continuous and discrete Noether theorems, we can restrict the flow of the EL and DEL equations to $J_L^{-1}(\mu)$ and $\mathbf{J}_d^{-1}(\mu)$, respectively. The flow on $J_L^{-1}(\mu)$ induces a reduced flow on $J_L^{-1}(\mu)/G \approx TS$, which is the flow of the Routh equations. Similarly, the discrete flow on $\mathbf{J}_d^{-1}(\mu)$ induces a reduced discrete flow on $\mathbf{J}_d^{-1}(\mu)/G \approx S \times S$, which is the flow of the discrete Routh equations. Since the DEL equations give us an integration algorithm for the EL equations, it follows that the DR equations give us an integration algorithm for the Routh equations.

To numerically integrate the Routh equations, we have two options.

(i) First solve the DEL equations to yield a discrete trajectory on Q, which can then be projected to a discrete trajectory on S.

(ii) Solve the DR equations to directly obtain a discrete trajectory on Q.

Either approach yields the same result, which we express by the commutative diagram:

$$\begin{array}{ccc}(J_L^{-1}(\mu), EL) & \dashrightarrow & (\mathbf{J}_d^{-1}(\mu), DEL) \\ \Big\downarrow{\scriptstyle \pi_{\mu,L}} & & \Big\downarrow{\scriptstyle \pi_{\mu,d}} \\ (TS, R) & \dashrightarrow & (S \times S, DR)\end{array} \tag{15}$$

3. Reduction of the symplectic Runge–Kutta algorithm

A well-studied class of numerical schemes for Hamiltonian and Lagrangian systems is the symplectic partitioned Runge–Kutta (SPRK) algorithms (see [10, 11] for history and details).

We will adopt a local trivialization to express the SPRK method in which the group action is addition. Given a connection $\mathfrak{A}$ on Q, it can be represented in local coordinates as $\mathfrak{A}(\theta, x)(\dot{\theta}, \dot{x}) = A(x)\dot{x} + \dot{\theta}$. By rewriting the symplectic partitioned Runge–Kutta algorithm in terms of this local trivialization, and using the local representation of the connection, we obtain the following algorithm on T^*S:

$$x_1 = x_0 + h \sum b_j \dot{X}_j \tag{16a}$$

$$s_1 = s_0 + h \sum_j \tilde{b}_j \dot{S}_j + \left[h \sum_j \left(\tilde{b}_j \mu \frac{\partial A}{\partial x}(X_j)\dot{X}_j \right) - (\mu A(x_1) - \mu A(x_0)) \right] \tag{16b}$$

$$X_i = x_0 + h \sum a_{ij} \dot{X}_j \tag{16c}$$

$$S_i = s_0 + h \sum_j \tilde{a}_{ij} \dot{S}_j + \left[h \sum_j \left(\tilde{a}_{ij} \mu \frac{\partial A}{\partial x}(X_j)\dot{X}_j \right) - (\mu A(X_i) - \mu A(x_0)) \right] \tag{16d}$$

$$S_j = \frac{\partial \hat{R}^\mu}{\partial \dot{x}}(X_j, \dot{X}_j) \tag{16e}$$

$$\dot{S}_j = \frac{\partial \hat{R}^\mu}{\partial x}(X_j, \dot{X}_j) - \mathrm{i}_{\dot{X}_j} \beta_\mu(X_j), \tag{16f}$$

This system of equations is called the *reduced symplectic partitioned Runge–Kutta (RSPRK)* algorithm. A detailed derivation can be found in section 2.5 of [21]. As we obtained this system by dropping the symplectic partitioned Runge–Kutta algorithm from $\mathbf{J}^{-1}(\mu)$ to T^*S, it follows that this algorithm preserves the reduced symplectic form $\Omega_\mu = \Omega_S - \pi^*_{T^*S,S}\beta_\mu$ on T^*S.

Since the SPRK algorithm is an integration algorithm for the Hamiltonian vector field X_H on T^*Q, the RSPRK algorithm is an integration algorithm for the reduced Hamiltonian vector field X_{H_μ} on T^*S. The relationship between the cotangent bundle reduction and the reduction of the SPRK algorithm can be represented by the following commutative diagram:

$$\begin{array}{ccc} (J^{-1}(\mu), X_H) & \dashrightarrow & (J^{-1}(\mu), SPRK) \\ \downarrow{\scriptstyle \pi_\mu} & & \downarrow{\scriptstyle \pi_\mu} \\ (T^*S, X_{H_\mu}) & \dashrightarrow & (T^*S, RSPRK) \end{array}$$

The dashed arrows here denote the corresponding discretization, as in (15). The SPRK algorithm can be obtained by pushing forward the DEL equations by the discrete Legendre transform. See, for example, [37]. By lemma 2.4, this implies that the RSPRK algorithm can be obtained by pushing forward the DR equations by the reduced discrete Legendre transform $\hat{\mathbb{F}} = D_2\hat{L}_d - \hat{\mathcal{A}}_2$. These relationships are shown in the following commutative diagram:

$$\begin{array}{ccc} (\mathbf{J}_d^{-1}(\mu), DEL) & \xrightarrow{\mathbb{F}L_d} & (J^{-1}(\mu), SPRK) \\ \downarrow{\scriptstyle \pi_{\mu,d}} & & \downarrow{\scriptstyle \pi_\mu} \\ (S \times S, DR) & \xrightarrow{\hat{\mathbb{F}}} & (T^*S, RSPRK) \end{array}$$

4. Putting everything together

The relationship between Routh reduction and cotangent bundle reduction can be represented by the following commutative diagram:

$$\begin{array}{ccc} (J_L^{-1}(\mu), EL) & \xrightarrow{\mathbb{F}L} & (J^{-1}(\mu), X_H) \\ \downarrow{\scriptstyle \pi_{\mu,L}} & & \downarrow{\scriptstyle \pi_\mu} \\ (TS, R) & \xrightarrow{\mathbb{F}\hat{R}^\mu} & (T^*S, X_{H_\mu}) \end{array}$$

We saw in section 2.6 that if L_d approximates the Jacobi solution of the Hamilton–Jacobi equation, discrete and continuous Routh reduction is related by the following diagram:

$$\begin{array}{ccc} (J_L^{-1}(\mu), EL) & \dashrightarrow & (J_d^{-1}(\mu), DEL) \\ \downarrow{\scriptstyle \pi_{\mu,L}} & & \downarrow{\scriptstyle \pi_{\mu,d}} \\ (TS, R) & \dashrightarrow & (S \times S, DR) \end{array}$$

The dashed arrows mean that the DEL equations are an integration algorithm for the EL equations, and that the DR equations are an integration algorithm for the Routh equations.

Pushing forward the DEL equation using the discrete Legendre transform $\mathbb{F}L_d$ yields the SPRK algorithm on T^*Q, which is an integration algorithm for X_H. This is depicted by

$$\begin{array}{ccc} (J_L^{-1}(\mu), EL) & \dashrightarrow & (J_d^{-1}(\mu), DEL) \\ \downarrow{\scriptstyle \mathbb{F}L} & & \downarrow{\scriptstyle \mathbb{F}L_d} \\ (J^{-1}(\mu), X_H) & \dashrightarrow & (J^{-1}(\mu), SPRK) \end{array}$$

The SPRK algorithm on $J^{-1}(\mu) \subset T^*Q$ induces the RSPRK algorithm on $J^{-1}(\mu)/G \approx T^*S$. As we saw in section 3, this reduction process is related to cotangent bundle reduction and to discrete Routh reduction as shown in the following diagram:

$$\begin{array}{ccccc} (J^{-1}(\mu), X_H) & \dashrightarrow & (J^{-1}(\mu), SPRK) & \xleftarrow{\mathbb{F}L_d} & (J_d^{-1}(\mu), DEL) \\ \downarrow{\scriptstyle \pi_\mu} & & \downarrow{\scriptstyle \pi_\mu} & & \downarrow{\scriptstyle \pi_{\mu,d}} \\ (T^*S, X_{H_\mu}) & \dashrightarrow & (T^*S, RSPRK) & \xleftarrow{\hat{\mathbb{F}}} & (S \times S, DR) \end{array}$$

Putting all the above commutative diagrams together into one diagram, we obtain figure 1.

5. Example: J_2 satellite dynamics

5.1. Configuration space and Lagrangian

An illustrative and important example of a system with an Abelian symmetry group is that of a single satellite in orbit about an oblate Earth. The general aspects and background for this problem are discussed in [40], and some interesting aspects of the geometry underlying it are discussed in [7].

The configuration manifold Q is $\mathbb{R}^3$, and the Lagrangian is

$$L(q, \dot{q}) = \tfrac{1}{2} M_s \|\dot{q}\|^2 - M_s V(q),$$

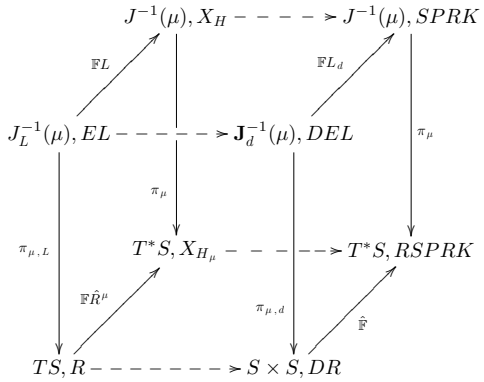


Figure 1. Complete commutative cube. The dashed arrows represent discretization from the continuous systems on the left face to the discrete systems on the right face. The vertical arrows represent reduction from the full systems on the top face to the reduced systems on the bottom face. The front and back faces represent Lagrangian and Hamiltonian viewpoints, respectively.

where M_s is the mass of the satellite and $V : \mathbb{R}^3 \to \mathbb{R}$ is the gravitational potential due to the Earth truncated at the first term in the expansion in the ellipticity

$$V(q) = \frac{GM_e}{\|q\|} + \frac{GM_eR_e^2J_2}{\|q\|^3}\left(\frac{3}{2}\frac{(q^3)^2}{\|q\|^2} - \frac{1}{2}\right).$$

Here, G is the gravitational constant, M_e is the mass of the Earth, R_e is the radius of the Earth, J_2 is a small non-dimensional parameter describing the degree of ellipticity and q^3 is the third component of q. In non-dimensional coordinates,

$$L(q, \dot{q}) = \frac{1}{2}\|\dot{q}\|^2 - \left[\frac{1}{\|q\|} + \frac{J_2}{\|q\|^3}\left(\frac{3}{2}\frac{(q^3)^2}{\|q\|^2} - \frac{1}{2}\right)\right]. \tag{17}$$

This corresponds to choosing space and time coordinates in which the radius of the Earth is 1 and the period of orbit at zero altitude is 2π when $J_2 = 0$ (spherical Earth).

5.2. Symmetry action

The symmetry of interest to us is that of rotation about the vertical (q^3) axis, so the symmetry group is the unit circle S^1. Using cylindrical coordinates $q = (r, \theta, z)$ for the configuration, the symmetry action is $\phi : (r, \theta, z) \mapsto (r, \theta + \phi, z)$. Since $\|q\|$, $\|\dot{q}\|$, and $q^3 = z$ are all invariant under this transformation, so too is the Lagrangian.

This action is clearly not free on all of $Q = \mathbb{R}^3$, as the z-axis is invariant for all group elements. This is not a serious obstacle as the lifted action is free on $T(Q\backslash(0, 0, 0))$ and this is enough to permit the application of the intrinsic Routh reduction theory. Alternatively, one can simply take $Q = \mathbb{R}^3\backslash\{(0, 0, z) \mid z \in \mathbb{R}\}$ and then the theory literally applies.

The shape space $S = Q/G$ is thus the half-plane $S = \mathbb{R}^+ \times \mathbb{R}$ and we will take coordinates (r, z) on S. In doing so, we are implicitly defining a global diffeomorphism $S \times G \to Q$ given by $((r, z), \theta) \mapsto (r, \theta, z)$.

The Lie algebra $\mathfrak{g}$ for $G = S^1$ is the real line $\mathfrak{g} = \mathbb{R}$, and we will identify the dual with the real line itself $\mathfrak{g}^* \cong \mathbb{R}$. For a Lie algebra element $\xi \in \mathfrak{g}$, the corresponding infinitesimal generator is given by $\xi_Q : (r, \theta, z) \mapsto ((r, \theta, z), (0, \xi, 0))$. The Lagrange momentum map $\mathbf{J}_L : TQ \to \mathfrak{g}^*$ is given by $\mathbf{J}_L(v_q) \cdot \xi = \langle \mathbb{F}L(v_q), \xi_Q(q)\rangle$, which in our case is a scalar quantity, the vertical component of the standard angular momentum $J_L((r, \theta, z), (\dot{r}, \dot{\theta}, \dot{z})) = r^2\dot{\theta}$.

Consider the Euclidean metric on $\mathbb{R}^3$, which corresponds to the kinetic energy norm in the Lagrangian. From this metric we define the mechanical connection $\mathfrak{A} : TQ \to \mathfrak{g}$ given by $\mathfrak{A}((r, \theta, z), (\dot{r}, \dot{\theta}, \dot{z})) = \dot{\theta}$. The 1-form $\mathfrak{A}_\mu$ on Q is thus given by $\mathfrak{A}_\mu = \mu\, \mathbf{d}\theta$. The exterior derivative of this expression gives $\mathbf{d}\mathfrak{A}_\mu = \mu\, \mathbf{d}^2\theta = 0$, and so the reduced 2-form is $\beta_\mu = 0$.

5.3. Equations of motion

The Euler–Lagrange equations for the Lagrangian (17) give the equations of motion,

$$\ddot{q} = -\nabla_q \left[\frac{1}{\|q\|} + \frac{J_2}{\|q\|^3} \left(\frac{3}{2} \frac{(q^3)^2}{\|q\|^2} - \frac{1}{2} \right) \right].$$

To calculate the reduced equations, we begin by calculating the Routhian

$$\begin{aligned} R^\mu(r, \theta, z, \dot{r}, \dot{\theta}, \dot{z}) &= L(r, \theta, z, \dot{r}, \dot{\theta}, \dot{z}) - \mathfrak{A}_\mu(r, \theta, z) \cdot (\dot{r}, \dot{\theta}, \dot{z}) \\ &= \frac{1}{2} \|(\dot{r}, \dot{\theta}, \dot{z})\|^2 - \left[\frac{1}{r} + \frac{J_2}{r^3} \left(\frac{3}{2} \frac{z^2}{r^2} - \frac{1}{2} \right) \right] - \mu\dot{\theta}. \end{aligned}$$

We choose a fixed value μ of the momentum and restrict ourselves to the space $J_L^{-1}(\mu)$, on which $\dot{\theta} = \mu$. The reduced Routhian $\hat{R}^\mu : TS \to \mathbb{R}$ is the restricted Routhian dropped to the tangent bundle of the shape space. In coordinates this is

$$\hat{R}^\mu(r, z, \dot{r}, \dot{z}) = \frac{1}{2} \|(\dot{r}, \dot{z})\|^2 - \left[\frac{1}{r} + \frac{J_2}{r^3} \left(\frac{3}{2} \frac{z^2}{r^2} - \frac{1}{2} \right) \right] - \frac{1}{2}\mu^2.$$

Recalling that $\beta_\mu = 0$, the Routh equations can be evaluated to give

$$(\ddot{r}, \ddot{z}) = -\nabla_{(r,z)} \left[\frac{1}{r} + \frac{J_2}{r^3} \left(\frac{3}{2} \frac{z^2}{r^2} - \frac{1}{2} \right) \right],$$

which describes the motion on the shape space.

To recover the unreduced Euler–Lagrange equations from the Routh equations one uses the procedure of reconstruction. This is covered in detail in [25, 26, 30].

5.4. Discrete Lagrangian system

We discretize this system with a high-order discrete Lagrangian. Recall that the pushforward discrete Lagrange map associated with this discrete Lagrangian is a symplectic partitioned Runge–Kutta method.

Given a point $(q_0, q_1) \in Q \times Q$, we will take (q_0, p_0) and (q_1, p_1) to be the associated discrete Legendre transforms. As the discrete momentum map is the pullback of the canonical momentum map, we have that $J_{L_d}(q_0, q_1) = (p_\theta)_0 = (p_\theta)_1$. Take a fixed momentum map value μ and restrict L_d to the set $J_{L_d}^{-1}(\mu)$. Dropping this to $S \times S$ now gives the reduced discrete Lagrangian $\hat{L}_d : S \times S \to \mathbb{R}$. More explicitly, L_d depends on (r_k, θ_k, z_k) and $(r_{k+1}, \theta_{k+1}, z_{k+1})$, but group invariance implies that the group variables only enter in the combination $(\theta_{k+1} - \theta_k)$. The condition $\mathbf{J}_d(q_k, q_{k+1}) = \mu$ can be inverted to eliminate the dependence of L_d on $(\theta_{k+1} - \theta_k)$, and L_d can be expressed in terms of μ, (r_k, z_k) and (r_{k+1}, z_{k+1}), which yields $\hat{L}_d$.

As discussed in section 3, the fact that we have taken coordinates in which the group action is addition in θ means that the pushforward discrete Lagrange map associated with the reduced discrete Lagrangian is the reduced method given by (16*a*)–(16*f*). In fact, as the mechanical connection has $A(r, z) = 0$ and $\beta_\mu = 0$, the pushforward discrete Lagrange map is exactly a partitioned Runge–Kutta method with Hamiltonian equal to the reduced Routhian. These are generically related by a momentum shift, rather than being equal.

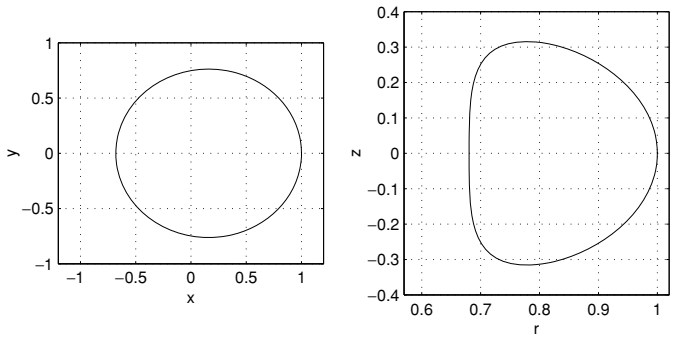


Figure 2. Unreduced (left) and reduced (right) views of an inclined elliptic trajectory for the continuous time system with $J_2 = 0$ (spherical Earth).

Given a trajectory of the reduced discrete system, we can reconstruct the unreduced discrete trajectory by solving for θ. Correspondingly, a trajectory of the unreduced discrete system can be projected onto shape space to give a trajectory of the reduced discrete system.

5.5. Example trajectories

We compute the unreduced trajectories using the fourth-order SPRK algorithm, and the reduced trajectories using the corresponding fourth-order RSPRK algorithm.

5.6. Solutions of the spherical Earth system

Consider initially the system with $J_2 = 0$. This corresponds to the case of a spherical Earth, and so the equations reduce to the standard Kepler problem. A slightly inclined circular trajectory is shown in figure 2, in both the unreduced and reduced pictures. Note that the graph of the reduced trajectory is a quadratic, as $\|q\| = \sqrt{r^2 + z^2}$ is a constant.

We will now investigate the effect of two different perturbations to the system, one due to taking non-zero J_2 and the other due to the numerical discretization.

5.7. The J_2 effect

Taking $J_2 = 0.05$ (which is close the actual value for the Earth), the system becomes near-integrable and experiences breakup of the KAM tori. This can be seen in figure 3, where the same initial condition is used as in figure 2.

Due to the fact that the reduced trajectory is no longer a simple curve, there is a geometric-phase-like effect which causes precession of the orbit. This precession can be seen in the thickening of the unreduced trajectory.

5.8. Solutions of the discrete system for a spherical Earth

We now consider the discrete system with $J_2 = 0$, for the second-order Gauss–Legendre discrete Lagrangian with the step size of $h = 0.3$. The trajectory with the same initial condition as above is given in figure 4.

As can be seen from the reduced trajectory, the discretization has caused a similar breakup of the periodic orbit as was produced by the non-zero J_2. This induces precession of the orbit in the unreduced trajectory, in a way which is difficult to distinguish from the perturbation

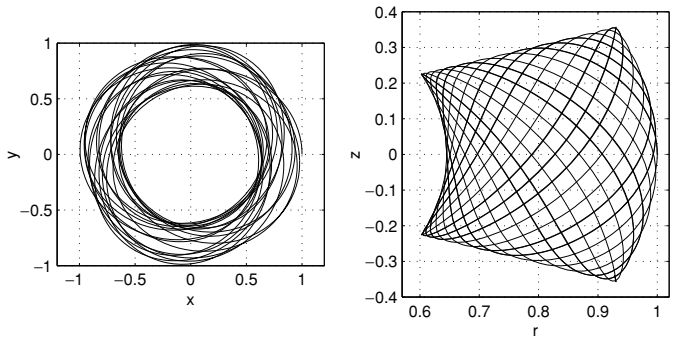


Figure 3. Unreduced (left) and reduced (right) views of an inclined elliptic trajectory for the continuous time system with $J_2 = 0.05$. Observe that the non-spherical terms introduce precession of the near-elliptic orbit in the symmetry direction.

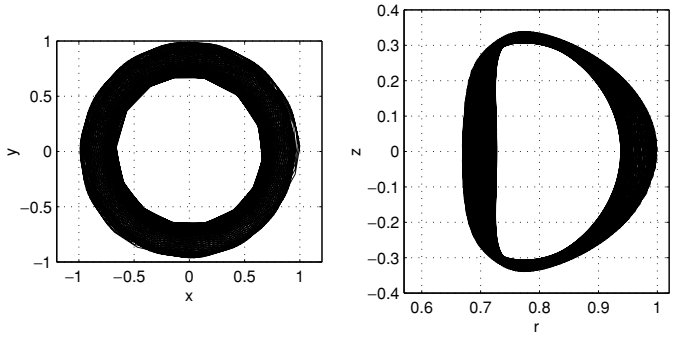


Figure 4. Unreduced (left) and reduced (right) views of an inclined trajectory of the discrete system with step size $h = 0.3$ and $J_2 = 0$. The initial condition is the same as that used in figure 2. The numerically introduced precession means that the unreduced picture looks similar to that of figure 3 with non-zero J_2, whereas only by considering the reduced picture we can see the correct resemblance to the $J_2 = 0$ case of figure 2.

above due to non-zero J_2 if only the unreduced picture is considered. If the reduced pictures are consulted, however, then it is immediately clear that the system is much closer to the continuous time system with $J_2 = 0$ than to the system with non-zero J_2.

5.9. Solutions of the discrete system with J_2 effect

Finally, we consider the discrete system with non-zero $J_2 = 0.05$. The resulting trajectory is shown in figure 5, and it is clearly not easy to determine from the unreduced picture whether the precession is due to the J_2 perturbation, the discretization, or some combination of the two.

Taking the reduced trajectories, however, immediately shows that this discrete time system is structurally much closer to the non-zero J_2 system than to the original $J_2 = 0$ system. This confusion arises because both the J_2 term and the discretization introduce perturbations which act in the symmetry direction.

While this system is sufficiently simple that one can run simulations with such small time steps that the discretization artefacts become negligible, this is certainly not generally possible. This example demonstrates how knowledge of the geometry of the system can be important in

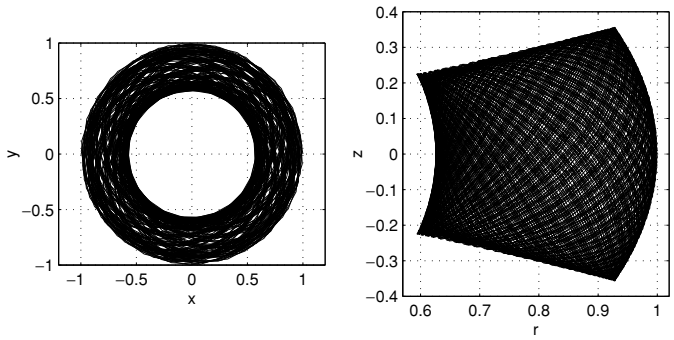


Figure 5. Unreduced (left) and reduced (right) views of an inclined trajectory of the discrete system with step size $h = 0.3$ and $J_2 = 0.05$. The initial condition is the same as that used in figure 3. The unreduced picture is similar to those of figures 3 and 4. By considering the reduced picture the correct resemblance to 3.

understanding the discretization process, and how this can give insight into the behaviour of numerical simulations. In particular, understanding how the discretization interacts with the symmetry action is extremely important.

5.10. Coordinate systems

In this example we have chosen cylindrical coordinates, thus making the group action addition in θ. One can always do this, as an Abelian Lie group is isomorphic to a product of copies of $\mathbb{R}$ and S^1, but it may sometimes be preferable to work in coordinates in which the group action is not addition. For example, Cartesian coordinates in the present example. Reasons for choosing a different coordinate system might include ease of computation, or simplicity of the expressions.

If we adopt a coordinate system wherein the group action is not expressed in terms of addition, the RSPRK method is not applicable, but we can still apply the Discrete Routh equations to obtain an integration scheme on $S \times S$. The push forward of this under $\hat{\mathbb{F}}$ yields an integration scheme on T^*S. The trajectories on the shape space that we obtain in this manner could be different from those we would get with the RSPRK method. However, in both cases we would have conservation of symplectic structure, momentum, and the order of accuracy would be the same. One could choose whichever approach is cheaper and easier.

6. Example: double spherical pendulum

6.1. Configuration space and Lagrangian

We consider the example of the double spherical pendulum which has a non-trivial magnetic term and constraints. The configuration manifold Q is $S^2 \times S^2$, and the embedding linear space V is $\mathbb{R}^3 \times \mathbb{R}^3$. The position vectors of each pendulum with respect to the pivot point are denoted by $\mathbf{q}_1$ and $\mathbf{q}_2$. These vectors are constrained to have lengths l_1 and l_2 respectively, and the pendula masses are denoted by m_1 and m_2. The Lagrangian is

$$L(\mathbf{q}_1, \mathbf{q}_2, \dot{\mathbf{q}}_1, \dot{\mathbf{q}}_2) = \tfrac{1}{2}m_1\|\dot{\mathbf{q}}_1\|^2 + \tfrac{1}{2}m_2\|\dot{\mathbf{q}}_1 + \dot{\mathbf{q}}_2\|^2 - m_1 g\mathbf{q}_1 \cdot \mathbf{k} - m_2 g(\mathbf{q}_1 + \mathbf{q}_2) \cdot \mathbf{k},$$

where g is the gravitational constant, and $\mathbf{k}$ is the unit vector in the z direction. The constraint function $c : V \to \mathbb{R}^2$ is given by $c(\mathbf{q}_1, \mathbf{q}_2) = (\|\mathbf{q}_1\| - l_1, \|\mathbf{q}_2\| - l_2)$. Using cylindrical

coordinates $\mathbf{q}_i = (r_i, \theta_i, z_i)$, L becomes

$$\begin{aligned} L(q,\dot{q}) = {} & \tfrac{1}{2}m_1\left(\dot{r}_1^2 + r_1^2\dot{\theta}_1^2 + \dot{z}_1^2\right) + \tfrac{1}{2}m_2\left\{\dot{r}_1^2 + r_1^2\dot{\theta}_1^2 + \dot{r}_2^2 + r_2^2\dot{\theta}_2^2\right. \\ & \left. + 2(\dot{r}_1\dot{r}_2 + r_1r_2\dot{\theta}_1\dot{\theta}_2)\cos\varphi + 2(r_1\dot{r}_2\dot{\theta}_1 - r_2\dot{r}_1\dot{\theta}_2)\sin\varphi + (\dot{z}_1+\dot{z}_2)^2\right\} \\ & - m_1gz_1 - m_2g(z_1+z_2), \end{aligned}$$

where $\varphi = \theta_2 - \theta_1$. Furthermore, we can automatically satisfy the constraints by performing the substitutions, $z_i = \left(l_i^2 - r_i^2\right)^{1/2}$, and $\dot{z}_i = -r_i\dot{r}_i\left(\sqrt{l_i^2 - r_i^2}\right)^{-1/2}$.

6.2. *Symmetry action*

The symmetry of interest to us is the simultaneous rotation of the two pendula about the vertical (z) axis, so the symmetry group is the unit circle S^1. Using cylindrical coordinates $\mathbf{q}_i = (r_i, \theta_i, z_i)$ for the configuration, the symmetry action is $\phi : (r_i, \theta_i, z_i) \mapsto (r_i, \theta_i + \phi, z_i)$. Since $\|\mathbf{q}_i\|$, $\|\dot{\mathbf{q}}_i\|$, $\|\dot{\mathbf{q}}_1 + \dot{\mathbf{q}}_2\|$ and $\mathbf{q}_i \cdot \mathbf{k}$ are all invariant under this transformation, so too is the Lagrangian.

This action is clearly not free on all of $V = \mathbb{R}^3 \times \mathbb{R}^3$, as the z-axis is invariant for all group elements. However, this does not pose a problem computationally, as long as the trajectories do not pass through the downward hanging configuration, corresponding to $r_1 = r_2 = 0$. To treat the downward handing configuration properly, we would need to develop a discrete Lagrangian analogue of the continuous theory of singular reduction described in [38].

We will now show the geometric structures involved in implementing the RSPRK algorithm. See section 2.10.2 of [21] for a detailed derivation. The Lie algebra $\mathfrak{g}$ for $G = S^1$ is the real line $\mathfrak{g} = \mathbb{R}$, and we will identify the dual with the real line itself $\mathfrak{g}^* \cong \mathbb{R}$. For a Lie algebra element $\xi \in \mathfrak{g}$, the corresponding infinitesimal generator is given by $\xi_Q : (r_1, \theta_1, z_1, r_2, \theta_2, z_2) \mapsto ((r_1, \theta_1, z_1, r_2, \theta_2, z_2), (0, \xi, 0, 0, \xi, 0))$. As such, the momentum map is given by

$$\begin{aligned} & J_L((r_1, \theta_1, z_1, r_2, \theta_2, z_2), (\dot{r}_1, \dot{\theta}_1, \dot{z}_1, \dot{r}_2, \dot{\theta}_2, \dot{z}_2)) \\ & \qquad = (m_1 + m_2)\, r_1^2\dot{\theta}_1 + m_2r_2^2\dot{\theta}_2 + m_2r_1r_2(\dot{\theta}_1 + \dot{\theta}_2)\cos\varphi + (r_1\dot{r}_2 - r_2\dot{r}_1)\sin\varphi, \end{aligned}$$

which is simply the vertical component of the standard angular momentum.

The locked inertia tensor is given by [25]

$$\mathbb{I}(\mathbf{q}_1\mathbf{q}_2) = m_1\left\|\mathbf{q}_1^{\perp}\right\|^2 + m_2\|(\mathbf{q}_1 + \mathbf{q}_2)^{\perp}\|^2 = m_1r_1^2 + m_2\left(r_1^2 + r_2^2 + 2r_1r_2\cos\varphi\right).$$

The mechanical connection as a 1-form is given by

$$\begin{aligned} \alpha(\mathbf{q}_1, \mathbf{q}_2) = {} & \left[m_1r_1^2 + m_2\left(r_1^2 + r_2^2 + 2r_1r_2\cos\varphi\right)\right]^{-1}\left[(m_1 + m_2)r_1^2\,\mathbf{d}\theta_1 + m_2r_2^2\,\mathbf{d}\theta_2\right. \\ & \left. + m_2r_1r_2(\mathbf{d}\theta_1 + \mathbf{d}\theta_2)\cos\varphi + (r_1\,\mathbf{d}r_2 - r_2\,\mathbf{d}r_1)\sin\varphi\right]. \end{aligned}$$

The μ-component of the mechanical connection is given by

$$\begin{aligned} \alpha_\mu(\mathbf{q}_1, \mathbf{q}_2) = {} & \mu\left[m_1r_1^2 + m_2\left(r_1^2 + r_2^2 + 2r_1r_2\cos\varphi\right)\right]^{-1} \\ & \times\left\{\left[(m_1 + m_2)r_1^2 + m_2r_1r_2\cos\varphi\right]\mathbf{d}\theta_1 + \left[m_2r_2^2 + m_2r_1r_2\cos\varphi\right]\mathbf{d}\theta_2\right\}. \end{aligned}$$

Taking the exterior derivative of this 1-form yields a non-trivial magnetic term on the reduced space, which drops to the quotient space to yield

$$\begin{aligned} \beta_\mu = {} & \mu m_2\left[2(m_1 + m_2)r_1r_2 + \left(m_1r_1^2 + m_2\left(r_1^2 + r_2^2\right)\right)\cos\varphi\right] \\ & \times\left[m_1r_1^2 + m_2\left(r_1^2 + r_2^2 + 2r_1r_2\cos\varphi\right)\right]^{-2}\mathbf{d}\varphi \wedge (r_2\,\mathbf{d}r_1 - r_1\,\mathbf{d}r_2). \end{aligned}$$

The local representation of the connection is given by

$$A(r_1, r_2, \varphi) = m_2\big(m_1 r_1^2 + m_2\big(r_1^2 + r_2^2 + 2r_1 r_2 \cos\varphi\big)\big)^{-1} \begin{bmatrix} -r_2 \sin\varphi \\ r_1 \sin\varphi \\ r_2^2 + r_1 r_2 \cos\varphi \end{bmatrix}^{\mathrm{T}},$$

and the amended potential V_μ has the form

$$V_\mu(q) = -m_1 g\big(l_1^2 - r_1^2\big)^{1/2} - m_2 g\big[\big(l_1^2 - r_1^2\big)^{1/2} + \big(l_2^2 - r_2^2\big)^{1/2}\big] + 2^{-1}\mu^2\big[m_1 r_1^2 + m_2\big(r_1^2 + r_2^2 + 2r_1 r_2 \cos\varphi\big)\big]^{-1}.$$

The Routhian on the momentum level set is given by $R^\mu = \frac{1}{2}\|\mathrm{hor}(q, v)\|^2 - V_\mu$. Recall that $\mathrm{hor}(v_q) = v_q - \xi_Q(v_q)$, where $\xi = \alpha(v_q)$ and $\xi_Q(v_q) = (0, \xi, 0, 0, \xi, 0)$. Then we obtain

$$\mathrm{hor}(v_q) = v_q - (0, \alpha(v_q), 0, 0, \alpha(v_q), 0) = (\dot r_1, \dot\theta_1 - \alpha(v_q), \dot z_1, \dot r_2, \dot\theta_2 - \alpha(v_q), \dot z_2).$$

The kinetic energy metric has the form

$$\begin{bmatrix} m_1 + m_2 & 0 & 0 & m_2\cos\varphi & -m_2 r_2 \sin\varphi & 0 \\ 0 & (m_1 + m_2) r_1^2 & 0 & m_2 r_1 \sin\varphi & m_2 r_1 r_2 \cos\varphi & 0 \\ 0 & 0 & m_1 + m_2 & 0 & 0 & 0 \\ m_2\cos\varphi & m_2 r_1 \sin\varphi & 0 & m_2 & 0 & 0 \\ -m_2 r_2 \sin\varphi & m_2 r_1 r_2 \cos\varphi & 0 & 0 & m_2 r_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & m_2 \end{bmatrix}.$$

This together with the expression for $\mathrm{hor}(v_q)$ allows us to compute $\frac{1}{2}\|\mathrm{hor}(q, v)\|^2$, and when combined with the formula for the amended potential V_μ gives the Routhian R^μ. Note that all our expressions are in terms of the reduced variables on TS, so they trivially drop to yield $\hat R^\mu$. These expressions can then be directly substituted into the RSPRK algorithm in equations (16*a*)–(16*f*) to obtain the example trajectories presented in the following subsection.

6.3. Example trajectories

We have computed the reduced trajectory of the double spherical pendulum using the fourth-order RSPRK algorithm on the Routh equations. We first consider the evolution of r_1, r_2 and φ using the RSPRK algorithm on the Routh equations, as well as the projection of the relative position of m_2 with respect to m_1 onto the xy plane as seen in figure 6.

Figure 7 illustrates that the energy behaviour of the trajectory is very good, as is typical of variational integrators, and does not exhibit a spurious drift. In contrast, the non-symplectic fourth-order Runge–Kutta applied to the unreduced dynamics has a systematic drift in the energy, even when using time steps that are smaller by a factor of 4.

7. Computational considerations

7.1. Reduced versus unreduced simulations

The reduced dynamics can either be computed directly, by using the discrete Routh or RSPRK equations, or by computing in the unreduced space, and projecting onto the shape space. We discuss the relative merits of these approaches.

Given a configuration space and symmetry group of dimensions n and m, respectively, we can either use a simpler algorithm in $2n$ dimensions, or a more geometrically involved algorithm in $2(n - m)$ dimensions. The reduced algorithm involves curvature terms that need to be symbolically precomputed, but these do not affect the sparsity of the system of equations,

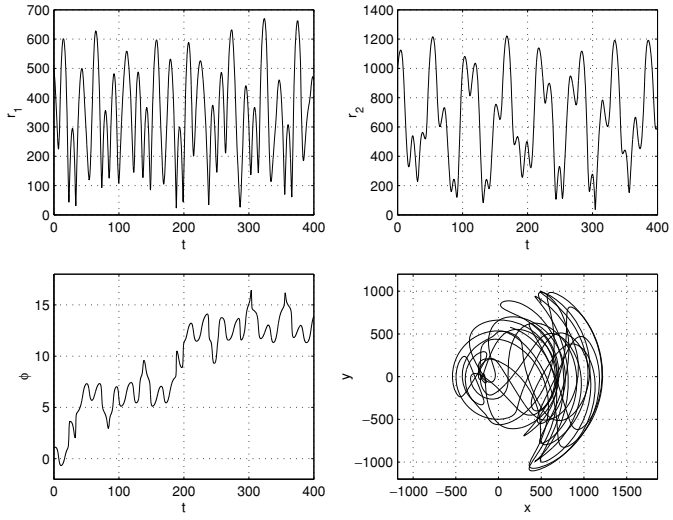


Figure 6. Time evolution of r_1, r_2, φ, and the trajectory of m_2 relative to m_1 using RSPRK.

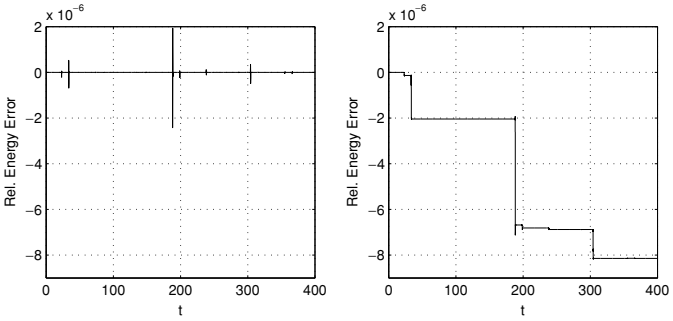


Figure 7. Relative energy drift $(E - E_0)/E_0$ using RSPRK (left) compared to the relative energy drift in a non-symplectic RK (right).

and the lower dimension results in computational saving that are particularly evident in the repeated, or long-time simulation of problems with a shape space of high codimension.

An example which is of current engineering interest is the dynamics of connected networks of systems with their own internal symmetries, such as coordinated clusters of satellites modelled as rigid bodies with internal rotors. If the systems to be connected are all identical, the geometric quantities that need to be computed, such as the mechanical connection, have a particularly simple repeated form, and the small additional upfront effort in implementing the reduced algorithm can result in substantial computational savings due to the lower dimension of the reduced system.

7.2. Intrinsic versus non-intrinsic methods

Non-intrinsic numerical schemes, such as SPRK applied to the classical Routh equations, can have undesirable numerical properties due to the need for coordinate-dependent local trivializations and the presence of coordinate singularities in these local trivializations, as is the case when using Euler angles for rigid body dynamics.

In the case of non-canonical symplectic forms, frequent computationally expensive coordinate changes are necessary when using standard non-intrinsic schemes, as documented in [39, 44], due to the need to repeatedly apply Darboux' theorem to put the symplectic structure into canonical form. In contrast, intrinsic methods do not depend on a particular choice of coordinate system, and allow for the use of global charts through the use of containing vector spaces with constraints enforced using Lagrange multipliers.

Coordinate singularities can affect the quality of the simulation in subtle ways that may depend on the numerical scheme. In the simulation of the double spherical pendulum, we notice spikes in the energy corresponding to times when r_1 or r_2 are close to 0. These errors accumulate in the non-symplectic method, but remain well behaved in the symplectic method. Alternatively, sharp spikes can be avoided altogether by evolving the equations on a constraint surface in $\mathbb{R}^3 \times \mathbb{R}^3$, as opposed to choosing local coordinates that automatically satisfy the constraints. Here, the increased cost of working in the containing linear space with constraints is offset by not having to transform between charts of $S^2_{l_1} \times S^2_{l_2}$, which can be significant if the trajectories are chaotic. An extensive discussion of the issue of representations and parametrizations of rotation groups and its implications for computation can be found in [5].

Methods which exhibit local conservation properties on each chart may still exhibit a drift in the conserved quantity across coordinate changes. As discussed in [1], only methods in which the local representatives of the algorithm commute with coordinate changes exhibit geometric conservation properties that are robust. In particular, intrinsic methods retain their conservation properties through coordinate changes.

Another promising approach is to use the exponential map to update the numerical solution, which is the basis of Lie group integrators, see [14]. The Lie group approach can yield rigid body integrators that are embedded in the space of 3×3 matrices but automatically evolve on the rotation group, without the use of constraints or reprojection. In [16, 17], analogues of the explicit Newmark and midpoint Lie algorithm are presented that automatically stay on the rotation group, and exhibit good energy and momentum stability. A method based on generating functions and the exponential on Riemannian manifolds is introduced in [19]. In the variational setting, a Lie group variational integrator for rigid body dynamics is described in [18], and higher-order generalizations can be found in [21].

8. Conclusions and future work

This paper derives the discrete Routh equations on the discrete shape space $S \times S$, which are symplectic with respect to a non-canonical symplectic form, and retains the good energy behaviour typically associated with variational integrators. Furthermore, when the group action can be expressed as addition, we obtain the reduced symplectic partitioned Runge–Kutta algorithm on T^*S, which can be considered as a discrete analogue of cotangent bundle reduction. By providing an understanding of how the reduced and unreduced formulations are related at a discrete level, we enable the user to freely choose whichever formulation is most appropriate, and provides the most insight to the problem at hand.

Certainly one of the obvious things to do in the future is to extend discrete reduction to the case of non-Abelian symmetry groups following the non-Abelian version of Routh reduction given in [15, 30]. There are also several problems, including the averaged J_2 problem, in which one can carry out discrete reduction by stages and relate it to the semidirect product work of [4]. This is motivated by the fact that the semidirect product reduction theory of [13] is a special case of reduction by stages (at least without the momentum map constraint), as was shown in [6].

In further developing discrete reduction theory, the discrete theory of connections on principal bundles developed in [22] is particularly relevant, as it provides an intrinsic method of representing the reduced space $(Q \times Q)/G$ as $(S \times S) \oplus \tilde{G}$, where $\tilde{G} = Q \times_G G$ with G acting by conjugation on G. Indeed, such a construction can be viewed as a connection on a Lie groupoid, and it is natural to express discrete mechanics on $Q \times Q$ in the language of pair groupoids, as originally proposed in [43]. Generalizations of this approach to arbitrary Lie groupoids, as well as a discussion of the role of discrete connections in yielding a discrete analogue of the Lagrange–Poincaré equations, can be found in [24].

Another component that is needed in this work is a good discrete version of the calculus of differential forms. Note that in our work we found, being directed by mechanics, that the right discrete version of the magnetic 2-form is the difference of two connection 1-forms. It is expected that we could recover such a magnetic 2-form by considering the discrete exterior derivative of a discrete connection form in a finite discretization of spacetime, and taking the continuum limit in the spatial discretization. Developing a discrete analogue of Stokes' theorem would also provide insight into the issue of discrete geometric phases. Some preliminary work on a discrete theory of exterior calculus can be found in [8].

Acknowledgments

We gratefully acknowledge helpful comments and suggestions of Alan Weinstein and the referees. SMJ was supported in part by ISRO and DRDO through the Nonlinear Studies Group, Indian Institute of Science, Bangalore. ML was supported in part by NSF Grant DMS-0504747, and a faculty grant and fellowship from the Rackham Graduate School, University of Michigan. JEM was supported in part by NSF ITR Grant ACI-0204932.

References

[1] Benettin G, Cherubini A M and Fassò F 2001 A changing-chart symplectic algorithm for rigid bodies and other Hamiltonian systems on manifolds *SIAM J. Sci. Comput.* **23** (4) 1189–203
[2] Bobenko A, Lorbeer B and Suris Y 1998 Integrable discretizations of the Euler top *J. Math. Phys.* **39** 6668–83
[3] Bobenko A and Suris Y 1999 Discrete lagrangian reduction, discrete Euler–Poincaré equations, and semidirect products *Lett. Math. Phys.* **49** 79–93
[4] Bobenko A I and Suris Y B 1999 Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top *Commun. Math. Phys.* **204** 147–88
[5] Borri B, Trainelli L and Bottasso C L 2000 On representations and parametrizations of motion *Multibody Syst. Dyn.* **4** 129–93
[6] Cendra H, Holm D, Marsden J and Ratiu T 1998 Lagrangian reduction, the Euler–Poincaré equations and semidirect products *Am. Math. Soc. Transl.* **186** 1–25
[7] Chang D E and Marsden J E 2003 Geometric derivation of the delaunay variables and geometric phases *Celest. Mech. Dyn. Astron.* **86** 185–208
[8] Desbrun M, Hirani A, Leok M and Marsden J 2003 Discrete exterior calculus *Preprint* math.DG/0508341
[9] Hairer E, Lubich C and Wanner G 2001 *Geometric Numerical Integration* (Berlin: Springer)
[10] Hairer E, Nørsett S and Wanner G 1993 *Solving Ordinary Differential Equations: I. Nonstiff Problems (Springer Series in Computational Mathematics* vol 8*)* 2nd edn (Berlin: Springer)
[11] Hairer E and Wanner G 1996 *Solving Ordinary Differential Equations: II. Stiff and Differential-Algebraic Problems (Springer Series in Computational Mathematics* vol 14*)* 2nd edn (Berlin: Springer)
[12] Holm D D and Lynch P 2002 Stepwise precession of the resonant swinging spring *SIAM J. Appl. Dyn. Syst.* **1** 44–64
[13] Holm D D, Marsden J E and Ratiu T S 1998 The Euler–Poincaré equations and semidirect products with applications to continuum theories *Adv. Math.* **137** 1–81
[14] Iserles A, Munthe-Kaas H, Nørsett S P and Zanna A 2000 Lie-group methods *Acta Numer.* **9** 215–365
[15] Jalnapurkar S and Marsden J 2000 Reduction of Hamilton's variational principle *Dyn. Stab. Syst.* **15** 287–318

[16] Krysl P 2005 Explicit momentum-conserving integrator for dynamics of rigid bodies approximating the midpoint Lie algorithm *Int. J. Numer. Methods Eng.* **63** (15) 2171–93
[17] Krysl P and Endres L 2005 Explicit Newmark/Verlet algorithm for time integration of the rotational dynamics of rigid bodies *Int. J. Numer. Methods Eng.* **62** (15) 2154–77
[18] Lee T Y, Leok M and McClamroch N H 2005 A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum *Proc. IEEE Conf. on Control Applications* pp 962–7
[19] Leimkuhler B and Patrick G 1996 A symplectic integrator for Riemannian manifolds *J. Nonlinear Sci.* **6** 367–84
[20] Leok M 2004 Foundations of computational geometric mechanics *Thesis* California Institute of Technology
[21] Leok M 2004 Generalized Galerkin variational integrators *Preprint* math.NA/0508360
[22] Leok M, Marsden J and Weinstein A 2004 A discrete theory of connections on principal bundles *Preprint* math.DG/0508338
[23] Lew A, Marsden J E, Ortiz M and West M 2003 Asynchronous variational integrators *Arch. Ration. Mech. Anal.* **167** 85–146
[24] Marrero J C, Martín de Diego D and Martínez E 2005 Discrete Lagrangian and Hamiltonian mechanics on Lie groupoids *Preprint* math.DG/0506299
[25] Marsden J 1992 *Lectures on Mechanics (London Mathematical Society Lecture Note Series* vol 174*)* (Cambridge: Cambridge University Press)
[26] Marsden J, Montgomery R and Ratiu T 1990 Reduction, symmetry, and phases in mechanics *Mem. Am. Math. Soc.* **88** 1–110
[27] Marsden J, Pekarsky S and Shkoller S 1999 Discrete Euler–Poincaré and Lie–Poisson equations *Nonlinearity* **12** 1647–62
[28] Marsden J, Pekarsky S and Shkoller S 2000 Symmetry reduction of discrete Lagrangian mechanics on Lie groups *J. Geom. Phys.* **36** 140–51
[29] Marsden J and Ratiu T 1999 *Introduction to Mechanics and Symmetry (Texts in Applied Mathematics* vol 17*)* 2nd edn (Berlin: Springer)
[30] Marsden J, Ratiu T and Scheurle J 2000 Reduction theory and the Lagrange–Routh equations *J. Math. Phys.* **41** 3379–429
[31] Marsden J E, Ratiu T S and Weinstein A 1984 Semi-direct products and reduction in mechanics *Trans. Am. Math. Soc.* **281** 147–77
[32] Marsden J and Scheurle J 1993a Lagrangian reduction and the double spherical pendulum *ZAMP* **44** 17–43
[33] Marsden J and Scheurle J 1993b The reduced Euler–Lagrange equations *Fields Inst. Commun.* **1** 139–64
[34] Marsden J E and Scheurle J 1995 Pattern evocation and geometric phases in mechanical systems with symmetry *Dyn. Stab. Syst.* **10** 315–38
[35] Marsden J E, Scheurle J and Wendlandt J 1996 Visualization of orbits and pattern evocation for the double spherical pendulum *ICIAM 95: Mathematical Research* ed K Kirchgässner, O Mahrenholtz and R Mennicken (Berlin: Academie Verlag) vol 87, pp 213–32
[36] Marsden J E and Weinstein A 1974 Reduction of symplectic manifolds with symmetry *Rep. Math. Phys.* **5** 121–30
[37] Marsden J and West M 2001 Discrete mechanics and variational integrators *Acta Numerica* vol 10 (Cambridge: Cambridge University Press)
[38] Ortega J-P and Ratiu T S 2004 *Momentum Maps and Hamiltonian Reduction (Progress in Mathematics* vol 222*)* (Boston, MA: Birkhäuser)
[39] Patrick G 1991 Two axially symmetric coupled rigid bodies: relative equilibria stability bifurcations and a momentum preserving symplectic integrator *PhD Thesis* University of California, Berkeley
[40] Prussing J and Conway B 1993 *Orbital Mechanics* (Oxford: Oxford University Press)
[41] Sanyal A, Bloch A J and McClamroch H N 2004 Dynamics of multibody systems in planar motion in a central gravitational field *Dyn. Syst. Int. J.* **19** (4) 303–43
[42] Sanyal A, Shen J and McClamroch H N 2003 Variational integrators for mechanical systems with cyclic generalized coordinates *Preprint*
[43] Weinstein A 1996 Lagrangian mechanics and groupoids *Mechanics Days (Waterloo, ON, 1992) (Fields Institute Communications* vol 7*)* (Providence, RI: American Mathematical Society) pp 207–31
[44] Wisdom J, Peale S and Mignard F 1984 The chaotic rotation of hyperion *Icarus* **58** 137–52

Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 196 (2007) 2907–2924
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Lie group variational integrators for the full body problem

Taeyoung Lee [a,*,1,2], Melvin Leok [b,1], N. Harris McClamroch [a,2]

[a] Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109, United States
[b] Department of Mathematics, Purdue University, West Lafayette, IN 47907, United States

Received 27 July 2005; accepted 11 January 2007

Abstract

We develop the equations of motion for full body models that describe the dynamics of rigid bodies, acting under their mutual gravity. The equations are derived using a variational approach where variations are defined on the Lie group of rigid body configurations. Both continuous equations of motion and variational integrators are developed in Lagrangian and Hamiltonian forms, and the reduction from the inertial frame to a relative frame is also carried out. The Lie group variational integrators are shown to be symplectic, to preserve conserved quantities, and to guarantee exact evolution on the configuration space. One of these variational integrators is used to simulate the dynamics of two rigid dumbbell bodies.

Keywords: Variational integrators; Lie group method; Full body problem

1. Introduction

1.1. Overview

The full body problem studies the dynamics of rigid bodies interacting under their mutual potential, and the mutual potential of distributed rigid bodies depends on both the position and the attitude of the bodies. Therefore, the translational and the rotational dynamics are coupled in the full body problem. The full body problem arises in numerous engineering and scientific fields. For example, in astrodynamics, the trajectory of a large spacecraft around the Earth is affected by the attitude of the spacecraft, and the dynamics of a binary asteroid pair is characterized by the non-spherical mass distributions of the bodies. In chemistry, the full rigid body model is used to study molecular dynamics. The importance of the full body problem is summarized in [1], along with a preliminary discussion from the point of view of geometric mechanics.

The full two body problem was studied by Maciejewski [2], and he presented equations of motion in inertial and relative frames and discussed the existence of relative equilibria in the system. However, he does not derive the equations of motion, nor does he discuss the reconstruction equations that allow the recovery of the inertial dynamics from the relative dynamics. Scheeres derived a stability condition for the full two body problem [3], and he studied the planar stability of an ellipsoid-sphere model [4]. Recently, interest in the full body problem has increased, as it is estimated that up to 16% of near-earth asteroids are binaries [5]. Spacecraft motion about binary asteroids have been discussed using the restricted three body model [6,7], and the four body model [8].

Conservation laws are important for studying the dynamics of the full body problem, because they describe qualitative characteristics of the system dynamics. The representation used for the attitude of the bodies should be globally defined since the complicated dynamics of such systems would require frequent coordinate changes when

* Corresponding author.
E-mail addresses: tylee@umich.edu (T. Lee), mleok@math.purdue.edu (M. Leok), nhm@engin.umich.edu (N.H. McClamroch).
[1] This research has been supported in part by a grant from the Rackham Graduate School, University of Michigan, and by NSF under grant DMS-0504747.
[2] This research has been supported in part by NSF under grant ECS-0244977 and CMS-0555797.

0045-7825/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2007.01.017

using representations that are only defined locally. General numerical integration methods, such as the Runge–Kutta scheme, do not preserve first integrals nor the exact geometry of the full body dynamics [9]. A more careful analysis of computational methods used to study the full body problem is crucial.

Variational integrators and Lie group methods provide a systematic method of constructing structure-preserving numerical integrators [9]. The idea of the variational approach is to discretize Hamilton's principle rather than the continuous equations of motion [10]. The numerical integrator obtained from the discrete Hamilton's principle exhibits excellent energy properties [11], conserves first integrals associated with symmetries by a discrete version of Noether's theorem, and preserves the symplectic structure. Many interesting differential equations, including full body dynamics, evolve on a Lie group. Lie group methods consist of numerical integrators that preserve the geometry of the configuration space by automatically remaining on the Lie group [12].

Moser and Vesselov [13], Wendlandt and Marsden [14] developed numerical integrators for a free rigid body by imposing an orthogonal constraint on the attitude variables and by using unit quaternions, respectively. The idea of using the Lie group structure and the exponential map to numerically compute rigid body dynamics arises in the work of Simo et al. [15], and in the work by Krysl [16]. A Lie group approach is explicitly adopted by Lee, Leok, and McClamroch in the context of a variational integrator for rigid body attitude dynamics with a potential dependent on the attitude, namely the 3D pendulum dynamics, in [17].

The motion of full rigid bodies depends essentially on the mutual gravitational potential, which in turn depends only on the relative positions and the relative attitudes of the bodies. Marsden et al. introduce discrete Euler–Poincaré and Lie–Poisson equations in [18,19]. They reduce the discrete dynamics on a Lie group to the dynamics on the corresponding Lie algebra. Sanyal, Shen and McClamroch develop variational integrators for mechanical systems with configuration dependent inertia and they perform discrete Routh reduction in [20]. A more intrinsic development of discrete Routh reduction is given in [21,22].

1.2. Contributions

The purpose of this paper is to provide a complete set of equations of motion for the full body problem. The equations of motion are categorized by three independent properties: continuous/discrete time, inertial/relative frame, and Lagrangian/Hamiltonian forms. Therefore, a total of eight types of equations of motion for the full body problem are given in this paper. The relationships between these equations of motion are shown in Fig. 1, and are further summarized in Fig. 7.

Continuous equations of motion for the full body problem are given in [2] without any formal derivation of the equations. We show, in this paper, that the equations of motion for the full body problem can be derived from Hamilton's principle. The proper form for the variations of Lie group elements in the configuration space lead to a systematic derivation of the equations of motion.

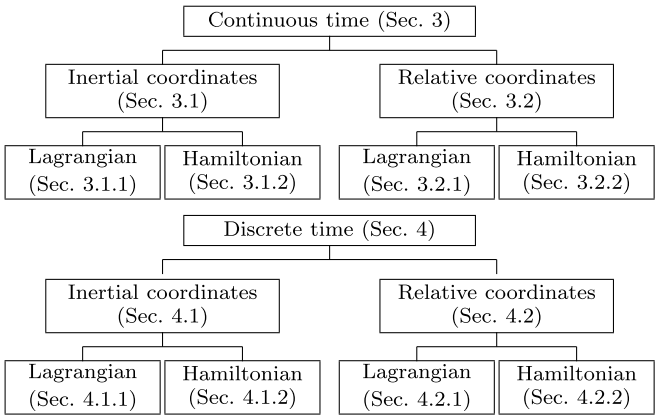


Fig. 1. Eight types of equations of motion for the full body problem.

This paper develops discrete variational equations of motion for the full body model following a similar variational approach but carried out within a discrete time framework. Since numerical integrators are derived from the discrete Hamilton's principle, they exhibit symplectic and momentum preserving properties, and good energy behavior. They are also constructed to conserve the Lie group geometry on the configuration space. Numerical simulation results for the interaction of two rigid dumbbell models are given.

This paper is organized as follows. The basic idea of the variational integrator is given in Section 2. The continuous equations of motion and variational integrators are developed in Section 3 and 4. Numerical simulation results are given in Section 5. An appendix contains several technical details that are utilized in the main development.

2. Background

2.1. Hamilton's principle and variational integrators

The procedure to derive the Euler–Lagrange equations and Hamilton's equations from Hamilton's principle is shown in Fig. 2. When deriving the equations of motion, we first choose generalized coordinates q, and a corresponding configuration space Q. We then construct a Lagrangian from the kinetic and the potential energy. An action integral $\mathfrak{G} = \int_{t_0}^{t_f} L(q,\dot{q})\,\mathrm{d}t$ is defined as the path integral of the Lagrangian along a time-parameterized trajectory. After taking the variation of the action integral, and requiring it to be stationary, we obtain the Euler–Lagrange equations. If we use the Legendre transformation defined as

$$p \cdot \delta\dot{q} = \mathbb{F}L(q,\dot{q}) \cdot \delta\dot{q} = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} L(q,\dot{q}+\epsilon\delta\dot{q}), \tag{1}$$

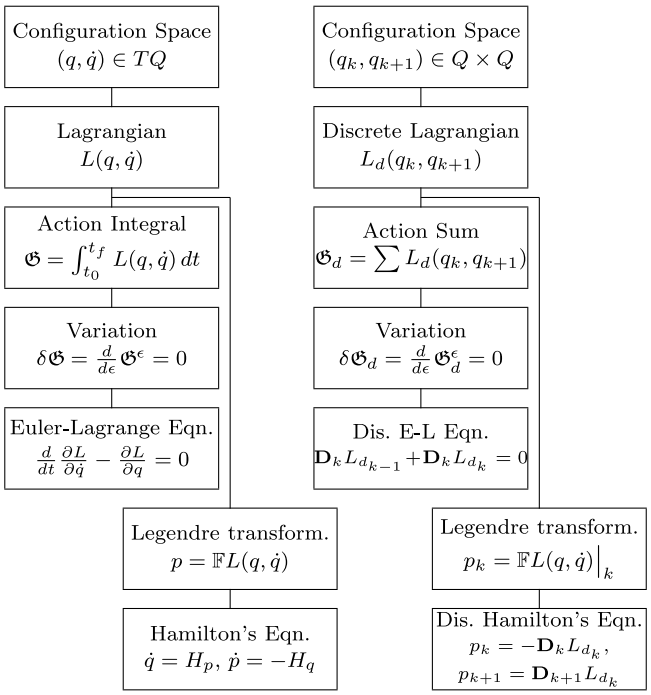


Fig. 2. Procedures to derive the continuous and discrete equations of motion.

where $\delta\dot{q} \in T_q Q$, then we obtain Hamilton's equations in terms of momenta variables $p = \mathbb{F}L(q, \dot{q}) \in T^*Q$. These equations are equivalent to the Euler–Lagrange equations [23].

There are two issues that arise in applying this procedure to the full body problem. The first is that the configuration space for each rigid body is the semi-direct product, $\mathbb{SE}(3) = \mathbb{R}^3 \circledS \mathbb{SO}(3)$, where SO(3) is the Lie group of orthogonal matrices with determinant 1. Therefore, variations should be carefully chosen such that they respect the geometry of the configuration space. For example, a varied rotation matrix $R^\epsilon \in \mathrm{SO}(3)$ can be expressed as

$$R^\epsilon = R\mathrm{e}^{\epsilon\eta},$$

where $\epsilon \in \mathbb{R}$, and $\eta \in \mathfrak{so}(3)$ is a variation represented by a skew-symmetric matrix. The variation of the rotation matrix δR is the part of R^ϵ that is linear in ϵ

$$R^\epsilon = R + \epsilon\delta R + \mathcal{O}(\epsilon^2).$$

More precisely, δR is given by

$$\delta R = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} R^\epsilon = R\eta. \tag{2}$$

The second issue is that reduced variables can be used to obtain equations of motion expressed in relative coordinates. The variations of the reduced variables are constrained as they must arise from the variations of the unreduced variables while respecting the geometry of the configuration space. The proper variations of Lie group elements and reduced quantities are computed while deriving the continuous equations of motion.

Generally, numerical integrators are obtained by approximating the continuous Euler–Lagrange equation using a finite difference rule such as $\dot{q}_k = (q_{k+1} - q_k)/h$, where q_k denotes the value of $q(t)$ at $t = hk$ for an integration step size $h \in \mathbb{R}$ and an integer k. A variational integrator is derived by following a procedure shown in the right column of Fig. 2. When deriving a variational integrator, the velocity phase space $(q, \dot{q}) \in TQ$ of the continuous Lagrangian is replaced by $(q_k, q_{k+1}) \in Q \times Q$, and the discrete Lagrangian L_d is chosen such that it approximates a segment of the action integral

$$L_d(q_k, q_{k+1}) \approx \int_0^h L(q_{k,k+1}(t), \dot{q}_{k,k+1}(t))\,\mathrm{d}t,$$

where $q_{k,k+1}(t)$ is the solution of the Euler–Lagrange equation satisfying boundary conditions $q_{k,k+1}(0) = q_k$ and $q_{k,k+1}(h) = q_{k+1}$. Then, the discrete action sum $\mathfrak{G}_d = \sum L_d(q_k, q_{k+1})$ approximates the action integral $\mathfrak{G}$. Taking the variations of the action sum, we obtain the discrete Euler–Lagrange equation

$$\mathbf{D}_{q_k} L_d(q_{k-1}, q_k) + \mathbf{D}_{q_k} L_d(q_k, q_{k+1}) = 0,$$

where $\mathbf{D}_{q_k} L_d$ denotes the partial derivative of L_d with respect to q_k. This yields a discrete Lagrangian map $F_{L_d} : (q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$. Using a discrete analogue of the Legendre transformation, referred to as a discrete fiber derivative $\mathbb{F}L_d : Q \times Q \to T^*Q$, variational integrators can be expressed in Hamiltonian form as

$$p_k = -\mathbf{D}_{q_k} L_d(q_k, q_{k+1}), \tag{3}$$

$$p_{k+1} = \mathbf{D}_{q_{k+1}} L_d(q_k, q_{k+1}). \tag{4}$$

This yields a discrete Hamiltonian map $\tilde{F}_{L_d} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$. A complete development of variational integrators can be found in [10].

2.2. Notation

Variables in the inertial and the body fixed frames are denoted by lower-case and capital letters, respectively. A subscript i is used for variables related to the ith body. The relative variables have no subscript and the kth discrete variables have the second level subscript k. The letters x, v, Ω and R are used to denote the position, velocity, angular velocity and rotation matrix, respectively.

The trace of $A \in \mathbb{R}^{n\times n}$ is denoted by

$$\mathrm{tr}[A] = \sum_{i=1}^{n} [A]_{ii},$$

where $[\mathrm{A}]_{ii}$ is the i, ith element of A. It can be shown that

$$\mathrm{tr}[AB] = \mathrm{tr}[BA] = \mathrm{tr}[B^{\mathrm{T}}A^{\mathrm{T}}] = \mathrm{tr}[A^{\mathrm{T}}B^{\mathrm{T}}], \tag{5}$$

$$\mathrm{tr}[A^{\mathrm{T}}B] = \sum_{p,q=1}^{3} [A]_{pq}[B]_{pq}, \tag{6}$$

$$\mathrm{tr}[PQ] = 0 \tag{7}$$

for matrices $A, B \in \mathbb{R}^{n\times n}$, a skew-symmetric matrix $P = -P^{\mathrm{T}} \in \mathbb{R}^{n\times n}$, and a symmetric matrix $Q = Q^{\mathrm{T}} \in \mathbb{R}^{n\times n}$.

The map $S : \mathbb{R}^3 \mapsto \mathbb{R}^{3\times 3}$ is defined by the condition that $S(x)y = x \times y$ for $x, y \in \mathbb{R}^3$. For $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, $S(x)$ is given by

$$S(x) = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}.$$

It can be shown that

$$S(x)^{\mathrm{T}} = -S(x), \tag{8}$$

$$S(x \times y) = S(x)S(y) - S(y)S(x) = yx^{\mathrm{T}} - xy^{\mathrm{T}}, \tag{9}$$

$$S(Rx) = RS(x)R^{\mathrm{T}}, \tag{10}$$

$$S(x)^{\mathrm{T}}S(x) = (x^{\mathrm{T}}x)I_{3\times 3} - xx^{\mathrm{T}} = \mathrm{tr}[xx^{\mathrm{T}}]I_{3\times 3} - xx^{\mathrm{T}} \tag{11}$$

for $x, y \in \mathbb{R}^3$ and $R \in \mathrm{SO}(3)$.

Using homogeneous coordinates, we can represent an element of SE(3) as follows:

$$\begin{bmatrix} R & x \\ 0 & 1 \end{bmatrix} \in \mathrm{SE}(3)$$

for $x \in \mathbb{R}^3$ and $R \in \mathrm{SO}(3)$. Then, an action on SE(3) is given by the usual matrix multiplication in $\mathbb{R}^{4\times 4}$. For example

$$\begin{bmatrix} R_1 & x_1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} R_2 & x_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_1R_2 & R_1x_2 + x_1 \\ 0 & 1 \end{bmatrix}$$

for $x_1, x_2 \in \mathbb{R}^3$ and $R_1, R_2 \in \mathrm{SO}(3)$.

The action of an element of SE(3) on $\mathbb{R}^3$ can be expressed using a matrix–vector product, once we identify $\mathbb{R}^3$ with $\mathbb{R}^3 \times \{1\} \subset \mathbb{R}^4$. In particular, we see from

$$\begin{bmatrix} R & x_1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_2 \\ 1 \end{bmatrix} = \begin{bmatrix} Rx_2 + x_1 \\ 1 \end{bmatrix}$$

that $x_2 \mapsto Rx_2 + x_1$.

3. Continuous time full body models

In this section, the continuous time equations of motion for the full body problem are derived in inertial and relative frames, and are expressed in both Lagrangian and in Hamiltonian form.

We define $O - e_1e_2e_3$ as an inertial frame, and $O_{\mathscr{B}_i} - E_{i_1}E_{i_2}E_{i_3}$ as a body fixed frame for the ith body, $\mathscr{B}_i$. The origin of the ith body fixed frame is located at the center of mass of body $\mathscr{B}_i$. The configuration space of the ith rigid body is $\mathrm{SE}(3) = \mathbb{R}^3 \circledS \mathrm{SO}(3)$. We denote the position of the center of mass of $\mathscr{B}_i$ in the inertial frame by $x_i \in \mathbb{R}^3$, and we denote the attitude of $\mathscr{B}_i$ by $R_i \in \mathrm{SO}(3)$, which is a rotation matrix from the ith body fixed frame to the inertial frame.

3.1. Inertial frame

Lagrangian: To derive the equations of motion, we first construct a Lagrangian for the full body problem. Given $(x_i, R_i) \in \mathrm{SE}(3)$, the inertial position of a mass element of $\mathscr{B}_i$ is given by $x_i + R_i\rho_i$, where $\rho_i \in \mathbb{R}^3$ denotes the position of the mass element in the body fixed frame. Then, the kinetic energy of $\mathscr{B}_i$ can be written as

$$T_i = \frac{1}{2}\int_{\mathscr{B}_i} \|\dot{x}_i + \dot{R}_i\rho_i\|^2 \mathrm{d}m_i.$$

Using the fact that $\int_{\mathscr{B}_i} \rho_i \mathrm{d}m_i = 0$ and the kinematic equation $\dot{R}_i = R_iS(\Omega_i)$, the kinetic energy T_i can be rewritten as

$$\begin{aligned} T_i(\dot{x}_i, \Omega_i) &= \frac{1}{2}\int_{\mathscr{B}_i} \|\dot{x}_i\|^2 + \|S(\Omega_i)\rho_i\|^2 \mathrm{d}m_i \\ &= \frac{1}{2}m_i\|\dot{x}_i\|^2 + \frac{1}{2}\mathrm{tr}[S(\Omega_i)J_{d_i}S(\Omega_i)^{\mathrm{T}}], \end{aligned} \tag{12}$$

where $m_i \in \mathbb{R}$ is the total mass of $\mathscr{B}_i$, $\Omega_i \in \mathbb{R}^3$ is the angular velocity of $\mathscr{B}_i$ in the body fixed frame, and $J_{d_i} = \int_{\mathscr{B}_i} \rho_i\rho_i^{\mathrm{T}} \mathrm{d}m_i \in \mathbb{R}^{3\times 3}$ is a nonstandard moment of inertia matrix. Using (11), it can be shown that the standard moment of inertia matrix $J_i = \int_{\mathscr{B}_i} S(\rho_i)^{\mathrm{T}}S(\rho_i)\mathrm{d}m_i \in \mathbb{R}^{3\times 3}$ is related to J_{d_i} by

$$J_i = \mathrm{tr}[J_{d_i}]I_{3\times 3} - J_{d_i},$$

and that the following condition holds for any $\Omega_i \in \mathbb{R}^3$

$$S(J_i\Omega_i) = S(\Omega_i)J_{d_i} + J_{d_i}S(\Omega_i). \tag{13}$$

We first derive equations using the nonstandard moment of inertia matrix J_{d_i}, and then express them in terms of the standard moment of inertia J_i by using (13).

The gravitational potential energy $U : \mathrm{SE}(3)^n \mapsto \mathbb{R}$ is given by

$$\begin{aligned} &U(x_1, x_2, \ldots, x_n, R_1, R_2, \ldots, R_n) \\ &\quad = -\frac{1}{2}\sum_{\substack{i,j=1 \\ i\neq j}}^{n} \int_{\mathscr{B}_i}\int_{\mathscr{B}_j} \frac{G\mathrm{d}m_i\mathrm{d}m_j}{\|x_i + R_i\rho_i - x_j - R_j\rho_j\|}, \end{aligned} \tag{14}$$

where G is the universal gravitational constant.

Then, the Lagrangian for n full bodies can be written as

$$L(\mathbf{x}, \dot{\mathbf{x}}, \mathbf{R}, \mathbf{\Omega}) = \sum_{i=1}^{n} \frac{1}{2}m_i\|\dot{x}_i\|^2 + \frac{1}{2}\mathrm{tr}[S(\Omega_i)J_{d_i}S(\Omega_i)^{\mathrm{T}}] - U(\mathbf{x}, \mathbf{R}), \tag{15}$$

where bold type letters denote ordered n-tuples of variables. For example, $\mathbf{x} \in (\mathbb{R}^3)^n$, $\mathbf{R} \in \mathrm{SO}(3)^n$, and $\mathbf{\Omega} \in (\mathbb{R}^3)^n$ are defined as $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{R} = (R_1, R_2, \ldots, R_n)$, and $\mathbf{\Omega} = (\Omega_1, \Omega_2, \ldots, \Omega_n)$, respectively.

Variations of variables: Since the configuration space is $\mathrm{SE}(3)^n$, the variations should be carefully chosen such that they respect the geometry of the configuration space. The variations of $x_i : [t_0, t_f] \mapsto \mathbb{R}^3$ and $\dot{x}_i : [t_0, t_f] \mapsto \mathbb{R}^3$ are trivial, namely

$$x_i^{\epsilon} = x_i + \epsilon\delta x_i + \mathscr{O}(\epsilon^2),$$
$$\dot{x}_i^{\epsilon} = \dot{x}_i + \epsilon\delta\dot{x}_i + \mathscr{O}(\epsilon^2),$$

where $\delta x_i : [t_0, t_f] \mapsto \mathbb{R}^3$, $\delta \dot{x}_i : [t_0, t_f] \mapsto \mathbb{R}^3$ vanish at the initial time t_0 and at the final time t_f. The variation of $R_i : [t_0, t_f] \mapsto \mathrm{SO}(3)$, as given in (2), is

$$\delta R_i = R_i \eta_i, \tag{16}$$

where $\eta_i : [t_0, t_f] \mapsto \mathfrak{so}(3)$ is a variation with values represented by a skew-symmetric matrix ($\eta_i^{\mathrm{T}} = -\eta_i$) vanishing at t_0 and t_f. The variation of Ω_i can be computed from the kinematic equation $\dot{R}_i = R_i S(\Omega_i)$ and (16) to be

$$\begin{aligned} S(\delta\Omega_i) &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} R_i^{\epsilon T} \dot{R}_i^{\epsilon} = \delta R_i^{\mathrm{T}} \dot{R}_i + R_i^{\mathrm{T}} \delta \dot{R}_i \\ &= -\eta_i S(\Omega_i) + S(\Omega_i)\eta_i + \dot{\eta}_i. \end{aligned} \tag{17}$$

3.1.1. Equations of motion: Lagrangian form

If we take variations of the Lagrangian using (16) and (17), we obtain the equations of motion from Hamilton's principle. We first take the variation of the kinetic energy of $\mathscr{B}_i$

$$\begin{aligned} \delta T_i &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} T_i(\dot{x}_i + \epsilon\delta\dot{x}_i, \Omega_i + \epsilon\delta\Omega_i) \\ &= m_i \dot{x}_i^{\mathrm{T}} \delta\dot{x}_i + \frac{1}{2}\mathrm{tr}[S(\delta\Omega_i) J_{d_i} S(\Omega_i)^{\mathrm{T}} + S(\Omega_i) J_{d_i} S(\delta\Omega_i)^{\mathrm{T}}]. \end{aligned}$$

Substituting (17) into the above equation and using (5), we obtain

$$\begin{aligned} \delta T_i = m_i \dot{x}_i^{\mathrm{T}} \delta\dot{x}_i &+ \frac{1}{2}\mathrm{tr}[-\dot{\eta}_i\{J_{d_i} S(\Omega_i) + S(\Omega_i) J_{d_i}\}] \\ &+ \frac{1}{2}\mathrm{tr}[\eta_i S(\Omega_i)\{J_{d_i} S(\Omega_i) + S(\Omega_i) J_{d_i}\}] \\ &- \frac{1}{2}\mathrm{tr}[\eta_i\{J_{d_i} S(\Omega_i) + S(\Omega_i) J_{d_i}\} S(\Omega_i)]. \end{aligned}$$

Using (9) and (13), δT_i is given by

$$\delta T_i = m_i \dot{x}_i^{\mathrm{T}} \delta\dot{x}_i + \frac{1}{2}\mathrm{tr}[-\dot{\eta}_i S(J_i\Omega_i) + \eta_i S(\Omega_i \times J_i\Omega_i)]. \tag{18}$$

The variation of the potential energy is given by

$$\delta U = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} U(\mathbf{x} + \epsilon\delta\mathbf{x}, \mathbf{R} + \epsilon\delta\mathbf{R}),$$

where $\delta\mathbf{x} = (\delta x_1, \delta x_2, \ldots, \delta x_n)$, $\delta\mathbf{R} = (\delta R_1, \delta R_2, \ldots, \delta R_n)$. Then, δU can be written as

$$\begin{aligned} \delta U &= \sum_{i=1}^{n} \left(\sum_{p=1}^{3} \frac{\partial U}{\partial [x_i]_p} [\delta x_i]_p + \sum_{p,q=1}^{3} \frac{\partial U}{\partial [R_i]_{pq}} [R_i \eta_i]_{pq} \right) \\ &= \sum_{i=1}^{n} \left(\frac{\partial U^{\mathrm{T}}}{\partial x_i} \delta x_i - \mathrm{tr}\left[\eta_i R_i^{\mathrm{T}} \frac{\partial U}{\partial R_i}\right] \right), \end{aligned} \tag{19}$$

where $[A]_{pq}$ denotes the p,qth element of a matrix A, and $\frac{\partial U}{\partial x_i}, \frac{\partial U}{\partial R_i}$ are given by $\left[\frac{\partial U}{\partial x_i}\right]_p = \frac{\partial U}{\partial [x_i]_p}$, $\left[\frac{\partial U}{\partial R_i}\right]_{pq} = \frac{\partial U}{\partial [R_i]_{pq}}$, respectively. The variation of the Lagrangian has the form

$$\delta L = \sum_{i=1}^{n} \delta T_i - \delta U, \tag{20}$$

which can be written more explicitly by using (18) and (19).

The action integral is defined to be

$$\mathfrak{G} = \int_{t_0}^{t_f} L(\mathbf{x}, \dot{\mathbf{x}}, \mathbf{R}, \mathbf{\Omega})\, \mathrm{d}t. \tag{21}$$

The variation of the action integral can be written as

$$\begin{aligned} \delta\mathfrak{G} = \sum_{i=1}^{n} \int_{t_0}^{t_f} & m_i \dot{x}_i^{\mathrm{T}} \delta\dot{x}_i - \frac{\partial U^{\mathrm{T}}}{\partial x_i}\delta x_i + \frac{1}{2}\mathrm{tr}[-\dot{\eta}_i S(J_i\Omega_i)] \\ &+ \frac{1}{2}\mathrm{tr}\left[\eta_i \left\{ S(\Omega_i \times J_i\Omega_i) + 2R_i^{\mathrm{T}} \frac{\partial U}{\partial R_i} \right\}\right] \mathrm{d}t. \end{aligned}$$

Using integration by parts,

$$\begin{aligned} \delta\mathfrak{G} = \sum_{i=1}^{n} & m_i \dot{x}_i^{\mathrm{T}} \delta x_i|_{t_0}^{t_f} - \frac{1}{2}\mathrm{tr}[\eta_i S(J_i\Omega_i)]|_{t_0}^{t_f} + \int_{t_0}^{t_f} -m_i \ddot{x}_i^{\mathrm{T}} \delta x_i \\ &+ \frac{1}{2}\mathrm{tr}[\eta_i S(J_i\dot{\Omega}_i)]\, \mathrm{d}t + \int_{t_0}^{t_f} -\frac{\partial U^{\mathrm{T}}}{\partial x_i}\delta x_i\, \mathrm{d}t \\ &+ \frac{1}{2}\mathrm{tr}\left[\eta_i \left\{ S(\Omega_i \times J_i\Omega_i) + 2R_i^{\mathrm{T}} \frac{\partial U}{\partial R_i} \right\}\right] \mathrm{d}t. \end{aligned}$$

Using the fact that δx_i and η_i vanish at t_0 and t_f, the first two terms of the above equation vanish. Then, $\delta\mathfrak{G}$ is given by

$$\begin{aligned} \delta\mathfrak{G} = \sum_{i=1}^{n} \int_{t_0}^{t_f} & -\delta x_i^{\mathrm{T}} \left\{ m_i \ddot{x}_i + \frac{\partial U}{\partial x_i} \right\} \\ &+ \frac{1}{2}\mathrm{tr}\left[\eta_i \left\{ S(J_i\dot{\Omega}_i + \Omega_i \times J_i\Omega_i) + 2R_i^{\mathrm{T}} \frac{\partial U}{\partial R_i} \right\}\right] \mathrm{d}t. \end{aligned}$$

From Hamilton's principle, $\delta\mathfrak{G}$ should be zero for all possible variations $\delta x_i : [t_0, t_f] \mapsto \mathbb{R}^3$ and $\eta_i : [t_0, t_f] \mapsto \mathfrak{so}(3)$. Therefore, the expression in the first brace should be zero. Furthermore, since η_i is skew symmetric, we have by (7), that the expression in the second brace should be symmetric. Then, we obtain the continuous equations of motion as

$$\begin{aligned} & m_i \ddot{x}_i = -\frac{\partial U}{\partial x_i}, \\ & S(J_i\dot{\Omega}_i + \Omega_i \times J_i\Omega_i) + 2R_i^{\mathrm{T}} \frac{\partial U}{\partial R_i} \\ & \quad = S(J_i\dot{\Omega}_i + \Omega_i \times J_i\Omega_i)^{\mathrm{T}} + 2\frac{\partial U^{\mathrm{T}}}{\partial R_i} R. \end{aligned} \tag{22}$$

Using (8), we can simplify (22) to be

$$S(J_i\dot{\Omega}_i + \Omega_i \times J_i\Omega_i) = \frac{\partial U^{\mathrm{T}}}{\partial R_i} R - R_i^{\mathrm{T}} \frac{\partial U}{\partial R_i}.$$

Note that the right hand side expression in the above equation is also skew symmetric. Then, the moment due to the gravitational potential on the ith body, $M_i \in \mathbb{R}^3$ is given by $S(M_i) = \frac{\partial U^{\mathrm{T}}}{\partial R_i} R_i - R_i^{\mathrm{T}} \frac{\partial U}{\partial R_i}$. This moment M_i can be expressed more explicitly as the following computation shows:

$$\begin{aligned} S(M_i) &= \frac{\partial U^{\mathrm{T}}}{\partial R_i} R_i - R_i^{\mathrm{T}} \frac{\partial U}{\partial R_i} \\ &= \begin{bmatrix} u_{ri_1}^{\mathrm{T}} & u_{ri_2}^{\mathrm{T}} & u_{ri_3}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} r_{i_1} \\ r_{i_2} \\ r_{i_3} \end{bmatrix} - \begin{bmatrix} r_{i_1}^{\mathrm{T}} & r_{i_2}^{\mathrm{T}} & r_{i_3}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} u_{ri_1} \\ u_{ri_2} \\ u_{ri_3} \end{bmatrix} \\ &= (u_{ri_1}^{\mathrm{T}} r_{i_1} - r_{i_1}^{\mathrm{T}} u_{ri_1}) + (u_{ri_2}^{\mathrm{T}} r_{i_2} - r_{i_2}^{\mathrm{T}} u_{ri_2}) + (u_{ri_3}^{\mathrm{T}} r_{i_3} - r_{i_3}^{\mathrm{T}} u_{ri_3}), \end{aligned}$$

where $r_{i_p}, u_{ri_p} \in \mathbb{R}^{1\times 3}$ are the pth row vectors of R_i and $\frac{\partial U}{\partial R_i}$, respectively. Substituting $x = r_{i_p}^{\mathrm{T}}$, $y = u_{ri_p}^{\mathrm{T}}$ into (9), we obtain

$$\begin{aligned} S(M_i) &= S(r_{i_1} \times u_{ri_1}) + S(r_{i_2} \times u_{ri_2}) + S(r_{i_3} \times u_{ri_3}) \\ &= S(r_{i_1} \times u_{ri_1} + r_{i_2} \times u_{ri_2} + r_{i_3} \times u_{ri_3}), \end{aligned} \tag{23}$$

Then, the gravitational moment M_i is given by

$$M_i = r_{i_1} \times u_{ri_1} + r_{i_2} \times u_{ri_2} + r_{i_3} \times u_{ri_3}. \tag{24}$$

In summary, *the continuous equations of motion for the full body problem, in Lagrangian form*, can be written for $i \in (1, 2, \ldots, n)$ as

$$\dot{v}_i = -\frac{1}{m_i}\frac{\partial U}{\partial x_i}, \tag{25}$$

$$J_i\dot{\Omega}_i + \Omega_i \times J_i\Omega_i = M_i, \tag{26}$$

$$\dot{x}_i = v_i, \tag{27}$$

$$\dot{R}_i = R_i S(\Omega_i), \tag{28}$$

where the translational velocity $v_i \in \mathbb{R}^3$ is defined as $v_i = \dot{x}_i$.

3.1.2. Equations of motion: Hamiltonian form

We denote the linear and angular momentum of the ith body $\mathscr{B}_i$ by $\gamma_i \in \mathbb{R}^3$, and $\Pi_i \in \mathbb{R}^3$, respectively. They are related to the linear and angular velocities by $\gamma_i = m_i v_i$, and $\Pi_i = J_i\Omega_i$. Then, the equations of motion can be rewritten in terms of the momenta variables. *The continuous equations of motion for the full body problem, in Hamiltonian form,* can be written for $i \in (1, 2, \ldots, n)$ as

$$\dot{\gamma}_i = -\frac{\partial U}{\partial x_i}, \tag{29}$$

$$\dot{\Pi}_i + \Omega_i \times \Pi_i = M_i, \tag{30}$$

$$\dot{x}_i = \frac{\gamma_i}{m_i}, \tag{31}$$

$$\dot{R}_i = R_i S(\Omega_i). \tag{32}$$

3.2. Relative frame

The motion of the full rigid bodies depends only on the relative positions and the relative attitudes of the bodies. This is a consequence of the property that the gravitational potential can be expressed using only these relative variables. Physically, this is related to the fact that the total linear momentum and the total angular momentum about the mass center of the bodies are conserved. Mathematically, the Lagrangian is invariant under a left action of an element of SE(3). So, it is natural to express the equations of motion in one of the body fixed frame. In this section, the equations of motion for the full two body problem are derived in the relative frame. This result can be readily generalized to the n body problem.

Reduction of variables: In [2], the reduction is carried out in stages, by first reducing position variables in $\mathbb{R}^3$, and then reducing attitude variables in SO(3). This is equivalent to directly reducing the position and the attitude variables in SE(3) in a single step, which is a result that can be explained by the general theory of Lagrangian reduction by stages [24]. The reduced position and the reduced attitude variables are the relative position and the relative attitude of the first body with respect to the second body. In other words, the variables are reduced by applying the inverse of $(R_2, x_2) \in \mathrm{SE}(3)$ given by $(R_2^{\mathrm{T}}, -R_2^{\mathrm{T}}x_2) \in \mathrm{SE}(3)$, to the following homogeneous transformations:

$$\begin{aligned} &\begin{bmatrix} R_2^{\mathrm{T}} & -R_2^{\mathrm{T}}x_2 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} R_1 & x_1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} R_2 & x_2 \\ 0 & 1 \end{bmatrix}\right) \\ &\quad = \left(\begin{bmatrix} R_2^{\mathrm{T}}R_1 & R_2^{\mathrm{T}}(x_1 - x_2) \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} R_2^{\mathrm{T}}R_2 & R_2^{\mathrm{T}}(x_2 - x_2) \\ 0 & 1 \end{bmatrix}\right) \\ &\quad = \left(\begin{bmatrix} R_2^{\mathrm{T}}R_1 & R_2^{\mathrm{T}}(x_1 - x_2) \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} I_{3\times 3} & 0 \\ 0 & 1 \end{bmatrix}\right). \end{aligned} \tag{33}$$

This motivates the definition of the reduced variables as

$$X = R_2^{\mathrm{T}}(x_1 - x_2), \tag{34}$$

$$R = R_2^{\mathrm{T}}R_1, \tag{35}$$

where $X \in \mathbb{R}^3$ is the relative position of the first body with respect to the second body expressed in the second body fixed frame, and $R \in \mathrm{SO}(3)$ is the relative attitude of the first body with respect to the second body. The corresponding linear and angular velocities are also defined as

$$V = R_2^{\mathrm{T}}(\dot{x}_1 - \dot{x}_2), \tag{36}$$

$$\Omega = R\Omega_1, \tag{37}$$

where $V \in \mathbb{R}^3$ represents the relative velocity of the first body with respect to the second body in the second body fixed frame, and $\Omega \in \mathbb{R}^3$ is the angular velocity of the first body expressed in the second body fixed frame. Here, the capital letters denote variables expressed in the second body fixed frame.

For convenience, we denote the inertial position and the inertial velocity of the second body, expressed in the second body fixed frame by $X_2, V_2 \in \mathbb{R}^3$

$$X_2 = R_2^{\mathrm{T}}x_2, \tag{38}$$

$$V_2 = R_2^{\mathrm{T}}\dot{x}_2. \tag{39}$$

Reduced lagrangian: The equations of motion in the relative frame are derived in the same way used to derive the equations in the inertial frame. We first construct a reduced Lagrangian. The reduced Lagrangian l is obtained by expressing the original Lagrangian (15) in terms of the reduced variables. The kinetic energy is given by

$$\begin{aligned} T_1 + T_2 &= \frac{1}{2}m_1\|\dot{x}_1\|^2 + \frac{1}{2}m_2\|\dot{x}_2\|^2 + \frac{1}{2}\mathrm{tr}[S(\Omega_1)J_{d_1}S(\Omega_1)^{\mathrm{T}}] \\ &\quad + \frac{1}{2}\mathrm{tr}[S(\Omega_2)J_{d_2}S(\Omega_2)^{\mathrm{T}}] \\ &= \frac{1}{2}m_1\|V + V_2\|^2 + \frac{1}{2}m_2\|V_2\|^2 \\ &\quad + \frac{1}{2}\mathrm{tr}[S(\Omega)J_{d_R}S(\Omega)^{\mathrm{T}}] + \frac{1}{2}\mathrm{tr}[S(\Omega_2)J_{d_2}S(\Omega_2)^{\mathrm{T}}], \end{aligned}$$

where (10) is used, and $J_{d_R} \in \mathbb{R}^{3\times 3}$ is defined as $J_{d_R} = RJ_{d_1}R^{\mathrm{T}}$, which is an expression of the nonstandard moment of inertia matrix of the first body with respect to the second body fixed frame. Note that J_{d_R} is not a constant matrix. Using (10), it can be shown that J_{d_R} also satisfies a property similar to (13), namely

$$S(J_R\Omega) = S(\Omega)J_{d_R} + J_{d_R}S(\Omega), \tag{40}$$

where $J_R = RJ_1R^{\mathrm{T}} \in \mathbb{R}^{3\times 3}$ is the standard moment of inertia matrix of the first body with respect to the second body fixed frame.

Using (14), the gravitational potential U of the full two bodies is given by

$$U(x_1,x_2,R_1,R_2) = -\int_{\mathscr{B}_1}\int_{\mathscr{B}_2} \frac{G\mathrm{d}m_1\mathrm{d}m_2}{\|x_1 + R_1\rho_1 - x_2 - R_2\rho_2\|},$$

and it is invariant under an action of an element of SE(3). Therefore, the gravitational potential can be written as a function of the relative variables only. By applying the inverse of $(R_2,x_2) \in \mathrm{SE}(3)$ as given in (33), we obtain

$$\begin{aligned} U(x_1,x_2,R_1,R_2) &= U(R_2^{\mathrm{T}}(x_1-x_2),0,R_2^{\mathrm{T}}R_1,I_{3\times 3}) \\ &= -\int_{\mathscr{B}_1}\int_{\mathscr{B}_2} \frac{G\mathrm{d}m_1\mathrm{d}m_2}{\|R_2^{\mathrm{T}}(x_1-x_2) + R_2^{\mathrm{T}}R_1\rho_1 - I_{3\times 3}\rho_2\|} \\ &= -\int_{\mathscr{B}_1}\int_{\mathscr{B}_2} \frac{G\mathrm{d}m_1\mathrm{d}m_2}{\|X + R\rho_1 - \rho_2\|} \triangleq U(X,R). \end{aligned}$$

Here we abuse notation slightly by using the same letter U to denote the gravitational potential as a function of the relative variables.

Then, the reduced Lagrangian l is given by

$$\begin{aligned} l(R,X,\Omega,V,\Omega_2,V_2) = {} & \frac{1}{2}m_1\|V+V_2\|^2 + \frac{1}{2}m_2\|V_2\|^2 \\ & + \frac{1}{2}\mathrm{tr}[S(\Omega)J_{d_R}S(\Omega)^{\mathrm{T}}] \\ & + \frac{1}{2}\mathrm{tr}[S(\Omega_2)J_{d_2}S(\Omega_2)^{\mathrm{T}}] - U(X,R). \end{aligned} \tag{41}$$

Variations of reduced variables: The variations of the reduced variables must be restricted to those that can arise from the variations of the original variables. For example, the variation of the relative attitude R is given by

$$\delta R = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} R_2^{\epsilon T}R_1^{\epsilon} = \delta R_2^{\mathrm{T}}R_1 + R_2^{\mathrm{T}}\delta R_1.$$

Substituting (16) into the above equation,

$$\delta R = -\eta_2 R_2^{\mathrm{T}}R_1 + R_2^{\mathrm{T}}R_1\eta_1 = -\eta_2 R + \eta R,$$

where a reduced variation $\eta \in \mathfrak{so}(3)$ is defined as $\eta = R\eta_1R^{\mathrm{T}}$. The variations of other reduced variables can be obtained in a similar way. The detailed derivations are given in A.1, and we summarize the results as follows:

$$\delta R = \eta R - \eta_2 R, \tag{42}$$

$$\delta X = \chi - \eta_2 X, \tag{43}$$

$$\begin{aligned} S(\delta\Omega) = {} & \dot\eta - S(\Omega)\eta + \eta S(\Omega) + S(\Omega)\eta_2 - \eta_2 S(\Omega) \\ & + S(\Omega_2)\eta - \eta S(\Omega_2), \end{aligned} \tag{44}$$

$$\delta V = \dot\chi + S(\Omega_2)\chi - \eta_2 V, \tag{45}$$

$$S(\delta\Omega_2) = \dot\eta_2 + S(\Omega_2)\eta_2 - \eta_2 S(\Omega_2), \tag{46}$$

$$\delta V_2 = \dot\chi_2 + S(\Omega_2)\chi_2 - \eta_2 V_2, \tag{47}$$

where $\chi,\chi_2 : [t_0,t_f] \mapsto \mathbb{R}^3$ and $\eta,\eta_2 : [t_0,t_f] \mapsto \mathfrak{so}(3)$ are variations that vanish at the end points. These Lie group variations are the key elements required to obtain the equations of motion in the relative frame.

3.2.1. Equations of motion: Lagrangian form

The reduced equations of motion can be computed from the reduced Lagrangian using the reduced Hamilton's principle. By taking the variation of the reduced Lagrangian (41) using the constrained variations given by (42)–(47), we can obtain the equations of motion in the relative frame.

Following a similar process to the derivation of $\delta T_i, \delta U$ as in (18) and (19), the variation of the reduced Lagrangian δl can be obtained as

$$\begin{aligned} \delta l = {} & \dot\chi^{\mathrm{T}}[m_1(V+V_2)] - \chi^{\mathrm{T}}[m_1\Omega_2\times(V+V_2)] \\ & + \dot\chi_2^{\mathrm{T}}[m_1(V+V_2) + m_2V_2] \\ & - \chi_2^{\mathrm{T}}[m_1\Omega_2\times(V+V_2) + m_2\Omega_2\times V_2] \\ & + \frac{1}{2}\mathrm{tr}[-\dot\eta S(J_R\Omega) + \eta S(\Omega_2\times J_R\Omega)] \\ & + \frac{1}{2}\mathrm{tr}[-\dot\eta_2 S(J_2\Omega_2) + \eta_2 S(\Omega_2\times J_2\Omega_2)] - \chi^{\mathrm{T}}\frac{\partial U}{\partial X} \\ & + \mathrm{tr}\left[\eta_2 X\frac{\partial U^{\mathrm{T}}}{\partial X}\right] + \mathrm{tr}\left[\eta_2 R\frac{\partial U^{\mathrm{T}}}{\partial R} - \eta R\frac{\partial U^{\mathrm{T}}}{\partial R}\right], \end{aligned} \tag{48}$$

where we used the identities (9), (13) and (40), and the constrained variations (42)–(47).

The action integral in terms of the reduced Lagrangian is

$$\mathfrak{G} = \int_{t_0}^{t_f} l(R,X,\Omega,V,\Omega_2,V_2)\,\mathrm{d}t. \tag{49}$$

Using integration by parts together with the fact that χ,χ_2,η and η_2 vanish at t_0 and t_f, the variation of the action integral can be expressed from (48) as (50).

$$\begin{aligned} \delta\mathfrak{G} = {} & -\int_{t_0}^{t_f} \chi^{\mathrm{T}}\left\{m_1(\dot V + \dot V_2) + m_1\Omega_2\times(V+V_2) + \frac{\partial U}{\partial X}\right\}\mathrm{d}t \\ & - \int_{t_0}^{t_f} \chi_2^{\mathrm{T}}\{m_1(\dot V + \dot V_2) + m_2\dot V_2 \\ & + m_1\Omega_2\times(V+V_2) + m_2\Omega_2\times V_2\}\,\mathrm{d}t \\ & + \frac{1}{2}\int_{t_0}^{t_f} \mathrm{tr}\left[\eta\left\{S((J_R\dot\Omega) + \Omega_2\times J_R\Omega) - 2R\frac{\partial U^{\mathrm{T}}}{\partial R}\right\}\right]\mathrm{d}t \\ & + \frac{1}{2}\int_{t_0}^{t_f} \mathrm{tr}\left[\eta_2\left\{S(J_2\dot\Omega_2 + \Omega_2\times J_2\Omega_2) + 2X\frac{\partial U^{\mathrm{T}}}{\partial X} + 2R\frac{\partial U^{\mathrm{T}}}{\partial R}\right\}\right]\mathrm{d}t. \end{aligned} \tag{50}$$

From the reduced Hamilton's principle, $\delta\mathfrak{G} = 0$ for all possible variations $\chi,\chi_2 : [t_0,t_f] \mapsto \mathbb{R}^3$ and $\eta,\eta_2 : [t_0,t_f] \mapsto \mathfrak{so}(3)$

that vanish at t_0 and t_f. Therefore, in (50), the expressions in the first two braces should be zero and the expressions in the last two braces should be symmetric since η, η_2 are skew symmetric. Then, we obtain the following equations of motion:

$$m_1(\dot{V} + \dot{V}_2) + m_1\Omega_2 \times (V + V_2) = -\frac{\partial U}{\partial X}, \tag{51}$$

$$m_2\dot{V}_2 + m_2\Omega_2 \times V_2 = \frac{\partial U}{\partial X}, \tag{52}$$

$$S((J_R\dot{\Omega}) + \Omega_2 \times J_R\Omega) = -S(M),$$

$$S(J_2\dot{\Omega}_2 + \Omega_2 \times J_2\Omega_2) = \frac{\partial U}{\partial X}X^{\mathrm{T}} - X\frac{\partial U^{\mathrm{T}}}{\partial X} + S(M), \tag{53}$$

where $M \in \mathbb{R}^3$ is defined by the relation $S(M) = \frac{\partial U}{\partial R}R^{\mathrm{T}} - R\frac{\partial U}{\partial R}^{\mathrm{T}}$. By a procedure analogous to the derivation of (23), M can be written as

$$M = r_1 \times u_{r_1} + r_2 \times u_{r_2} + r_3 \times u_{r_3}, \tag{54}$$

where $r_p, u_{r_p} \in \mathbb{R}^3$ are the pth column vectors of R and $\frac{\partial U}{\partial R}$, respectively.

Eq. (51) can be simplified using (52) as

$$\dot{V} + \Omega_2 \times V = -\frac{m_1 + m_2}{m_1 m_2}\frac{\partial U}{\partial X}.$$

For reconstruction of the motion of the second body, it is natural to express the motion of the second body in the inertial frame. Since $\dot{V}_2 = \dot{R}_2^{\mathrm{T}}\dot{x}_2 + R_2^{\mathrm{T}}\ddot{x}_2 = -S(\Omega_2)V + R_2^{\mathrm{T}}\dot{v}_2$, (52) can be written as

$$m_2 R_2^{\mathrm{T}}\dot{v}_2 = \frac{\partial U}{\partial X}.$$

Eq. (53) can be simplified using the property $\frac{\partial U}{\partial X}X^{\mathrm{T}} - X\frac{\partial U}{\partial X}^{\mathrm{T}} = S(X \times \frac{\partial U}{\partial X})$ from (9). The kinematics equations for $\dot{R}$ and $\dot{X}$ can be derived in a similar way.

In summary, *the continuous equations of relative motion for the full two body problem, in Lagrangian form*, can be written as

$$\dot{V} + \Omega_2 \times V = -\frac{1}{m}\frac{\partial U}{\partial X}, \tag{55}$$

$$(J_R\dot{\Omega}) + \Omega_2 \times J_R\Omega = -M, \tag{56}$$

$$J_2\dot{\Omega}_2 + \Omega_2 \times J_2\Omega_2 = X \times \frac{\partial U}{\partial X} + M, \tag{57}$$

$$\dot{X} + \Omega_2 \times X = V, \tag{58}$$

$$\dot{R} = S(\Omega)R - S(\Omega_2)R, \tag{59}$$

where $m = \frac{m_1 m_2}{m_1 + m_2}$. The following equations can be used for reconstruction of the motion of the second body in the inertial frame:

$$\dot{v}_2 = \frac{1}{m_2}R_2\frac{\partial U}{\partial X}, \tag{60}$$

$$\dot{x}_2 = v_2, \tag{61}$$

$$\dot{R}_2 = R_2 S(\Omega_2). \tag{62}$$

These equations are equivalent to those given in [2]. However, (60) is not given in [2]. (55)–(62) give a complete set of equations for the reduced dynamics and reconstruction. Furthermore, they are derived systematically in the context of geometric mechanics using proper variational formulas given in (42)–(47). This result can be readily generalized for n bodies.

3.2.2. Equations of motion: Hamiltonian form

Define the linear momenta $\Gamma, \gamma_2 \in \mathbb{R}^3$, and the angular momenta $\Pi, \Pi_2 \in \mathbb{R}^3$ as

$$\Gamma = mV,$$
$$\gamma_2 = mv_2,$$
$$\Pi = J_R\Omega = RJ_1\Omega_1,$$
$$\Pi_2 = J_2\Omega_2.$$

Then, the equations of motion can be rewritten in terms of these momenta variables. *The continuous equations of relative motion for the full two body problem, in Hamiltonian form*, can be written as

$$\dot{\Gamma} + \Omega_2 \times \Gamma = -\frac{\partial U}{\partial X}, \tag{63}$$

$$\dot{\Pi} + \Omega_2 \times \Pi = -M, \tag{64}$$

$$\dot{\Pi}_2 + \Omega_2 \times \Pi_2 = X \times \frac{\partial U}{\partial X} + M, \tag{65}$$

$$\dot{X} + \Omega_2 \times X = \frac{\Gamma}{m}, \tag{66}$$

$$\dot{R} = S(\Omega)R - S(\Omega_2)R, \tag{67}$$

where $m = \frac{m_1 m_2}{m_1 + m_2}$. The following equations can be used to reconstruct the motion of the second body in the inertial frame:

$$\dot{\gamma}_2 = R_2\frac{\partial U}{\partial X}, \tag{68}$$

$$\dot{x}_2 = \frac{\gamma_2}{m_2}, \tag{69}$$

$$\dot{R}_2 = R_2 S(\Omega_2). \tag{70}$$

4. Lie group variational integrators

A variational integrator discretizes Hamilton's principle rather than the continuous equations of motion. Taking variations of the discretization of the action integral leads to the discrete Euler–Lagrange or discrete Hamilton's equations. The discrete Euler–Lagrange equations can be interpreted as a discrete Lagrangian map that updates the variables in the configuration space, which are the positions and the attitudes of the bodies. A discrete Legendre transformation relates the configuration variables with the linear and angular momenta variables, and yields a discrete Hamiltonian map, which is equivalent to the discrete Lagrangian map.

In this section, we derive both a Lagrangian and Hamiltonian form of variational integrators for the full body

problem in inertial and relative frames. The second level subscript k denotes the value of variables at $t = kh + t_0$ for an integration step size $h \in \mathbb{R}$ and an integer k. The integer N satisfies $t_f = kN + t_0$, so N is the number of time-steps of length h to go from the initial time t_0 to the final time t_f.

4.1. Inertial frame

Discrete Lagrangian: In continuous time, the structure of the kinematics Eqs. (28), (59) and (62) ensure that R_i, R and R_2 evolve on SO(3) automatically. Here, we introduce a new variable $F_{i_k} \in \mathrm{SO}(3)$ defined such that $R_{i_{k+1}} = R_{i_k} F_{i_k}$, i.e.

$$F_{i_k} = R_{i_k}^{\mathrm{T}} R_{i_{k+1}}. \tag{71}$$

Thus, F_{i_k} represents the relative attitude between two integration steps, and by requiring that $F_{i_k} \in \mathrm{SO}(3)$, we guarantee that R_{i_k} evolves in SO(3).

Using the kinematic equation $\dot{R}_i = R_i S(\Omega_i)$, the skew-symmetric matrix $S(\Omega_k)$ can be approximated as

$$S(\Omega_{i_k}) = R_{i_k}^{\mathrm{T}} \dot{R}_{i_k} \approx R_{i_k}^{\mathrm{T}} \frac{R_{i_{k+1}} - R_{i_k}}{h} = \frac{1}{h}(F_{i_k} - I_{3\times 3}). \tag{72}$$

The velocity, $\dot{x}_{i_k}$ can be approximated simply by $(x_{i_{k+1}} - x_{i_k})/h$. Using these approximations of the angular and linear velocity, the kinetic energy of the ith body given in (12) can be approximated as

$$\begin{aligned} T_i(\dot{x}_i, \Omega_i) &\approx T_i\left(\frac{1}{h}(x_{i_{k+1}} - x_{i_k}), \frac{1}{h}(F_{i_k} - I_{3\times 3})\right) \\ &= \frac{1}{2h^2} m_i \|x_{i_{k+1}} - x_{i_k}\|^2 \\ &\quad + \frac{1}{2h^2} \mathrm{tr}[(F_{i_k} - I_{3\times 3}) J_{d_i} (F_{i_k} - I_{3\times 3})^{\mathrm{T}}] \\ &= \frac{1}{2h^2} m_i \|x_{i_{k+1}} - x_{i_k}\|^2 + \frac{1}{h^2} \mathrm{tr}[(I_{3\times 3} - F_{i_k}) J_{d_i}], \end{aligned}$$

where (5) is used. A discrete Lagrangian $L_d(\mathbf{x}_k, \mathbf{x}_{k+1}, \mathbf{R}_k, \mathbf{F}_k)$ is constructed such that it approximates a segment of the action integral (21),

$$\begin{aligned} L_d &= \frac{h}{2} L\left(\mathbf{x}_k, \frac{1}{h}(\mathbf{x}_{k+1} - \mathbf{x}_k), \mathbf{R}_k, \frac{1}{h}(\mathbf{F}_k - \mathbf{I})\right) \\ &\quad + \frac{h}{2} L\left(\mathbf{x}_{k+1}, \frac{1}{h}(\mathbf{x}_{k+1} - \mathbf{x}_k), \mathbf{R}_{k+1}, \frac{1}{h}(\mathbf{F}_k - \mathbf{I})\right) \\ &= \sum_{i=1}^{n} \frac{1}{2h} m_i \|x_{i_{k+1}} - x_{i_k}\|^2 + \frac{1}{h} \mathrm{tr}[(I_{3\times 3} - F_{i_k}) J_{d_i}] \\ &\quad - \frac{h}{2} U(\mathbf{x}_k, \mathbf{R}_k) - \frac{h}{2} U(\mathbf{x}_{k+1}, \mathbf{R}_{k+1}), \end{aligned} \tag{73}$$

where $\mathbf{x}_k \in (\mathbb{R}^3)^n$, $\mathbf{R}_k \in \mathrm{SO}(3)^n$, and $\mathbf{F}_k \in (\mathbb{R}^3)^n$, and $\mathbf{I} \in (\mathbb{R}^{3\times 3})^n$ are defined as $\mathbf{x}_k = (x_{1_k}, x_{2_k}, \ldots, x_{n_k})$, $\mathbf{R}_k = (R_{1_k}, R_{2_k}, \ldots, R_{n_k})$, $\mathbf{F}_k = (F_{1_k}, F_{2_k}, \ldots, F_{n_k})$, and $\mathbf{I} = (I_{3\times 3}, I_{3\times 3}, \ldots, I_{3\times 3})$, respectively.

This discrete Lagrangian is self-adjoint [9], and self-adjoint numerical integration methods have even order, so we are guaranteed that the resulting integration method is at least second-order accurate.

Variations of discrete variables: The variations of the discrete variables are chosen to respect the geometry of the configuration space SE(3). The variation of x_{i_k} is given by

$$x_{i_k}^{\epsilon} = x_{i_k} + \epsilon \delta x_{i_k} + \mathcal{O}(\epsilon^2),$$

where $\delta x_{i_k} \in \mathbb{R}^3$ and vanishes at $k = 0$ and $k = N$. The variation of R_{i_k} is given by

$$\delta R_{i_k} = R_{i_k} \eta_{i_k}, \tag{74}$$

where $\eta_{i_k} \in \mathfrak{so}(3)$ is a variation represented by a skew-symmetric matrix and vanishes at $k = 0$ and $k = N$. The variation of F_{i_k} can be computed from the definition $F_{i_k} = R_{i_k}^{\mathrm{T}} R_{i_{k+1}}$ to give

$$\begin{aligned} \delta F_{i_k} &= \delta R_{i_k}^{\mathrm{T}} R_{i_{k+1}} + R_{i_k}^{\mathrm{T}} \delta R_{i_{k+1}} \\ &= -\eta_{i_k} R_{i_k}^{\mathrm{T}} R_{i_{k+1}} + R_{i_k}^{\mathrm{T}} R_{i_{k+1}} \eta_{i_{k+1}} \\ &= -\eta_{i_k} F_{i_k} + F_{i_k} \eta_{i_{k+1}}. \end{aligned} \tag{75}$$

4.1.1. Discrete equations of motion: Lagrangian form

To obtain the discrete equations of motion in Lagrangian form, we compute the variation of the discrete Lagrangian from (19), (74) and (75), to give

$$\begin{aligned} \delta L_d &= \sum_{i=1}^{n} \frac{1}{h} m_i (x_{i_{k+1}} - x_{i_k})^{\mathrm{T}} (\delta x_{i_{k+1}} - \delta x_{i_k}) \\ &\quad + \frac{1}{h} \mathrm{tr}[(\eta_{i_k} F_{i_k} - F_{i_k} \eta_{i_{k+1}}) J_{d_i}] \\ &\quad - \frac{h}{2} \left(\frac{\partial U_k^{\mathrm{T}}}{\partial x_{i_k}} \delta x_{i_k} + \frac{\partial U_{k+1}^{\mathrm{T}}}{\partial x_{i_{k+1}}} \delta x_{i_{k+1}} \right) \\ &\quad + \frac{h}{2} \mathrm{tr}\left[\eta_{i_k} R_{i_k}^{\mathrm{T}} \frac{\partial U_k}{\partial R_{i_k}} + \eta_{i_{k+1}} R_{i_{k+1}}^{\mathrm{T}} \frac{\partial U_{k+1}}{\partial R_{i_{k+1}}} \right], \end{aligned} \tag{76}$$

where $U_k = U(\mathbf{x}_k, \mathbf{R}_k)$ denotes the value of the potential at $t = kh + t_0$.

Define the action sum as

$$\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(\mathbf{x}_k, \mathbf{x}_{k+1}, \mathbf{R}_k, \mathbf{F}_k). \tag{77}$$

The discrete action sum $\mathfrak{G}_d$ approximates the action integral (21), because the discrete Lagrangian approximates a segment of the action integral.

Substituting (76) into (77), the variation of the action sum is given by

$$\begin{aligned} \delta \mathfrak{G}_d &= \sum_{k=0}^{N-1} \sum_{i=1}^{n} \delta x_{i_{k+1}}^{\mathrm{T}} \left\{ \frac{1}{h} m_i (x_{i_{k+1}} - x_{i_k}) - \frac{h}{2} \frac{\partial U_{k+1}}{\partial x_{i_{k+1}}} \right\} \\ &\quad + \delta x_{i_k}^{\mathrm{T}} \left\{ -\frac{1}{h} m_i (x_{i_{k+1}} - x_{i_k}) - \frac{h}{2} \frac{\partial U_k}{\partial x_{i_k}} \right\} \\ &\quad + \mathrm{tr}\left[\eta_{i_{k+1}} \left\{ -\frac{1}{h} J_{d_i} F_{i_k} + \frac{h}{2} R_{i_{k+1}}^{\mathrm{T}} \frac{\partial U_{k+1}}{\partial R_{i_{k+1}}} \right\} \right] \\ &\quad + \mathrm{tr}\left[\eta_{i_k} \left\{ \frac{1}{h} F_{i_k} J_{d_i} + \frac{h}{2} R_{i_k}^{\mathrm{T}} \frac{\partial U_k}{\partial R_{i_k}} \right\} \right]. \end{aligned}$$

Using the fact that δx_{i_k} and η_{i_k} vanish at $k=0$ and $k=N$, we can reindex the summation, which is the discrete analogue of integration by parts, to yield

$$\delta\mathfrak{G}_d = \sum_{k=1}^{N-1}\sum_{i=1}^{n} -\delta x_{i_k}\left\{\frac{1}{h}m_i(x_{i_{k+1}} - 2x_{i_k} + x_{i_{k-1}}) + h\frac{\partial U_k}{\partial x_{i_k}}\right\} + \mathrm{tr}\left[\eta_{i_k}\left\{\frac{1}{h}(F_{i_k}J_{d_i} - J_{d_i}F_{i_{k-1}}) + hR_{i_k}^{\mathrm{T}}\frac{\partial U_k}{\partial R_{i_k}}\right\}\right].$$

Hamilton's principle states that $\delta\mathfrak{G}_d$ should be zero for all possible variations $\delta x_{i_k} \in \mathbb{R}^3$ and $\eta_{i_k} \in \mathfrak{so}(3)$ that vanish at the endpoints. Therefore, the expression in the first brace should be zero, and since η_{i_k} is skew-symmetric, the expression in the second brace should be symmetric. Thus, we obtain *the discrete equations of motion for the full body problem, in Lagrangian form*, for $i \in (1,2,\ldots,n)$ as

$$\frac{1}{h}(x_{i_{k+1}} - 2x_{i_k} + x_{i_{k-1}}) = -h\frac{\partial U_k}{\partial x_{i_k}}, \tag{78}$$

$$\frac{1}{h}(F_{i_{k+1}}J_{d_i} - J_{d_i}F_{i_{k+1}}^{\mathrm{T}} - J_{d_i}F_{i_k} + F_{i_k}^{\mathrm{T}}J_{d_i}) = hS(M_{i_{k+1}}), \tag{79}$$

$$R_{i_{k+1}} = R_{i_k}F_{i_k}, \tag{80}$$

where $M_{i_k} \in \mathbb{R}^3$ is defined in (24) as

$$M_{i_k} = r_{i_1} \times u_{ri_1} + r_{i_2} \times u_{ri_2} + r_{i_3} \times u_{ri_3}, \tag{81}$$

where $r_{i_p}, u_{ri_p} \in \mathbb{R}^{1\times 3}$ are pth row vectors of R_{i_k} and $\frac{\partial U_k}{\partial R_{i_k}}$, respectively. Given the initial conditions $(x_{i_0}, R_{i_0}, x_{i_1}, R_{i_1})$, we can obtain x_{i_2} from (78). Then, F_{i_0} is computed from (80), and F_{i_1} can be obtained by solving the implicit Eq. (79). Finally, R_{i_2} is found from (80). This yields an update map $(x_{i_0}, R_{i_0}, x_{i_1}, R_{i_1}) \mapsto (x_{i_1}, R_{i_1}, x_{i_2}, R_{i_2})$, and this process can be repeated.

4.1.2. Discrete equations of motion: Hamiltonian form

As discussed above, (78)–(80) defines a discrete Lagrangian map that updates x_{i_k} and R_{i_k}. The discrete Legendre transformation given in (3) and (4) relates the configuration variables x_{i_k}, R_{i_k} and the corresponding momenta. This induces a discrete Hamiltonian map that is equivalent to the discrete Lagrangian map. The discrete Hamiltonian map is particularly convenient if the initial conditions are given in terms of the positions and momenta at the initial time $(x_{i_0}, v_{i_0}, R_{i_0}, \Omega_{i_0})$.

Before deriving the variational integrator in Hamiltonian form, consider the momenta conjugate to x_i and R_i, namely $P_{v_i} \in \mathbb{R}^3$ and $P_{\Omega_i} \in \mathbb{R}^{3\times 3}$. From the definition (1), $\mathbb{F}_{v_i}L$ is obtained by taking the derivative of L, given in (15), with $\dot{x}_i$ while holding other variables fixed.

$$\delta\dot{x}_i^{\mathrm{T}}P_{v_i} = \mathbb{F}_{v_i}L(\mathbf{x},\dot{\mathbf{x}},\mathbf{R},\boldsymbol{\Omega}) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} L(\mathbf{x},\dot{\mathbf{x}}+\epsilon\delta\dot{\mathbf{x}}_i,\mathbf{R},\boldsymbol{\Omega}) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} T_i(\dot{x}_i+\epsilon\delta\dot{x}_i,\Omega_i) = \delta\dot{x}_i^{\mathrm{T}}(m_i\dot{x}_i),$$

where $\delta\dot{\mathbf{x}}_i \in (\mathbb{R}^3)^n$ denotes $(0,0,\ldots,\delta\dot{x}_i,\ldots,0)$, and T_i is given in (12). Then, we obtain

$$P_{v_i} = m_i v_i = \gamma_i, \tag{82}$$

which is equal to the linear momentum of $\mathscr{B}_i$. Similarly,

$$\mathrm{tr}[S(\delta\Omega_i)^{\mathrm{T}}P_{\Omega_i}] = \mathbb{F}_{\Omega_i}L(\mathbf{x},\dot{\mathbf{x}},\mathbf{R},\boldsymbol{\Omega}) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} T_i(\dot{x}_i,\Omega_i+\epsilon\delta\Omega_i) = \frac{1}{2}\mathrm{tr}[S(\delta\Omega_i)^{\mathrm{T}}S(J_i\Omega_i)],$$

where (5) and (13) are used. Now, we obtain

$$\mathrm{tr}\left[S(\delta\Omega_i)^{\mathrm{T}}\left\{P_{\Omega_i} - \frac{1}{2}S(J_i\Omega_i)\right\}\right] = 0.$$

Since $S(\Omega_i)$ is skew-symmetric, the expression in the braces should be symmetric. This implies that

$$P_{\Omega_i} - P_{\Omega_i}^{\mathrm{T}} = S(J_i\Omega_i) = S(\Pi_i). \tag{83}$$

Eqs. (82) and (83) give expressions for the momenta conjugate to x_i and R_i. Consider the discrete Legendre transformations given in (3) and (4). Then,

$$\delta x_{i_k}^{\mathrm{T}}\mathbf{D}_{x_{i,k}}L_d(\mathbf{x}_k,\mathbf{x}_{k+1},\mathbf{R}_k,\mathbf{F}_k) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} L_d(\mathbf{x}_k+\epsilon\delta\mathbf{x}_{i_k},\mathbf{x}_{k+1},\mathbf{R}_k,\mathbf{F}_k) = -\delta x_{i_k}^{\mathrm{T}}\left[\frac{1}{h}m_i(x_{i_{k+1}} - x_{i_k}) + \frac{h}{2}\frac{\partial U_k}{\partial x_{i_k}}\right], \tag{84}$$

where $\delta\mathbf{x}_{i_k} \in (\mathbb{R}^3)^n$ denotes $(0,0,\ldots,\delta x_{i_k},\ldots,0)$. Therefore, we have

$$\mathbf{D}_{x_{i,k}}L_d(\mathbf{x}_k,\mathbf{x}_{k+1},\mathbf{R}_k,\mathbf{F}_k) = -\frac{1}{h}m_i(x_{i_{k+1}} - x_{i_k}) - \frac{h}{2}\frac{\partial U_k}{\partial x_{i_k}}. \tag{85}$$

From the discrete Legendre transformation given in (3), $P_{v_{i,k}} = -\mathbf{D}_{x_{i,k}}L_d$. Using (82) and (85), we obtain

$$\gamma_{i_k} = \frac{1}{h}m_i(x_{i_{k+1}} - x_{i_k}) + \frac{h}{2}\frac{\partial U_k}{\partial x_{i_k}}. \tag{86}$$

Using the discrete Legendre transformation given in (4), $P_{v_{i,k+1}} = \mathbf{D}_{x_{i,k+1}}L_d$, we can derive the following equation similarly:

$$\gamma_{i_{k+1}} = \frac{1}{h}m_i(x_{i_{k+1}} - x_{i_k}) - \frac{h}{2}\frac{\partial U_{k+1}}{\partial x_{i_{k+1}}}. \tag{87}$$

Eqs. (86) and (87) define the variational integrator in Hamiltonian form for the translational motion. Now, consider the rotational motion. We have

$$\mathrm{tr}[\eta_{i_k}\mathbf{D}_{R_{i,k}}L_d^{\mathrm{T}}] = \mathrm{tr}\left[\eta_{i_k}\left\{\frac{1}{h}F_{i_k}J_{d_i} + \frac{h}{2}R_{i_k}^{\mathrm{T}}\frac{\partial U_k}{\partial R_{i_k}}\right\}\right], \tag{88}$$

where the right side is obtained by taking the variation of L_d with respect to R_{i_k}, while holding other variables fixed. Since η_{i_k} is skew-symmetric,

$$-\mathbf{D}_{R_{i,k}}L_d + \mathbf{D}_{R_{i,k}}L_d^{\mathrm{T}} = \frac{1}{h}(F_{i_k}J_{d_i} - J_{d_i}F_{i_k}^{\mathrm{T}}) - \frac{h}{2}S(M_{i_k}), \tag{89}$$

where $M_{i_k} \in \mathbb{R}^3$ is defined in (81).

From the discrete Legendre transformation given in (3), $P_{\Omega_{i,k}} = -\mathbf{D}_{R_{i,k}} L_d$, we obtain the following equation by using (83) and (89),

$$S(\Pi_{i_k}) = \frac{1}{h}(F_{i_k} J_{d_i} - J_{d_i} F_{i_k}^{\mathrm{T}}) - \frac{h}{2} S(M_{i_k}). \tag{90}$$

Using the discrete Legendre transformation given in (4), $P_{\Omega_{i,k+1}} = \mathbf{D}_{R_{i,k+1}} L_d$, we can obtain the following equation:

$$S(\Pi_{i_{k+1}}) = \frac{1}{h} F_{i_k}^{\mathrm{T}}(F_{i_k} J_{d_i} - J_{d_i} F_{i_k}^{\mathrm{T}}) F_{i_k} + \frac{h}{2} S(M_{i_{k+1}}). \tag{91}$$

By using (10) and substituting (90), we can reduce (91) to the following equation in vector form.

$$\Pi_{i_{k+1}} = F_{i_k}^{\mathrm{T}} \Pi_{i_k} + \frac{h}{2} F_{i_k}^{\mathrm{T}} M_{i_k} + \frac{h}{2} M_{i_{k+1}}. \tag{92}$$

Eqs. (90) and (92) define the variational integrator in Hamiltonian form for the rotational motion.

In summary, using (86), (87), (90) and (92), *the discrete equations of motion for the full body problem, in Hamiltonian form,* can be written for $i \in (1, 2, \ldots, n)$ as

$$x_{i_{k+1}} = x_{i_k} + \frac{h}{m_i}\gamma_{i_k} - \frac{h^2}{2m_i}\frac{\partial U_k}{\partial x_{i_k}}, \tag{93}$$

$$\gamma_{i_{k+1}} = \gamma_{i_k} - \frac{h}{2}\frac{\partial U_k}{\partial x_{i_k}} - \frac{h}{2}\frac{\partial U_{k+1}}{\partial x_{i_{k+1}}}, \tag{94}$$

$$hS\left(\Pi_{i_k} + \frac{h}{2} M_{i_k}\right) = F_{i_k} J_{d_i} - J_{d_i} F_{i_k}^{\mathrm{T}}, \tag{95}$$

$$\Pi_{i_{k+1}} = F_{i_k}^{\mathrm{T}} \Pi_{i_k} + \frac{h}{2} F_{i_k}^{\mathrm{T}} M_{i_k} + \frac{h}{2} M_{i_{k+1}}, \tag{96}$$

$$R_{i_{k+1}} = R_{i_k} F_{i_k}. \tag{97}$$

Given $(x_{i_0}, \gamma_{i_0}, R_{i_0}, \Pi_{i_0})$, we can find x_{i_1} from (93). Solving the implicit Eq. (95) yields F_{i_0}, and R_{i_1} is computed from (97). Then, (94) and (96) gives γ_{i_1}, and Π_{i_1}. This defines the discrete Hamiltonian map, $(x_{i_0}, \gamma_{i_0}, R_{i_0}, \Pi_{i_0}) \mapsto (x_{i_1}, \gamma_{i_1}, R_{i_1}, \Pi_{i_1})$, and this process can be repeated.

4.2. Relative frame

In this section, we derive the variational integrator for the full two body problem in the relative frame by following the procedure given before. This result can be readily generalized to n bodies.

Reduction of discrete variables: The discrete reduced variables are defined in the same way as the continuous reduced variables, which are given in (34)–(39). We introduce $F_k \in \mathrm{SO}(3)$ such that $R_{k+1} = R_{2_{k+1}}^{\mathrm{T}} R_{1_{k+1}} = F_{2_k}^{\mathrm{T}} F_k R_k$. i.e.

$$F_k = R_k F_{1_k} R_k^{\mathrm{T}}. \tag{98}$$

Discrete reduced Lagrangian: The discrete reduced Lagrangian is obtained by expressing the original discrete Lagrangian given in (73) in terms of the discrete reduced variables.

From the definition of the discrete reduced variables given in (34) and (38), we have

$$\begin{aligned} x_{1_{k+1}} - x_{1_k} &= R_{2_{k+1}}(X_{k+1} + X_{2_{k+1}}) - R_{2_k}(X_k + X_{2_k}) \\ &= R_{2_k}\{F_{2_k}(X_{k+1} + X_{2_{k+1}}) - (X_k + X_{2_k})\}, \end{aligned} \tag{99}$$

$$x_{2_{k+1}} - x_{2_k} = R_{2_k}\{F_{2_k} X_{2_{k+1}} - X_{2_k}\}. \tag{100}$$

From (72), $S(\Omega_{1_k})$ and $S(\Omega_{2_k})$ are expressed as

$$\begin{aligned} S(\Omega_{1_k}) &= \frac{1}{h}(F_{1_k} - I_{3\times 3}) \\ &= \frac{1}{h} R_k^{\mathrm{T}}(F_k - I_{3\times 3}) R_k, \end{aligned} \tag{101}$$

$$S(\Omega_{2_k}) = \frac{1}{h}(F_{2_k} - I_{3\times 3}). \tag{102}$$

Substituting (99)–(102) into (73), we obtain the discrete reduced Lagrangian

$$\begin{aligned} l_{d_k} &= l_d(X_k, X_{k+1}, X_{2_k}, X_{2_{k+1}}, R_k, F_k, F_{2_k}) \\ &= \frac{1}{2h} m_1 \|F_{2_k}(X_{k+1} + X_{2_{k+1}}) - (X_k + X_{2_k})\|^2 \\ &\quad + \frac{1}{2h} m_2 \|F_{2_k} X_{2_{k+1}} - X_{2_k}\|^2 + \frac{1}{h}\mathrm{tr}[(I_{3\times 3} - F_k) J_{dR_k}] \\ &\quad + \frac{1}{h}\mathrm{tr}[(I_{3\times 3} - F_{2_k}) J_{d_2}] - \frac{h}{2} U(X_k, R_k) - \frac{h}{2} U(X_{k+1}, R_{k+1}), \end{aligned} \tag{103}$$

where $J_{dR_k} \in \mathbb{R}^{3\times 3}$ is defined to be $J_{dR_k} = R_k J_{d_1} R_k^{\mathrm{T}}$, which gives the nonstandard moment of inertia matrix of the first body with respect to the second body fixed frame at $t = kh + t_0$.

Variations of discrete reduced variables: The variations of the discrete reduced variables can be derived from those of the original variables. The variations of R_k, X_k, and F_{2_k} are the same as given in (42), (43), and (75), respectively. The variation of F_k is computed in (A.2).

In summary, the variations of discrete reduced variables are given by

$$\delta R_k = \eta_k R_k - \eta_{2_k} R_k, \tag{104}$$

$$\delta X_k = \chi_k - \eta_{2_k} X_k, \tag{105}$$

$$\delta F_k = -\eta_{2_k} F_k + F_{2_k}\eta_{k+1} F_{2_k}^{\mathrm{T}} F_k + F_k(-\eta_k + \eta_{2_k}), \tag{106}$$

$$\delta X_{2_k} = \chi_{2_k} - \eta_{2_k} X_{2_k}, \tag{107}$$

$$\delta F_{2_k} = -\eta_{2_k} F_{2_k} + F_{2_k}\eta_{2_{k+1}}. \tag{108}$$

These Lie group variations are the main elements required to derive the variational integrator equations.

4.2.1. Discrete equations of motion: Lagrangian form

As before, we can obtain the discrete equations of motion in Lagrangian form by computing the variation of the discrete reduced Lagrangian which, by using (104)–(108), is given as (109).

$$\begin{aligned}\delta l_{d_k} &= \frac{1}{h}\chi_{k+1}^{\mathrm{T}}[m_1(X_{k+1}+X_{2_{k+1}}) - m_1 F_{2_k}^{\mathrm{T}}(X_k+X_{2_k})]\\ &+\frac{1}{h}\chi_k^{\mathrm{T}}[m_1(X_k+X_{2_k}) - m_1 F_{2_k}(X_{k+1}+X_{2_{k+1}})]\\ &+\frac{1}{h}\chi_{2_{k+1}}^{\mathrm{T}}[m_1(X_{k+1}+X_{2_{k+1}}) - m_1 F_{2_k}^{\mathrm{T}}(X_k+X_{2_k})\\ &+m_2 X_{2_{k+1}} - m_2 F_{2_k}^{\mathrm{T}} X_{2_k}] + \frac{1}{h}\chi_{2_k}^{\mathrm{T}}[m_1(X_k+X_{2_k})\\ &-m_1 F_{2_k}(X_{k+1}+X_{2_{k+1}}) + m_2 X_{2_k} - m_2 F_{2_k} X_{2_{k+1}}]\\ &-\frac{1}{h}\mathrm{tr}[\eta_{k+1} F_{2_k}^{\mathrm{T}} F_k J_{dR_k} F_{2_k}] + \frac{1}{h}\mathrm{tr}[\eta_k F_k J_{dR_k}]\\ &-\frac{1}{h}\mathrm{tr}[\eta_{2_{k+1}} J_{d_2} F_{2_k}] + \frac{1}{h}\mathrm{tr}[\eta_{2_k} F_{2_k} J_{d_2}]\\ &-\frac{h}{2}\chi_k^{\mathrm{T}}\frac{\partial U_k}{\partial X_k} + \frac{h}{2}\mathrm{tr}\left[\eta_{2_k} X_k \frac{\partial U_k^{\mathrm{T}}}{\partial X_k}\right] - \frac{h}{2}\chi_{k+1}^{\mathrm{T}}\frac{\partial U_{k+1}}{\partial X_{k+1}}\\ &+\frac{h}{2}\mathrm{tr}\left[\eta_{2_{k+1}} X_{k+1}\frac{\partial U_{k+1}^{\mathrm{T}}}{\partial X_{k+1}}\right] + \frac{h}{2}\mathrm{tr}\left[\eta_{2_k} R_k \frac{\partial U_k^{\mathrm{T}}}{\partial R_k} - \eta_k R_k \frac{\partial U_k^{\mathrm{T}}}{\partial R_k}\right]\\ &+\frac{h}{2}\mathrm{tr}\left[\eta_{2_{k+1}} R_{k+1}\frac{\partial U_{k+1}^{\mathrm{T}}}{\partial R_{k+1}} - \eta_{k+1} R_{k+1}\frac{\partial U_{k+1}^{\mathrm{T}}}{\partial R_{k+1}}\right]. \end{aligned} \tag{109}$$

The action sum expressed in terms of the discrete reduced Lagrangian has the form

$$\mathfrak{G}_d = \sum_{k=0}^{N-1} l_d(X_k, X_{k+1}, X_{2_k}, X_{2_{k+1}}, R_k, F_k, F_{2_k}). \tag{110}$$

The discrete action sum $\mathfrak{G}_d$ approximates the action integral (49), because the discrete Lagrangian approximates a piece of the integral. Using the fact that the variations $\chi_k, \chi_{2_k}, \eta_k, \eta_{2_k}$ vanish at $k=0$ and $k=N$, the variation of the discrete action sum can be expressed as (111).

$$\begin{aligned}\delta\mathfrak{G}_d &= \sum_{k=1}^{N-1}\frac{1}{h}\chi_k^{\mathrm{T}}\left\{-m_1 F_{2_{k-1}}^{\mathrm{T}}(X_{k-1}+X_{2_{k-1}}) + 2m_1(X_k+X_{2_k})\right.\\ &\left.-m_1 F_{2_k}(X_{k+1}+X_{2_{k+1}}) - h^2\frac{\partial U_k}{\partial X_k}\right\}\\ &+\sum_{k=1}^{N-1}\frac{1}{h}\chi_{2_k}^{\mathrm{T}}\left\{-m_1 F_{2_{k-1}}^{\mathrm{T}}(X_{k-1}+X_{2_{k-1}}) + 2m_1(X_k+X_{2_k})\right.\\ &\left.-m_1 F_{2_k}(X_{k+1}+X_{2_{k+1}}) - m_2 F_{2_{k-1}}^{\mathrm{T}} X_{2_{k-1}} + 2m_2 X_{2_k} - m_2 F_{2_k} X_{2_{k+1}}\right\}\\ &+\sum_{k=1}^{N-1}\mathrm{tr}\left[\eta_k\left\{\frac{1}{h}(-F_{2_{k-1}}^{\mathrm{T}} F_{k-1} R_{k-1} J_{d_1} R_{k-1}^{\mathrm{T}} F_{2_{k-1}} + F_k R_k J_{d_1} R_k^{\mathrm{T}}) - hR_k\frac{\partial U_k^{\mathrm{T}}}{\partial R_k}\right\}\right]\\ &+\sum_{k=1}^{N-1}\mathrm{tr}\left[\eta_{2_k}\left\{\frac{1}{h}(-J_{d_2} F_{2_{k-1}} + F_{2_k} J_{d_2}) + hX_k\frac{\partial U_k^{\mathrm{T}}}{\partial X_k} + hR_k\frac{\partial U_k^{\mathrm{T}}}{\partial R_k}\right\}\right]. \end{aligned} \tag{111}$$

From Hamilton's principle, $\delta\mathfrak{G}_d$ should be zero for all possible variations $\chi_k, \chi_{2_k} \in \mathbb{R}^3$ and $\eta_k, \eta_{2_k} \in \mathfrak{so}(3)$ which vanish at the endpoints. Therefore, in (111), the expressions in the first two braces should be zero, and the expressions in the last two braces should be symmetric since η_k, η_{2_k} are skew-symmetric. After some simplification, we obtain *the discrete equations of relative motion for the full two body problem, in Lagrangian form*, as

$$F_{2_k} X_{k+1} - 2X_k + F_{2_{k-1}}^{\mathrm{T}} X_{k-1} = -\frac{h^2}{m}\frac{\partial U_k}{\partial X_k}, \tag{112}$$

$$\begin{aligned}&F_{k+1} J_{dR_{k+1}} - J_{dR_{k+1}} F_{k+1}^{\mathrm{T}}\\ &\quad = F_{2_k}^{\mathrm{T}}(F_k J_{dR_k} - J_{dR_k} F_k^{\mathrm{T}}) F_{2_k} - h^2 S(M_{k+1}),\end{aligned} \tag{113}$$

$$\begin{aligned}F_{2_{k+1}} J_{d_2} - J_{d_2} F_{2_{k+1}}^{\mathrm{T}} &= F_{2_k}^{\mathrm{T}}(F_{2_k} J_{d_2} - J_{d_2} F_{2_k}^{\mathrm{T}}) F_{2_k}\\ &+ h^2 X_{k+1}\times\frac{\partial U_{k+1}}{\partial X_{k+1}} + h^2 S(M_{k+1}),\end{aligned} \tag{114}$$

$$R_{k+1} = F_{2_k}^{\mathrm{T}} F_k R_k, \tag{115}$$

$$R_{2_{k+1}} = R_{2_k} F_{2_k}. \tag{116}$$

It is natural to express equations of motion for the second body in the inertial frame.

$$x_{2_{k+1}} - 2x_{2_k} + x_{2_{k-1}} = \frac{h^2}{m_2} R_k \frac{\partial U_k}{\partial X_k}. \tag{117}$$

Given $(X_0, R_0, R_{2_0}, X_1, R_1, R_{2_1})$, we can determine F_0 and F_{2_0} from (115) and (116). Solving the implicit equations (113) and (114) gives F_1 and F_{2_1}. Then X_2, R_2 and R_{2_2} are found from (112), (115) and (116), respectively. This yields the discrete Lagrangian map $(X_0, R_0, R_{2_0}, X_1, R_1, R_{2_1}) \mapsto (X_1, R_1, R_{2_1}, X_2, R_2, R_{2_2})$ and this process can be repeated. We can separately reconstruct x_{2_k} using (117).

4.2.2. Discrete equations of motion: Hamiltonian form

Using the discrete Legendre transformation, we can obtain the Hamiltonian map, in terms of reduced variables, that is equivalent to the Lagrangian map given in (112)–(117). We will only sketch the procedure as it is analogous to the approach of the previous section. First, we find expressions for the conjugate momenta variables corresponding to (82) and (83). We compute the discrete Legendre transformation by taking the variation of the discrete reduced Lagrangian as in (84) and (88). Then, we obtain the discrete equations of motion in Hamiltonian form using (3) and (4).

The discrete equations of relative motion for the full two body problem, in Hamiltonian form, can be written as

$$X_{k+1} = F_{2_k}^{\mathrm{T}}\left(X_k + h\frac{\Gamma_k}{m} - \frac{h^2}{2m}\frac{\partial U_k}{\partial X_k}\right), \tag{118}$$

$$\Gamma_{k+1} = F_{2k}^{\mathrm{T}}\left(\Gamma_k - \frac{h}{2}\frac{\partial U_k}{\partial X_k}\right) - \frac{h}{2}\frac{\partial U_{k+1}}{\partial X_{k+1}}, \tag{119}$$

$$\Pi_{k+1} = F_{2_k}^{\mathrm{T}}\left(\Pi_k - \frac{h}{2}M_k\right) - \frac{h}{2}M_{k+1}, \tag{120}$$

$$\begin{aligned}\Pi_{2_{k+1}} &= F_{2_k}^{\mathrm{T}}\left(\Pi_{2_k} + \frac{h}{2}X_k\times\frac{\partial U_k}{\partial X_k} + \frac{h}{2}M_k^{\mathrm{T}}\right)\\ &+\frac{h}{2}X_{k+1}\times\frac{\partial U_{k+1}}{\partial X_{k+1}} + \frac{h}{2}M_{k+1},\end{aligned} \tag{121}$$

$$R_{k+1} = F_{2_k}^{\mathrm{T}} F_k R_k, \tag{122}$$

$$hS\left(\Pi_k - \frac{h}{2}M_k\right) = F_k J_{dR_k} - J_{dR_k} F_k^{\mathrm{T}}, \tag{123}$$

$$hS\left(\Pi_{2_k} + \frac{h}{2}X_k\times\frac{\partial U}{\partial X_k} + \frac{h}{2}M_k\right) = F_{2_k} J_{d_2} - J_{d_2} F_{2_k}^{\mathrm{T}}. \tag{124}$$

It is natural to express equations of motion for the second body in the inertial frame for reconstruction:

$$x_{2_{k+1}} = x_{2_k} + h\frac{\gamma_{2_k}}{m_2} + \frac{h^2}{2m_2}R_k\frac{\partial U_k}{\partial X_k}, \tag{125}$$

$$\gamma_{2_{k+1}} = \gamma_{2_k} + \frac{h}{2}R_k\frac{\partial U_k}{\partial X_k} + \frac{h}{2}R_{k+1}\frac{\partial U_{k+1}}{\partial X_{k+1}}, \tag{126}$$

$$R_{2_{k+1}} = R_{2_k}F_{2_k}. \tag{127}$$

Given $(R_0, X_0, \Pi_0, \Gamma_0, \Pi_{2_0})$, we can determine F_0 and F_{2_0} by solving the implicit Eqs. (123) and (124). Then, X_1 and R_1 are found from (118) and (122), respectively. After that, we can compute Γ_1, Π_1, and Π_{2_1} from (119)–(121). This yields a discrete Hamiltonian map $(R_0, X_0, \Pi_0, \Gamma_0, \Pi_{2_0}) \mapsto (R_1, X_1, \Pi_1, \Gamma_1, \Pi_{2_1})$, and this process can be repeated. x_{2_k}, γ_{2_k} and R_{2_k} can be updated separately using (125)–(127), respectively, for reconstruction.

4.3. Numerical considerations

Properties of the variational integrators: Variational integrators exhibit a discrete analogue of Noether's theorem [10], and symmetries of the discrete Lagrangian result in conservation of the corresponding momentum maps. Our choice of discrete Lagrangian is such that it inherits the symmetries of the continuous Lagrangian. Therefore, all the conserved momenta in the continuous dynamics are preserved by the discrete dynamics.

The proposed variational integrators are expressed in terms of Lie group computations [12]. During each integration step, $F_{i_k} \in \mathrm{SO}(3)$ is obtained by solving an implicit equation, and R_{i_k} is updated by multiplication with F_{i_k}. Since SO(3) is closed under matrix multiplication, the attitude matrix $R_{i_{k+1}}$ remains in SO(3). We make this more explicit in Section 4.4 by expressing F_{i_k} as the exponential function of an element of the Lie algebra $\mathfrak{so}(3)$.

An adjoint integration method is the inverse map of the original method with reversed time-step. An integration method is called self-adjoint or symmetric if it is identical to its adjoint; a self-adjoint method always has even order. Our discrete Lagrangian is chosen to be self-adjoint, and therefore the corresponding variational integrators are second-order accurate.

Higher-order methods: While the numerical methods we present in this paper are second-order, it is possible to apply the symmetric composition methods, introduced in [25], to construct higher-order versions of the Lie group variational integrators introduced here. Given a basic numerical method represented by the flow map Φ_h, the composition method is obtained by applying the basic method using different step sizes,

$$\Psi_h = \Phi_{\lambda_s h} \circ \cdots \circ \Phi_{\lambda_1 h},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{R}$. In particular, the Yoshida symmetric composition method for composing a symmetric method of order 2 into a symmetric method of order 4 is obtained when $s = 3$, and

$$\lambda_1 = \lambda_3 = \frac{1}{2 - 2^{1/3}}, \qquad \lambda_2 = -\frac{2^{1/3}}{2 - 2^{1/3}}.$$

Alternatively, by adopting the formalism of higher-order Lie group variational integrators introduced in [21] in conjunction with the Rodrigues formula, one can directly construct higher-order generalizations of the Lie group methods presented here.

Reduction of orthogonality loss due to roundoff error: In the Lie group variational integrators, the numerical solution is made to automatically remain on the rotation group by requiring that the numerical solution is updated by matrix multiplication with the exponential of a skew symmetric matrix.

Since the exponential of the skew symmetric matrix is orthogonal to machine precision, the numerical solution will only deviate from orthogonality due to the accumulation of roundoff error in the matrix multiplication, and this orthogonality loss grows linearly with the number of time-steps taken.

One possible method of addressing this issue is to use the Baker–Campbell–Hausdorff (BCH) formula to track the updates purely at the level of skew symmetric matrices (the Lie algebra). This allows us to find a matrix $C(t)$, such that,

$$\exp(tA)\exp(tB) = \exp C(t).$$

This matrix $C(t)$ satisfies the following differential equation,

$$\dot{C} = A + B + \frac{1}{2}[A - B, C] + \sum_{k \geqslant 2}\frac{B_k}{k!}\mathrm{ad}_C^k(A + B),$$

with initial value $C(0) = 0$, and where B_k denotes the Bernoulli numbers, and $\mathrm{ad}_C A = [C, A] = CA - AC$.

The problem with this approach is that the matrix $C(t)$ is not readily computable for arbitrary A and B, and in practice, the series is truncated, and the differential equation is solved numerically.

An error is introduced in truncating the series, and numerical errors are introduced in numerically integrating the differential equations. Consequently, while the BCH formula could be used solely at the reconstruction stage to ensure that the numerical attitude always remains in the rotation group to machine precision, the truncation error would destroy the symplecticity and momentum preserving properties of the numerical scheme.

However, by combining the BCH formula with the Rodrigues formula in constructing the discrete variational principle, it might be possible to construct a Lie group variational integrator that tracks the reconstructed trajectory on the rotation group at the level of a curve in the Lie algebra, while retaining its structure-preservation properties.

4.4. Computational approach

The structure of the discrete equations of motion given in (79), (95), (113), (114), (123), and (124) suggests

a specific computational approach. For a given $g \in \mathbb{R}^3$, we have to solve the following Lyapunov-like equation to find $F \in \mathrm{SO}(3)$ at each integration step.

$$FJ_d - J_d F^{\mathrm{T}} = S(g). \tag{128}$$

We now introduce an iterative approach to solve (128) numerically. An element of a Lie group can be expressed as the exponential of an element of its Lie algebra, so $F \in \mathrm{SO}(3)$ can be expressed as an exponential of $S(f) \in \mathfrak{so}(3)$ for some vector $f \in \mathbb{R}^3$. The exponential can be written in closed form, using Rodrigues' formula,

$$F = \mathrm{e}^{S(f)} = I_{3\times 3} + \frac{\sin\|f\|}{\|f\|} S(f) + \frac{1-\cos\|f\|}{\|f\|^2} S(f)^2. \tag{129}$$

Substituting (129) into (128), we obtain

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

where (9) and (13) are used. Thus, (128) is converted into the equivalent vector equation $g = G(f)$, where $G : \mathbb{R}^3 \mapsto \mathbb{R}^3$ is

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

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

$$f_{i+1} = f_i + \nabla G(f_i)^{-1}(g - G(f_i)). \tag{130}$$

We iterate until $\|g - G(f_i)\| < \epsilon$ for a small tolerance $\epsilon > 0$. The Jacobian $\nabla G(f)$ in (130) can be expressed as

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

Numerical simulations show that 3 or 4 iterations are sufficient to achieve a tolerance of $\epsilon = 10^{-15}$.

5. Numerical simulations

The variational integrator in Hamiltonian form given in (118)–(127) is used to simulate the dynamics of two simple dumbbell bodies acting under their mutual gravity.

5.1. Full body problem defined by two dumbbell bodies

Each dumbbell model consists of two equal rigid spheres and a massless rod as shown in Fig. 3. The gravitational potential of the two dumbbell models is given by

$$U(X,R) = -\sum_{p,q=1}^{2} \frac{Gm_1m_2/4}{\|X + \rho_{2_p} + R\rho_{1_q}\|},$$

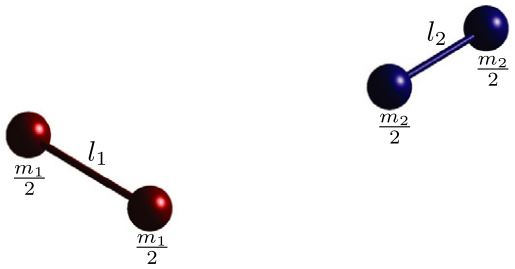


Fig. 3. Dumbbell model of the full body problem.

where G is the universal gravitational constant, $m_i \in \mathbb{R}$ is the total mass of the ith dumbbell, and $\rho_{i_p} \in \mathbb{R}^3$ is a vector from the origin of the body fixed frame to the pth sphere of the ith dumbbell in the ith body fixed frame. The vectors $\rho_{i_1} = [l_i/2, 0, 0]^{\mathrm{T}}$, $\rho_{i_2} = -\rho_{i_1}$, where l_i is the length between the two spheres.

Normalization: Mass, length and time dimensions are normalized as follows:

$$\bar{m}_i = \frac{m_i}{m},$$
$$\overline{X}_i = \frac{X_i}{l},$$
$$\bar{t} = \sqrt{\frac{G(m_1+m_2)}{l^3}}\, t,$$

where $m = \frac{m_1 m_2}{m_1+m_2}$, and l is chosen as the initial horizontal distance between the center of mass of the two dumbbells. The time is normalized so that the orbital period is of order unity. Over-bars denote normalized variables. We can expresses the equations of motion in terms of the normalized variables. For example, (55) can be written as

$$\overline{V}' + \overline{\Omega}_2 \times \overline{V} = -\frac{\partial \overline{U}}{\partial \overline{X}},$$

where $'$ denotes a derivative with respect to $\bar{t}$. The normalized gravitational potential and its partial derivatives are given by

$$\overline{U} = -\frac{1}{4}\sum_{p,q=1}^{2} \frac{1}{\|\overline{X} + \bar{\rho}_{2_p} + R\bar{\rho}_{1_q}\|},$$
$$\frac{\partial \overline{U}}{\partial \overline{X}} = \frac{1}{4}\sum_{p,q=1}^{2} \frac{\overline{X} + \bar{\rho}_{2_p} + R\bar{\rho}_{1_q}}{\|\overline{X} + \bar{\rho}_{2_p} + R\bar{\rho}_{1_q}\|^3},$$
$$\frac{\partial \overline{U}}{\partial R} = \frac{1}{4}\sum_{p,q=1}^{2} \frac{(\overline{X} + \bar{\rho}_{2_p})\bar{\rho}_{1_q}^{\mathrm{T}}}{\|\overline{X} + \bar{\rho}_{2_p} + R\bar{\rho}_{1_q}\|^3}.$$

Conserved quantities: The total energy E is conserved:

$$\begin{aligned}E = {} & \frac{1}{2}m_1\|V + V_2\|^2 + \frac{1}{2}m_2\|V_2\|^2 + \frac{1}{2}\mathrm{tr}[S(\Omega)J_{d_R}S(\Omega)^{\mathrm{T}}] \\ & + \frac{1}{2}\mathrm{tr}[S(\Omega_2)J_{d_2}S(\Omega_2)^{\mathrm{T}}] + U(X,R).\end{aligned}$$

The total linear momentum $\gamma_T \in \mathbb{R}^3$, and the total angular momentum about the mass center of the system $\pi_T \in \mathbb{R}^3$, in the inertial frame, are also conserved:

$$\gamma_T = R_2\{m_1(V + V_2) + m_2 V_2\},$$
$$\pi_T = R_2\{mX \times V + J_R\Omega + J_2\Omega_2\}.$$

5.2. Simulation results

The properties of the two dumbbell bodies are chosen to be

$$\bar{m}_1 = 1.5, \quad \bar{l}_1 = 0.25, \quad \overline{J}_1 = \mathrm{diag}[0.0004, 0.0238, 0.0238],$$
$$\bar{m}_2 = 3, \quad \bar{l}_2 = 0.5, \quad \overline{J}_2 = \mathrm{diag}[0.0030, 0.1905, 0.1905].$$

The mass and length of the second dumbbell are twice that of the first dumbbell. The initial conditions are chosen such that the total linear momentum in the inertial frame is zero and the total energy is positive.

$$\overline{X}_0 = [1, 0, 0.3], \quad \overline{V}_0 = [0, 1, 0],$$
$$\overline{\Omega}_{1_0} = [0, 0, 9], \quad R_0 = I_{3\times 3},$$
$$\bar{x}_{2_0} = [-0.33, 0, -0.1], \quad \bar{v}_{2_0} = [0, -0.33, 0],$$
$$\overline{\Omega}_{2_0} = [0, 0, 0], \quad R_{2_0} = I_{3\times 3}.$$

Simulation results obtained using the Lie group variational integrator are given in Figs. 4 and 5. Fig. 4 shows the trajectory of the two dumbbells in the inertial frame. Fig. 5(a) shows the evolution of the normalized energy, where the upper figure gives the history of the translational kinetic energy and the rotational kinetic energy, and the lower figure shows the interchange between the total kinetic energy and the gravitational potential energy. Fig. 5(b) shows the evolution of the theoretically conserved quantities, where the upper figure is the history of the total energy, and the lower figure is the error in the rotation matrix.

Initially, the first dumbbell rotates around the vertical e_3 axis, and the second dumbbell does not rotate. Since the angular velocity of the first dumbbell is relatively large, the rotational kinetic energy initially exceeds the translational kinetic energy. As the two dumbbells orbit around each other, the second dumbbell starts to rotate, the rotational kinetic energy increases, and the translational kinetic energy decreases slightly for about 6 normalized units of time. At 9 units of time, the distance between the two dumbbells reaches its minimal separation, and the potential energy is transformed into kinetic energy, especially translational kinetic energy. After that, two dumbbells continue to move apart, and the translational energy and the rotational energy equalize. (A simple animation of this motion can be found at http://www.umich.edu/~tylee.) This shows some of the interesting dynamics that the full body problem can exhibit. The non-trivial interchange between rotational kinetic energy, translational kinetic energy, and potential energy may yield complicated motions that cannot be observed in the classical two body problem.

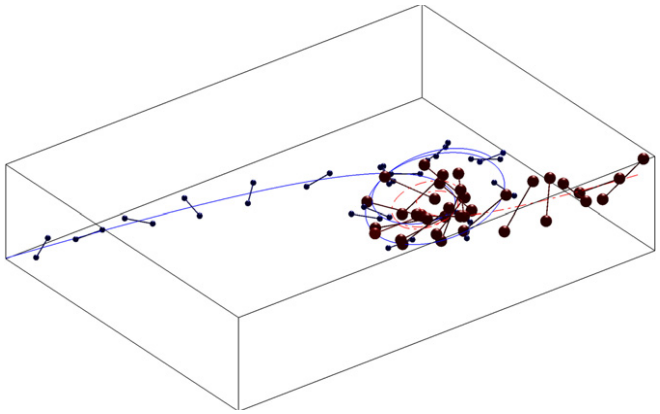


Fig. 4. Trajectory in the inertial frame.

The Lie group variational integrator preserves the total energy and the geometry of the configuration space. The maximum deviation of the total energy is 2.6966×10^{-7}, and the maximum value of the rotation matrix error $\|I - R^{\mathrm{T}}R\|$ is 2.8657×10^{-13}.

As a comparison, Fig. 6 shows simulation results obtained by numerically integrating the continuous equations of motion (63)–(70) using a standard Runge–Kutta

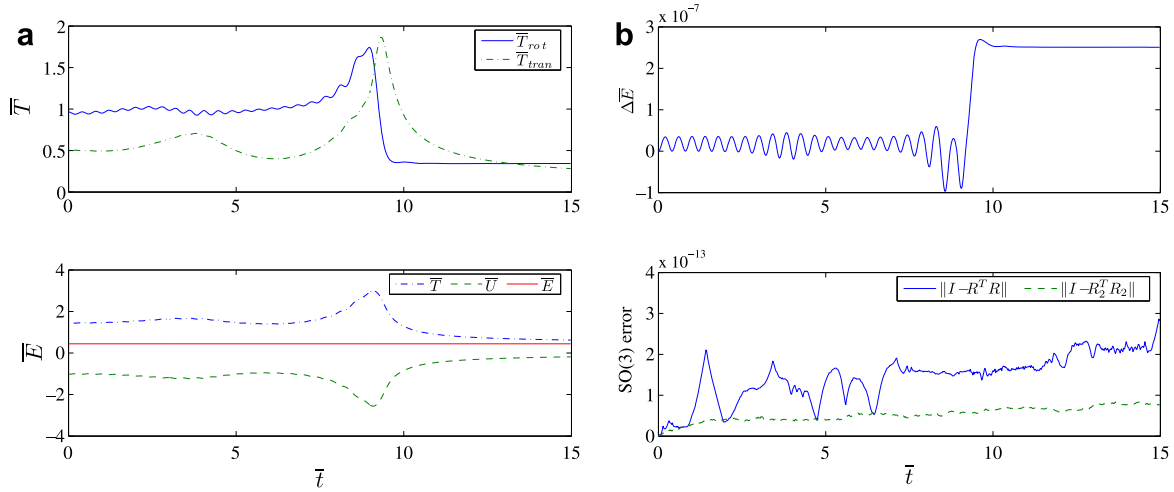


Fig. 5. Lie group variational integrator: (a) interchange of energy and (b) conserved quantities.

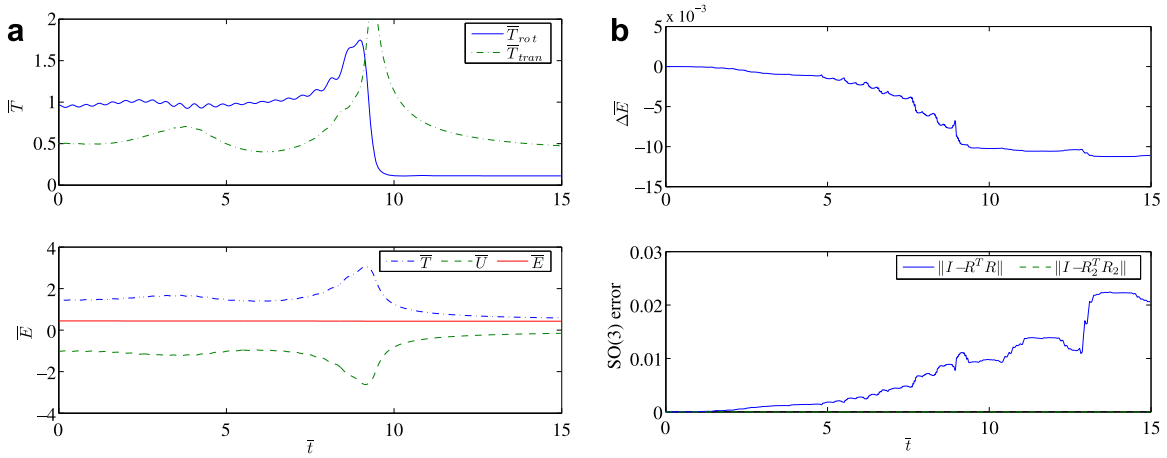


Fig. 6. Runge–Kutta method: (a) interchange of energy and (b) conserved quantities.

method. The rotational and the translational kinetic energy responses are similar to those given in Fig. 5 prior to the close encounter. However, it fails to simulate the rapid interchange of the energy near the minimal separation of the two dumbbells. The deviation of the total energy is relatively large, with a maximum deviation of 1.1246×10^{-2}. Also, the energy transfer is quite different from that given in Fig. 5(a). The Runge–Kutta method does not preserve the geometry of the configuration space, as the discrete trajectory rapidly drifts off the rotation group to give a maximum rotation matrix error of 2.2435×10^{-2}. As the gravity and momentum between the two dumbbells depend on the relative attitude, the errors in the rotation matrix limits the applicability of standard techniques to long time simulations.

An extensive computational comparison between the Lie group variational integrator to other geometric integrators such as symplectic Runge–Kutta method and Lie group method can be found in [26].

6. Conclusions

Eight different forms of the equations of motion for the full body problem are derived. The continuous equations of motion and variational integrators are derived both in the inertial and relative frames, and each set of equations of motion is expressed in both Lagrangian and Hamiltonian form. The relationships between these equations of motion are summarized in Fig. 7. This commutative cube was originally given in [22]. In the figure, dashed arrows represent discretization from the continuous systems on the left face of the cube to the discrete systems on the right face. Vertical arrows represent reduction from the full (inertial) equations on the top face to the reduced (relative) equations on the bottom face. Front and back faces represent Lagrangian and Hamiltonian forms, respectively. The corresponding equation numbers are also indicated in parentheses.

It is shown that the equations of motion for the full body problem can be derived systematically, using proper

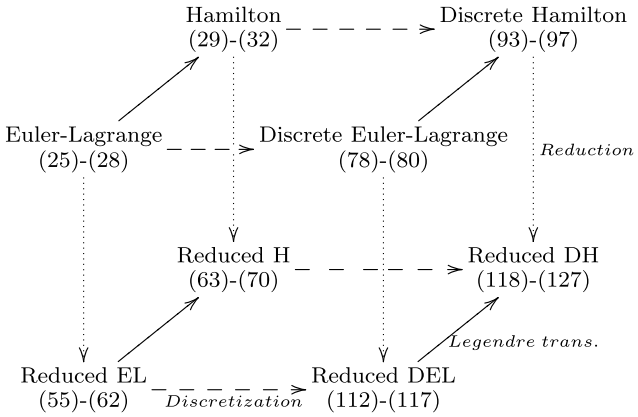


Fig. 7. Commutative cube of the equations of motion.

Lie group variations, from Hamilton's principle. The proposed variational integrators preserve the momenta and symplectic form of the continuous dynamics, exhibit good energy properties, and they also conserve the geometry of the configuration space since they are based on Lie group computations. The main contribution of this paper is the combination of variational integrators and Lie group computations, developed for the full body problem. Hence, the resulting numerical integrators conserve the first integrals as well as the geometry of the configuration space of the full body dynamics.

Appendix A. Variations of reduced variables

A.1. Continuous variables

The variations of the reduced variables given in (43)–(47) are derived in this section. The variations of the reduced variables can be obtained from the definitions of the reduced variables, and the variations of the original variables.

The variation of $X = R_2^{\mathrm{T}}(x_1 - x_2)$ is given by

$$\delta X = \delta R_2^{\mathrm{T}}(x_1 - x_2) + R_2(\delta x_1 - \delta x_2).$$

Substituting (16) into the above equation, we obtain

$$\delta X = -\eta_2 R_2^{\mathrm{T}}(x_1 - x_2) + R_2(\delta x_1 - \delta x_2) = -\eta_2 X + \chi,$$

where the reduced variation $\chi : [t_0, t_f] \mapsto \mathbb{R}^3$ is defined to be $\chi = R_2(\delta x_1 - \delta x_2)$.

From the definition of $\Omega = R\Omega_1$ and (10), $S(\delta\Omega)$ is given by

$$\begin{aligned} S(\delta\Omega) &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} S(R^\epsilon \Omega_1^\epsilon) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} R^\epsilon S(\Omega_1^\epsilon) R^{\epsilon T} \\ &= \delta R S(\Omega_1) R^{\mathrm{T}} + R S(\delta\Omega_1) R^{\mathrm{T}} + R S(\Omega_1) \delta R^{\mathrm{T}}. \end{aligned}$$

Substituting (42) and (17) into the above equation, we obtain

$$\begin{aligned} S(\delta\Omega) &= \{\eta - \eta_2\} R S(\Omega_1) R^{\mathrm{T}} + R\{\dot\eta_1 + S(\Omega_1)\eta_1 - \eta_1 S(\Omega_1)\} R^{\mathrm{T}} \\ &\quad + R S(\Omega_1) R^{\mathrm{T}}\{-\eta + \eta_2\} \\ &= \{\eta - \eta_2\} S(R\Omega_1) + R\dot\eta_1 R^{\mathrm{T}} + S(R\Omega_1) R\eta_1 R^{\mathrm{T}} \\ &\quad - R\eta_1 R^{\mathrm{T}} S(R\Omega_1) + S(R\Omega_1)\{-\eta + \eta_2\}. \end{aligned}$$

Since $\eta = R\eta_1 R^{\mathrm{T}}$ and $\Omega = R\Omega_1$, the above equation reduces to

$$S(\delta\Omega) = -\eta_2 S(\Omega) + R\dot\eta_1 R^{\mathrm{T}} + S(\Omega)\eta_2. \tag{A.1}$$

From the definition of $R = R_2^{\mathrm{T}} R_1$, $\dot R$ is given by

$$\dot R = \dot R_2^{\mathrm{T}} R_1 + R_2^{\mathrm{T}} \dot R_1 = -S(\Omega_2) R + S(\Omega) R. \tag{A.2}$$

Then, $\dot\eta$ can be written as

$$\begin{aligned} \dot\eta &= R\dot\eta_1 R^{\mathrm{T}} + \dot R \eta_1 R^{\mathrm{T}} + R\eta_1 \dot R^{\mathrm{T}} \\ &= R\dot\eta_1 R^{\mathrm{T}} + \{S(\Omega) - S(\Omega_2)\}\eta - \eta\{S(\Omega) - S(\Omega_2)\}. \end{aligned} \tag{A.3}$$

Substituting (A.3) into (A.1), we obtain $S(\delta\Omega)$ in terms of η, η_2 as

$$\begin{aligned} S(\delta\Omega) &= \dot\eta - S(\Omega)\eta + \eta S(\Omega) + S(\Omega)\eta_2 - \eta_2 S(\Omega) + S(\Omega_2)\eta \\ &\quad - \eta S(\Omega_2), \end{aligned}$$

which is equivalent to (44).

The variation of $V = R_2^{\mathrm{T}}(\dot x_1 - \dot x_2)$ is given by

$$\begin{aligned} \delta V &= \delta R_2^{\mathrm{T}}(\dot x_1 - \dot x_2) + R_2^{\mathrm{T}}(\delta\dot x_1 - \delta\dot x_2) \\ &= -\eta_2 V + R_2^{\mathrm{T}}(\delta\dot x_1 - \delta\dot x_2). \end{aligned} \tag{A.4}$$

From the definition of $\chi = R_2^{\mathrm{T}}(\delta x_1 - \delta x_2)$, $\dot\chi$ is given by

$$\begin{aligned} \dot\chi &= \dot R_2^{\mathrm{T}}(\delta x_1 - \delta x_2) + R_2^{\mathrm{T}}(\delta\dot x_1 - \delta\dot x_2) \\ &= -S(\Omega_2)\chi + R_2^{\mathrm{T}}(\delta\dot x_1 - \delta\dot x_2). \end{aligned} \tag{A.5}$$

Substituting (A.5) into (A.4), we obtain

$$\delta V = -\eta_2 V + \dot\chi + S(\Omega_2)\chi,$$

which is equivalent to (45). The variation δV_2 can be derived in the same way, and $S(\delta\Omega_2)$ is given in (17).

A.2. Discrete variables

The variation of the reduced variables δF_k given in (106) is derived in this section. From (75) and (98), the variation δF_{1_k} is written as

$$\begin{aligned} \delta F_{1_k} &= -\eta_{1_k} F_{1_k} + F_{1_k}\eta_{1_{k+1}} \\ &= -R_k^{\mathrm{T}}\eta_k F_k R_k + R_k^{\mathrm{T}} F_k R_k R_{k+1}^{\mathrm{T}}\eta_{k+1} R_{k+1}, \end{aligned}$$

where $\eta_k \in \mathfrak{so}(3)$ is defined as $\eta_k = R_k\eta_{1_k}R_k^{\mathrm{T}}$. Since $F_k R_k R_{k+1}^{\mathrm{T}} = F_k R_k (R_k^{\mathrm{T}} F_k^{\mathrm{T}} F_{2_k}) = F_{2_k}$, we have

$$\delta F_{1_k} = R_k^{\mathrm{T}}(-\eta_k F_k + F_{2_k}\eta_{k+1} F_{2_k}^{\mathrm{T}} F_k) R_k.$$

Then, the variation δF_k is given by

$$\begin{aligned} \delta F_k &= \delta R_k F_{1_k} R_k^{\mathrm{T}} + R_k \delta F_{1_k} R_k^{\mathrm{T}} + R_k F_{1_k} \delta R_k^{\mathrm{T}} \\ &= -\eta_{2_k} F_k + F_{2_k}\eta_{k+1} F_{2_k}^{\mathrm{T}} F_k + F_k(-\eta_k + \eta_{2_k}), \end{aligned}$$

which is equivalent to (106).

References

[1] W.S. Koon, J.E. Marsden, S. Ross, M. Lo, D.J. Scheeres, Geometric mechanics and the dynamics of asteroid pairs, in: New Trends in Astrodynamics Conference, 2003.

[2] A.J. Maciejewski, Reduction, relative equilibria and potential in the two rigid bodies problem, Celest. Mech. Dynam. Astron. 63 (1995) 1–28.

[3] D.J. Scheeres, Stability in the full two-body problem, Celest. Mech. Dynam. Astron. 83 (2002) 155–169.

[4] D.J. Scheeres, Stability of relative equilibria in the full two-body problem, in: New Trends in Astrodynamics Conference, 2003.

[5] J.L. Margot, M.C. Nolan, L.A.M. Benner, S.J. Ostro, R.F. Jurgens, J.D. Giorgini, M.A. Slade, D.B. Campbell, Binary asteroids in the near-earth object population, Science 296 (2002) 1445–1448.

[6] D.J. Scheeres, S. Augenstein, Spacecraft motion about binary asteroids, in: Astrodynamics Specialist Conferences, 2003.

[7] F. Gabern, W.S. Koon, J.E. Marsden, Spacecraft dynamics near a binary asteroid, in: Proceedings of the Fifth International Conference on Dynamical Systems and Differential Equations, 2004.

[8] D.J. Scheeres, J. Bellerose, The restricted hill full 4-body problem: application to spacecraft motion about binary asteroids, Dynam. Syst.: Int. J. 20 (1) (2005) 23–44.

[9] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration, Springer, 2000.

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

[11] E. Hairer, Backward analysis of numerical integrators and symplectic methods, Ann. Numer. Math. 1 (1–4) (1994) 107–132, scientific computation and differential equations (Auckland, 1993).

[12] A. Iserles, H.Z. Munthe-Kaas, S.P. Nørsett, A. Zanna, Lie-group methods, Acta Numer. 9 (2000) 215–365.

[13] J. Moser, A.P. Veselov, Discrete versions of some classical integrable systems and factorization of matrix polynomials, Commun. Math. Phys. 139 (1991) 217–243.

[14] J.M. Wendlandt, J.E. Marsden, Mechanical integrator derived from a discrete variational principle, Physica D 106 (1997) 223–246.

[15] J.C. Simo, N. Tarnow, K.K. Wong, Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics, Comput. Methods Appl. Mech. Engrg. 100 (1992) 63–116.

[16] P. Krysl, Explicit momentum-conserving integrator for dynamics of rigid bodies approximating the midpoint Lie algorithm, Int. J. Numer. Methods Engrg. 63 (15) (2005) 2171–2193.

[17] T. Lee, M. Leok, N.H. McClamroch, A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum, in: Proceedings of the IEEE Conference on Control Applications, 2005, pp. 962–967.

[18] J.E. Marsden, S. Pekarsky, S. Shkoller, Discrete Euler–Poincaré and Lie–Poisson equations, Nonlinearity 12 (1999) 1647–1662.

[19] J.E. Marsden, S. Pekarsky, S. Shkoller, Symmetry reduction of discrete Lagrangian mechanics on Lie groups, J. Geometry Phys. 36 (2000) 139–150.

[20] A.K. Sanyal, J. Shen, N.H. McClamroch, Variational integrators for mechanical systems with configuration dependent inertia, Numer. Methods Engrg., submitted for publication.
[21] M. Leok, Foundations of Computational Geometric Mechanics, Ph.D. thesis, California Institute of Technology, 2004.
[22] S.M. Jalnapurkar, M. Leok, J.E. Marsden, M. West, Discrete Routh reduction, J. Phys. A: Math. General 39 (2006) 5521–5544.
[23] J.E. Marsden, T.S. Ratiu, Introduction to Mechanics and Symmetry, Springer, 1999.
[24] H. Cendra, J.E. Marsden, T.S. Ratiu, Lagrangian reduction by stages, Memoirs of the American Mathematical Society 152 (722) (2001).
[25] H. Yoshida, Construction of high order symplectic integrators, Phys. Lett. A 150 (1990) 262–268.
[26] T. Lee, M. Leok, N.H. McClamroch, Lie group variational integrators for the full body problem in orbital mechanics, Celest. Mech. Dynam. Astron., in press, doi:10.1007/S10569-007-9073-X.

Celestial Mech Dyn Astr (2007) 98:121–144
DOI 10.1007/s10569-007-9073-x

ORIGINAL ARTICLE

Lie group variational integrators for the full body problem in orbital mechanics

Taeyoung Lee · Melvin Leok · N. Harris McClamroch

Received: 12 October 2006 / Revised: 1 February 2007 / Accepted: 15 February 2007 /
Published online: 12 April 2007

Abstract Equations of motion, referred to as full body models, are developed to describe the dynamics of rigid bodies acting under their mutual gravitational potential. Continuous equations of motion and discrete equations of motion are derived using Hamilton's principle. These equations are expressed in an inertial frame and in relative coordinates. The discrete equations of motion, referred to as a Lie group variational integrator, provide a geometrically exact and numerically efficient computational method for simulating full body dynamics in orbital mechanics; they are symplectic and momentum preserving, and they exhibit good energy behavior for exponentially long time periods. They are also efficient in only requiring a single evaluation of the gravity forces and moments per time step. The Lie group variational integrator also preserves the group structure without the use of local charts, reprojection, or constraints. Computational results are given for the dynamics of two rigid dumbbell bodies acting under their mutual gravity; these computational results demonstrate the superiority of the Lie group variational integrator compared with integrators that are not symplectic or do not preserve the Lie group structure.

Keywords Symplectic integrator · Variational integrator · Lie group method · Full rigid body problem

T. Lee (✉)· N. H. McClamroch
Department of Aerospace Engineering,
The University of Michigan, Ann Arbor,
MI 48109 USA
e-mail: tylee@umich.edu

M. Leok
Department of Mathematics, Purdue University,
West Lafayette, IN 47907 USA

1 Introduction

The full body problem in orbital mechanics treats the dynamics of non-spherical rigid bodies in space interacting under their mutual potential. Since the mutual gravitational potential of distributed rigid bodies depends on both the position and the attitude of the bodies, the translational and the rotational dynamics are coupled in the full body problem. For example, the orbital motion and the attitude dynamics of a very large spacecraft in the Earth's gravity field are coupled, and the dynamics of a binary asteroid pair, with non-spherical mass distributions of the bodies, involves coupled orbital and attitude dynamics. Recently, interest in the full body problem has increased, as it is estimated that up to 16% of near-earth asteroids are binaries (Margot et al. 2002).

These full body dynamics arise from Lagrangian and Hamiltonian mechanics; they are characterized by symplectic, momentum and energy preserving properties. These geometric features determine the qualitative behavior of the full body dynamics, for example stability conditions (Scheeres 2002), and they can serve as a basis for further theoretical study of the full body problem. The configuration space of the full body dynamics has a Lie group structure referred to as the special Euclidean group, SE(3). The representation used for the attitude of the bodies should be globally defined since the complicated dynamics would require frequent changes of coordinates when using representations that are only locally defined.

However, general numerical integration methods, including the widely used explicit (non-symplectic) Runge-Kutta schemes, neither preserve the Lie group structure nor these geometric properties. (Hairer et al 2006). They fail to preserve the conserved quantities such as total energy and angular momentum, which determine the qualitative behavior of the full body dynamics. Attitude errors tend to accumulate, and this attitude degradation causes significant errors in the computation of gravitational forces and moments. The accuracy of such general purpose integrators also rapidly degrades as the simulation time increases (Fahnestock et al 2006).

Moser and Veselov (1991), Wendlandt and Marsden (1997) developed numerical integrators for a free rigid body by imposing an orthogonal constraint on the attitude variables, and by using unit quaternions, respectively. The idea of using the Lie group structure and the exponential map to numerically compute rigid body dynamics arises in the work of Simo et al. (1992), and in the work by Krysl (2005). Marsden et al. (1999) and Marsden et al. (2000) introduce discrete Euler–Poincaré and Lie–Poisson equations, where the discrete dynamics on a Lie group are reduced to the dynamics on the corresponding Lie algebra in the absence of potential. Lie-Poisson integrators have been developed by splitting the Hamiltonian into separate integrable terms for an elliptical body (Touma and Wisdom, 1994; Breiter et al. 2005a) and for the secular spin dynamics of a rigid body (Breiter et al. 2005b).

Variational integrators and Lie group methods provide a systematic method of constructing structure-preserving numerical integrators. The idea of the variational approach is to discretize Hamilton's principle (Marsden and West 2001). The numerical integrator obtained from discrete Hamilton's principle exhibits excellent energy properties (Hairer 1994), conserves first integrals associated with symmetries by a discrete version of Noether's theorem, and preserves the symplectic structure. Lie group methods consist of numerical integrators that preserve the geometry of the configuration space by automatically remaining on the Lie group (Iserles et al. 2000).

In this paper, the Lie group approach is explicitly adopted in the context of a variational integrator for the full rigid bodies problem. This is an extension of the geometric integrator for rigid body attitude dynamics on the rotation group SO(3) by Lee et al. (2005) to attitude and translational dynamics on the special Euclidean group SE(3), and here reduced equations of motion are also developed in relative coordinates. This unified integrator, hereafter referred to as the Lie Group Variational Integrator (or LGVI for short), is symplectic and momentum preserving, and it exhibits good total energy behavior for exponentially long time periods. It also preserves the Euclidian Lie group structure without the use of local charts, reprojection, or constraints. The exact geometric properties of the discrete flow not only provides improved qualitative behavior, but also results in accurate long-time simulation. This provides a uniform method that can be applied to rigid bodies acting under any type of potential that depends on the position and the attitude, but we focus on the application to astrodynamics problems with the gravitational potential in this paper. This development has been presented in Lee et al. (2007); the present paper emphasizes the development of this approach for full body problems in orbital mechanics and the special computational features of this approach for full body problems in orbital mechanics. In addition, we make a computational comparison between the Lie group variational integrator to other geometric integrators such as symplectic Runge-Kutta method and Lie group method (Hairer et al. 2006).

Numerical simulation of the full body problem involves a large computational burden in computing mutual gravitational forces and moments, which are usually represented by a finite series approximation for the double volume integration (Werner and Scheeres 2005). The forces and moments must be reevaluated for any position change or any orientation change, not only at each time step but at each sub-step involved in the differencing scheme behind any general purpose numerical integrator. Therefore the choice of numerical integrator can significantly amplify the burden of computing the gravity forces and moments. The LGVI minimizes the computational burden in the sense that it requires only one force and torque evaluation per integration step for second order accuracy. It has been shown that the LGVI yields a numerically efficient computational algorithm for the full body problem (Fahnestock et al. 2006).

This paper is organized as follows. The continuous equations of motion and Lie group variational integrators are derived in Sects. 2 and 3, respectively. Numerical simulations for two rigid dumbbell bodies are presented in Sect. 4.

2 Continuous time full body models

Maciejewski (1995) presented the continuous equations of motion for the full body problem in Hamiltonian form without providing a formal derivation. Here, we show that the equations can be derived from Hamilton's variational principle using the Lagrangian formalism. The proper form for the variations of Lie group elements in the configuration space leads to a systematic derivation of the equations of motion. In this section, we summarize those procedures; the Lie group variational integrators presented in Sect. 3 are obtained by following a similar procedure using the discrete Hamilton's principle. Additional details in this development can be found in (Lee et al. 2007).

2.1 Inertial coordinates

The configuration space of a rigid body is $\mathrm{SE}(3) = \mathbb{R}^3 \circledS \mathrm{SO}(3)$, where $\mathrm{SO}(3)$ denotes the group of 3×3 orthogonal matrices with unit determinant, and $\circledS$ represents a semi-direct product. We derive continuous equations of motion for n rigid bodies. We define an inertial frame and a body-fixed frame for each body, and assume that the origin of the ith body-fixed frame is located at the center of mass of the ith body.

For the ith body, the position of the center of mass in the inertial frame, and the attitude, which is a rotation matrix from the body-fixed frame to the inertial frame, are represented by $(x_i, R_i) \in \mathrm{SE}(3)$. The translational velocity in the inertial frame and the angular velocity in the body-fixed frame are represented by $v_i, \Omega_i \in \mathbb{R}^3$. The subscript i denotes the ith rigid body. The kinematic equations are given by

$$\dot{x}_i = v_i \tag{1}$$

$$\dot{R}_i = R_i S(\Omega_i), \tag{2}$$

where $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is the isomorphism between the Lie algebra $\mathfrak{so}(3)$, which represents 3×3 skew-symmetric matrices, and $\mathbb{R}^3$ defined by: $S(x)y = x \times y$ for any $x, y \in \mathbb{R}^3$. The mass and the moment of inertia matrix of the ith body is denoted by $m_i \in \mathbb{R}$ and $J_i \in \mathbb{R}^{3\times3}$, respectively. We construct a nonstandard moment of inertia matrix $J_{d_i} \in \mathbb{R}^{3\times3}$ by

$$J_{d_i} = \int_{\mathcal{B}_i} \rho_i \rho_i^T \, dm_i, \tag{3}$$

where $\rho_i \in \mathbb{R}^3$ is the position of a mass element of the ith body in its body-fixed frame. It can be shown that the standard moment of inertia matrix $J_i = \int_{\mathcal{B}_i} S(\rho_i)^T S(\rho_i) dm_i \in \mathbb{R}^{3\times3}$ is related to the nonstandard moment of inertia matrix by the following properties.

$$J_i = \mathrm{tr}\big[J_{d_i}\big] I_{3\times3} - J_{d_i}, \tag{4}$$

$$S(J_i\Omega_i) = S(\Omega_i)J_{d_i} + J_{d_i}S(\Omega_i)\,, \tag{5}$$

for any $\Omega_i \in \mathbb{R}^3$. Conversely, one can obtain the nonstandard moment of inertia from the standard momentum of inertia from the following relation,

$$J_{d_i} = \frac{1}{2}\mathrm{tr}[J_i]\, I_{3\times3} - J_i\,. \tag{6}$$

The linear momentum in the inertial frame and the angular momentum in the body-fixed frame are denoted by $\gamma_i = m_i v_i$ and $\Pi_i = J_i\Omega_i \in \mathbb{R}^3$, respectively, for the ith body.

The procedure for deriving the continuous equations of motion is shown in Fig. 1. For the given configuration space, we find an expression for the Lagrangian and the action integral. Hamilton's principle, which involves taking the variation of the action integral, yields the Euler–Lagrange equation. The Legendre transformation gives Hamilton's equation, which is equivalent to the Euler–Lagrange equation.

Lagrangian: Given $(x_i, R_i) \in \mathrm{SE}(3)$, the inertial position of a mass element of the ith body is given by $x_i + R_i\rho_i$, where $\rho_i \in \mathbb{R}^3$ denotes the position of the mass element in the body-fixed frame. Then, the kinetic energy of the ith body $\mathcal{B}_i$ can be written as

$$T_i = \tfrac{1}{2}\int_{\mathcal{B}_i} \|\dot{x}_i + \dot{R}_i\rho_i\|^2 \, dm_i.$$

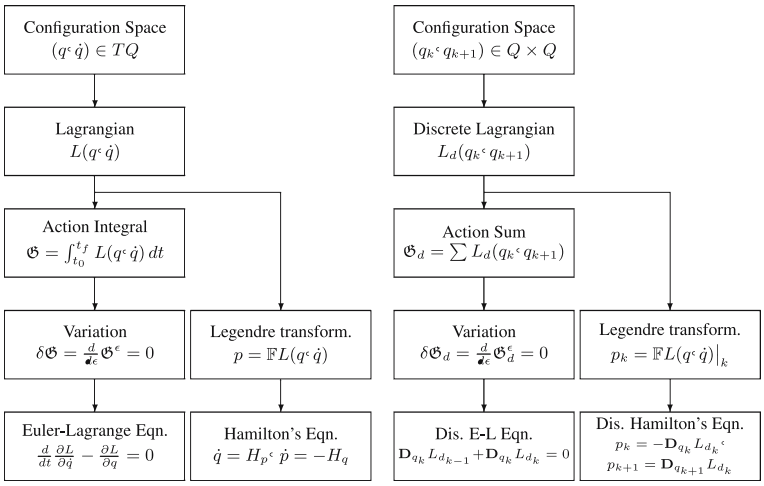


Fig. 1 Procedures to derive continuous and discrete equations of motion (Lee et al. 2007)

Using the fact that $\int_{\mathcal{B}_i} \rho_i dm_i = 0$ and (2), the kinetic energy T_i can be rewritten in terms of the nonstandard moment of inertia matrix as

$$\begin{aligned} T_i(\dot{x}_i, \Omega_i) &= \frac{1}{2}\int_{\mathcal{B}_i} \|\dot{x}_i\|^2 + \|S(\Omega_i)\rho_i\|^2\, dm_i, \\ &= \frac{1}{2} m_i \|\dot{x}_i\|^2 + \frac{1}{2}\mathrm{tr}\Big[S(\Omega_i) J_{d_i} S(\Omega_i)^T\Big]. \end{aligned} \tag{7}$$

The gravitational potential energy $U : \mathrm{SE}(3)^n \mapsto \mathbb{R}$ is given by

$$U(x_1, \ldots, x_n, R_1, \ldots, R_n) = -\frac{1}{2}\sum_{\substack{i,j=1 \\ i\neq j}}^{n} \int_{\mathcal{B}_i}\int_{\mathcal{B}_j} \frac{G dm_i dm_j}{\big\|x_i + R_i\rho_i - x_j - R_j\rho_j\big\|}, \tag{8}$$

where G is the universal gravitational constant.

Then, the Lagrangian for n rigid bodies, $L : \mathrm{TSE}(3)^n \mapsto \mathbb{R}$, is given by

$$\begin{aligned} L(x_1, \dot{x}_1, R_1, \Omega_1, \ldots, x_n, \dot{x}_n, R_n, \Omega_n) = &\sum_{i=1}^{n} \frac{1}{2} m_i \|\dot{x}_i\|^2 + \frac{1}{2}\mathrm{tr}\Big[S(\Omega_i) J_{d_i} S(\Omega_i)^T\Big] \\ &- U(x_1, \ldots, x_n, R_1 \ldots R_n). \end{aligned} \tag{9}$$

Variations of variables: Since the configuration space is $\mathrm{SE}(3)^n$, the variations should be carefully chosen so as to respect the geometry of the configuration space. The variations of $x_i, \dot{x}_i$ are trivial, namely

$$x_i^\epsilon = x_i + \epsilon\delta x_i + \mathcal{O}(\epsilon^2), \tag{10}$$

$$\dot{x}_i^\epsilon = \dot{x}_i + \epsilon\delta\dot{x}_i + \mathcal{O}(\epsilon^2), \tag{11}$$

where $\delta x_i, \delta\dot{x}_i \in \mathbb{R}^3$ are infinitesimal variations that vanish at the initial time t_0 and at the final time t_f. The infinitesimal variation of a rotation matrix R_i can be expressed in terms of a Lie algebra element $\eta_i \in \mathfrak{so}(3)$ and the exponential map as

$$\delta R_i = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_i \exp \epsilon\eta_i = R_i\eta_i, \tag{12}$$

where η_i vanish at the initial time t_0 and at the final time t_f. The infinitesimal variation of Ω_i can be computed from (2) and (12) to be

$$\begin{aligned} S(\delta\Omega_i) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_i^{\epsilon\,T}\dot{R}_i^{\epsilon} = \delta R_i^T\dot{R}_i + R_i^T\delta\dot{R}_i, \\ &= -\eta_i S(\Omega_i) + S(\Omega_i)\eta_i + \dot{\eta}_i. \end{aligned} \tag{13}$$

Hamilton's principle: The action integral is defined to be

$$\mathfrak{G} = \int_{t_0}^{t_f} L(x_1, \dot{x}_1, R_1, \Omega_1, \ldots, x_n, \dot{x}_n, R_n, \Omega_n)\,\mathrm{d}t. \tag{14}$$

Using the variational expression (10)–(13), the variation of the action integral can be written as

$$\delta\mathfrak{G} = \sum_{i=1}^{n}\int_{t_0}^{t_f} m_i\dot{x}_i^T\delta\dot{x}_i - \tfrac{\partial U}{\partial x_i}^T\delta x_i + \tfrac{1}{2}\mathrm{tr}\Big[-\dot{\eta}_i S(J_i\Omega_i) + \eta_i\left\{S(\Omega_i\times J_i\Omega_i) + 2R_i^T\tfrac{\partial U}{\partial R_i}\right\}\Big]\,\mathrm{d}t,$$

where $\frac{\partial U}{\partial R_i} \in \mathbb{R}^{3\times 3}$ is determined by the relationship, $[\frac{\partial U}{\partial R_i}]_{p,q} = \frac{\partial U}{\partial [R_i]_{p,q}}$. Here $[A]_{p,q}$ denotes the (p,q)th element of a matrix A. Using integration by parts and the fact that δx_i and η_i vanish at t_0 and t_f, $\delta\mathfrak{G}$ is given by

$$\begin{aligned} \delta\mathfrak{G} = \sum_{i=1}^{n}\int_{t_0}^{t_f} &-\delta x_i^T\left\{m_i\ddot{x}_i + \frac{\partial U}{\partial x_i}\right\} \\ &+\frac{1}{2}\mathrm{tr}\Big[\eta_i\left\{S(J_i\dot{\Omega}_i + \Omega_i\times J_i\Omega_i) + 2R_i^T\frac{\partial U}{\partial R_i}\right\}\Big]\,\mathrm{d}t. \end{aligned} \tag{15}$$

From Hamilton's principle, $\delta\mathfrak{G}$ should be zero for all possible variations $\delta x_i \in \mathbb{R}^3$ and $\eta_i \in \mathfrak{so}(3)$. Therefore, the expression in the first brace should be zero, and the expression in the second brace should be symmetric, since η_i is skew-symmetric. Then, we obtain the continuous equations of motion as

$$\begin{aligned} m_i\ddot{x}_i &= -\frac{\partial U}{\partial x_i}, \\ S(J_i\dot{\Omega}_i + \Omega_i\times J_i\Omega_i) &= \frac{\partial U}{\partial R_i}^T R - R_i^T\frac{\partial U}{\partial R_i}. \end{aligned}$$

Note that the right hand side expression in the second equation is skew-symmetric. The moment due to the gravitational potential on the ith body, $M_i \in \mathbb{R}^3$, can be expressed explicitly as the following computation shows.

$$\begin{aligned} S(M_i) &= \frac{\partial U}{\partial R_i}^T R_i - R_i^T\frac{\partial U}{\partial R_i}, \\ &= \begin{bmatrix} u_{i_1}^T & u_{i_2}^T & u_{i_2}^T \end{bmatrix}\begin{bmatrix} r_{i_1} \\ r_{i_2} \\ r_{i_3} \end{bmatrix} - \begin{bmatrix} r_{i_1}^T & r_{i_2}^T & r_{i_3}^T \end{bmatrix}\begin{bmatrix} u_{i_1} \\ u_{i_2} \\ u_{i_3} \end{bmatrix}, \\ &= \left(u_{i_1}^T r_{i_1} - r_{i_1}^T u_{i_1}\right) + \left(u_{i_2}^T r_{i_2} - r_{i_2}^T u_{i_2}\right) + \left(u_{i_3}^T r_{i_3} - r_{i_3}^T u_{i_3}\right), \end{aligned}$$

where $r_{i_p}, u_{i_p} \in \mathbb{R}^{1\times 3}$ are the pth row vectors of R_i and $\frac{\partial U}{\partial R_i}$, respectively. Since $S(x\times y) = yx^T - xy^T$ for any column vectors $x, y \in \mathbb{R}^3$, we obtain

$$M_i = r_{i_1} \times u_{i_1} + r_{i_2} \times u_{i_2} + r_{i_3} \times u_{i_3}. \tag{16}$$

Equations of motion: In summary, *the continuous equations of motion for the full body problem, in Lagrangian form,* can be written for bodies $i \in (1, 2, \cdots, n)$ as

$$\dot{v}_i = -\frac{1}{m_i}\frac{\partial U}{\partial x_i}, \tag{17}$$

$$J_i\dot{\Omega}_i + \Omega_i \times J_i\Omega_i = M_i, \tag{18}$$

$$\dot{x}_i = v_i, \tag{19}$$

$$\dot{R}_i = R_i S(\Omega_i). \tag{20}$$

where the gravitational moment M_i is obtained by (16). The above equations can be readily rewritten in Hamiltonian form, using the definition of the linear momentum and the angular momentum, $\gamma_i = m_i v_i$, and $\Pi_i = J_i\Omega_i$.

2.2 Relative coordinates

The motion of the full rigid bodies depends only on the relative positions and the relative attitudes of the bodies. This is a consequence of the property that the gravitational potential can be expressed in terms of only these relative variables. Physically, this is related to the fact that the total linear momentum and the total angular momentum about the mass center of the bodies are conserved. Mathematically, the Lagrangian is invariant under the lifted left action of an element of SE(3). So, it is natural to express the equations of motion in one of the body-fixed frames. In this section, the equations of motion for the full two body problem are derived in relative coordinates. This result can be readily generalized to the n body problem.

Reduction of variables: In (Maciejewski 1995), the reduction is carried out in stages, by first reducing position variables in $\mathbb{R}^3$, and then reducing attitude variables in SO(3). This is equivalent to directly reducing the position and the attitude variables in SE(3) in a single step, which is a consequence of the general theory of Lagrangian reduction by stages (Cendra et al. 2001).

We express each variable with respect to the second body-fixed frame. The reduced position and the reduced attitude variables are the relative position and the relative attitude of the first body with respect to the second body. In other words, the variables are reduced by applying the inverse of $(x_2, R_2) \in \mathrm{SE}(3)$, given by $(-R_2^T x_2, R_2^T) \in \mathrm{SE}(3)$, to each variable, which can be written using homogeneous coordinates:

$$\begin{aligned}\begin{bmatrix} R_2^T & -R_2^T x_2 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} R_1 & x_1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} R_2 & x_2 \\ 0 & 1 \end{bmatrix}\right) &= \left(\begin{bmatrix} R_2^T R_1 & R_2^T(x_1 - x_2) \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} R_2^T R_2 & R_2^T(x_2 - x_2) \\ 0 & 1 \end{bmatrix}\right), \\ &= \left(\begin{bmatrix} R_2^T R_1 & R_2^T(x_1 - x_2) \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} I_{3\times 3} & 0 \\ 0 & 1 \end{bmatrix}\right). \end{aligned} \tag{21}$$

This motivates the definition of the reduced variables as

$$X = R_2^T(x_1 - x_2), \tag{22}$$

$$R = R_2^T R_1, \tag{23}$$

where $X \in \mathbb{R}^3$ is the relative position of the first body with respect to the second body expressed in the second body-fixed frame, and $R \in \mathrm{SO}(3)$ is the relative attitude of the first body with respect to the second body. The corresponding linear and angular velocities are also defined as

$$V = R_2^T(\dot{x}_1 - \dot{x}_2), \tag{24}$$

$$\Omega = R\Omega_1, \tag{25}$$

where $V \in \mathbb{R}^3$ represents the relative velocity of the first body with respect to the second body in the second body-fixed frame, and $\Omega \in \mathbb{R}^3$ is the angular velocity of the first body expressed in the second body-fixed frame. Here, the capital letters denote variables expressed in the second body-fixed frame. For convenience, we denote the inertial position and the inertial velocity of the second body, expressed in the second body-fixed frame by $X_2, V_2 \in \mathbb{R}^3$:

$$X_2 = R_2^T x_2, \tag{26}$$

$$V_2 = R_2^T \dot{x}_2. \tag{27}$$

The moment of inertia matrices of the first body are expressed with respect to the second body-fixed frame. We define $J_R = RJ_1R^T, J_{d_R} = RJ_{d_1}R^T \in \mathbb{R}^{3\times3}$. Note that J_R and J_{d_R} are not constant. It can be shown that J_{d_R} also satisfies a property similar to (5), namely

$$S(J_R\Omega) = S(\Omega)J_{d_R} + J_{d_R}S(\Omega) \tag{28}$$

for any $\Omega \in \mathbb{R}^3$. Define the linear momenta $\Gamma, \gamma_2 \in \mathbb{R}^3$, and the angular momenta $\Pi, \Pi_2 \in \mathbb{R}^3$ as

$$\Gamma = mV, \tag{29}$$

$$\gamma_2 = mv_2, \tag{30}$$

$$\Pi = J_R\Omega = RJ_1\Omega_1, \tag{31}$$

$$\Pi_2 = J_2\Omega_2. \tag{32}$$

The equations of motion in relative coordinates are derived in the same way used to derive the equations in the inertial frame. Here, the Lagrangian is expressed in terms of the reduced variables, and the expressions for the reduced variations are derived.

Reduced Lagrangian: The reduced Lagrangian l is obtained by expressing the original Lagrangian (9) for two bodies in terms of the reduced variables. The kinetic energy is given by

$$T_1+T_2 = \tfrac{1}{2}m_1\|V+V_2\|^2+\tfrac{1}{2}m_2\|V_2\|^2+\tfrac{1}{2}\mathrm{tr}\big[S(\Omega)J_{d_R}S(\Omega)^T\big]+\tfrac{1}{2}\mathrm{tr}\big[S(\Omega_2)J_{d_2}S(\Omega_2)^T\big],$$

The gravitational potential can be written as a function of the relative variables only. By applying the inverse of $(R_2, x_2) \in \mathrm{SE}(3)$ as given in (21), we obtain

$$\begin{aligned} U(x_1, x_2, R_1, R_2) &= U(R_2^T(x_1 - x_2), 0, R_2^T R_1, I_{3\times3}), \\ &= -\int_{\mathcal{B}_1}\int_{\mathcal{B}_2} \frac{Gdm_1dm_2}{\|X + R\rho_1 - \rho_2\|}, \\ &\triangleq U(X, R). \end{aligned}$$

Here, we abuse notation slightly by using the same letter U to denote the gravitational potential as a function of the relative variables. Then, the reduced Lagrangian l is given by

$$\begin{aligned} l(R, X, \Omega, V, \Omega_2, V_2) = \; & \frac{1}{2} m_1 \|V + V_2\|^2 + \frac{1}{2} m_2 \|V_2\|^2 \\ & + \frac{1}{2} \mathrm{tr}\Big[S(\Omega) J_{d_R} S(\Omega)^T\Big] + \frac{1}{2} \mathrm{tr}\Big[S(\Omega_2) J_{d_2} S(\Omega_2)^T\Big] \\ & - U(X, R). \end{aligned} \tag{33}$$

Variations of reduced variables: The variations of the reduced variables must be restricted to those that arise from variations of the original variables. For example, the variation of the relative attitude R is given by

$$\delta R = \frac{d}{d\epsilon}\Big|_{\epsilon=0} R_2^{\epsilon T} R_1^{\epsilon} = \delta R_2^T R_1 + R_2^T \delta R_1.$$

Substituting (12) into the above equation,

$$\begin{aligned} \delta R &= -\eta_2 R_2^T R_1 + R_2^T R_1 \eta_1, \\ &= -\eta_2 R + \eta R, \end{aligned}$$

where a reduced variation $\eta \in \mathfrak{so}(3)$ is defined as $\eta = R\eta_1 R^T$. The variations of other reduced variables can be obtained in a similar way:

$$\delta R = \eta R - \eta_2 R, \tag{34}$$

$$\delta X = \chi - \eta_2 X, \tag{35}$$

$$S(\delta\Omega) = \dot{\eta} - S(\Omega)\eta + \eta S(\Omega) + S(\Omega)\eta_2 - \eta_2 S(\Omega) + S(\Omega_2)\eta - \eta S(\Omega_2), \tag{36}$$

$$\delta V = \dot{\chi} + S(\Omega_2)\chi - \eta_2 V, \tag{37}$$

$$S(\delta\Omega_2) = \dot{\eta}_2 + S(\Omega_2)\eta_2 - \eta_2 S(\Omega_2), \tag{38}$$

$$\delta V_2 = \dot{\chi}_2 + S(\Omega_2)\chi_2 - \eta_2 V_2, \tag{39}$$

where $\chi, \chi_2 \in \mathbb{R}^3$ and $\eta, \eta_2 \in \mathfrak{so}(3)$ are variations that vanish at the end points. These Lie group variations are the key elements required to obtain the equations of motion in relative coordinates.

Reduced equations of motion: The reduced equations of motion can be computed from the reduced Lagrangian using the reduced Hamilton's principle. By taking the variation of the reduced Lagrangian (33) using the constrained variations given by (34) through (39), we obtain the variation of the action integral similar to (15). The reduced Hamilton's principle yields the following *continuous equations of relative motion for the full two body problem, in Lagrangian form*

$$\dot{V} + \Omega_2 \times V = -\frac{1}{m}\frac{\partial U}{\partial X}, \tag{40}$$

$$(J_R \dot{\Omega}) + \Omega_2 \times J_R \Omega = -M, \tag{41}$$

$$J_2 \dot{\Omega}_2 + \Omega_2 \times J_2 \Omega_2 = X \times \frac{\partial U}{\partial X} + M, \tag{42}$$

$$\dot{X} + \Omega_2 \times X = V, \tag{43}$$

$$\dot{R} = S(\Omega)R - S(\Omega_2)R, \tag{44}$$

where $m = \frac{m_1 m_2}{m_1+m_2} \in \mathbb{R}$, and the moment due to the gravity potential $M \in \mathbb{R}^3$ is obtained by

$$M = r_1 \times u_{r_1} + r_2 \times u_{r_2} + r_3 \times u_{r_3}, \tag{45}$$

where $r_p, u_{r_p} \in \mathbb{R}^3$ are the pth column vectors of R and $\frac{\partial U}{\partial R}$, respectively. The following equations can be used for reconstruction of the motion of the second body in the inertial frame:

$$\dot{v}_2 = \frac{1}{m_2} R_2 \frac{\partial U}{\partial X}, \tag{46}$$

$$\dot{x}_2 = v_2, \tag{47}$$

$$\dot{R}_2 = R_2 S(\Omega_2). \tag{48}$$

These equations are equivalent to those given by (Maciejewski 1995), although he omitted the reconstruction equations. Equations (40) though (48) give a complete set of equations for the reduced dynamics and reconstruction. Furthermore, they are derived systematically in the context of geometric mechanics using proper variational formulas given in (34) through (39). The above equations can be readily rewritten in Hamiltonian form using (29)–(32).

3 Lie group variational integrators

General purpose numerical integration methods, including the popular explicit (non-symplectic) Runge-Kutta methods, fail to preserve the geometric characteristics of the full body problem. Integration formulas are obtained by approximating the continuous equations of motion by directly discretizing them with respect to time. With each integration step, the updates involve additive operations, so that the underlying Lie group structure is not necessarily preserved as time progresses. This is caused by the fact that the Euclidean Lie group is not closed under addition. For example, if we use a Runge–Kutta method for numerical integration of (44), then the rotation matrices inevitably drift from the orthogonal rotation group SO(3); the quantity $R^T R$ drifts from the identity matrix. Then, the attitudes of the rigid bodies are not determined accurately, resulting in significant errors in computation of the gravitational forces and moments that depend on the attitude, and consequently errors in the entire simulation. It is often proposed to parameterize (44) by Euler angles or unit quaternions. However, Euler angles are not global expressions of the attitude since they have associated singularities. Unit quaternions do not exhibit singularities, but they are constrained to lie on the unit three-sphere $\mathbb{S}^3$; general purpose numerical integration methods do not preserve the unit length constraint. Therefore, quaternions lead to the same numerical drift problem. Re-normalizing the quaternion vector at each step destroys the conservation properties. Furthermore, unit quaternions double cover SO(3), so that there are inevitable ambiguities in expressing the attitude.

One might instead attempt to apply a symplectic Runge–Kutta algorithm to a rotation matrix based formulation of the problem. But even if it were possible to reproject the numerical solution onto SO(3) while preserving the energy, momentum, and symplectic properties, this would still introduce a drift in the energy. This is because the symplectic integrator does not exactly preserve the energy; instead, the numerical solution evolves on the isoenergy surface of a modified Hamiltonian

(Hairer et al. 2006), which is close to the isoenergy surface of the original Hamiltonian. This is why symplectic integrators exhibit bounded energy fluctuations, as the two isoenergy surfaces are always close, but do not coincide. Since explicit expressions for the modified Hamiltonian do not exist, the reprojection invariably changes the modified Hamiltonian associated with the discrete flow, which in turn introduces a drift in the energy if reprojection is performed repeatedly. Indeed, it was shown in (Ge and Marsden 1988) that a fixed time-step numerical algorithm cannot simultaneously preserve the energy, momentum, and symplecticity, unless it samples the exact trajectory of the system. It is however possible to construct variable time-step methods that are symplectic-energy-momentum preserving integrators (Kane et al. 1999).

In contrast, the Lie Group Variational Integrator has desirable properties such as symplecticity, momentum preservation, and good energy stability for exponentially long time periods, while simultaneously preserving the Euclidian Lie group structure without the use of local charts, reprojection, or constraints. The LGVI is obtained by discretizing Hamilton's principle as shown in Fig. 1; the velocity phase space of the continuous Lagrangian is replaced by discrete variables, and a discrete Lagrangian is chosen such that it approximates a segment of the action integral. Taking the variation of the resulting action sum, we obtain discrete equations of motion referred to as a variational integrator. Since the discrete variables are updated by Lie group operations, the group structure is preserved automatically.

In this section, we derive both a Lagrangian and Hamiltonian form of variational integrators for the full body problem in inertial and relative coordinates. The second level subscript k denotes the value of variables at $t = kh + t_0$ for an integration step size $h \in \mathbb{R}$ and an integer k. The integer N satisfies $t_f = kN + t_0$, so N is the number of time-steps of length h to go from the initial time t_0 to the final time t_f.

3.1 Inertial coordinates

Discrete Lagrangian: In continuous time, the structure of the kinematic equations (20), (44) and (48) ensure that R_i, R and R_2 evolve on SO(3) automatically. Here, we introduce a new variable $F_{i_k} \in \mathrm{SO}(3)$ defined such that $R_{i_{k+1}} = R_{i_k} F_{i_k}$, i.e.

$$F_{i_k} = R_{i_k}^T R_{i_{k+1}}. \tag{49}$$

Thus, F_{i_k} represents the relative attitude between two integration steps, and by requiring that $F_{i_k} \in \mathrm{SO}(3)$, we guarantee that R_{i_k} evolves on SO(3) automatically. This is a consequence of the fact that the Lie group is closed under the group operation of matrix multiplication.

Using the kinematic equation $\dot{R}_i = R_i S(\Omega_i)$, the skew-symmetric matrix $S(\Omega_k)$ can be approximated as

$$S(\Omega_{i_k}) = R_{i_k}^T \dot{R}_{i_k} \approx R_{i_k}^T \frac{R_{i_{k+1}} - R_{i_k}}{h} = \frac{1}{h}(F_{i_k} - I_{3\times 3}). \tag{50}$$

The velocity $\dot{x}_{i_k}$ can be approximated simply by $(x_{i_{k+1}} - x_{i_k})/h$. Using these approximations of the angular and linear velocity, the kinetic energy of the ith body given in (7) can be approximated as

$$
\begin{aligned}
T_i(\dot{x}_i, \Omega_i) &\approx T_i\left(\frac{1}{h}(x_{i_{k+1}} - x_{i_k}), \frac{1}{h}(F_{i_k} - I_{3\times 3})\right), \\
&= \frac{1}{2h^2} m_i \left\| x_{i_{k+1}} - x_{i_k} \right\|^2 + \frac{1}{2h^2}\mathrm{tr}\Big[(F_{i_k} - I_{3\times 3}) J_{d_i} (F_{i_k} - I_{3\times 3})^T\Big], \\
&= \frac{1}{2h^2} m_i \left\| x_{i_{k+1}} - x_{i_k} \right\|^2 + \frac{1}{h^2}\mathrm{tr}\big[(I_{3\times 3} - F_{i_k}) J_{d_i}\big].
\end{aligned}
$$

A discrete Lagrangian $L_d : \mathrm{SE}(3)^n \times \mathrm{SE}(3)^n \mapsto \mathbb{R}$ is constructed such that it approximates a segment of the action integral (14),

$$
\begin{aligned}
L_d = \sum_{i=1}^{n} & \frac{1}{2h} m_i \left\| x_{i_{k+1}} - x_{i_k} \right\|^2 + \frac{1}{h}\mathrm{tr}\big[(I_{3\times 3} - F_{i_k}) J_{d_i}\big] \\
& - \frac{h}{2} U(x_{1_k}, \ldots, R_{n_k}) - \frac{h}{2} U(x_{1_{k+1}}, \ldots, R_{n_{k+1}}).
\end{aligned} \tag{51}
$$

This discrete Lagrangian is self-adjoint (Hairer et al. 2006), and self-adjoint numerical integration methods have even order, so we are guaranteed that the resulting integration method is at least second-order accurate.

Variations of discrete variables: The variations of the discrete variables are chosen to respect the geometry of the configuration space SE(3). The variation of x_{i_k} is given by

$$
x_{i_k}^{\epsilon} = x_{i_k} + \epsilon \delta x_{i_k} + \mathcal{O}(\epsilon^2),
$$

where $\delta x_{i_k} \in \mathbb{R}^3$ and vanishes at $k = 0$ and $k = N$. The variation of R_{i_k} is given by

$$
\delta R_{i_k} = R_{i_k} \eta_{i_k}, \tag{52}
$$

where $\eta_{i_k} \in \mathfrak{so}(3)$ is a variation represented by a skew-symmetric matrix and vanishes at $k = 0$ and $k = N$. The variation of F_{i_k} can be computed from the definition $F_{i_k} = R_{i_k}^T R_{i_{k+1}}$ to give

$$
\begin{aligned}
\delta F_{i_k} &= \delta R_{i_k}^T R_{i_{k+1}} + R_{i_k}^T \delta R_{i_{k+1}}, \\
&= -\eta_{i_k} R_{i_k}^T R_{i_{k+1}} + R_{i_k}^T R_{i_{k+1}} \eta_{i_{k+1}}, \\
&= -\eta_{i_k} F_{i_k} + F_{i_k} \eta_{i_{k+1}}.
\end{aligned} \tag{53}
$$

Discrete Hamilton's principle: To obtain the discrete equations of motion in Lagrangian form, we compute the variation of the discrete Lagrangian from (52) and (53) to give

$$
\begin{aligned}
\delta L_d = \sum_{i=1}^{n} & \frac{1}{h} m_i (x_{i_{k+1}} - x_{i_k})^T (\delta x_{i_{k+1}} - \delta x_{i_k}) + \frac{1}{h}\mathrm{tr}\big[\big(\eta_{i_k} F_{i_k} - F_{i_k} \eta_{i_{k+1}}\big) J_{d_i}\big] \\
& - \frac{h}{2}\left(\frac{\partial U_k}{\partial x_{i_k}}^T \delta x_{i_k} + \frac{\partial U_{k+1}}{\partial x_{i_{k+1}}}^T \delta x_{i_{k+1}}\right) + \frac{h}{2}\mathrm{tr}\left[\eta_{i_k} R_{i_k}^T \frac{\partial U_k}{\partial R_{i_k}} + \eta_{i_{k+1}} R_{i_{k+1}}^T \frac{\partial U_{k+1}}{\partial R_{i_{k+1}}}\right],
\end{aligned} \tag{54}
$$

where $U_k = U(x_{1_k}, \ldots, R_{n_k})$ denotes the value of the potential at $t = kh + t_0$.

Define the action sum as

$$
\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(x_{1_k}, x_{1_{k+1}}, R_{1_k}, F_{1_k}, \ldots, x_{n_k}, x_{n_{k+1}}, R_{n_k}, F_{n_k}). \tag{55}
$$

The discrete action sum $\mathfrak{G}_d$ approximates the action integral (14), because the discrete Lagrangian approximates a segment of the action integral. Substituting (54) into (55), the variation of the action sum is given by

$$\delta\mathfrak{G}_d = \sum_{k=0}^{N-1}\sum_{i=1}^{n} \delta x_{i_{k+1}}^T \left\{\tfrac{1}{h}m_i(x_{i_{k+1}} - x_{i_k}) - \tfrac{h}{2}\tfrac{\partial U_{k+1}}{\partial x_{i_{k+1}}}\right\} + \delta x_{i_k}^T \left\{-\tfrac{1}{h}m_i(x_{i_{k+1}} - x_{i_k}) - \tfrac{h}{2}\tfrac{\partial U_k}{\partial x_{i_k}}\right\}$$
$$+\mathrm{tr}\left[\eta_{i_{k+1}}\left\{-\tfrac{1}{h}J_{d_i}F_{i_k} + \tfrac{h}{2}R_{i_{k+1}}^T\tfrac{\partial U_{k+1}}{\partial R_{i_{k+1}}}\right\}\right] + \mathrm{tr}\left[\eta_{i_k}\left\{\tfrac{1}{h}F_{i_k}J_{d_i} + \tfrac{h}{2}R_{i_k}^T\tfrac{\partial U_k}{\partial R_{i_k}}\right\}\right].$$

Using the fact that δx_{i_k} and η_{i_k} vanish at $k = 0$ and $k = N$, we can reindex the summation, which is the discrete analogue of integration by parts, to yield

$$\delta\mathfrak{G}_d = \sum_{k=1}^{N-1}\sum_{i=1}^{n} -\delta x_{i_k}\left\{\tfrac{1}{h}m_i\left(x_{i_{k+1}} - 2x_{i_k} + x_{i_{k-1}}\right) + h\tfrac{\partial U_k}{\partial x_{i_k}}\right\}$$
$$+\mathrm{tr}\left[\eta_{i_k}\left\{\tfrac{1}{h}\left(F_{i_k}J_{d_i} - J_{d_i}F_{i_{k-1}}\right) + hR_{i_k}^T\tfrac{\partial U_k}{\partial R_{i_k}}\right\}\right].$$

Hamilton's principle states that $\delta\mathfrak{G}_d$ should be zero for all possible variations $\delta x_{i_k} \in \mathbb{R}^3$ and $\eta_{i_k} \in \mathfrak{so}(3)$ that vanish at the endpoints. Therefore, the expression in the first brace should be zero, and since η_{i_k} is skew-symmetric, the expression in the second brace should be symmetric.

Discrete equations of motion: We obtain *the discrete equations of motion for the full body problem, in Lagrangian form,* for bodies $i \in (1, 2, \cdots, n)$ as

$$\frac{1}{h}\left(x_{i_{k+1}} - 2x_{i_k} + x_{i_{k-1}}\right) = -h\frac{\partial U_k}{\partial x_{i_k}}, \tag{56}$$

$$\frac{1}{h}\left(F_{i_{k+1}}J_{d_i} - J_{d_i}F_{i_{k+1}}^T - J_{d_i}F_{i_k} + F_{i_k}^T J_{d_i}\right) = hS(M_{i_{k+1}}), \tag{57}$$

$$R_{i_{k+1}} = R_{i_k}F_{i_k}, \tag{58}$$

where $M_{i_k} \in \mathbb{R}^3$ is defined in (16) as

$$M_{i_k} = r_{i_1} \times u_{i_1} + r_{i_2} \times u_{i_2} + r_{i_3} \times u_{i_3}, \tag{59}$$

where $r_{i_p}, u_{i_p} \in \mathbb{R}^{1\times 3}$ are pth row vectors of R_{i_k} and $\frac{\partial U_k}{\partial R_{i_k}}$, respectively. Given initial conditions $(x_{i_0}, R_{i_0}, x_{i_1}, R_{i_1})$, we can obtain x_{i_2} from (56). Then, F_{i_0} is computed from (58), and F_{i_1} can be obtained by solving the implicit equation (57). Finally, R_{i_2} is found from (58). This yields an update map $(x_{i_0}, R_{i_0}, x_{i_1}, R_{i_1}) \mapsto (x_{i_1}, R_{i_1}, x_{i_2}, R_{i_2})$, and this process can be repeated.

As discussed above, Eqs. (56) through (58) defines a discrete Lagrangian map that updates x_{i_k} and R_{i_k}. The discrete Legendre transformation relates the configuration variables x_{i_k}, R_{i_k} and the corresponding momenta γ_{i_k}, Π_{i_k}. This induces a discrete Hamiltonian map that is equivalent to the discrete Lagrangian map. The detailed development of the discrete Legendre transformation can be found in (Lee et al. 2007); here we summarize the result.

The discrete equations of motion for the full body problem, in Hamiltonian form, can be written for bodies $i \in (1, 2, \cdots, n)$ as

$$x_{i_{k+1}} = x_{i_k} + \frac{h}{m_i}\gamma_{i_k} - \frac{h^2}{2m_i}\frac{\partial U_k}{\partial x_{i_k}}, \tag{60}$$

$$\gamma_{i_{k+1}} = \gamma_{i_k} - \frac{h}{2}\frac{\partial U_k}{\partial x_{i_k}} - \frac{h}{2}\frac{\partial U_{k+1}}{\partial x_{i_{k+1}}}, \tag{61}$$

$$hS(\Pi_{i_k} + \frac{h}{2}M_{i_k}) = F_{i_k}J_{d_i} - J_{d_i}F_{i_k}^T, \tag{62}$$

$$\Pi_{i_{k+1}} = F_{i_k}^T\Pi_{i_k} + \frac{h}{2}F_{i_k}^T M_{i_k} + \frac{h}{2}M_{i_{k+1}}., \tag{63}$$

$$R_{i_{k+1}} = R_{i_k}F_{i_k}. \tag{64}$$

Given $(x_{i_0}, \gamma_{i_0}, R_{i_0}, \Pi_{i_0})$, we can find x_{i_1} from (60). Solving the implicit equation (62) yields F_{i_0}, and R_{i_1} is computed from (64). Then, (61) and (63) gives γ_{i_1}, and Π_{i_1}. This defines the discrete Hamiltonian map, $(x_{i_0}, \gamma_{i_0}, R_{i_0}, \Pi_{i_0}) \mapsto (x_{i_1}, \gamma_{i_1}, R_{i_1}, \Pi_{i_1})$, and this process can be repeated.

3.2 Relative coordinates

In this section, we derive the variational integrator for the full two body problem in relative coordinates by expressing the discrete Lagrangian in relative coordinates, and then computing the constrained variations of the discrete reduced variables. This result can be readily generalized to n bodies. A more intrinsic development of discrete Routh reduction can be found in Jalnapurkar et al. (2006).

Reduction of discrete variables: The discrete reduced variables are defined in the same way as the continuous reduced variables, which are given in (22) through (32). We introduce $F_k \in \mathrm{SO}(3)$ such that $R_{k+1} = R_{2_{k+1}}^T R_{1_{k+1}} = F_{2_k}^T F_k R_k$, i.e.

$$F_k = R_k F_{1_k} R_k^T. \tag{65}$$

Discrete reduced Lagrangian: The discrete reduced Lagrangian is obtained by expressing the original discrete Lagrangian given in (51) in terms of the discrete reduced variables.

From the definition of the discrete reduced variables given in (22) and (26), we have

$$\begin{aligned} x_{1_{k+1}} - x_{1_k} &= R_{2_{k+1}}(X_{k+1} + X_{2_{k+1}}) - R_{2_k}(X_k + X_{2_k}), \\ &= R_{2_k}\left\{F_{2_k}(X_{k+1} + X_{2_{k+1}}) - (X_k + X_{2_k})\right\}, \end{aligned} \tag{66}$$

$$x_{2_{k+1}} - x_{2_k} = R_{2_k}\left\{F_{2_k}X_{2_{k+1}} - X_{2_k}\right\}. \tag{67}$$

From (50), $S(\Omega_{1_k})$ and $S(\Omega_{2_k})$ are expressed as

$$S(\Omega_{1_k}) = \frac{1}{h}\left(F_{1_k} - I_{3\times 3}\right) = \frac{1}{h}R_k^T\left(F_k - I_{3\times 3}\right)R_k, \tag{68}$$

$$S(\Omega_{2_k}) = \frac{1}{h}\left(F_{2_k} - I_{3\times 3}\right). \tag{69}$$

Substituting (66) through (69) into (51), we obtain the discrete reduced Lagrangian.

$$\begin{aligned} l_{d_k} &= l_d(X_k, X_{k+1}, X_{2_k}, X_{2_{k+1}}, R_k, F_k, F_{2_k}) \\ &= \frac{1}{2h} m_1 \left\| F_{2_k}(X_{k+1} + X_{2_{k+1}}) - (X_k + X_{2_k}) \right\|^2 + \frac{1}{2h} m_2 \left\| F_{2_k} X_{2_{k+1}} - X_{2_k} \right\|^2 \\ &\quad + \frac{1}{h} \mathrm{tr}\left[(I_{3\times 3} - F_k) J_{dR_k} \right] + \frac{1}{h} \mathrm{tr}\left[(I_{3\times 3} - F_{2_k}) J_{d_2} \right] \\ &\quad - \frac{h}{2} U(X_k, R_k) - \frac{h}{2} U(X_{k+1}, R_{k+1}), \end{aligned} \tag{70}$$

where $J_{dR_k} \in \mathbb{R}^{3\times 3}$ is defined to be $J_{dR_k} = R_k J_{d_1} R_k^T$, which gives the nonstandard moment of inertia matrix of the first body with respect to the second body-fixed frame at $t = kh + t_0$.

Variations of discrete reduced variables: The variations of the discrete reduced variables can be derived from those of the original variables. The variations of R_k, X_k, and F_{2_k} are the same as those given in (34), (35), and (53), respectively. The variation of F_k is computed in a similar fashion to (53). In summary, the variations of discrete reduced variables are given by

$$\delta R_k = \eta_k R_k - \eta_{2_k} R_k, \tag{71}$$
$$\delta X_k = \chi_k - \eta_{2_k} X_k, \tag{72}$$
$$\delta F_k = -\eta_{2_k} F_k + F_{2_k} \eta_{k+1} F_{2_k}^T F_k + F_k \left(-\eta_k + \eta_{2_k} \right), \tag{73}$$
$$\delta X_{2_k} = \chi_{2_k} - \eta_{2_k} X_{2_k}, \tag{74}$$
$$\delta F_{2_k} = -\eta_{2_k} F_{2_k} + F_{2_k} \eta_{2_{k+1}}. \tag{75}$$

These Lie group variations are the main elements required to derive the variational integrator equations.

Discrete reduced equations of motion: Define the action sum in terms of the discrete reduced Lagrangian

$$\mathfrak{G}_d = \sum_{k=0}^{N-1} l_d(X_k, X_{k+1}, X_{2_k}, X_{2_{k+1}}, R_k, F_k, F_{2_k}). \tag{76}$$

Using the expressions for the reduced variations, (71)–(75), we obtain the variation of the action sum. From Hamilton's principle, it should be zero for all possible variations $\chi_k, \chi_{2_k} \in \mathbb{R}^3$ and $\eta_k, \eta_{2_k} \in \mathfrak{so}(3)$ which vanish at the endpoints.

As a result, *the discrete equations of relative motion for the full two body problem, in Lagrangian form,* are obtained as

$$F_{2_k} X_{k+1} - 2X_k + F_{2_{k-1}}^T X_{k-1} = -\frac{h^2}{m} \frac{\partial U_k}{\partial X_k}, \tag{77}$$

$$F_{k+1} J_{dR_{k+1}} - J_{dR_{k+1}} F_{k+1}^T = F_{2_k}^T \left(F_k J_{dR_k} - J_{dR_k} F_k^T \right) F_{2_k} - h^2 S(M_{k+1}), \tag{78}$$

$$\begin{aligned} F_{2_{k+1}} J_{d_2} - J_{d_2} F_{2_{k+1}}^T &= F_{2_k}^T \left(F_{2_k} J_{d_2} - J_{d_2} F_{2_k}^T \right) F_{2_k} \\ &\quad + h^2 X_{k+1} \times \frac{\partial U}{\partial X_{k+1}} + h^2 S(M_{k+1}), \end{aligned} \tag{79}$$

$$R_{k+1} = F_{2_k}^T F_k R_k, \tag{80}$$

$$R_{2_{k+1}} = R_{2_k} F_{2_k}. \tag{81}$$

It is natural to express the equations of motion for the second body in the inertial frame.

$$x_{2_{k+1}} - 2x_{2_k} + x_{2_{k-1}} = \frac{h^2}{m_2} R_k \frac{\partial U_k}{\partial X_k}. \tag{82}$$

Given $(X_0, R_0, R_{2_0}, X_1, R_1, R_{2_1})$, we can determine F_0 and F_{2_0} from (80) and (81). Solving the implicit equations (78) and (79) gives F_1 and F_{2_1}. Then X_2, R_2 and R_{2_2} are found from (77), (80) and (81), respectively. This yields the discrete Lagrangian map $(X_0, R_0, R_{2_0}, X_1, R_1, R_{2_1}) \mapsto (X_1, R_1, R_{2_1}, X_2, R_2, R_{2_2})$ and this process can be repeated. We can separately reconstruct x_{2_k} using (82).

Alternatively, once we have obtained F_0 and F_{2_0} from (80) and (81), we can view Eqs. (78), (79), (77) and (80) as defining an implicit update map $(X_0, R_0, F_0, F_{2_0}, X_1) \mapsto (X_1, R_1, F_1, F_{2_1}, X_2)$. As a post-processing step, R_{2_k} and x_{2_k} can be reconstructed using (81) and (82), respectively.

The discrete Legendre transformation yields *the discrete equations of relative motion for the full two body problem, in Hamiltonian form,*

$$X_{k+1} = F_{2_k}^T \left(X_k + h\frac{\Gamma_k}{m} - \frac{h^2}{2m}\frac{\partial U_k}{\partial X_k} \right), \tag{83}$$

$$\Gamma_{k+1} = F_{2_k}^T \left(\Gamma_k - \frac{h}{2}\frac{\partial U_k}{\partial X_k} \right) - \frac{h}{2}\frac{\partial U_{k+1}}{\partial X_{k+1}}, \tag{84}$$

$$\Pi_{k+1} = F_{2_k}^T \left(\Pi_k - \frac{h}{2} M_k \right) - \frac{h}{2} M_{k+1}, \tag{85}$$

$$\begin{aligned} \Pi_{2_{k+1}} = {} & F_{2_k}^T \left(\Pi_{2_k} + \frac{h}{2} X_k \times \frac{\partial U}{\partial X_k} + \frac{h}{2} M_k \right) \\ & + \frac{h}{2} X_{k+1} \times \frac{\partial U}{\partial X_{k+1}} + \frac{h}{2} M_{k+1}, \end{aligned} \tag{86}$$

$$R_{k+1} = F_{2_k}^T F_k R_k, \tag{87}$$

$$hS\left(\Pi_k - \frac{h}{2} M_k \right) = F_k J_{dR_k} - J_{dR_k} F_k^T, \tag{88}$$

$$hS\left(\Pi_{2_k} + \frac{h}{2} X_k \times \frac{\partial U}{\partial X_k} + \frac{h}{2} M_k \right) = F_{2_k} J_{d_2} - J_{d_2} F_{2_k}^T. \tag{89}$$

It is natural to express the equations of motion for the second body in the inertial frame for reconstruction:

$$x_{2_{k+1}} = x_{2_k} + h\frac{\gamma_{2_k}}{m_2} + \frac{h^2}{2m_2} R_k \frac{\partial U_k}{\partial X_k}, \tag{90}$$

$$\gamma_{2_{k+1}} = \gamma_{2_k} + \frac{h}{2} R_k \frac{\partial U_k}{\partial X_k} + \frac{h}{2} R_{k+1} \frac{\partial U_{k+1}}{\partial X_{k+1}}, \tag{91}$$

$$R_{2_{k+1}} = R_{2_k} F_{2_k}. \tag{92}$$

Given $(R_0, X_0, \Pi_0, \Gamma_0, \Pi_{2_0})$, we can determine F_0 and F_{2_0} by solving the implicit equations (88) and (89). Then, X_1 and R_1 are found from (83) and (87), respectively. After that, we can compute Γ_1, Π_1, and Π_{2_1} from (84), (85) and (86). This yields a discrete Hamiltonian map $(R_0, X_0, \Pi_0, \Gamma_0, \Pi_{2_0}) \mapsto (R_1, X_1, \Pi_1, \Gamma_1, \Pi_{2_1})$, and this process can

be repeated. The discrete evolution of x_{2_k}, γ_{2_k} and R_{2_k} can be obtained as a post-processing step by using the reconstruction equations (90), (91) and (92), respectively.

3.3 Numerical considerations

Properties of the variational integrators: Since the LGVI is obtained by discretizing Hamilton's principle, it is symplectic and preserves the structure of the configuration space, SE(3), as well as the relevant geometric features of the full two rigid body problem, and the conserved first integrals of total linear and angular momenta and total energy. The total energy oscillates around its initial value with small bounds on a comparatively short timescale, but there is no tendency for the mean of the oscillation in the total energy to drift (increase or decrease) from the initial value for exponentially long time.

The LGVI preserves the group structure. By using the computational approach described in Sect. 3.4, the matrices F_{i_k} representing the change in relative attitude are guaranteed to be rotation matrices. The group operation of the Lie group SO(3) is matrix multiplication. Hence rotation matrices R_{i_k} are updated by the group operation, so that they evolve on SO(3) automatically without constraints or reprojection. Therefore, the orthogonal structure of the rotation matrices is preserved, and the attitude of each rigid body is determined accurately and globally without the need to use local charts (parameterizations) such as Euler angles or quaternions. These exact geometric properties of the discrete flow not only generate improved qualitative behavior, but also allow for accurate long-time simulation.

This geometrically exact numerical integration method yields a highly efficient and accurate computational algorithm for the full rigid body problem. For arbitrary shaped rigid bodies such as binary asteroids, there is a large burden in computing the mutual gravitational forces and moments, so the number of force and moment evaluations should be minimized. We have seen that the LGVI requires only one such evaluation per integration step, the minimum number of evaluations consistent with the presented LGVI having second order accuracy (because it is a self-adjoint method). Within the LGVI, implicit equations must be solved at each time step to determine the matrix-multiplication updates for rotation matrices. However the LGVI is only weakly implicit in the sense that the iteration for each implicit equation is independent of the much more costly gravitational force and moment computation. The computational load to solve each implicit equation is negligible; only two or three iterations are typically required. We make this more explicit in Sect. 3.4 by expressing F_{i_k} as the exponential function of an element of the Lie algebra $\mathfrak{so}(3)$. Altogether, the entire method could be considered *almost explicit*.

The LGVI is a fixed step size integrator, but all of the properties above are independent of the step size. Consequently, we can achieve the same level of accuracy while choosing a larger step size as compared to other numerical integrators of the same order.

All of these features are revealed by numerical simulations in Sect. 4 and in the work by Fahnestock et al. (2006). In Sect. 4, the LGVI is compared with other second order geometric integrators: a symplectic Runge-Kutta method and a Lie group method. In Fahnestock et al. (2006), the LGVI is directly compared with the 7(8)th order Runge-Kutta-Fehlberg method (RK78) for two octahedral rigid bodies. It is shown that the LGVI requires 8 times less computational load than RK78 for similar error measures, and the accuracy of the LGVI is maintained for exponentially

long time. The trajectories computed using RK78 are unreliable for the long time simulation of the full two rigid body dynamics.

Higher-order methods: While the numerical methods we present in this paper are second order, it is possible to apply symmetric composition methods, introduced by Yoshida (1990), to construct higher-order versions of the Lie group variational integrators introduced here. Given a basic numerical method represented by the flow map Φ_h, the composition method is obtained by applying the basic method using different step sizes,

$$\Psi_h = \Phi_{\lambda_s h} \circ \ldots \circ \Phi_{\lambda_1 h},$$

where $\lambda_1, \lambda_2, \cdots, \lambda_s \in \mathbb{R}$. In particular, the Yoshida symmetric composition method for composing a symmetric method of order 2 into a symmetric method of order 4 is obtained when $s = 3$, and

$$\lambda_1 = \lambda_3 = \frac{1}{2 - 2^{1/3}}, \qquad \lambda_2 = -\frac{2^{1/3}}{2 - 2^{1/3}}.$$

Alternatively, by adopting the formalism of higher-order Lie group variational integrators introduced by Leok (2004) in conjunction with the Rodrigues formula, one can directly construct higher-order generalizations of the Lie group methods presented here.

Reduction of orthogonality loss due to roundoff error: In Lie group variational integrators, the numerical solution is made to automatically remain on the rotation group by requiring that the numerical solution is updated by matrix multiplication with the exponential of a skew-symmetric matrix.

Since the exponential of a skew-symmetric matrix is orthogonal to machine precision, the numerical solution will only deviate from orthogonality due to the accumulation of roundoff error in the matrix multiplication, and this orthogonality loss grows linearly with the number of timesteps taken. The phenomena can be observed in the numerical simulations described in Fig. 3(c) of Sect. 4, wherein the orthogonality error for the Lie group method and the Lie group variational integrator increases as the step size decreases, due to the roundoff error accumulation as the number of matrix multiplications increase.

One possible method of addressing this issue is to use the Baker-Campbell-Hausdorff (BCH) formula to track the updates as skew-symmetric matrices (the Lie algebra). This allows us to find a matrix $C(t)$, such that,

$$\exp(tA)\exp(tB) = \exp C(t).$$

This matrix $C(t)$ satisfies the following differential equation,

$$\dot{C} = A + B + \frac{1}{2}[A - B, C] + \sum_{k \geq 2} \frac{B_k}{k!} \operatorname{ad}_C^k (A + B),$$

with initial value $C(0) = 0$, where B_k denotes the Bernoulli numbers and $\operatorname{ad}_C A = [C, A] = CA - AC$.

The problem with this approach is that the matrix $C(t)$ is not readily computable for arbitrary A and B, and in practice, the series is truncated, and the differential equation is solved numerically.

An error is introduced in truncating the series, and numerical errors are introduced in numerically integrating the differential equations. Consequently, while the BCH

formula could be used solely at the reconstruction stage to ensure that the numerical attitude always remains in the rotation group to machine precision, the truncation error would destroy the symplecticity and momentum preserving properties of the numerical scheme.

However, by combining the BCH formula with the Rodrigues formula in constructing the discrete variational principle, it might be possible to construct a Lie group variational integrator that tracks the reconstructed trajectory on the rotation group at the level of a curve in the Lie algebra, while retaining its structure-preservation properties.

3.4 Computational approach

The structure of the discrete equations of motion given in (57), (62), (78), (79), (88), and (89) suggests a specific computational approach. For a given $g \in \mathbb{R}^3$, we have to solve the following Lyapunov-like equation to find $F \in \mathrm{SO}(3)$ at each integration step.

$$FJ_d - J_d F^T = S(g). \tag{93}$$

This equation is linear in F, but it is implicit due to the nonlinear constraint $F^T F = I_{3\times 3}$. We now introduce two iterative approaches to solve (93) numerically.

Exponential map: An element of a Lie group can be expressed as the exponential of an element of its Lie algebra, so $F \in \mathrm{SO}(3)$ can be expressed as an exponential of $S(f) \in \mathfrak{so}(3)$ for some vector $f \in \mathbb{R}^3$. The exponential can be written in closed form, using Rodrigues' formula,

$$F = \exp S(f) = I_{3\times 3} + \tfrac{\sin\|f\|}{\|f\|} S(f) + \tfrac{1-\cos\|f\|}{\|f\|^2} S(f)^2. \tag{94}$$

Substituting (94) into (93), we obtain

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

Thus, (93) is converted into the equivalent vector equation $g = G(f)$, where $G : \mathbb{R}^3 \mapsto \mathbb{R}^3$ is given by

$$G(f) = \tfrac{\sin\|f\|}{\|f\|} Jf + \tfrac{1-\cos\|f\|}{\|f\|^2} f \times Jf. \tag{95}$$

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

$$f_{i+1} = f_i + \nabla G(f_i)^{-1} \left(g - G(f_i)\right). \tag{96}$$

We iterate until $\|g - G(f_i)\| < \epsilon$ for a small tolerance $\epsilon > 0$. The Jacobian $\nabla G(f)$ in (96) can be expressed as

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

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

$$F = \mathrm{cay}\, S(f_c) = (I_{3\times 3} + S(f_c))(I_{3\times 3} - S(f_c))^{-1}. \tag{97}$$

Substituting (97) into (93), we obtain a vector equation $G_c(f_c) = 0$ equivalent to (93)

$$G_c(f_c) = g + g \times f_c + (g^T f_c) f_c - 2J f_c = 0, \tag{98}$$

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

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

Then, (98) is solved by using Newton's iteration (96), and the rotation matrix is obtained by the Cayley transformation.

For both methods, numerical experiments show that 2 or 3 iterations are sufficient to achieve a tolerance of $\epsilon = 10^{-15}$. Numerical iteration with the Cayley transformation is a little faster due to the simpler expressions. It should be noted that since $F = \exp S(f)$ or $F = \mathrm{cay}\, S(f_c)$, it is automatically a rotation matrix, even when the equation $g = G(f)$ is not satisfied to machine precision.

These computational approaches are distinguished from solving the implicit equation (93) with 9 variables and 6 constraints. Due to their numerical efficiency, the Lie group variational integrator can be considered an almost explicit computational method as demonstrated in the next section.

4 Numerical simulations

We simulate the dynamics of two simple dumbbell bodies acting under their mutual gravity. Each dumbbell model consists of two equal rigid spheres and a rigid massless connecting rod. This dumbbell rigid body model has a simple closed form for the mutual gravitational potential given by

$$U(X, R) = -\sum_{p,q=1}^{2} \frac{G m_1 m_2 / 4}{\left\| X + \rho_{2_p} + R\rho_{1_q} \right\|},$$

where G is the universal gravitational constant, $m_i \in \mathbb{R}$ is the total mass of the ith dumbbell, and $\rho_{i_p} \in \mathbb{R}^3$ is a vector from the origin of the body-fixed frame to the pth sphere of the ith dumbbell in the ith body-fixed frame. The vectors $\rho_{i_1} = [l_i/2, 0, 0]^T$, $\rho_{i_2} = -\rho_{i_1}$, where l_i is the length between the two spheres. Mass, length and time dimensions are normalized.

The mass and length of the second dumbbell are twice that of the first dumbbell. The other simulation parameters are chosen such that the total linear momentum in the inertial frame is zero and the relative motion between two bodies are near-elliptic orbits. The trajectories of dumbbell bodies are shown in Fig. 2.

We compare the computational properties of the Lie group variational integrator (LGVI) with other second order numerical integration methods; an explicit Runge-Kutta method (RK), a symplectic Runge-Kutta method (SRK), and a Lie group method (LGM). One of the distinct features of the LGVI is that it preserves both the symplectic property and the Lie group structure of the full rigid body dynamics. A comparison can be made between the LGVI and other integration methods that

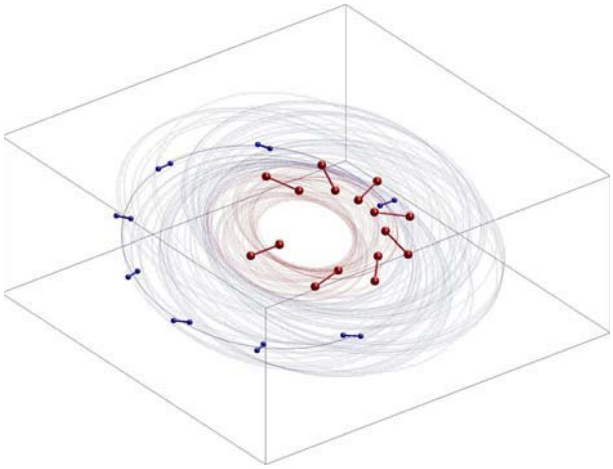
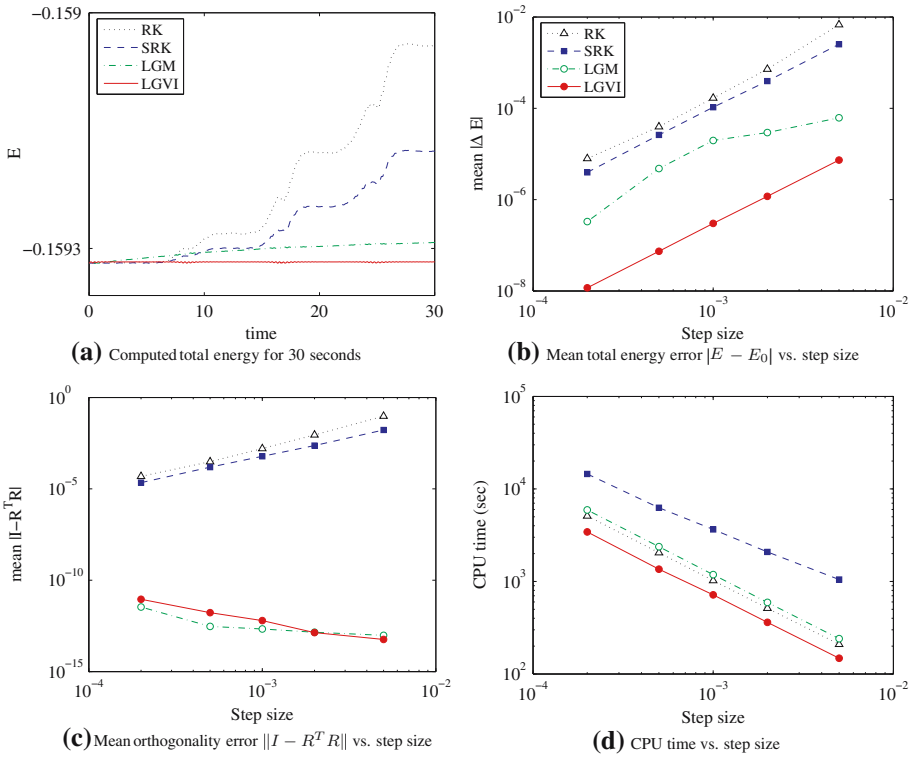


Fig. 2 Trajectories of two dumbbell bodies in the inertial frame (The initial orbit is shown with solid lines and snapshots of dumbbell body maneuver. A simple animation is available at http://www.umich.edu/ tylee.)

preserve either none or one of these properties: an integrator that does not preserve any of these properties (RK), a symplectic integrator that does not preserve the Lie group structure (SRK), and a Lie group integrator that does not preserve symplecticity (LGM). These methods are implemented by an explicit mid-point rule, an implicit mid-point rule, and the Crouch-Grossman method presented in (Hairer et al. 2006) for the continuous equations of motion (40)–(48), respectively. For the LGVI, the discrete equations of motion given by (83) through (92) are used. All of these integrators are second order accurate. A comparison with a higher-order integrator can be found in (Fahnestock et al. 2006).

Fig. 3(a) shows the computed total energy response over 30 seconds with an integration step size $h = 0.002$ sec. For the LGVI, the total energy is nearly constant, and there is no tendency to drift, while the other integrators fail to preserve the total energy. This can be observed in Fig. 3(b), where the mean total energy deviations are shown for varying integration step sizes. It is seen that the total energy errors of the SRK method is close to the RK method, but the total energy error of the LGVI is smaller by several orders. Fig. 3(c) shows the mean orthogonality errors. The LGVI and the LGM conserve the orthogonal structure at an error level of 10^{-10}, while the RK and the SRK do not.

These computational comparisons suggest that for numerical integration of Hamiltonian systems evolving on a Lie group, such as full body problems, it is critical to preserve both the symplectic property and the Lie group structure. For the RK and the SRK, the orthogonality error in the rotation matrix corrupts the attitude of the rigid bodies. The accumulation of this attitude degradation causes significant errors in the computation of the gravitational forces and moments dependent upon the position and the attitude, which affect the accuracy of the entire numerical simulation. The LGM conserves the orthogonal structure of rotation matrices numerically, but it does not respect the characteristics of the Hamiltonian dynamics properly as a non-symplectic integrator; this causes a drift of the computed total energy. The LGVI is a geometrically exact integration method in the sense that it preserves all of the



(a) Computed total energy for 30 seconds
(b) Mean total energy error $|E - E_0|$ vs. step size
(c) Mean orthogonality error $\|I - R^T R\|$ vs. step size
(d) CPU time vs. step size

Fig. 3 Computational properties of explicit Runge-Kutta (RK), symplectic Runge-Kutta (SRK), Lie group method (LGM), and Lie group variational integrator (LGVI)

features of the full rigid body dynamics concurrently. This verifies the superiority of the LGVI in terms of computational accuracy. The performance advantages of the LGVI becomes even more dramatic as the simulation time is increased.

Computational efficiency is compared in Fig. 3(d), where CPU times of all methods are shown for varying step sizes. The SRK has the largest CPU time requiring solution of an implicit equation in 36 variables at each integration step. The RK and the LGM require similar CPU times since both are explicit. It is interesting to see that the implicit LGVI actually requires less CPU time than the explicit methods RK and LGM. This follows from the fact that the second order explicit methods RK and LGM require two evaluations of (40)–(48), including the expensive force and moment computations at each step. The LGVI requires only one evaluation at each step in addition to the solution the implicit equation. The computational approach described in Sect. Fig. 3 is efficient for solving the implicit equation (93) and hence it takes less time than the evaluation of (40)–(48). The difference is further increased as the rigid body model becomes more complicated since it involves a larger computation burden in computing the gravitational forces and moments. Based on these properties, we claim that the LGVI is *almost explicit*. This comparison demonstrates the higher computational efficiency of the LGVI.

In summary, comparing both Fig. Fig. 3(b) and (d), we see that the LGVI requires 16 times less CPU time than the LGM, 35 times less CPU time than the RK, and 98

times less CPU time than the SRK for similar total energy error in this computational example for the full body problem.

5 Conclusions

Eight different forms of the equations of motion for the full body problem are derived. The continuous equations of motion and variational integrators are derived both in inertial coordinates and in relative coordinates, and each set of equations of motion is expressed in both Lagrangian and Hamiltonian form.

It is shown that both of the continuous equations and the discrete equations of motion for the full body problem can be derived systematically, using proper variations of Lie group elements, according to Hamilton's principle. The proposed Lie group variational integrators are geometrically exact; they preserve the momenta and symplectic form of the continuous dynamics, exhibit good energy properties, and they also conserve the geometry of the configuration space. They provide a numerically efficient computational approach especially for the full body problems in the sense that they require only one evaluation of mutual gravity forces and moments per step. The exact geometric properties of the discrete flow not only yields improved qualitative behavior, but also allow for accurate long-time simulation. The numerical example verifies the substantial superiority of the Lie group variational integrator compared with other geometric integrators in terms of computational accuracy and efficiency.

Acknowledgements We gratefully acknowledge helpful comments and suggestions of the referees. TL and ML have been supported in part by NSF Grant DMS-0504747, and a grant from the Rackham Graduate School, University of Michigan. TL and NHM have been supported in part by NSF Grant ECS-0244977 and CMS-0555797.

References

Breiter, S., Melendo, B., Bartczak, P., Wytrzyszczak, I.: Synchronous motion in the Kinoshita problem. Astron. Astrophys. **437**, 753–764 (2005a)

Breiter, S., Nesvorný, D., Vokrouhlický, D.: Efficient Lie-Poisson integrator for secular spin dynamics of rigid bodies. Astron. J. **130**, 1267–1277 (2005b)

Cendra, H., Marsden, J.E., Ratiu, T.S.: Lagrangian reduction by stages. Mem. Am. Math. Soc. 152 (2001)

Fahnestock, E., Lee, T., Leok, M., McClamroch, N.H., Scheeres, D.J.: Polyhedral potential and variational integrator computation of the full two body problem. In: AIAA/AAS Astrodynamics Specialist Meeting. AIAA-2006-6289 (2006)

Ge, Z., Marsden, J.E.: Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators. Phys. Lett. A **133**(3), 134–139 (1988)

Hairer, E. : Backward analysis of numerical integrators and symplectic methods. Ann. Numer. Math. **1**, 107–132 scientific computation and differential equations (Auckland, 1993). (1994)

Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration, 2nd edn. Springer (2006)

Iserles, A., Munthe-Kaas, H.Z., Nrsett, S.P., Zanna, A.: Lie-group methods. Acta Numerica **9**, 215–365 (2000)

Jalnapurkar, S.M., Leok, M., Marsden, J.E., West, M.: Discrete Routh reduction. J. Physica A **39**, 5521–5544 (2006)

Kane, C., Marsden, J.E., Ortiz, M.: Symplectic-energy-momentum preserving variational integrators. J. Math. Phys. **40**(7), 3353–3371 (1999)

Krysl, P.: Explicit momentum-conserving integrator for dynamics of rigid bodies approximating the midpoint Lie algorithm. Int. J. Numer. Methods Eng. **63**, 2171–2193 (2005)

Lee, T., Leok, M., McClamroch, N.H.: A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum. In: Proceedings of the IEEE Conference on Control Applications, pp. 962–967 (2005)
Lee, T., Leok, M., McClamroch, N.H.: Lie group variational integrators for the full body problem. Comput. Methods App. Mech. Eng. in press (2007)
Leok, M.: Foundations of Computational Geometric Mechanics. Ph.D. thesis, California Instittute of Technology (2004)
Maciejewski, A.J.: Reduction, relative equilibria and potential in the two rigid bodies problem. Celestial Mech. Dyn. Astr. **63**, 1–28 (1995)
Margot, J.L., Nolan, M.C., Benner, L.A.M., Ostro, S.J., Jurgens, R.F., Giorgini, J.D., Slade, M.A., Campbell, D.B.: Binary asteroids in the Near-Earth object population. Science **296**, 1445–1448 (2002)
Marsden, J.E., Pekarsky, S., Shkoller, S.: Discrete Euler–Poincaré and Lie–Poisson equations. Nonlinearity **12**, 1647–1662 (1999)
Marsden, J.E., Pekarsky, S., Shkoller, S.: Symmetry reduction of discrete Lagrangian mechanics on Lie groups. J. Geom. Phys. **36**, 139–150 (2000)
Marsden, J.E., West, M.: Discrete mechanics and variational integrators. Acta Numerica **10**, 357–514 (2001)
Moser, J., Veselov, A.P.: Discrete versions of some classical integrable systems and factorization of matrix polynomials. Commun. Math. Phys. **139**, 217–243 (1991)
Scheeres, D.J.: Stability in the Full Two-Body problem. Celestial Mech. Dyn. Astr. **83**, 155–169 (2002)
Simo, J.C., Tarnow, N., Wong, K.K.: Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics. Comput. Methods Appl. Mech. Eng. **100**, 63–116 (1992)
Touma, J., Wisdom, J.: Lie-Poisson integrator for rigid body dynamics in the solar system. Astron. J. **107**, 1189–1202 (1994)
Wendlandt, J.M., Marsden, J.E.: Mechanical integrator derived from a discrete variational principle. Physica D **106**, 223–246 (1997)
Werner, R.A., Scheeres, D.J.: Mutual potential of homogenous polyhedra. Celest. Mech. and Dyn. Astr. **91**, 337–349 (2005)
Yoshida, H.: Construction of high order symplectic integrators. Phys. Lett. A **150**, 262–268 (1990)

Available online at www.sciencedirect.com
ScienceDirect
Systems & Control Letters 57 (2008) 236–245
SYSTEMS & CONTROL LETTERS
www.elsevier.com/locate/sysconle

Global optimal attitude estimation using uncertainty ellipsoids

Amit K. Sanyal[a,*], Taeyoung Lee[b], Melvin Leok[c], N. Harris McClamroch[c]

[a]*Department of Mechanical Engineering, University of Hawaii, Honolulu, HI 96822, USA*
[b]*Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109-2140, USA*
[c]*Department of Mathematics, Purdue University, West Lafayette, IN 47907-2067, USA*

Received 4 June 2006; received in revised form 21 April 2007; accepted 31 August 2007
Available online 22 October 2007

Abstract

A deterministic attitude estimation problem for a rigid body in a potential field, with bounded attitude and angular velocity measurement errors is considered. An attitude estimation algorithm that globally minimizes the attitude estimation error is obtained. Assuming that the initial attitude, the initial angular velocity and measurement noise lie within given ellipsoidal bounds, an uncertainty ellipsoid that bounds the attitude and the angular velocity of the rigid body is obtained. The center of the uncertainty ellipsoid provides point estimates, and the size of the uncertainty ellipsoid measures the accuracy of the estimates. The point estimates and the uncertainty ellipsoids are propagated using a Lie group variational integrator and its linearization, respectively. The attitude and angular velocity estimates are optimal in the sense that the sizes of the uncertainty ellipsoids are minimized.
Published by Elsevier B.V.

Keywords: Global attitude representation; Deterministic estimation; Uncertainty ellipsoids

1. Introduction

Attitude estimation is often a prerequisite for controlling aerospace and underwater vehicles, mobile robots, and other mechanical systems moving in space. In this paper, we study the attitude estimation problem for the uncontrolled dynamics of a rigid body in an attitude-dependent force potential (like uniform gravity). The estimation scheme we present has the following important features: (1) the attitude is globally represented without using any coordinate system, (2) the filter obtained is not a Kalman or extended Kalman filter and (3) the attitude and angular velocity measurement errors are assumed to be bounded, with known ellipsoidal uncertainty bounds. The static attitude estimation scheme presented here is based on [22]. The attitude is represented globally using proper orthogonal matrices and the exponential map on the set of 3×3 skew-symmetric matrices. Such a global representation has been recently used for partial attitude estimation with a linear dynamics model in [20]. The uncertainty ellipsoids are represented in terms of local charts that are based at the current estimated attitude, and as such do not suffer from the coordinate singularities that arise when attempting to represent uncertainties using only a single chart based at the initial attitude, as is typical for competing approaches. The estimation scheme presented here is deterministic, based on known measurement uncertainty bounds propagated by the dynamic flow.

The attitude determination problem for a rigid body from vector measurements was first posed in [28]. A sample of the literature in spacecraft attitude estimation can be found in [7,2,17,24,25]. Applications of attitude estimation to unmanned vehicles and robots can be found in [20,3,21,27]. Most existing attitude estimation schemes use generalized coordinate representations of the attitude. As is well known, minimal coordinate representations of the rotation group, like Euler angles, Rodrigues parameters, and modified Rodrigues parameters (see [8]), usually lead to geometric or kinematic singularities. Non-minimal coordinate representations, like the unit quaternions used in the quaternion estimation (QUEST) algorithm [24] and its several variants [2,19,25], have their own associated problems. Besides the extra constraint of unit norm that one needs to impose, the quaternion vector itself can be defined in one

* Corresponding author. Tel.: +1 808 956 2142.
E-mail addresses: aksanyal@hawaii.edu (A.K. Sanyal), tylee@umich.edu (T. Lee), mleok@math.purdue.edu (M. Leok), nhm@umich.edu (N.H. McClamroch).

0167-6911/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.sysconle.2007.08.014

of two ways, depending on the sense of rotation used to define the principal angle.

A brief outline of this paper is given here. In Section 2, the attitude determination problem for vector measurements with measurement noise is introduced, and a global attitude determination algorithm which minimizes the attitude estimation error is presented. The attitude dynamics and dynamic estimation problem is formulated for a rigid body in an attitude-dependent potential. Section 3 presents the attitude estimation scheme assuming that both attitude and angular velocity measurements are available simultaneously. Sufficient conditions for convergence of the estimates are given. This attitude estimation scheme has also been extended and applied to the case where only attitude but no angular velocity measurements are available, and recently reported in [14]. Section 4 presents concluding remarks and observations.

2. Attitude estimation from vector observations

Attitude of a rigid body is defined as the orientation of a body fixed frame with respect to an inertial reference frame; it is represented by a rotation matrix that is a 3×3 orthogonal matrix with determinant $+1$. Rotation matrices have a group structure denoted by SO(3), and its action on $\mathbb{R}^3$ takes a vector represented in body fixed frame into its representation in the reference frame by matrix multiplication.

2.1. Attitude determination procedure

We assume that there are m fixed points in the spatial reference frame, no two of which are co-linear, that are measured in the body frame. We denote the known direction of the ith point in the spatial reference frame as $e^i \in \mathbb{S}^2$ where $\mathbb{S}^2$ denotes the two sphere (embedded in $\mathbb{R}^3$). The corresponding vector is represented in the body fixed frame as $b^i \in \mathbb{S}^2$. Since we only measure directions, we normalize e^i and b^i so that they have unit lengths. The e^i and b^i are related by a rotation matrix $C \in \mathrm{SO}(3)$ that defines the attitude of the rigid body;

$$e^i = Cb^i$$

for all $i \in \{1, 2, \ldots, m\}$. We assume that e^i is known accurately and we assume that b^i is measured in the body fixed frame. Let the measured direction vector be $\tilde{b}^i \in \mathbb{S}^2$, which contains measurement errors, and let an estimate of the rotation matrix be $\widehat{C} \in \mathrm{SO}(3)$. The vector estimation errors are given by

$$e^i - \widehat{C}\tilde{b}^i, \quad i = 1, \ldots, m.$$

The weighted 2 norm of these errors is given by the error functional,

$$\begin{aligned}\mathscr{J}(\widehat{C}) &= \frac{1}{2}\sum_{i=1}^{m} w_i (e^i - \widehat{C}\tilde{b}^i)^{\mathrm{T}}(e^i - \widehat{C}\tilde{b}^i),\\ &= \frac{1}{2}\mathrm{tr}[(E - \widehat{C}\tilde{B})^{\mathrm{T}} W (E - \widehat{C}\tilde{B})],\end{aligned}$$

where $E = [e^1, e^2, \ldots, e^m] \in \mathbb{R}^{3\times m}$, $\tilde{B} = [\tilde{b}^1, \tilde{b}^2, \ldots, \tilde{b}^m] \in \mathbb{R}^{3\times m}$, and $W = \mathrm{diag}[w^1, w^2, \ldots, w^m] \in \mathbb{R}^{m\times m}$ has a weighting factor for each measurement. We assume that $m \geqslant 3$ in this paper. If $m = 2$, we can take the cross-product of the two measured unit vectors $\tilde{b}^1 \times \tilde{b}^2 = \tilde{b}^3$ and treat that as a third measured direction, with the corresponding unit vector in the inertial frame taken to be $e^3 = e^1 \times e^2$. The attitude determination problem then consists of finding $\widehat{C} \in \mathrm{SO}(3)$ such that the error functional $\mathscr{J}$ is minimized:

$$\widehat{C} = \arg \min_{\tilde{C}\in \mathrm{SO}(3)} \mathscr{J}(\tilde{C}). \tag{1}$$

The problem (1) is known as Wahba's problem [28]. The original solution of Wahba's problem is given in [11], and a solution expressed in terms of quaternions, known as the QUEST algorithm, is presented in [24]. A solution without using generalized attitude coordinates is given in [24]. A necessary condition for optimality of (1) is given in [22] as

$$L^{\mathrm{T}}\widehat{C} = \widehat{C}^{\mathrm{T}}L, \tag{2}$$

where $L = EW\tilde{B}^{\mathrm{T}} \in \mathbb{R}^{3\times 3}$.

The following result, proved in [22], gives an unique estimate $\widehat{C} \in \mathrm{SO}(3)$ of the attitude matrix that satisfies (2) and solves the attitude determination problem (1).

Theorem 2.1. *The unique minimizing solution to the attitude determination problem* (1) *is given by*

$$\widehat{C} = SL, \quad S = Q\sqrt{(RR^{\mathrm{T}})^{-1}}Q^{\mathrm{T}}, \tag{3}$$

where

$$L = QR, \quad Q \in \mathrm{SO}(3) \tag{4}$$

and R is upper triangular and invertible; *this is the QR decomposition of L. The symmetric positive definite* (*principal*) *square root is used in* (3).

Proof. Using the QR decomposition of L given by (4), we can express the orthogonality condition of $\widehat{C}$ as follows:

$$\widehat{C}\widehat{C}^{\mathrm{T}} = SQRR^{\mathrm{T}}Q^{\mathrm{T}}S = I_3.$$

Due to the necessary condition (2), S is symmetric since $L^{\mathrm{T}}\widehat{C} = L^{\mathrm{T}}SL$ is symmetric. The sufficient condition for $\widehat{C}$ to minimize cost function $\mathscr{J}$, is obtained by taking its second variation with respect to $\widehat{C}$. The following sufficient condition is obtained in [22]:

$$\delta^2_{\widehat{C}}\mathscr{J} = -\mathrm{tr}[(L^{\mathrm{T}}\widehat{C}U^2)] > 0. \tag{5}$$

It is shown in [22] that the sufficient condition (5) is equivalent to $D = L^{\mathrm{T}}\widehat{C}$ being positive definite, i.e., $v\mathrm{T}Dv > 0$ for any $v \in \mathbb{R}^3$. This in turn is equivalent to $S = (L^{\mathrm{T}})^{-1}DL^{-1}$ being positive definite.

Therefore, S is given by

$$S = \sqrt{Q(RR^{\mathrm{T}})^{-1}Q^{\mathrm{T}}} = Q\sqrt{(RR^{\mathrm{T}})^{-1}}Q^{\mathrm{T}},$$

where the principal (positive definite) square root is taken, as given by Eq. (3). This makes S positive definite, as required by

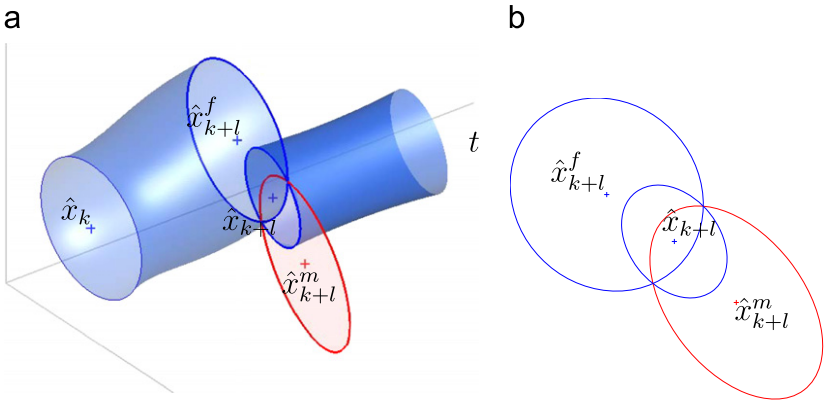


Fig. 1. Uncertainty ellipsoids. (a) Propagation of uncertainty ellipsoid. (b) Filtering procedure.

Eq. (5). By construction, $\widehat{C} = SL$ satisfies $\widehat{C}\widehat{C}^{\mathrm{T}} = I_3$. Now we check the determinant of $\widehat{C} = SL$:

$$\begin{aligned}\det \widehat{C} &= \det S \det L \\ &= \left(\det Q^{\mathrm{T}} \det \sqrt{(RR^{\mathrm{T}})^{-1}} \det Q\right)(\det Q \det R) \\ &= \frac{(\det Q)^2}{\sqrt{\det R \det R^{\mathrm{T}}}} \det Q \det R \\ &= \frac{1}{\sqrt{(\det R)^2}} \det Q \det R = \det Q = 1,\end{aligned}$$

since $Q \in \mathrm{SO}(3)$. Since $\mathscr{J}$ is a Morse function, its critical points are non-degenerate. Due to the Morse lemma [16], these critical points are isolated, and hence the estimate given by (3) is uniquely minimizing. □

It would be pertinent to mention here that there are a total of eight critical points of the function $\mathscr{J} : \mathrm{SO}(3) \to \mathbb{R}$ depending on the different sign permutations of the three real eigenvalues of the symmetric matrix S; these permutations are $(-,-,-)$, $(-,-,+)$, $(-,+,-)$, $(+,-,-)$, $(+,+,-)$, $(+,-,+)$, $(-,+,+)$, and $(+,+,+)$. The last of these permutations gives rise to a positive definite S that results in the absolute minimum of $\mathscr{J}$ over SO(3).

2.2. State bounding estimation

Here we discuss the general idea of deterministic state bounding estimation, using ellipsoidal sets to describe state uncertainty and measurement noise. A stochastic state estimator requires probabilistic models for the state uncertainty and the noise. However, statistical properties of the uncertainty and the noise are often not available. We usually make statistical assumptions on disturbance and noise in order to make the estimation problem mathematically tractable. In many practical situations such idealized assumptions are not appropriate, and this may cause poor estimation performance [26].

An alternative deterministic approach is to specify bounds on the uncertainty and the measurement noise without any assumption on its distribution. Noise bounds are available in many cases, and deterministic estimation is robust to the noise distribution. An efficient but flexible way to describe the bounds is using ellipsoidal sets, referred to as uncertainty ellipsoids.

The deterministic estimation process using uncertainty ellipsoids has the same structure as the Kalman filter. This deterministic estimation procedure is illustrated in Fig. 1, where the figure on the left shows evolution of an uncertainty ellipsoid in time, and the figure on the right shows a cross-section at a fixed time when the state is measured. Suppose that the time interval between two sets of measurements is divided into l equal time steps for discrete integration, and the subscript k denotes the kth discrete integration time step. At the previous measurement instant, corresponding to the kth time step, the state is bounded by an uncertainty ellipsoid centered at the estimated state $\hat{x}_k$. This initial ellipsoid and the estimated state at its center are propagated through time. Depending on the dynamics of the system, the size and the shape of the tube are changed. The new set of measurements are taken at the $(k+l)$th time step. At the $(k+l)$th time step, the predicted uncertainty ellipsoid is centered at $\hat{x}^f_{k+l}$. Another ellipsoidal bound on the state is obtained from the measurements. The measured uncertainty ellipsoid is centered at $\hat{x}^m_{k+l}$. The state lies in the intersection of the two ellipsoids. In the estimation procedure, we find the smallest ellipsoid that contains this intersection, which is shown in the right figure. The center of the new ellipsoid, $\hat{x}_{k+l}$ is considered as a point estimate at time step $k+l$, and the magnitude of the new uncertainty ellipsoid measures the accuracy of the estimation. If the size of the uncertainty ellipsoid is small, then we can conclude that the estimated state is accurate. The deterministic estimation is optimal in the sense that the size of the new ellipsoid is minimized.

The deterministic estimation process is based on the state estimation techniques developed in [23]. Optimal deterministic state or parameter estimation is considered in [9,15,18], where an analytic solution for the minimum ellipsoid that contains a union or an intersection of ellipsoids is given. Parameter estimation in the presence of bounded noise is dealt with in [1,4].

2.3. Attitude estimation problem formulation

2.3.1. Equations of motion

We consider estimation of the attitude dynamics of a rigid body in the presence of an attitude dependent potential. We assume that the potential $U(\cdot) : \mathrm{SO}(3) \mapsto \mathbb{R}$ is determined by the attitude of the rigid body, $C \in \mathrm{SO}(3)$. A spacecraft on a circular orbit including gravity gradient effects [13], or a 3D pendulum [12] can be modeled in this way. The continuous equations of motion are given by

$$J\dot{\omega} + \omega \times J\omega = M, \tag{6}$$

$$\dot{C} = CS(\omega), \tag{7}$$

where $J \in \mathbb{R}^{3\times3}$ is the moment of inertia matrix of the rigid body, $\omega \in \mathbb{R}^3$ is the angular velocity of the body expressed in the body fixed frame, and $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is a skew mapping defined such that $S(x)y = x \times y$ for all $x, y \in \mathbb{R}^3$. $M \in \mathbb{R}^3$ is the moment due to the potential. The moment is determined by the relationship, $S(M) = \frac{\partial U}{\partial C}^{\mathrm{T}} C - C^{\mathrm{T}} \frac{\partial U}{\partial C}$, or more explicitly,

$$M = r_1 \times v_{r_1} + r_2 \times v_{r_2} + r_3 \times v_{r_3}, \tag{8}$$

where $r_i, v_{r_i} \in \mathbb{R}^{1\times3}$ are the ith row vectors of C and $\partial U/\partial C$, respectively. Here the matrix $\partial U/\partial C \in \mathbb{R}^{3\times3}$ is defined such that $[\partial U/\partial C]_{ij} = \partial U/\partial [C]_{ij}$ for $i, j \in (1, 2, 3)$, where $[A]_{ij}$ denote the (i, j)th element of a matrix A. The detailed description of this rigid body model and the derivation of the above equations can be found in [12].

General numerical integration methods, including the popular Runge–Kutta schemes, typically preserve neither first integrals nor the characteristics of the configuration space, SO(3). In particular, the orthogonal structure of the rotation matrices is not preserved numerically. It is often proposed to parameterize (7) by Euler angles or quaternions instead of integrating (7) directly. However, Euler angles have singularities. The numerical simulation process has to be monitored and switching between Euler angle charts is necessary in order to avoid singularities. Quaternions are free of singularities, but the quaternion representing the attitude is required to have unit length. The matrix corresponding to a quaternion which is not of unit length is not orthogonal, and hence does not represent a rotation.

To resolve these problems, a Lie group variational integrator for the attitude dynamics of a rigid body is proposed in [12]. This Lie group variational integrator is described by the discrete time equations.

$$hS\left(J\omega_k + \frac{h}{2}M_k\right) = F_k J_d - J_d F_k^{\mathrm{T}}, \tag{9}$$

$$C_{k+1} = C_k F_k, \tag{10}$$

$$J\omega_{k+1} = F_k^{\mathrm{T}} J\omega_k + \frac{h}{2}F_k^{\mathrm{T}} M_k + \frac{h}{2}M_{k+1}, \tag{11}$$

where $J_d \in \mathbb{R}^3$ is a non-standard moment of inertia matrix defined by $J_d = \frac{1}{2}\mathrm{tr}[J]I_{3\times3} - J$, and $F_k \in \mathrm{SO}(3)$ is the relative attitude over an integration step. The constant $h \in \mathbb{R}^+$ is the integration step size. This integrator yields a map $(C_k, \omega_k) \mapsto (C_{k+1}, \omega_{k+1})$ by solving (9) to obtain $F_k \in \mathrm{SO}(3)$ and substituting it into (10) and (11) to obtain C_{k+1} and ω_{k+1}. Numerically, we ensure that F_k remains on SO(3) by requiring that $F_k = \exp(S(f_k))$, where $f_k \in \mathbb{R}^3$. This allows us to express the discrete equations in terms of $f_k \in \mathbb{R}^3$ as opposed to $F_k \in \mathrm{SO}(3)$.

Since this integrator does not use a local parameterization, the attitude is defined globally without singularities. It preserves the orthogonal structure of SO(3) because the rotation matrix is updated by a multiplication of two rotation matrices in (10), which is a group operation of SO(3). This integrator is obtained from a discrete variational principle, and it exhibits the characteristic symplectic and momentum preservation properties, and good energy behavior characteristic of variational integrators. We use (9), (10), and (11) in the following development of the attitude estimator.

2.3.2. Uncertainty ellipsoid

An uncertainty ellipsoid in $\mathbb{R}^n$ is defined as

$$\mathscr{E}_{\mathbb{R}^n}(\hat{x}, P) = \{x \in \mathbb{R}^n | (x - \hat{x})^{\mathrm{T}} P^{-1} (x - \hat{x}) \leqslant 1\}, \tag{12}$$

where $\hat{x} \in \mathbb{R}^n$ and $P \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. We call $\hat{x}$ the center of the uncertainty ellipsoid, and we call P the uncertainty matrix that determines the size and the shape of the uncertainty ellipsoid. The size of an uncertainty ellipsoid is measured by $\mathrm{tr}[P]$. It equals the sum of the squares of the semi principal axes of the ellipsoid.

The configuration space of the attitude dynamics is SO(3), so the state evolves in the 6 D tangent bundle, TSO(3). The corresponding uncertainty ellipsoid is a submanifold of TSO(3).

We obtain the diffeomorphism $\mathrm{TSO}(3) \simeq \mathrm{SO}(3) \times \mathfrak{so}(3)$ by left trivialization, and the diffeomorphism $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ identifies $\mathfrak{so}(3)$ and $\mathbb{R}^3$. Therefore, $\mathrm{TSO}(3) \simeq \mathrm{SO}(3) \times \mathbb{R}^3$. For the remainder of the paper, we will represent the state as an element of $\mathrm{SO}(3) \times \mathbb{R}^3$. An uncertainty ellipsoid centered at $(\hat{C}, \hat{\omega})$ is induced from an uncertainty ellipsoid in $\mathbb{R}^6$, using the Lie algebra $\mathfrak{so}(3)$;

$$\mathscr{E}(\hat{C}, \hat{\omega}, P) = \left\{ C \in \mathrm{SO}(3),\ \omega \in \mathbb{R}^3 \middle| \begin{bmatrix} \zeta \\ \delta\omega \end{bmatrix} \in \mathscr{E}_{\mathbb{R}^6}(0_6, P) \right\}, \tag{13}$$

where $S(\zeta) = \log \mathrm{m}(\hat{C}^{\mathrm{T}} C) \in \mathfrak{so}(3)$, $\delta\omega = \omega - \hat{\omega} \in \mathbb{R}^3$, and $P \in \mathbb{R}^{6\times6}$ is a symmetric positive definite matrix. Equivalently, an element $(C, \omega) \in \mathscr{E}(\hat{C}, \hat{\omega}, P)$ can be written as

$$C = \hat{C}\exp(S(\zeta)),$$

$$\omega = \hat{\omega} + \delta\omega$$

for some $x = [\zeta^{\mathrm{T}},\ \delta\omega^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^6$ satisfying $x^{\mathrm{T}} P^{-1} x \leqslant 1$.

2.3.3. Uncertainty model

We describe the measurement error models for the measured direction vectors and the angular velocity. The direction vector $b^i \in \mathbb{S}^2$ is measured in the body fixed frame, and let $\tilde{b}^i \in \mathbb{S}^2$ denote the measured direction. Since we only measure directions, we normalize b^i and $\tilde{b}^i$ so that they have unit lengths.

Therefore it is inappropriate to express the measurement error by a vector difference. The measurement error is modeled by rotation of the measured direction;

$$b^i = \exp(S(v^i))\tilde{b}^i \simeq \tilde{b}^i + S(v^i)\tilde{b}^i, \tag{14}$$

where $v^i \in \mathbb{R}^3$ is the measurement error, which represents the Euler axis of rotation vector from $\tilde{b}^i$ to b^i, and $\|v^i\|$ is the corresponding rotation angle in radians. We assume that the measurement error is small to obtain the second equality.

The angular velocity measurement errors are modeled as

$$\omega = \tilde{\omega} + \upsilon, \tag{15}$$

where $\tilde{\omega} \in \mathbb{R}^3$ is the measured angular velocity and $\upsilon \in \mathbb{R}^3$ is an additive measurement error.

We assume that the initial conditions and the measurement error are bounded by prescribed uncertainty ellipsoids.

$$(C_0, \omega_0) \in \mathscr{E}(\hat{C}_0, \hat{\omega}_0, P_0), \tag{16}$$

$$v^i \in \mathscr{E}_{\mathbb{R}^3}(0, S^i), \tag{17}$$

$$\upsilon \in \mathscr{E}_{\mathbb{R}^3}(0, T), \tag{18}$$

where $P_0 \in \mathbb{R}^{6\times 6}$ and $S^i, T \in \mathbb{R}^{3\times 3}$ are symmetric positive definite matrices that define the shape and the size of the uncertainty ellipsoids.

3. Attitude estimation with angular velocity measurements

In this section, we develop a deterministic estimator for the attitude and the angular velocity of a rigid body assuming that both the attitude measurement and the angular velocity measurements are available. The estimation process consists of three stages; flow update, measurement update, and filtering. The flow update predicts the uncertainty ellipsoid in the future. The measurement update finds an uncertainty ellipsoid in the state space using the measurements and the measurement error model. The filtering stage obtains a new uncertainty ellipsoid compatible with the predicted uncertainty ellipsoid and the measured uncertainty ellipsoid.

The superscript i denotes the ith directional measurement. The superscript f denotes variables related to the flow update, while the superscript m denotes variables related to the measurement update. The notation $\tilde{\cdot}$ denotes a measured variable, while $\hat{\cdot}$ denotes an estimated variable.

3.1. Flow update

Suppose that the attitude and the angular momentum at the kth step, which corresponds to the previous measurement instant, lie in a given uncertainty ellipsoid:

$$(C_k, \omega_k) \in \mathscr{E}(\hat{C}_k, \hat{\omega}_k, P_k).$$

The flow update gives us the center and the uncertainty matrix that define the uncertainty ellipsoid at the $(k+l)$th step (the current measurement instant) using the given uncertainty ellipsoid at the kth step. Since the attitude dynamics of a rigid body is nonlinear, the boundary of the state at the $(k+l)$th step is not an ellipsoid in general. We assume that the given uncertainty ellipsoid at the kth step is sufficiently small that the states in the uncertainty ellipsoids can be approximated by linearized equations of motion. Then we can guarantee that the boundary of the state at the $(k+l)$th step is an ellipsoid, and we can compute the center and the uncertainty matrix at the $(k+l)$th step separately.

Center: For the given center, $(\hat{C}_k, \hat{\omega}_k)$, the center of the uncertainty ellipsoid due to flow propagation is denoted $(\hat{C}^f_{k+1}, \hat{\omega}^f_{k+1})$. This center is obtained from the discrete equations of motion, (9), (10), and (11) applied to $(\hat{C}_k, \hat{\omega}_k)$:

$$hS\left(J\hat{\omega}_k + \frac{h}{2}\hat{M}_k\right) = \hat{F}_k J_d - J_d \hat{F}_k^{\mathrm{T}}, \tag{19}$$

$$\hat{C}^f_{k+1} = \hat{C}_k \hat{F}_k, \tag{20}$$

$$J\hat{\omega}^f_{k+1} = \hat{F}_k^{\mathrm{T}} J\hat{\omega}_k + \frac{h}{2}\hat{F}_k^{\mathrm{T}}\hat{M}_k + \frac{h}{2}\hat{M}_{k+1}. \tag{21}$$

This integrator yields a map $(\hat{C}_k, \hat{\omega}_k) \mapsto (\hat{C}^f_{k+1}, \hat{\omega}^f_{k+1})$, and this process can be repeated to find the center at the $(k+l)$th step, $(\hat{C}^f_{k+l}, \hat{\omega}^f_{k+l})$.

Uncertainty matrix: We assume that an uncertainty ellipsoid contains small perturbations from its center. Then the uncertainty matrix is obtained by linearizing the above discrete equations of motion. At the $(k+1)$th step, the uncertainty ellipsoid is represented by perturbations from the center $(\hat{C}^f_{k+l}, \hat{\omega}^f_{k+l})$ as

$$C_{k+1} = \hat{C}^f_{k+1}\exp(S(\zeta^f_{k+1})),$$

$$\omega_{k+1} = \hat{\omega}^f_{k+1} + \delta\omega^f_{k+1}$$

for some $\zeta^f_{k+1}, \delta\omega^f_{k+1} \in \mathbb{R}^3$. The uncertainty matrix at the $(k+1)$th step is obtained by finding a bound on $\zeta^f_{k+1}, \delta\omega^f_{k+1} \in \mathbb{R}^3$. Assume that the uncertainty ellipsoid at the kth step is sufficiently small. Then, $\zeta^f_{k+1}, \delta\omega^f_{k+1}$ are represented by the following linear equations (using the results presented in [13]):

$$x^f_{k+1} = A^f_k x_k,$$

where $x_k = [\zeta_k^{\mathrm{T}}, \delta\omega_k^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^6$, and $A^f_k \in \mathbb{R}^{6\times 6}$ can be suitably defined. Since $(C_k, \omega_k) \in \mathscr{E}(\hat{C}_k, \hat{\omega}_k, P_k)$, $x_k \in \mathscr{E}_{\mathbb{R}^6}(0, P_k)$ by the definition of the uncertainty ellipsoid given in (13). Then we can show that $A^f_k x_k$ lies in the following uncertainty ellipsoid.

$$A^f_k x_k \in \mathscr{E}_{\mathbb{R}^6}(0, A^f_k P_k (A^f_k)^{\mathrm{T}}).$$

Thus, the uncertainty matrix at the $(k+1)$th step is given by

$$P^f_{k+1} = A^f_k P_k (A^f_k)^{\mathrm{T}}. \tag{22}$$

The above equation can be applied repeatedly to find the uncertainty matrix at the $(k+l)$th step.

We have obtained expressions to predict the center and the uncertainty matrix in the future from the current uncertainty

ellipsoid using the discrete flow. In summary, the uncertainty ellipsoid at the $(k+l)$th step is computed using (19), (20), (21), and (22) as:

$$(C_{k+l}, \omega_{k+l}) \in \mathscr{E}(\hat{C}^f_{k+l}, \hat{\omega}^f_{k+l}, P^f_{k+l}),$$
$$P^f_{k+l} = A^f P_k (A^f_k)^{\mathrm{T}}, \tag{23}$$

where $A^f = A^f_{k+l-1} A^f_{k+l-2} \cdots A^f_k \in \mathbb{R}^{6\times 6}$.

3.2. Measurement update

We assume that the attitude and the angular velocity of a rigid body are measured simultaneously. The measured attitude and the measured angular velocity have uncertainties since the measurements contain measurement errors. However, we can find bounds for the actual state because the measurement errors are bounded by known uncertainty ellipsoids given by (17) and (18). The measurement update stage finds an uncertainty ellipsoid in the state space using the measurements and the measurement error models. The measured attitude and the measured angular velocity are the center of the measured uncertainty ellipsoid, and the measurement error models are used to find the uncertainty matrix.

Center: The center of the uncertainty ellipsoid, $(\hat{C}^m_{k+l}, \hat{\omega}^m_{k+l})$ is obtained by measurements. The attitude is determined by measuring the directions to the known points in the inertial frame. Let the measured directions to the known points be $\tilde{B}_{k+l} = [\tilde{b}^1, \tilde{b}^2, \ldots, \tilde{b}^m] \in \mathbb{R}^{3\times m}$. Then, the attitude $\hat{C}^m_{k+l}$ satisfies the following necessary condition given in (2):

$$(\hat{C}^m_{k+l})^{\mathrm{T}} \tilde{L}_{k+l} - \tilde{L}^{\mathrm{T}}_{k+l} \hat{C}^m_{k+l} = 0, \tag{24}$$

where $\tilde{L}_{k+l} = E_{k+l} W_{k+l} \tilde{B}^{\mathrm{T}}_{k+l} \in \mathbb{R}^{3\times 3}$. The solution of (24) is obtained by a QR factorization of $\tilde{L}_{k+l}$ as given in Theorem 2.1

$$\hat{C}^m_{k+l} = \left(Q_q \sqrt{(Q_r Q_r^{\mathrm{T}})^{-1}} Q_q^{\mathrm{T}} \right) \tilde{L}_{k+l}, \tag{25}$$

where $Q_q \in \mathrm{SO}(3)$ is an orthogonal matrix and $Q_r \in \mathbb{R}^{3\times 3}$ is an upper triangular matrix satisfying $\tilde{L}_{k+l} = Q_q Q_r$.

The angular velocity is measured directly,

$$\hat{\omega}^m_{k+l} = \tilde{\omega}_{k+l}. \tag{26}$$

Uncertainty matrix: We can represent the actual state at the $(k+l)$th step using the measured center and perturbations as follows:

$$C_{k+l} = \hat{C}^m_{k+l} \exp(S(\zeta^m_{k+l})), \tag{27}$$

$$\omega_{k+l} = \hat{\omega}^m_{k+l} + \delta\omega^m_{k+l} \tag{28}$$

for $\zeta^m_{k+l}, \delta\omega^m_{k+l} \in \mathbb{R}^3$. The uncertainty matrix is obtained by finding an ellipsoid containing $\zeta^m_{k+l}, \delta\omega^m_{k+l}$.

We determine the attitude indirectly by comparing the known directions in the reference frame with measurements in the body frame. So, we need to transform the uncertainties in the direction measurements into the uncertainties in the rotation matrix by (24). Using the measurement error model defined in (14), the actual direction matrix to the known point B_{k+l} is given by

$$B_{k+l} = \tilde{B}_{k+l} + \delta\tilde{B}_{k+l}, \tag{29}$$

where $\delta B_{k+l} = [S(v^1)\tilde{b}^1, S(v^2)\tilde{b}^2, \ldots, S(v^m)\tilde{b}^m] \in \mathbb{R}^{3\times m}$.

The actual matrix giving the known directions B_{k+l} and the actual attitude C_{k+l} at the $(k+l)$th step also satisfy (25);

$$C^{\mathrm{T}}_{k+l} L_{k+l} - L^{\mathrm{T}}_{k+l} C_{k+l} = 0, \tag{30}$$

where $L_{k+l} = E_{k+l} W_{k+l} B^{\mathrm{T}}_{k+l} \in \mathbb{R}^{3\times 3}$. Substituting (27) and (29) into (30), and assuming that the size of the measurement error is sufficiently small, the above equation can be written as

$$\tilde{L}^{\mathrm{T}}_{k+l} \hat{C}^m_{k+l} S(\zeta^m_{k+l}) + S(\zeta^m_{k+l})(\hat{C}^m_{k+l})^{\mathrm{T}} \tilde{L}_{k+l}$$
$$= (\hat{C}^m_{k+l})^{\mathrm{T}} E_{k+l} W_{k+l} \delta B^{\mathrm{T}}_{k+l} - \delta B_{k+l} W_{k+l} E^{\mathrm{T}}_{k+l} \hat{C}^m_{k+l}.$$

Using the identity, $S(x)A + A^{\mathrm{T}} S(x) = S(\{\mathrm{tr}[A] I_{3\times 3} - A\}x)$ for $A \in \mathbb{R}^{3\times 3}$, $x \in \mathbb{R}^3$, the above equation can be written in a vector form.

$$\{\mathrm{tr}[(\hat{C}^m_{k+l})^{\mathrm{T}} \tilde{L}_{k+l}] - (\hat{C}^m_{k+l})^{\mathrm{T}} \tilde{L}_{k+l}\} \zeta^m_{k+l}$$
$$= -\sum_{i=1}^{m} w_i \{\mathrm{tr}[\tilde{b}^i (e^i)^{\mathrm{T}} \hat{C}^m_{k+l}] I_{3\times 3} - \tilde{b}^i (e^i)^{\mathrm{T}} \hat{C}^m_{k+l}\} v^i.$$

Then, we obtain

$$\zeta^m_{k+l} = \sum_{i=1}^{m} \mathscr{A}^{m,i}_{k+l} v^i, \tag{31}$$

where

$$\mathscr{A}^{m,i}_{k+l} = -\{\mathrm{tr}[(\hat{C}^m_{k+l})^{\mathrm{T}} \tilde{L}_{k+l}] - (\hat{C}^m_{k+l})^{\mathrm{T}} \tilde{L}_{k+l}\}^{-1}$$
$$w_i \{\mathrm{tr}[\tilde{b}^i (e^i)^{\mathrm{T}} \hat{C}^m_{k+l}] I_{3\times 3} - \tilde{b}^i (e^i)^{\mathrm{T}} \hat{C}^m_{k+l}\}. \tag{32}$$

This equation expresses the uncertainty in the measured attitude as a linear combination of the directional measurement errors.

The perturbation of the angular velocity $\delta\omega^m_{k+l}$ is equal to the angular velocity measurement error υ_{k+l}, since we measure the angular velocity directly. Substituting (28) into (15), we obtain

$$\delta\omega^m_{k+l} = \upsilon_{k+l}. \tag{33}$$

Define $x^m_{k+l} = [(\zeta^m_{k+l})^{\mathrm{T}}, \ (\delta\omega^m_{k+l})^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^6$. Using (31) and (33),

$$x^m_{k+l} = H_1 \sum_{i=1}^{m} \mathscr{A}^{m,i}_{k+l} v^i_{k+l} + H_2 \upsilon_{k+l},$$

where $H_1 = [I_{3\times 3}, \ 0_{3\times 3}]^{\mathrm{T}}$, $H_2 = [0_{3\times 3}, \ I_{3\times 3}]^{\mathrm{T}} \in \mathbb{R}^{6\times 3}$. Now x^m_{k+l} is expressed as a linear combination of the measurement errors v^i and υ. Using (17) and (18), we can show that each term in the right-hand side of the above equation is in the following uncertainty ellipsoids:

$$H_1 \mathscr{A}^{m,i}_{k+l} v^i_{k+l} \in \mathscr{E}_{\mathbb{R}^6}(0, H_1 \mathscr{A}^{m,i}_{k+l} S^i_{k+l} (\mathscr{A}^{m,i}_{k+l})^{\mathrm{T}} H_1^{\mathrm{T}}),$$

$$H_2 \upsilon_{k+l} \in \mathscr{E}_{\mathbb{R}^6}(0, H_2 T_{k+l} H_2^{\mathrm{T}}).$$

Thus, the uncertainty ellipsoid for x^m_{k+l} is obtained as the vector sum of the above uncertainty ellipsoids. The measurement

update procedure is to find a minimal ellipsoid that contains the vector sum of those uncertainty ellipsoids. Expressions for a minimal ellipsoid containing the vector sum of multiple ellipsoids are presented in [15] and [9]. Using the results, we obtain

$$P^m_{k+l} = \left\{ \sum_{i=1}^{m} \sqrt{\mathrm{tr}[H_1 \mathscr{A}^{m,i}_{k+l} S^i_{k+l} (\mathscr{A}^{m,i}_{k+l})^{\mathrm{T}} H_1^{\mathrm{T}}]} + \sqrt{\mathrm{tr}[H_2 T_{k+l} H_2^{\mathrm{T}}]} \right\} \times \left\{ \sum_{i=1}^{m} \frac{H_1 \mathscr{A}^{m,i}_{k+l} S^i_{k+l} (\mathscr{A}^{m,i}_{k+l})^{\mathrm{T}} H_1^{\mathrm{T}}}{\sqrt{\mathrm{tr}[H_1 \mathscr{A}^{m,i}_{k+l} S^i_{k+l} (\mathscr{A}^{m,i}_{k+l})^{\mathrm{T}} H_1^{\mathrm{T}}]}} + \frac{H_2 T_{k+l} H_2^{\mathrm{T}}}{\sqrt{\mathrm{tr}[H_2 T_{k+l} H_2^{\mathrm{T}}]}} \right\}. \tag{34}$$

In summary, the measured uncertainty ellipsoid at the $(k+l)$th step is defined by (25), (26), and (34);

$$(C_{k+l}, \omega_{k+l}) \in \mathscr{E}(\hat{C}^m_{k+l}, \hat{\omega}^m_{k+l}, P^m_{k+l}). \tag{35}$$

3.3. Filtering procedure

The filtering procedure is to find a new uncertainty ellipsoid compatible with both the predicted uncertainty ellipsoid and the measured uncertainty ellipsoid. From (23) and (35), we know that the state at $(k+l)$th step lies in the intersection of the two uncertainty ellipsoids:

$$(C_{k+l}, \omega_{k+l}) \in \mathscr{E}(\hat{C}^f_{k+l}, \hat{\omega}^f_{k+l}, P^f_{k+l}) \cap \mathscr{E}(\hat{C}^m_{k+l}, \hat{\omega}^m_{k+l}, P^m_{k+l}). \tag{36}$$

However, the intersection of two ellipsoids is not generally an ellipsoid. We find a minimal uncertainty ellipsoid containing this intersection. We first obtain equivalent uncertainty ellipsoids in $\mathbb{R}^6$, and convert them to uncertainty ellipsoids in TSO(3). We omit the subscript $(k+l)$ in this subsection for convenience.

The uncertainty ellipsoid obtained from the measurements, $\mathscr{E}(\hat{C}^m, \hat{\omega}^m, P^m)$, is identified by its center $(\hat{C}^m, \hat{\omega}^m)$, and the uncertainty ellipsoid in $\mathbb{R}^6$:

$$(\zeta^m, \delta\omega^m) \in \mathscr{E}_{\mathbb{R}^6}(0_{6\times 1}, P^m), \tag{37}$$

where $S(\zeta^m) = \log \mathrm{m}(\hat{C}^{m,T} C) \in \mathfrak{so}(3)$, $\delta\omega^m = \omega - \hat{\omega}^m \in \mathbb{R}^3$. Similarly, the uncertainty ellipsoid obtained from the flow update, $\mathscr{E}(\hat{C}^f, \hat{\omega}^f, P^f)$, is identified by its center $(\hat{C}^f, \hat{\omega}^f)$, and the uncertainty ellipsoid in $\mathbb{R}^6$.

$$(\zeta^f, \delta\omega^f) \in \mathscr{E}_{\mathbb{R}^6}(0_{6\times 1}, P^f), \tag{38}$$

where $S(\zeta^f) = \log \mathrm{m}(\hat{C}^{f,T} C) \in \mathfrak{so}(3)$, $\delta\omega^f = \omega - \hat{\omega}^f \in \mathbb{R}^3$. Equivalently, an element $(C^f, \omega^f) \in \mathscr{E}(\hat{C}^f, \hat{\omega}^f, P^f)$, is given by

$$C^f = \hat{C}^f \exp(S(\zeta^f)), \tag{39}$$

$$\omega^f = \hat{\omega}^f + \delta\omega^f, \tag{40}$$

for some $(\zeta^f, \delta\omega^f) \in \mathscr{E}_{\mathbb{R}^6}(0_{6\times 1}, P^f)$. We find an equivalent expression for (38) based on the center $(\hat{C}^m, \hat{\omega}^m)$ obtained from the measurements.

Define $\hat{\zeta}^{mf}, \delta\hat{\omega}^{mf} \in \mathbb{R}^3$ such that

$$\hat{C}^f = \hat{C}^m \exp(S(\hat{\zeta}^{mf})), \tag{41}$$

$$\hat{\omega}^f = \hat{\omega}^m + \delta\hat{\omega}^{mf}. \tag{42}$$

Thus, $\hat{\zeta}^{mf}, \delta\hat{\omega}^{mf}$ represent the difference between the centers of the two ellipsoids.

Substituting (41), (42) into (39), (40), we obtain

$$C^f = \hat{C}^m \exp(S(\hat{\zeta}^{mf})) \exp(S(\zeta^f)), \simeq \hat{C}^m \exp(S(\hat{\zeta}^{mf} + \zeta^f)), \tag{43}$$

$$\omega^f = \hat{\omega}^m + (\delta\hat{\omega}^{mf} + \delta\omega^f), \tag{44}$$

where we assumed that $\hat{\zeta}^{mf}, \zeta^f$ are sufficiently small to obtain the second equality. Thus, the uncertainty ellipsoid obtained by the flow update, $\mathscr{E}(\hat{C}^f, \hat{\omega}^f, P^f)$, is described by the center $(\hat{C}^m, \hat{\omega}^m)$ obtained from the measurement and the following uncertainty ellipsoid in $\mathbb{R}^6$:

$$\mathscr{E}_{\mathbb{R}^6}(\hat{x}^{mf}, P^f), \tag{45}$$

where $\hat{x}^{mf} = [(\hat{\zeta}^{mf})^{\mathrm{T}}, (\delta\hat{\omega}^{mf})^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^6$.

We seek a minimal ellipsoid that contains the intersection of two uncertainty ellipsoids in $\mathbb{R}^6$.

$$\mathscr{E}_{\mathbb{R}^6}(0_{6\times 1}, P^m) \cap \mathscr{E}_{\mathbb{R}^6}(\hat{x}^{mf}, P^f) \subset \mathscr{E}_{\mathbb{R}^6}(\hat{x}, P), \tag{46}$$

where $\hat{x} = [\hat{\zeta}^{\mathrm{T}}, \delta\hat{\omega}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^6$. Using the result presented in [15], $\hat{x}$ and P can be written as

$$\hat{x} = L\hat{x}^{mf}, \quad P = \beta(q)(I - L)P^m, \tag{47}$$

where

$$\beta(q) = 1 + q - (\hat{x}^{mf})^{\mathrm{T}} (P^m)^{-1} L \hat{x}^{mf}, \tag{48}$$

$$L = P^m (P^m + q^{-1} P^f)^{-1}. \tag{49}$$

The constant $q \geqslant 0$ is chosen such that $\mathrm{tr}[P]$ is minimized. We convert $\hat{x}$ to points in TSO(3) using the common center $(\hat{C}^m, \hat{\omega}^m)$.

In summary, the attitude estimation filter algorithm is given by the following statement.

Proposition 3.1. *The attitude and angular velocity estimates and the new uncertainty ellipsoid at the $(k+l)$th step are given by*

$$\hat{C}_{k+l} = \hat{C}^m_{k+l} \exp(S(\hat{\zeta})), \tag{50}$$

$$\hat{\omega}_{k+l} = \hat{\omega}^m_{k+l} + \delta\hat{\omega}, \tag{51}$$

$$P_{k+l} = P, \tag{52}$$

where $\hat{x} = [\hat{\zeta}^{\mathrm{T}}, \delta\hat{\omega}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^6$ *and* $P \in \mathbb{R}^{6\times 6}$ *are given by Eqs.* (47)–(49). *The actual state lies in the ellipsoid*

$$(C_{k+l}, \omega_{k+l}) \in \mathscr{E}(\hat{C}_{k+l}, \hat{\omega}_{k+l}, P_{k+l}), \tag{53}$$

centered at the estimated attitude and angular velocity states.

Remark. For the approximation given in (43), we assume that $\hat{\zeta}^{mf}, \zeta^f$ are sufficiently small. This assumption can be avoided if we use the Baker–Campbell–Hausdorff formula given by

$$\exp(S(\hat{\zeta}^{mf}))\exp(S(\zeta^f)) = \exp(S(\mathrm{BCH}(\hat{\zeta}^{mf}, \zeta^f))),$$

where the explicit expression for $\mathrm{BCH} : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$ can be found in [10]. Then, instead of (45), the uncertainty ellipsoid obtained by the flow update is described by the center $(\hat{C}^m, \hat{\omega}^m)$ obtained from the measurement and the uncertainty ellipsoid in $\mathbb{R}^6$, $\mathscr{E}_{\mathbb{R}^6}(\hat{y}^{mf}, Q^f)$, where the expressions for $\hat{y}^{mf} \in \mathbb{R}^6$ and a symmetric positive definite $Q^f \in \mathbb{R}^{6\times 6}$ are given by the solution of the following optimization problem.

for given $\hat{\zeta}^{mf}, \delta\hat{\omega}^{mf}, P^f$,

$$\min_{\hat{y}^{mf}, Q^f} \mathrm{tr}[Q^f],$$

$$\text{such that } [\mathrm{BCH}(\hat{\zeta}^{mf}, \zeta^f)^{\mathrm{T}}, (\delta\hat{\omega}^{mf} + \delta\omega^f)^{\mathrm{T}}]^{\mathrm{T}} \in \mathscr{E}_{\mathbb{R}^6}(\hat{y}^{mf}, Q^f),$$

for any $[\zeta^f, \delta\omega^f] \in \mathscr{E}_{\mathbb{R}^6}(0_6, P^f)$.

This optimization problem can be solved numerically using computational methods for finding minimum volume ellipsoids covering a finite set, given in [6], by discretizing the uncertainty ellipsoid $\mathscr{E}_{\mathbb{R}^6}(0_6, P^f)$. However, the resulting computational burden could decrease the speed of the estimation algorithm, which would be detrimental for real-time implementation.

The center of the new uncertainty ellipsoid is the estimated state, considered as point estimates of the attitude and the angular velocity at the $(k+l)$th step. The uncertainty matrix represents the bounds on the uncertainty of the estimated state. The size of the uncertainty matrix characterizes the accuracy of the estimate. If the size of the uncertainty ellipsoid is small, we conclude that the estimation is accurate. This estimation is optimal in the sense that the size of the new uncertainty ellipsoid is minimized. The uncertainty matrix can also be used to predict the distribution of the uncertainty. The eigenvector of the uncertainty matrix corresponding to the maximum eigenvalue shows the direction of the maximum uncertainty.

3.4. Convergence of filter

We now present a sufficient condition under which this estimation algorithm converges, i.e., the size of the uncertainty matrix decreases monotonically with measurements. The trace of the positive definite uncertainty matrix P is the measure of size used in this analysis.

Theorem 3.1. *The estimation algorithm given by Proposition* 3.1 *is convergent if there exists a constant* $c \in (0, 1)$ *such that the following inequality holds for every measurement*:

$$q\,\mathrm{tr}[P^f] + \mathrm{tr}[P^m] < c(1+q)\mathrm{tr}[P_0], \tag{54}$$

where P_0 *denotes the uncertainty matrix of the filtered estimate at the previous measurement instant*.

Proof. For convergence, it is sufficient that the filtering process is a contraction mapping, which is to say that $\mathrm{tr}[P] < c\,\mathrm{tr}[P_0]$ where $c \in (0, 1)$ and P_0 denotes the uncertainty matrix of the filtered estimate at the previous measurement instant.

Using the matrix inversion lemma, we can express the uncertainty matrix P given by (47)–(49) as

$$P = \beta(q)(q(P^f)^{-1} + (P^m)^{-1})^{-1}.$$

Now we have

$$\mathrm{tr}[P] = \beta(q)\mathrm{tr}[(q(P^f)^{-1} + (P^m)^{-1})^{-1}].$$

From Eqs. (48)–(49), we have $\beta(q) \leqslant 1 + q$. Thus,

$$\mathrm{tr}[P] \leqslant (1+q)\mathrm{tr}[(q(P^f)^{-1} + (P^m)^{-1})^{-1}].$$

From Fact 8.10.15 in [5], we know that

$$\mathrm{tr}[(\alpha(P^f)^{-1} + (1-\alpha)(P^m)^{-1})^{-1}] \leqslant \mathrm{tr}[\alpha P^f + (1-\alpha)P^m],$$

for $\alpha \in [0, 1]$, since $(P^f)^{-1}$ and $(P^m)^{-1}$ are positive definite. Since the parameter $q \geqslant 0$, we can substitute $q = \alpha/(1-\alpha)$ for some $\alpha \in [0, 1]$, so that

$$\begin{aligned}
&\mathrm{tr}[(q(P^f)^{-1} + (P^m)^{-1})^{-1}] \\
&\quad = \mathrm{tr}\left[\left\{\frac{1}{1-\alpha}(\alpha(P^f)^{-1} + (1-\alpha)(P^m)^{-1})\right\}^{-1}\right] \\
&\quad \leqslant (1-\alpha)\mathrm{tr}[\alpha P^f + (1-\alpha)P^m] \\
&\quad = \alpha(1-\alpha)\mathrm{tr}[P^f] + (1-\alpha)^2\mathrm{tr}[P^m] \\
&\quad = \frac{q}{(1+q)^2}\mathrm{tr}[P^f] + \frac{1}{(1+q)^2}\mathrm{tr}[P^m].
\end{aligned}$$

A sufficient condition for convergence for the estimation algorithm is then obtained from the above as:

$$\begin{aligned}
\mathrm{tr}[P] &\leqslant (1+q)\mathrm{tr}[(q(P^f)^{-1} + (P^m)^{-1})^{-1}] \\
&\leqslant \frac{q}{(1+q)}\mathrm{tr}[P^f] + \frac{1}{(1+q)}\mathrm{tr}[P^m] < c\,\mathrm{tr}[P_0],
\end{aligned}$$

which is equivalent to the sufficient condition (54). □

The rate of convergence is determined by the contraction constant $c \in (0, 1)$ for which the inequality (54) is satisfied for all measurements. This condition shows that having small uncertainties in both measurement and flow propagation facilitates convergence of the filter estimates, as can be expected. It also shows that there is a trade-off between the effects of

the flow induced uncertainty and the measurement uncertainty, given by the parameter q. When q is large, the effect of the flow uncertainty is larger than the measurement uncertainty. The reverse is true when q is small. The limiting cases are (i) when $q = 0$ and (ii) when $q \to \infty$. In the first case, the measurement ellipsoid contains the flow ellipsoid (see [15]) and Theorem 3.1 states that it is sufficient to have $\mathrm{tr}[P^m] < c\,\mathrm{tr}[P_0]$ for convergence. In the second case, the flow ellipsoid contains the measurement ellipsoid (see [15]) and Theorem 3.1 states that it is sufficient to have $\mathrm{tr}[P^f] < c\,\mathrm{tr}[P_0]$ for convergence. For intermediate values of the parameter q, if knowledge of absolute bounds on the size of the flow and measurement uncertainties is available, one can check to see if condition (54) is satisfied, in which case the estimates are guaranteed to converge. However, such knowledge may not always be available a priori. For the special case when the measurement ellipsoid P^m is known a priori or may be assumed constant (i.e., the sensor accuracy remains constant), the sufficient condition (54) provides a bound on the flow uncertainty ellipsoid P^f (or equivalently, a bound on the linear flow matrix A^f). With prior knowledge of the bounds on P^f and P^m and after determining q during the filter step of the estimation scheme, one could also check this condition in real time. One of the important factors that the rate of convergence depends upon is obtained from consideration of the geometry of the intersecting ellipsoids. The rate of convergence depends upon the relative orientation of the major axes of the non-degenerate flow and measurement ellipsoids in the 6D state space. The volume of intersection of these two ellipsoids when their major axes are orthogonal, is smaller than thevolume of intersection when their major axes are closely aligned.

4. Conclusion

A deterministic estimation scheme for the attitude dynamics of a rigid body in an attitude dependent potential field is presented, with an assumption of bounded measurement errors. The properties of the proposed attitude estimation scheme are as follows. This attitude estimator has no singularities since the attitude is represented by a rotation matrix, and the structure of the rotation matrix is preserved since it is updated by group operations in SO(3) using a Lie group variational integrator. The proposed attitude estimator is robust to the distribution of the uncertainty in initial conditions and the measurement noise, since it is a deterministic scheme based on knowledge of the bounds in these uncertainties. A sufficient condition for convergence of this filter has been obtained. These results can be extended in a number of different directions, like: relaxing the assumption that the number of attitude measurements at each measurement step is $m=3$, eliminating the assumption of velocity measurements at each measurement update step, and adding the effects of process noise.

Acknowledgments

We gratefully acknowledge helpful comments and suggestions of the referee. AKS supported in part by a Faculty Development Grant from the University of Hawaii. TL and ML supported in part by NSF under Grants DMS-0504747, DMS-0726263 and a Grant from the Rackham Graduate School, University of Michigan. TL and NHM supported in part by NSF under Grant ECS-0244977 and CMS-0555797.

References

[1] E.-W. Bai, K. Nagpal, R. Tempo, Bounded-error parameter estimation: noise models and recursive algorithms, Automatica 32 (7) (1996) 985–999.
[2] I.Y. Bar-Itzhack, Y. Oshman, Attitude determination from vector observations: quaternion estimation, IEEE Trans. Aerospace Electron. Systems 21 (1) (1985) 128–136.
[3] B. Barshan, H.F. Durrant-Whyte, Inertial navigation systems for mobile robots, IEEE Trans. Robotics Automat. 11 (3) (1995) 328–342.
[4] G. Belforte, P. Gay, Optimal worst case estimation for lpv-fir models with bounded errors, Systems Control Lett. 53 (4) (2004) 259–268.
[5] D.S. Bernstein, Matrix Mathematics, Princeton University Press, Princeton, NJ, 2005.
[6] S. Boyd, L.E. Ghaoui, E. Feron, V. Balakrishnan, Linear Matrix Inequalities in System and Control Theory, SIAM, 1994.
[7] J.L. Crassidis, F.L. Markley, A minimum model error approach for attitude estimation, J. Guidance Control Dynamics 20 (6) (1997) 1241–1247.
[8] J.L. Crassidis, F.L. Markley, Unscented filtering for spacecraft attitude estimation, AIAA J. Guidance Control Dynamics 26 (4) (2003) 536–542.
[9] C. Durieu, E. Walter, B. Polyak, Multi-input multi-output ellipsoidal state bounding, J. Optim. Theory Appl. 111 (2) (2001) 273–303.
[10] K. Engø, On the BCH-formula in $\mathfrak{so}(3)$, BIT 41 (3) (2001) 629–631.
[11] J.L. Farrell, J.C. Stuelpnagel, R.H. Wessner, J.R. Velman, J.E. Brock, A least squares estimate of satellite attitude, solution 65-1, SIAM Rev. 8 (3) (1966) 384–386.
[12] T. Lee, M. Leok, N.H. McClamroch, A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum, in: Proceedings of the IEEE Conference on Control Applications, 2005, pp. 962–967.
[13] T. Lee, M. Leok, N.H. McClamroch, Attitude maneuvers of a rigid spacecraft in a circular orbit, in: Proceedings of the Americal Control Conference, Minneapolis, MN, 2006, pp. 1742–1747.
[14] T. Lee, A.K. Sanyal, M. Leok, N.H. McClamroch, Deterministic global attitude estimation, in: Proceedings of the 45th IEEE Conference on Decision and Control, San Diego, CA, 2006, pp. 3174–3179.
[15] D.G. Maksarov, J.P. Norton, State bounding with ellipsoidal set description of the uncertainty, Internat. J. Control 65 (5) (1996) 847–866.
[16] J. Milnor, Morse Theory, Princeton University Press, Princteon, NJ, 1963.
[17] F.L. Markley, Attitude determination and parameter estimation using vector observations: theory, J. Astronaut. Sci. 37 (6) (1989) 41–58.
[18] B.T. Polyak, S.A. Nazin, C. Durieu, E. Walter, Ellipsoidal parameter or state estimation under model uncertainty, Automatica 40 (7) (2004) 1171–1179.
[19] M.L. Psiaki, Attitude-determination filtering via extended quaternion estimation, AIAA J. Guidance Control Dynamics 23 (2) (2000) 206–214.
[20] H. Rehbinder, X. Hu, Drift-free attitude estimation for accelerated rigid bodies, Automatica 40 (4) (2004) 653–659.
[21] S.I. Roumeliotis, G.S. Sukhatme, G.A. Bekey, Smoother based 3d attitude estimation for mobile robot localization, in: Proceedings of 1999 IEEE International Conference on Robotics and Automation, vol. 3, Detroit, MI, 1999, pp. 1979–1986.
[22] A.K. Sanyal, Optimal attitude estimation and filtering without using local coordinates, Part I: uncontrolled and deterministic attitude dynamics, in: Proceedings of the American Control Conference, Minneapolis, MN, 2006, pp. 5734–5739.

[23] F.C. Schweppe, Recursive state estimation: unknown but bounded errors and system inputs, IEEE Trans. Automat. Control 13 (1) (1968) 22–28.
[24] M.D. Shuster, S.D. Oh, Three-axis attitude determination from vector observations, AIAA J. Guidance Control 4 (1) (1981) 70–77.
[25] M.D. Shuster, Kalman filtering of spacecraft attitude and the quest model, J. Astronaut. Sci. 38 (3) (1990) 377–393.
[26] Y. Theodor, U. Shaked, C.E. de Souza, A game theory approach to robust discrete-time H_∞-estimation, IEEE Trans. Signal Process. 42 (6) (1994) 1486–1495.
[27] J. Vaganay, M.J. Aldon, A. Fournier, Mobile robot attitude estimation by fusion of inertial data, in: Proceedings of the 1993 IEEE International Conference on Robotics and Automation, vol. 1, Atlanta, GA, 1993, pp. 277–282.
[28] G. Wahba, A least squares estimate of satellite attitude, problem 65-1, SIAM Rev. 7 (5) (1965) 409.

OPTIMAL ATTITUDE CONTROL OF A RIGID BODY USING GEOMETRICALLY EXACT COMPUTATIONS ON SO(3)

TAEYOUNG LEE, MELVIN LEOK, AND N. HARRIS MCCLAMROCH

Abstract. An efficient and accurate computational approach is proposed for a non-convex optimal attitude control for a rigid body. The problem is formulated directly as a discrete time optimization problem using a Lie group variational integrator. Discrete time necessary conditions for optimality are derived, and an efficient computational approach is proposed to solve the resulting two-point boundary value problem. This formulation wherein the optimal control problem is solved based on discretization of the attitude dynamics and derivation of discrete time necessary conditions, rather than development and discretization of continuous time necessary conditions, is shown to have significant advantages. In particular, the use of geometrically exact computations on SO(3) guarantees that this optimal control approach has excellent convergence properties even for highly nonlinear large angle attitude maneuvers.

1. Introduction

A discrete optimal control problem for attitude dynamics of a rigid body in the presence of an attitude dependent potential is considered. The objective is to minimize the square of the l_2 norm of external control torques which transfer a given initial attitude and an initial angular momentum of the rigid body to a desired terminal attitude and a terminal angular momentum during a fixed maneuver time. The attitude of the rigid body is defined by the orientation of a body fixed frame with respect to a reference frame; the attitude is represented by a rotation matrix that is a 3×3 orthogonal matrix with determinant of 1. Rotation matrices have a Lie group structure denoted by SO(3).

The dynamics of a rigid body has fundamental invariance properties. In the absence of nonconservative forces, total energy is preserved. A consequence of Noether's theorem is that symmetries in the Lagrangian result in conservation of the associated momentum map. Furthermore, the configuration space of the rigid body has the orthogonal structure of the Lie group SO(3). General-purpose numerical integration methods, including the popular Runge–Kutta schemes, typically preserve neither first integrals nor the geometric

1991 *Mathematics Subject Classification.* 49J15, 37M15.
Key words and phrases. Optimal control, symplectic integrator.
The first author and the second author have been supported in part by NSF Grant DMS-0504747, and a grant from the Rackham Graduate School, University of Michigan.
The first author and the third author have been supported in part by NSF Grant ECS-0244977 and CMS-0555797.

characteristics of the configuration space. In particular, the orthogonal structure of rotation matrices is not preserved numerically with standard schemes.

A Lie group variational integrator that preserves those geometric features is presented in [1], and integrators on the configuration space SO(3) and SE(3) are developed in [2] and [3, 4], respectively. These integrators are obtained from a discrete variational principle, and they exhibit the symplectic and momentum preservation properties, and good energy behavior characteristic of variational integrators [5]. Since the rotation matrices are updated by a group operation, they automatically evolve on the rotation group without the need for reprojection techniques or constraints [6].

Optimal attitude control problems are studied in [7]. The angular velocity of a rigid body is treated as a control input; an optimal angular velocity that steers the rigid body is derived from the attitude kinematics. Continuous time optimal control problems on a Riemannian manifold are studied in [8], where necessary conditions for optimality are derived from a variational principle. An optimal control problem based on discrete mechanics is studied in [9]. The discrete equations of motion and the boundary conditions are imposed as constraints, and the optimal control problem is solved by a general-purpose parameter optimization tool. This approach requires large computation time, since the number of optimization parameters is proportional to the number of discrete time steps. Discrete necessary conditions for optimal control of the attitude dynamics of a rigid body are presented in [10].

This paper proposes an exact and efficient computational approach to solve an optimal control problem associated with the attitude dynamics of a rigid body that evolves on the configuration space SO(3). We assume that the control inputs are parameterized by their value at each time step. A Lie group variational integrator on SO(3) that includes the effects of external control inputs is developed using the discrete Lagrange-d'Alembert principle. Discrete necessary conditions for optimality are obtained using a variational principle, while imposing the Lie group variational integrator as dynamic constraints.

The necessary conditions are expressed as a two-point boundary value problem on $T^*SO(3)$ and its dual. Sensitivity derivatives along an extremal solution are developed by following the procedures presented in [11], and they are used to construct an algorithm that solves the boundary value problem efficiently. Since the attitude of the rigid body is represented by a rotation matrix, and the orthogonal structure of rotation matrices is preserved by the Lie group variational integrator, the discretization of the optimal control problem does not exhibit singularities. The extension of this work to the configuration space of SE(3), and a discrete optimal control problem with symmetry can be found in [12] and [13], respectively.

This paper is organized as follows. In Section 2, a Lie group variational integrator is developed using the discrete Lagrange-d'Alembert principle. Necessary conditions for optimality and a proposed approach to solve the two-point boundary problem are presented in Section 3. Numerical results for the attitude control of an underactuated 3D pendulum, and for a fully actuated spacecraft in a circular orbit are given in Section 4.

2. Equations of Motion for the Attitude Dynamics of a Rigid Body

In this section, we define a rigid body model in a potential field and we develop discrete equations of motion for the attitude dynamics of the rigid body, referred to as a Lie group variational integrator. These discrete equations of motion are used as dynamic constraints for the optimal control problem presented in Section 3.

2.1. **Rigid body model.**

We consider the attitude dynamics of a rigid body in the presence of an attitude dependent potential. The configuration space is the Lie group, SO(3). We assume that the potential $U(\cdot) : \mathrm{SO}(3) \mapsto \mathbb{R}$ is determined by the attitude of the rigid body, $R \in \mathrm{SO}(3)$. External control inputs generate moments about the mass center of the rigid body. A spacecraft on a circular orbit including gravity gradient effects [11], a 3D pendulum [2], or a free rigid body can be modeled in this way. The continuous equations of motion are given by

$$\dot{\Pi} + \Omega \times \Pi = M + Bu, \tag{1}$$

$$\dot{R} = RS(\Omega), \tag{2}$$

where $\Omega \in \mathbb{R}^3$ is the angular velocity of the body expressed in the body fixed frame, and $\Pi = J\Omega \in \mathbb{R}^3$ is the angular momentum of the body for a moment of inertia matrix $J \in \mathbb{R}^{3\times3}$. $M \in \mathbb{R}^3$ is the moment due to the potential, $u \in \mathbb{R}^m$ is the external control input, and $B \in \mathbb{R}^{3\times m}$ is an input matrix. If the rank of the input matrix is less than 3, then the rigid body is underactuated. The matrix valued function, $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is an isomorphism defined such that $S(x)y = x \times y$ for all $x, y \in \mathbb{R}^3$. The moment due to the potential is determined by the relationships, $S(M) = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}$, or more explicitly,

$$M = r_1 \times v_{r_1} + r_2 \times v_{r_2} + r_3 \times v_{r_3}, \tag{3}$$

where $r_i, v_{r_i} \in \mathbb{R}^{1\times3}$ are the ith row vectors of R and $\frac{\partial U}{\partial R}$, respectively. A detailed derivation of the above equations can be found in [2].

2.2. **Lie group variational integrator.** We describe a Lie group variational integrator that is obtained from a discrete variational approach, and therefore it exactly preserves the momentum and symplectic form, while exhibiting good energy behavior over exponentially long times. Since a Lie group numerical method [6] is explicitly adopted, the rotation matrix automatically remains on SO(3).

The Lie group variational integrator is obtained by following procedures commonly adopted in Lagrangian mechanics. The variational approach is based on discretizing Hamilton's principle rather than discretizing the continuous equations of motion. The velocity phase space of the continuous Lagrangian is replaced by discrete variables, and a discrete Lagrangian is chosen. Taking a variation of the action sum defined as the summation of the discrete Lagrangian, we obtain a Lagrangian form of the discrete equations of motion using the Lagrange-d'Alembert principle. A discrete version of the Legendre transformation yields discrete equations in Hamiltonian form.

The detailed derivation is presented in [2] and [3]. In this paper, we extend these results to include the effects of external control inputs. Consider the fixed integration step size $h \in \mathbb{R}$. Let $R_k \in \mathrm{SO}(3)$ denote the attitude of the rigid body at time $t = kh$. We introduce a new variable $F_k \in \mathrm{SO}(3)$ defined by

$$F_k = R_k^T R_{k+1}, \tag{4}$$

which represents a relative attitude between integration steps. If we find $F_k \in \mathrm{SO}(3)$ at each integration step, the rotation matrix is updated by multiplication of two rotation matrices, i.e. $R_{k+1} = R_k F_k$, which is the group operation on SO(3). In particular, we ensure that F_k is a rotation matrix by expressing it as $F_k = \exp(\xi_k)$, where $\xi_k \in \mathfrak{so}(3)$ is a skew-symmetric matrix and $\exp(\cdot)$ denotes the matrix exponential. This guarantees that the rotation matrix evolves on SO(3) automatically. This is the approach of Lie group methods [6]. The following procedure provides an expression for F_k using the discrete Lagrange-d'Alembert principle.

Using the kinematic relationship (2), $S(\Omega_k)$ can be approximated as

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

Using the above equation, we can show that the kinetic energy of the rigid body is given by

$$T = \frac{1}{2} \mathrm{tr}\big[S(\Omega_k) J_d S(\Omega_k)^T \big] = \frac{1}{h^2} \mathrm{tr}\big[(I_{3\times 3} - F_k)\, J_d \big],$$

where $J_d \in \mathbb{R}^{3\times3}$ is a nonstandard moment of inertia matrix of the rigid body defined in terms of the standard moment of inertia matrix $J \in \mathbb{R}^{3\times3}$ as $J_d = \frac{1}{2}\text{tr}[J]\, I_{3\times3} - J$. Define a discrete Lagrangian L_d as

$$L_d(R_k, F_k) = \frac{1}{h}\text{tr}[(I_{3\times3} - F_k)\, J_d] - hU(R_{k+1}). \tag{5}$$

This discrete Lagrangian is a first-order approximation of the integral of the continuous Lagrangian over one integration step. Therefore, the following action sum, defined as the summation of the discrete Lagrangian, approximates the action integral; $\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(R_k, F_k)$. Taking a variation of the action sum, we obtain the discrete equations of motion using the discrete Lagrange-d'Alembert principle. The variation of a rotation matrix can be expressed using the exponential of a Lie algebra element:

$$R_k^\epsilon = R_k e^{\epsilon\eta_k},$$

where $\epsilon \in \mathbb{R}$ and $\eta_k \in \mathfrak{so}(3)$ is the variation expressed as a skew-symmetric matrix. Thus the infinitesimal variation is given by

$$\delta R_k = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} R_k^\epsilon = R_k\eta_k. \tag{6}$$

The Lagrange-d'Alembert principle states that the following equation is satisfied for all possible variations $\eta_k \in \mathfrak{so}(3)$.

$$\delta\sum_{k=0}^{N-1}\frac{1}{h}\text{tr}[(I_{3\times3} - F_k)\, J_d] - hU(R_{k+1}) - \sum_{k=0}^{N-1}\frac{h}{2}\text{tr}[\eta_{k+1}S(Bu_{k+1})] = 0. \tag{7}$$

Using the expression of the infinitesimal variation of a rotation matrix (6) and using the fact that the variations vanish at the end points, the above equation can be written as

$$\sum_{k=1}^{N-1}\text{tr}\left[\eta_k\left\{\frac{1}{h}\left(F_kJ_d - J_dF_{k-1}\right) + hR_k^T\frac{\partial U}{\partial R_k} - \frac{h}{2}S(Bu_k)\right\}\right] = 0.$$

Since the above expression should be zero for all possible variations $\eta_k \in \mathfrak{so}(3)$, the expression in the braces should be symmetric. Then, *the discrete equations of motion in Lagrangian form* are given by

$$\frac{1}{h}\left(F_{k+1}J_d - J_dF_k - J_dF_{k+1}^T + F_k^TJ_d\right) = hS(M_{k+1}) + hS(Bu_{k+1}), \tag{8}$$

$$R_{k+1} = R_kF_k. \tag{9}$$

Using the discrete version of the Legendre transformation, *the discrete equations of motion in Hamiltonian form* are given by

$$hS(\Pi_k) = F_k J_d - J_d F_k^T, \tag{10}$$

$$R_{k+1} = R_k F_k, \tag{11}$$

$$\Pi_{k+1} = F_k^T \Pi_k + h\left(M_{k+1} + B u_{k+1}\right). \tag{12}$$

Given (R_k, Π_k), we can obtain F_k by solving (10), and R_{k+1} is obtained by (11). The moment due to the potential M_{k+1} can be calculated by (3). Finally, Π_{k+1} is updated by (12). This yields a map $(R_k, \Pi_k) \mapsto (R_{k+1}, \Pi_{k+1})$, and this process can be repeated. The only implicit part is solving (10). We can express (10) in terms of a Lie algebra element $S(f_k) = \mathrm{logm}(F_k) \in \mathfrak{so}(3)$, and find $f_k \in \mathbb{R}^3$ numerically by a Newton iteration. The relative attitude F_k is obtained by the exponential map: $F_k = e^{S(f_k)}$. Therefore we are guaranteed that F_k is a rotation matrix.

The order of the variational integrator is equal to the order of the corresponding discrete Lagrangian. Consequently, the above Lie group variational integrator is of first-order since (5) is a first-order approximation. While higher-order variational integrators can be obtained by modifying (5), we use the first-order integrator because it yields a compact form for the necessary conditions that preserves the geometry; these necessary conditions are developed in Section 3.

It is worthwhile noting that while the Lie group variational integrator adopted in this paper is formally first-order, it is the Lie group analogue of symplectic Euler, which is symplectically equivalent [14, 15] to the second-order accurate Störmer–Verlet method [16]. This means that one can obtain the Störmer–Verlet method by conjugating the symplectic Euler method with a symplectic transformation. In particular, numerical trajectories of symplectic Euler will shadow numerical trajectories obtained using Störmer–Verlet.

In practice, the shadowing result imparts the symplectic Euler method with the same desirable qualitative properties as Störmer–Verlet. Since on an appropriate choice of charts, our Lie symplectic Euler method reduces to symplectic Euler in coordinates, it follows that there is a corresponding second-order Lie Störmer–Verlet method that our integrator is symplectically equivalent to, and whose discrete trajectories are shadowed by our method.

3. Discrete Optimal Control of the Attitude Dynamics of a Rigid Body

We formulate a discrete optimal control problem for the attitude dynamics of a rigid body, and we derive necessary conditions for optimality using a variational principle. The necessary conditions are expressed as

a two-point boundary value problem, and a computational approach to solve the boundary value problem is proposed using sensitivity derivatives.

3.1. **Problem formulation.** A discrete time optimal control problem in SO(3) is formulated as a maneuver of a rigid body from a given initial attitude $R_0 \in \mathrm{SO}(3)$ and an initial angular momentum $\Pi_0 \in \mathbb{R}^3$ to a desired terminal attitude $R_N^d \in \mathrm{SO}(3)$ and a terminal angular momentum $\Pi_N^d \in \mathbb{R}^3$ during a given maneuver time N. The performance index is the square of the l_2 norm of the control input; the discrete equations of motion developed in the previous section are imposed as constraints.

$$\text{given: } R_0, \Pi_0,\ R_N^d, \Pi_N^d,\ N,$$

$$\min_{u_{k+1}} \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2} \left\| u_{k+1} \right\|^2,$$

$$\text{such that } R_N = R_N^d,\ \Pi_N = \Pi_N^d,$$

$$\text{subject to (10), (11) and (12).}$$

In [9], an optimal control problem based on discrete mechanics is considered. The control inputs at each discrete step are considered as optimization parameters, and the discrete equations of motion and the boundary conditions are imposed as constraints. The optimization problem is solved numerically by a general-purpose parameter optimization tool such as Sequential Quadratic Programming (SQP). The same approach can be applied to the above optimization problem. However, it has a large computational burden since the number of optimization parameters, $m \times N$, is proportional to the number of integration steps. Usually, a large time step size is chosen to make the number of integration steps small, or the control inputs are approximated by collocation points. The resulting control inputs tends to be under-resolved and sub-optimal.

We derive necessary conditions for optimality using the standard calculus of variations. We assume that the control inputs are parameterized by their value at each time step. The necessary conditions are expressed as a two-point boundary value problem.

3.2. **Necessary conditions of optimality.** Define an augmented performance index as

$$\begin{aligned} \mathcal{J}_a = \sum_{k=0}^{N-1} \frac{h}{2} \left\| u_{k+1} \right\|^2 &+ \lambda_k^{1,T} S^{-1} \big(\mathrm{logm}(F_k - R_k^T R_{k+1}) \big) \\ &+ \lambda_k^{2,T} \left\{ -\Pi_{k+1} + F_k^T \Pi_k + h \left(M_{k+1} + B u_{k+1} \right) \right\}, \end{aligned} \tag{13}$$

where $\lambda_k^1, \lambda_k^2 \in \mathbb{R}^3$, are Lagrange multipliers corresponding to the discrete equations of motion. The augmented performance index is chosen such that the dimension of the multipliers is equal to the dimension of the rotation matrix and the angular momentum vector. The discrete kinematics equation (11) is transformed into a matrix logarithm form. The constraints arising from the discrete kinematics equation (11) and the angular momentum equation (12) are explicitly applied. Equation (10) appears in the discrete equations of motion because we introduce the auxiliary variable $F_k \in \mathrm{SO}(3)$. The constraint (10) is considered implicitly when taking a variation of the performance index.

Consider small variations from a given trajectory denoted by Π_k, R_k, F_k, u_k:

$$\Pi_k^\epsilon = \Pi_k + \epsilon \delta \Pi_k, \tag{14}$$

$$\begin{aligned} R_k^\epsilon &= R_k e^{\epsilon S(\zeta_k)}, \\ &= R_k + \epsilon R_k S(\zeta_k) + \mathcal{O}(\epsilon^2), \end{aligned} \tag{15}$$

$$\begin{aligned} F_k^\epsilon &= F_k e^{\epsilon S(\xi_k)}, \\ &= F_k + \epsilon F_k S(\xi_k) + \mathcal{O}(\epsilon^2), \end{aligned} \tag{16}$$

where $\zeta_k, \xi_k \in \mathbb{R}^3 \simeq \mathfrak{so}(3)$. The real space $\mathbb{R}^3$ is isomorphic to the Lie algebra $\mathfrak{so}(3)$ according to the skew mapping $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$. The variations of the rotation matrices are expressed using the exponential of the Lie algebra elements. The corresponding infinitesimal variations of Π_k, R_k, and F_k are given by $\delta\Pi_k$, $\delta R_k = R_k S(\zeta_k)$, and $\delta F_k = F_k S(\xi_k)$, respectively.

The variation of the augmented performance index is obtained from the above expressions. Instead of taking a variation of the matrix logarithm in (13), we take a variation of the kinematics equation, (11) and we use it as a constrained variation. Since $F_k = R_k^T R_{k+1}$ by (11), the variation δF_k is given by

$$\delta F_k = \delta R_k^T R_{k+1} + R_k^T \delta R_{k+1}.$$

Substituting the expression for the infinitesimal variation of R_k, we obtain

$$F_k S(\xi_k) = -S(\zeta_k) F_k + F_k S(\zeta_{k+1}).$$

Multiplying both sides of the above equation by F_k^T and using the property $S(R^T x) = R^T S(x) R$ for all $R \in \mathrm{SO}(3)$ and $x \in \mathbb{R}^3$, we obtain

$$\xi_k = -F_k^T \zeta_k + \zeta_{k+1}. \tag{17}$$

We use (17) as a constrained variation equivalent to (11).

Now we develop another expression for a constrained variation using (10). Since we do not use (10) explicitly as a constraint in (13), $\delta\Pi_k$ and δF_k are not independent. Taking a variation of (10), we obtain

$$hS(\delta\Pi_k) = F_k S(\xi_k) J_d + J_d S(\xi_k) F_k^T.$$

Using the properties, $S(Rx) = RS(x)R^T$ and $S(x)A + A^T S(x) = S(\{\mathrm{tr}[A]\, I_{3\times 3} - A\}\, x)$ for all $x \in \mathbb{R}^3$, $A \in \mathbb{R}^{3\times 3}$, and $R \in \mathrm{SO}(3)$, the above equation can be written as

$$\begin{aligned} hS(\delta\Pi_k) &= S(F_k\xi_k)F_k J_d + J_d F_k^T S(F_k\xi_k), \\ &= S(\{\mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d\}\, F_k\xi_k). \end{aligned}$$

Thus, ξ_k is given by

$$\xi_k = \mathcal{B}_k \delta\Pi_k, \tag{18}$$

where $\mathcal{B}_k = hF_k^T \left\{\mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d\right\}^{-1} \in \mathbb{R}^{3\times 3}$. Equation (18) shows the relationship between $\delta\Pi_k$ and δF_k.

Since the moment due to the potential M_k is dependent on the attitude of the rigid body, the variation of the moment δM_k can be written using a variation of the rotation matrix:

$$\delta M_k = \mathcal{M}_k \zeta_k, \tag{19}$$

where $\mathcal{M}_k \in \mathbb{R}^{3\times 3}$ is expressed in terms of the the attitude of the rigid body, and the expression is determined by the potential field. We present detailed expressions of $\mathcal{M}_k$ in Section 4 for a 3D pendulum and for a spacecraft in a circular orbit. Using (17) and (18), δM_{k+1} is given by

$$\begin{aligned} \delta M_{k+1} &= \mathcal{M}_{k+1}\zeta_{k+1}, \\ &= \mathcal{M}_{k+1}F_k^T\zeta_k + \mathcal{M}_{k+1}\mathcal{B}_k\delta\Pi_k. \end{aligned} \tag{20}$$

Now, we take a variation of the augmented performance index (13) using the constrained variations (17), (18), and (20). Using (17), the variation of the performance index is given by

$$\begin{aligned}\delta \mathcal{J}_a = \sum_{k=0}^{N-1} & h\delta u_{k+1}^T u_{k+1} + \lambda_k^{1,T} \left\{ \xi_k + F_k^T \zeta_k - \zeta_{k+1} \right\} \\ & + \lambda_k^{2,T} \left\{ -\delta\Pi_{k+1} + \delta F_k^T \Pi_k + F_k^T \delta\Pi_k + h\delta M_{k+1} + hB\delta u_{k+1} \right\}. \end{aligned} \tag{21}$$

Substituting (20) into (21) and rearranging, we obtain

$$\begin{aligned}\delta \mathcal{J}_a = \sum_{k=0}^{N-1} & h\delta u_{k+1}^T \left\{ u_{k+1} + B^T \lambda_k^2 \right\} - \zeta_{k+1}^T \lambda_k^1 + \zeta_k^T \left\{ F_k \lambda_k^1 + hF_k \mathcal{M}_{k+1}^T \lambda_k^2 \right\} \\ & - \delta\Pi_{k+1}^T \lambda_k^2 + \delta\Pi_k^T \left\{ F_k \lambda_k^2 + h\mathcal{B}_k^T \mathcal{M}_{k+1}^T \lambda_k^2 \right\} \\ & + \xi_k^T \left\{ -S(F_k^T \Pi_k)\lambda_k^2 + \lambda_k^1 \right\}. \end{aligned} \tag{22}$$

Substituting (18) into (22) and using the fact that the variations ζ_k, $\delta\Pi_k$ vanish at $k = 0, N$, we obtain

$$\begin{aligned}\delta \mathcal{J}_a = \sum_{k=1}^{N-1} & h\delta u_k^T \left\{ u_k + B^T \lambda_{k-1}^2 \right\} + \zeta_k^T \left\{ -\lambda_{k-1}^1 + F_k \lambda_k^1 + hF_k \mathcal{M}_{k+1}^T \lambda_k^2 \right\} \\ & + \delta\Pi_k^T \left\{ -\lambda_{k-1}^2 + (F_k - \mathcal{B}_k^T S(F_k^T \Pi_k) + h\mathcal{B}_k^T \mathcal{M}_{k+1}^T)\lambda_k^2 + \mathcal{B}_k^T \lambda_k^1 \right\}. \end{aligned} \tag{23}$$

Since $\delta \mathcal{J}_a = 0$ for all variations of $\delta u_k, \zeta_k, \delta\Pi_k$ which are independent, the expression in the braces are zero. Thus we obtain necessary conditions for optimality as follows.

$$\Pi_{k+1} = F_k^T \Pi_k + h \left(M_{k+1} + B u_{k+1} \right), \tag{24}$$

$$hS(\Pi_k) = F_k J_d - J_d F_k^T, \tag{25}$$

$$R_{k+1} = R_k F_k, \tag{26}$$

$$u_{k+1} = -B^T \lambda_k^2, \tag{27}$$

$$\begin{bmatrix} \lambda_k^1 \\ \lambda_k^2 \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{k+1}^T & \mathcal{C}_{k+1}^T \\ \mathcal{B}_{k+1}^T & \mathcal{D}_{k+1}^T \end{bmatrix} \begin{bmatrix} \lambda_{k+1}^1 \\ \lambda_{k+1}^2 \end{bmatrix}, \tag{28}$$

where

$$\mathcal{A}_k = F_k^T, \tag{29}$$

$$\mathcal{B}_k = hF_k^T \left\{\mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d\right\}^{-1}, \tag{30}$$

$$\mathcal{C}_k = h\mathcal{M}_{k+1} F_k^T, \tag{31}$$

$$\mathcal{D}_k = F_k^T + S(F_k^T \Pi_k)\mathcal{B}_k + h\mathcal{M}_{k+1}\mathcal{B}_k. \tag{32}$$

In the above equations, the only implicit part is (25). For a given initial condition $(R_0, \Pi_0, \lambda_0^1, \lambda_0^2)$, we can find F_0 by solving (25). Then, R_1 is obtained from (26). Since $u_1 = -B^T\lambda_0^2$ by (27), and M_1 is a function of R_1, Π_1 can be obtained using (24). We solve (25) to obtain F_1 using Π_1. Finally, λ_1^1, λ_1^2 are obtained from (28), since $\mathcal{A}_1, \mathcal{B}_1, \mathcal{C}_1, \mathcal{D}_1$ are functions of R_1, Π_1, F_1. This yields a map $(R_0, \Pi_0, \lambda_0^1, \lambda_0^2) \mapsto (R_1, \Pi_1, \lambda_1^1, \lambda_1^2)$, and this process can be repeated.

3.3. **Two-point boundary value problem.** The necessary conditions for optimality are given as a 12 dimensional two-point boundary value problem on $T^*\mathrm{SO}(3)$ and its dual space. This problem is to find

$$\text{Attitude and angular momentum :} R_k, \Pi_k,$$

$$\text{Multiplier variables :} \lambda_k^1, \lambda_k^2,$$

$$\text{Control inputs :} u_k,$$

for $k = \{0, 1, \cdots, N\}$, to satisfy simultaneously,

$$\text{Equations of motion :}(24), (25), (26),$$

$$\text{Multiplier equations :}(28),$$

$$\text{Optimality condition :}(27),$$

$$\text{Boundary conditions :} R_0, \Pi_0, R_N, \Pi_N.$$

An iterative numerical method for the two-point boundary value problem is presented. A nominal solution that satisfies some of the above conditions is chosen, and this nominal solution is updated by successive linearization so that the remaining conditions are also satisfied as the process converges.

We use a neighboring extremal method [17]. A nominal solution satisfies all of the necessary conditions except the boundary conditions. The neighboring extremal method is characterized as an iterative algorithm

for improving estimates of the unspecified multiplier initial conditions so as to satisfy the specified terminal boundary conditions in the limit. This is sometimes referred to as a shooting method. The optimality condition (27) is substituted into the equations of motion and the multiplier equations. The sensitivities of the specified terminal boundary conditions with respect to the unspecified initial multiplier conditions can be calculated by direct numerical differentiation, or they can be obtained by a linear analysis. The main advantage of the neighboring extremal method is that the number of iteration variables is small. It is equal to the dimension of the equations of motion. The difficulty is that the extremal solutions are sensitive to small changes in the unspecified initial multiplier values. Therefore, it is important to compute the sensitivities accurately.

We use linear analysis to compute the sensitivities. The sensitivity model is defined at the Lie algebra level as presented in [11]. It is natural to define the sensitivity model in the Lie algebra, since the Lie algebra is a linear vector space. The resulting sensitivity model is global, and it has the same dimension as the Lie group. The sensitivity derivatives in the Lie algebra are related to the original Lie group by the exponential map.

Using the perturbation models defined in (14), (15), the linearized equations of motion for the attitude dynamics can be written as

$$\begin{bmatrix} \zeta_{k+1} \\ \delta\Pi_{k+1} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{bmatrix} \begin{bmatrix} \zeta_k \\ \delta\Pi_k \end{bmatrix} - \begin{bmatrix} 0_{3\times 3} \\ hBB^T \end{bmatrix} \delta\lambda_k^2. \tag{33}$$

Note that the homogeneous part of (33) is equivalent to equations that are the dual of (28). The variation of the equations of motion is equivalent to the dual of the multiplier equations, so the variation of the multiplier equations is equivalent to the second variation of the attitude dynamics equations. Together with (33), a tedious but straightforward application of (14), (15) to the multiplier equation gives the following expression for the sensitivity derivatives of the attitude, angular momentum, and the multipliers

$$\begin{bmatrix} x_{k+1} \\ \delta\lambda_{k+1} \end{bmatrix} = A_k \begin{bmatrix} x_k \\ \delta\lambda_k \end{bmatrix}, \tag{34}$$

where $x_k = [\zeta_k^T, \delta\Pi_k^T]^T$, $\delta\lambda_k = [\delta\lambda_k^{1,T}, \delta\lambda_k^{2,T}]^T \in \mathbb{R}^6$, and $A_k \in \mathbb{R}^{12\times 12}$ can be suitably defined in terms of R_k, Π_k. The solution of (34) is given by

$$\begin{aligned} \begin{bmatrix} x_N \\ \delta\lambda_N \end{bmatrix} &= \left(\prod_{k=0}^{N-1} A_k\right) \begin{bmatrix} x_0 \\ \delta\lambda_0 \end{bmatrix}, \\ &\triangleq \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} x_0 \\ \delta\lambda_0 \end{bmatrix}. \end{aligned} \tag{35}$$

For the given two-point boundary value problem $x_0 = 0$ since the initial attitude and the initial angular momentum are given, and λ_N is free. Then, we obtain

$$x_N = \Phi_{12}\delta\lambda_0 \tag{36}$$

The unspecified initial multipliers are λ_0, and the specified terminal boundary conditions are the terminal attitude R_N^d and the terminal angular momentum Π_N^d. Thus Φ_{12} represents the sensitivity of the specified terminal boundary conditions with respect to the unspecified initial multipliers. Using this sensitivity, an initial guess of the unspecified initial conditions is iterated to satisfy the specified terminal conditions in the limit.

TABLE 1. Newton-Armijo iteration procedures

1: Guess an initial multiplier λ_0.
2: Find $\Pi_k, R_k, \lambda_k^1, \lambda_k^2$ for $k = 1, 2, \cdots, N$ using the initial conditions and (24)–(28).
3: Compute the error in satisfaction of the terminal boundary condition;
   $\zeta_N = S^{-1}\big(\mathrm{logm}\big(R_N^T R_N^d\big)\big)$, $\delta\Pi_N = \Pi_N^d - \Pi_N$.
   Error $= \|[\zeta_N; \delta\Pi_N]\|$.
4: Set Error^t = Error, $i = 1$.
5: **while** Error $> \epsilon_S$.
6: Find a line search direction; $D = \Phi_{12}^{-1}$.
7: Set $c = 1$.
8: **while** Error$^t > (1 - 2\alpha c)$Error
9: Choose a trial initial condition $\lambda_0^t = \lambda_0 + cD[\zeta_N; \delta\Pi_N]$.
10: Find $\Pi_k, R_k, \lambda_k^1, \lambda_k^2$ for $k = 1, 2, \cdots, N$ using the trial initial conditions and (24)–(28).
11: Compute the error in satisfaction of the terminal boundary condition
   $\zeta_N^t = S^{-1}\big(\mathrm{logm}\big(R_N^T R_N^d\big)\big)$, $\delta\Pi_N^t = \Pi_N^d - \Pi_N$.
   Error$^t = \|[\zeta_N^t; \delta\Pi_N^t]\|$.
12: Set $c = c/2$, $i = i + 1$.
13: **end while**
14: Set $\lambda_0 = \lambda_0^t$, Error = Error^t. (accept the trial)
15: **end while**

Any type of Newton iteration can be applied to this problem using the sensitivity derivative as a gradient. The procedure uses a line search with backtracking algorithm, referred to as Newton-Armijo iteration in [18]. The procedure is summarized in Table 1, where i is the iteration index, and $\epsilon_S, \alpha \in \mathbb{R}$ are a stopping criterion and a scaling factor, respectively. The outer loop finds a search direction by computing the sensitivity derivatives, and the inner loop performs a line search along the obtained direction; the error in satisfaction of the terminal boundary condition is determined on each iteration.

4. Numerical Computations

Numerical results are given for two optimal attitude control problems; optimal attitude control of an underactuated 3D pendulum and optimal attitude control of a fully actuated spacecraft on a circular orbit.

(a) 3D Pendulum

(b) Spacecraft on a circular orbit

Figure 1. Rigid body models

4.1. **3D Pendulum.** A 3D pendulum is a rigid body supported at a frictionless pivot acting under the influence of uniform gravity [19]. The gravity potential, acting in the vertical or e_3 direction, is given by

$$U(R) = -mge_3^T R\rho, \tag{37}$$

where $m \in \mathbb{R}$ is the mass of the pendulum, $g \in \mathbb{R}$ is the gravitational acceleration, and $\rho \in \mathbb{R}^3$ represents a vector from the pivot point to the mass center of the pendulum in the body fixed frame. The pendulum model is shown in Fig. 1.(a) with the pivot located at the origin, and we assume that the pendulum is axially

symmetric. The gravity moment and its variations are given by

$$M = mg\rho \times R^T e_3,$$

$$\delta M = \mathcal{M}\zeta = mgS(\rho)S(R^T e_3)\zeta.$$

There are two equilibrium manifolds; a hanging equilibrium manifold when $R\rho = \|\rho\| e_3$, an inverted equilibrium manifold when $R\rho = -\|\rho\| e_3$.

The properties of the axially symmetric pendulum are given by $J = \text{diag}\,[0.156,\, 0.156,\, 0.3]\,\text{kg}\,\text{m}^2$, $m = 1\text{kg}$, and $\rho = \left[0, 0, \frac{3}{4}\right]\text{m}$. We assume that the component of the control input along the axis of symmetry is zero; this corresponds to an underactuated 3D pendulum. The corresponding input matrix is given by

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Two types of boundary conditions are considered. The first maneuver is to transfer the 3D pendulum from a hanging equilibrium to an inverted equilibrium. The second maneuver is a 180 degree rotation about the uncontrolled axis of symmetry starting in a hanging equilibrium. The terminal attitude also lies in the hanging equilibrium manifold. Each maneuver is completed in 1 sec. The time step size is $h = 0.001$ sec and the number of integration steps is $N = 1000$. The corresponding boundary conditions are given by

(i) Rotation from a hanging equilibrium to an inverted equilibrium.

$$R_0 = I_{3\times 3}, \quad R_N^d = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

$$\Pi_0 = 0_{3\times 1}, \quad \Pi_N^d = 0_{3\times 1}.$$

(ii) Rotation from one hanging equilibrium to another hanging equilibrium.

$$R_0 = I_{3\times 3}, \quad R_N^d = \text{diag}[-1, -1, 1],$$

$$\Pi_0 = 0_{3\times 1}, \quad \Pi_N^d = 0_{3\times 1}.$$

The optimized performance index and the violation of the constraints are given in Table 2. The terminal boundary conditions are satisfied at the level of machine precision for both cases. Figures 2 and 3 show

TABLE 2. Optimized performance index and violation of constraints

Case	Model	$\mathcal{J}$	$\left\|\text{logm}(R_N^{d,T} R_N)\right\|$	$\left\|\Pi_N^d - \Pi_N\right\|$
(i)	3D Pendulum	1.52	1.77×10^{-14}	7.08×10^{-15}
(ii)		40.22	2.22×10^{-16}	2.55×10^{-14}
(iii)	Spacecraft	23.35	2.90×10^{-15}	5.13×10^{-15}
(iv)		70.74	7.31×10^{-15}	1.48×10^{-14}

snapshots of the attitude maneuvers, the control input history, and the angular velocity response. (Simple animations which show these optimal attitude maneuvers of the 3D pendulum can be found at `http://www.umich.edu/~tylee`.) The third component of the angular velocity is constant; this is a conservation property of the controlled axially symmetric 3D pendulum.

The optimized attitude maneuver of the first case is an eigen-axis rotation about the fixed axis; $[\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0]$. In the second case, the rotation about the axis of symmetry is induced from control moments about the first and second body fixed axes. The resulting attitude maneuver is more complicated, and it requires larger control inputs.

Figures 2.(d) and 3.(d) show the violation of the terminal boundary condition according to the number of iterations in a logarithm scale. The circles denote outer iterations to compute the sensitivity derivatives. The inner iterations correspond to backtracking to decrease the step length along the search direction. For all cases, the initial guesses of the unspecified initial multiplier are arbitrarily chosen such that the initial trial of control inputs is close to zero throughout the maneuver time. The error in satisfaction of the terminal boundary condition of the first case converges quickly to machine precision; only 7 iterations are required. A longer number of iterations is required in the second case, but the error converges exponentially to machine precision after the solution is close to the local minimum at the 55th iteration. These convergence rates are consistent with the quadratic rate of convergence of the Newton iteration, and demonstrate that the sensitivity derivatives are being computed accurately.

4.2. **Spacecraft on a Circular Orbit.** We consider a spacecraft on a circular orbit about a large central body, including gravity gradient effects [20]. The spacecraft model is shown at Fig. 1.(b). The attitude of the spacecraft is represented with respect to the local vertical local horizontal (LVLH) axes. The gravity potential is given by

$$U(R) = -\frac{GM}{r_0} - \frac{1}{2}\omega_0^2 \left(\text{tr}[J] - 3e_3^T R J R^T e_3\right),$$

where $G \in \mathbb{R}$ is the gravitational constant, $M \in \mathbb{R}$ is the mass of the central body, $r_0 \in \mathbb{R}$ is the orbital radius, and $\omega_0 = \sqrt{\frac{GM}{r_0^3}} \in \mathbb{R}$ is the orbital angular velocity. The gravity moment and its variations are given by

$$
\begin{aligned}
M &= 3\omega_0^2 R^T e_3 \times J R^T e_3, \\
\delta M &= \mathcal{M}\zeta = 3\omega_0^2 \Big[- S(JR^T e_3) S(R^T e_3) + S(R^T e_3) J S(R^T e_3) \Big] \zeta
\end{aligned}
$$

There are 24 distinct relative equilibria for which the principal axes are exactly aligned with the LVLH axes, and the spacecraft angular velocity is identical to the orbital angular velocity of the LVLH coordinate frame.

We assume that the spacecraft is fully actuated. The corresponding input matrix is $B = I_{3\times 3}$. The mass, length and time dimensions are normalized by the mass of the spacecraft, a size scale factor of the spacecraft, and the orbital angular velocity $\omega_0 \in \mathbb{R}$, respectively. The mass property of the spacecraft is chosen as $J = \mathrm{diag}\,[1,\, 2.8,\, 2]$.

Two boundary conditions are considered. Each maneuver is a large attitude change completed in a quarter of the orbit, $T_f = \frac{\pi}{2}$. The time step size is $h = 0.001$ and the number of integration steps is $N = 1571$. The terminal angular momentum is chosen such that the terminal attitude is maintained after the maneuver.

(iii) Rotation maneuver about the LVLH axis e_1:

$$
\begin{aligned}
R_0 = I_{3\times 3}, \quad R_N^d = \mathrm{diag}\,[1, -1, -1]\,, \\
\Pi_0 = \omega_0 J R_0^T e_2, \quad \Pi_N^d = \omega_0 J R_N^{d,T} e_2.
\end{aligned}
$$

(iv) Rotation maneuver about the LVLH axes e_1 and e_2:

$$
\begin{aligned}
R_0 = \mathrm{diag}\,[1, -1, -1]\,, \quad R_N^d = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \\
\Pi_0 = \omega_0 J R_0^T e_2, \quad \Pi_N^d = \omega_0 J R_N^{d,T} e_2.
\end{aligned}
$$

The optimized performance index and the violation of the constraints are given in Table 2. Figures 4 and 5 show the attitude maneuver of the spacecraft (clockwise direction), the control inputs, the angular velocity response, and the violation of the terminal boundary condition according to the number of iterations.

4.3. **Numerical properties.** The neighboring extremal method or the shooting method are numerically efficient in the sense that the number of optimization parameters is minimized. But, this approach may be prone to numerical ill-conditioning [21]. A small change in the initial multiplier can cause highly nonlinear behavior of the terminal attitude and angular momentum. It is difficult to compute the Jacobian matrix for Newton iterations accurately, and consequently, the numerical error may not converge to zero within machine precision.

However, the numerical examples presented in this section show excellent numerical convergence properties. They exhibit a quadratic rate of convergence. This is because the proposed computational algorithms on SO(3) are geometrically exact and numerically accurate. The attitude dynamics of a rigid body arises from Hamiltonian mechanics, which have neutral stability. The adjoint system is also neutrally stable. The proposed Lie group variational integrators and the discrete multiplier equations, obtained from variations expressed in the Lie algebra, can preserve the neutrally stability property. Therefore the sensitivity derivatives are computed accurately.

5. Conclusions

A discrete optimal control problem for the attitude dynamics of a rigid body in the presence of an attitude dependent potential is studied. The performance index is the l_2 norm of external control inputs and boundary conditions on the attitude and the angular momentum are prescribed. The attitude is represented by a rotation matrix in the Lie group, SO(3). This paper proposes three levels of geometrically exact computations on SO(3) to solve the optimal control problem; Lie group variational integrator, discrete-time necessary conditions for optimality, and discrete-time sensitivity derivatives.

The Lie group variational integrator obtained from a discrete variational principle preserves the geometric features of the attitude dynamics of the rigid body. It exhibits symplectic and momentum preservation properties, and good energy behavior characteristic of variational integrators. Since the rotation matrices are updated by a group operation, the Lie group structure is also preserved.

The necessary conditions of optimality are derived by a variational principle. The Lie group variational integrators are imposed as constraints, and the variation of the rotation matrices are expressed in terms of Lie algebra elements. The proposed discrete optimality conditions are the basis for a numerically efficient computational algorithms for the optimal attitude control problem, since the implicit part of the optimality conditions occurs in a single equation of one variable. This implicit condition can be solved easily by Newton iteration. Other algorithms require iteration on the entire discrete time trajectory simultaneously.

The necessary conditions are expressed as a two-point boundary value problem on $T^*SO(3)$ and its dual space. The sensitivity derivatives are developed in the Lie algebra, and the two-point boundary value problem is solved using a neighboring extremal method. The neighboring extremal method is efficient for this class of optimal control problems because the resulting problem of satisfying the terminal boundary conditions has a small number of variables. The main disadvantage is that a small change in the initial multipliers can produce a very large change in the terminal condition. This can result in numerical ill-conditioning. The nonlinearity also makes it hard to construct an accurate estimate of the Jacobian matrix that is needed for a Newton iteration. We address this issue in the paper by accurately computing the sensitivity derivatives in the Lie algebra of SO(3). This yields an efficient method for solving the two-point boundary problem, and the error in the terminal boundary conditions converges exponentially to machine precision.

Numerical results for an optimal attitude control problem involving an underactuated axially symmetric 3D pendulum and for an optimal attitude control problem involving a fully actuated spacecraft on a circular orbit are given. The boundary conditions are chosen such that the resulting maneuvers are large angle attitude maneuver. It is shown that the proposed numerical computations on SO(3) are geometrically exact and highly efficient.

References

[1] M. Leok, "Foundations of Computational Geometric Mechanics," Ph.D. dissertation, California Instittute of Technology, 2004.

[2] T. Lee, M. Leok, and N. H. McClamroch, "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum," in *Proceedings of the IEEE Conference on Control Applications*, 2005, pp. 962–967.

[3] ——, "Lie group variational integrators for the full body problem," *Computer Methods in Applied Mechanics and Engineering*, 2007, accepted. [Online]. Available: http://arxiv.org/math.NA/0508365

[4] ——, "Lie group variational integrators for the full body problem in orbital mechanics," *Celestial Mechanics and Dynamical Astronomy*, 2007, accepted.

[5] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[6] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, "Lie-group methods," *Acta Numerica*, vol. 9, pp. 215–365, 2000.

[7] K. Spindler, "Optimal control on Lie groups with applications to attitude control," *Mathematics of Control, Signals, and Systems*, vol. 11, pp. 197–219, 1998.

[8] I. I. Hussein and A. M. Bloch, "Optimal control on Riemannian manifolds with potential fields," in *Proceedings of IEEE Conference on Decision and Control*, 2004, pp. 1982–1987.

[9] O. Junge, J. E. Marsden, and S. Ober-Blöbaum, "Discrete mechanics and optimal control," in *IFAC Congress*, Praha, 2005.

[10] I. I. Hussein, M. Leok, A. K. Sanyal, and A. M. Bloch, "A discrete variational integrator for optimal control problems on SO(3)," in *Proceedings of IEEE Conference on Decision and Control*, 2006, pp. 6636–6641.

[11] T. Lee, M. Leok, and N. H. McClamroch, "Attitude maneuvers of a rigid spacecraft in a circular orbit," in *Proceedings of the American Control Conference*, 2005, pp. 1742–1747.

[12] ——, "Optimal control of a rigid body using geometrically exact computations on SE(3)," in *Proceedings of the IEEE Conference on Decision and Control*, 2006, pp. 2170–2175. [Online]. Available: http://arxiv.org/abs/math.OC/0602588

[13] ——, "Optimal attitude control for a rigid body with symmetry," in *Proceedings of the American Control Conference*, 2007, accepted. [Online]. Available: http://arxiv.org/abs/math.OC/06009482

[14] M. Suzuki, "Improved Trotter-like formula," *Phys. Lett. A*, vol. 180, no. 3, pp. 232–234, 1993.

[15] T. Littell, R. Skeel, and M. Zhang, "Error analysis of symplectic multiple time stepping," *SIAM J. Numer. Anal.*, vol. 34, no. 5, pp. 1792–1807, 1997.

[16] E. Hairer, C. Lubich, and G. Wanner, "Geometric numerical integration illustrated by the Störmer-Verlet method," *Acta Numer.*, vol. 12, pp. 399–450, 2003.

[17] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control*. Hemisphere Publishing Corporation, 1975.

[18] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.

[19] J. Shen, A. K. Sanyal, N. A. Chaturvedi, D. Bernstein, and N. H. McClamroch, "Dynamics and control of a 3D pendulum," in *Proceedings of 43rd IEEE Conference on Decision and Control*, Dec. 2004, pp. 323–328.

[20] B. Wie, *Space Vehicle Dynamics and Control*. AIAA, 1998.

[21] J. T. Betts, *Practical Methods for Optimal Control Using Nonlinear Programming*. SIAM, 2001.

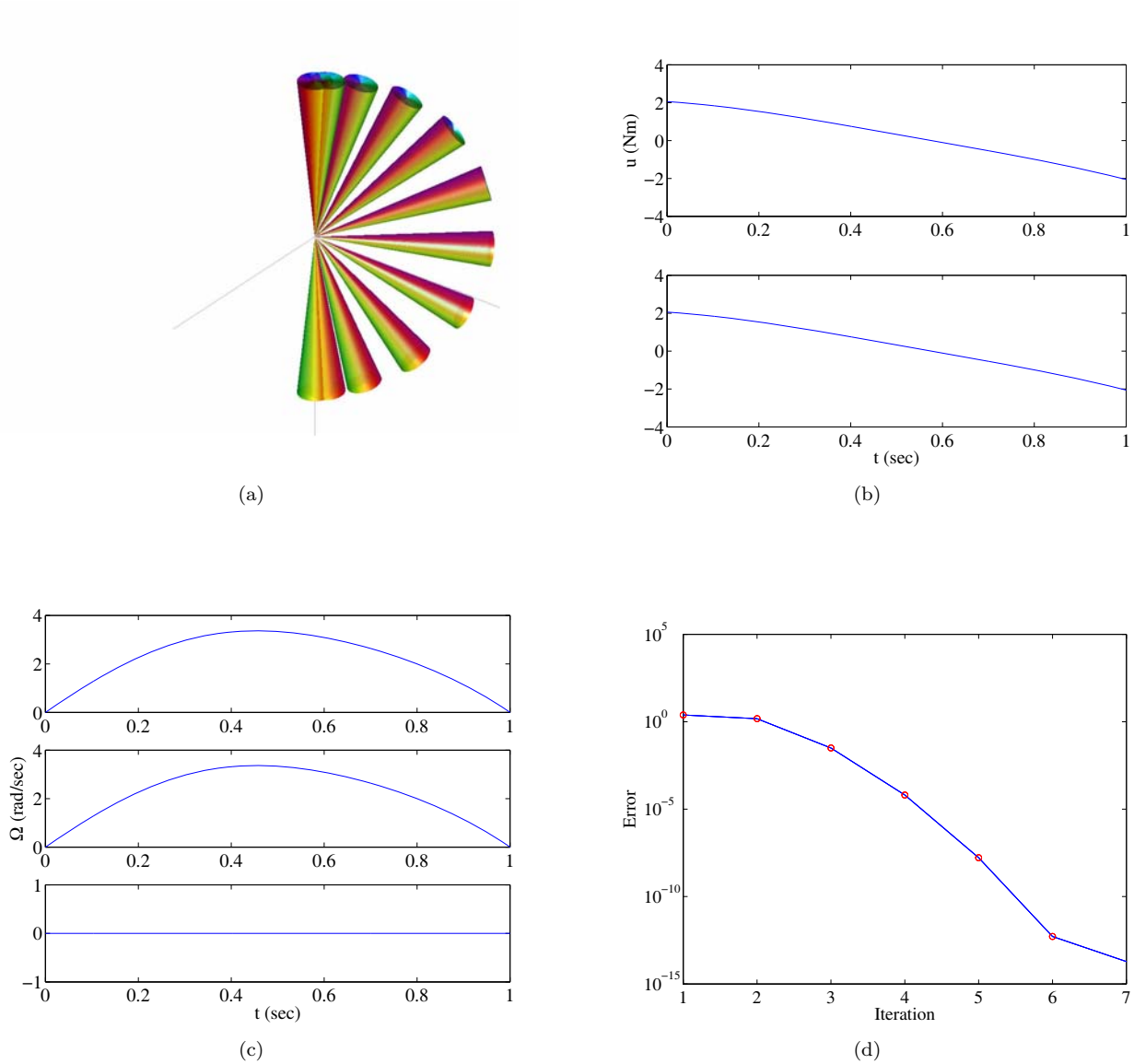
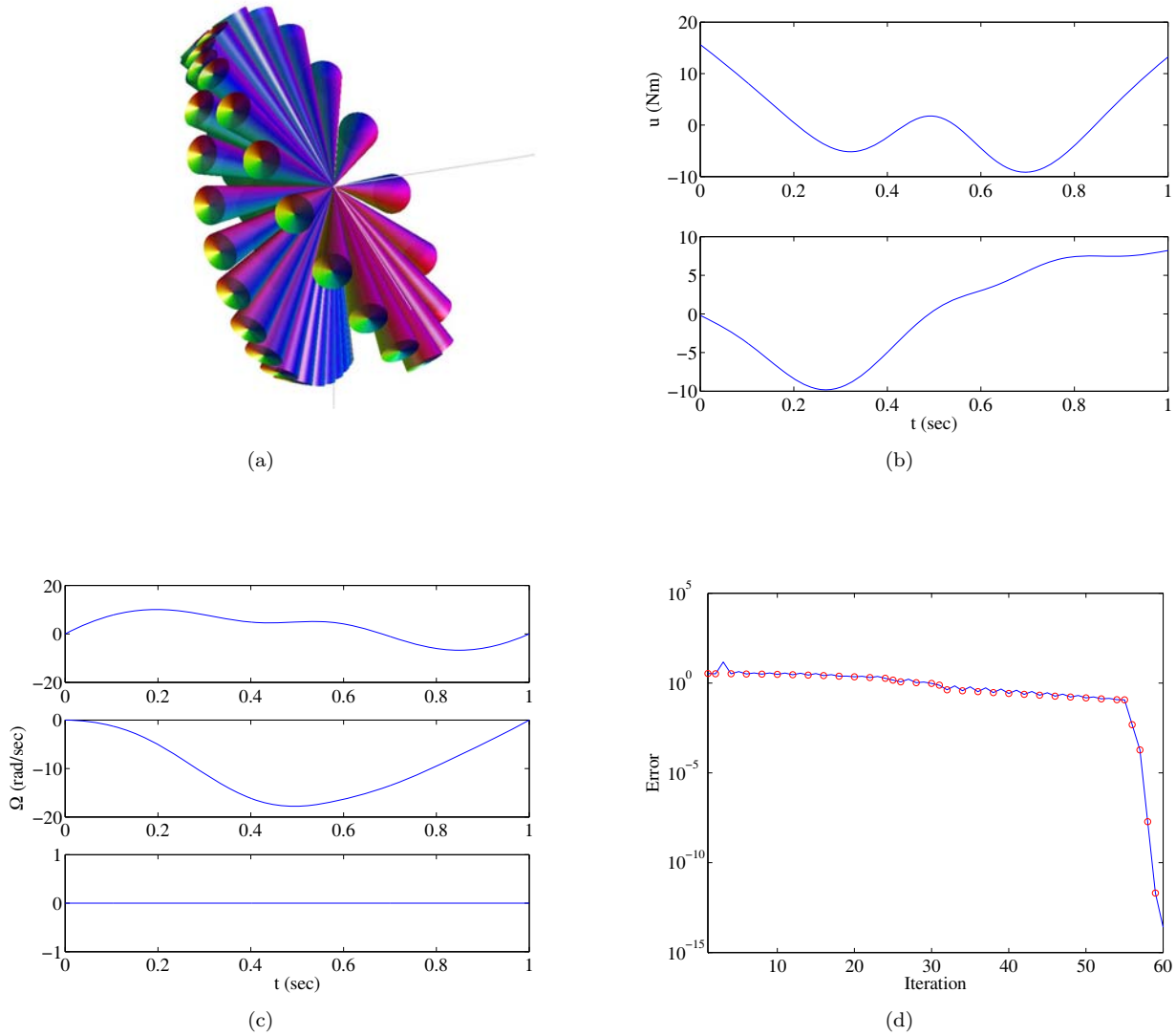
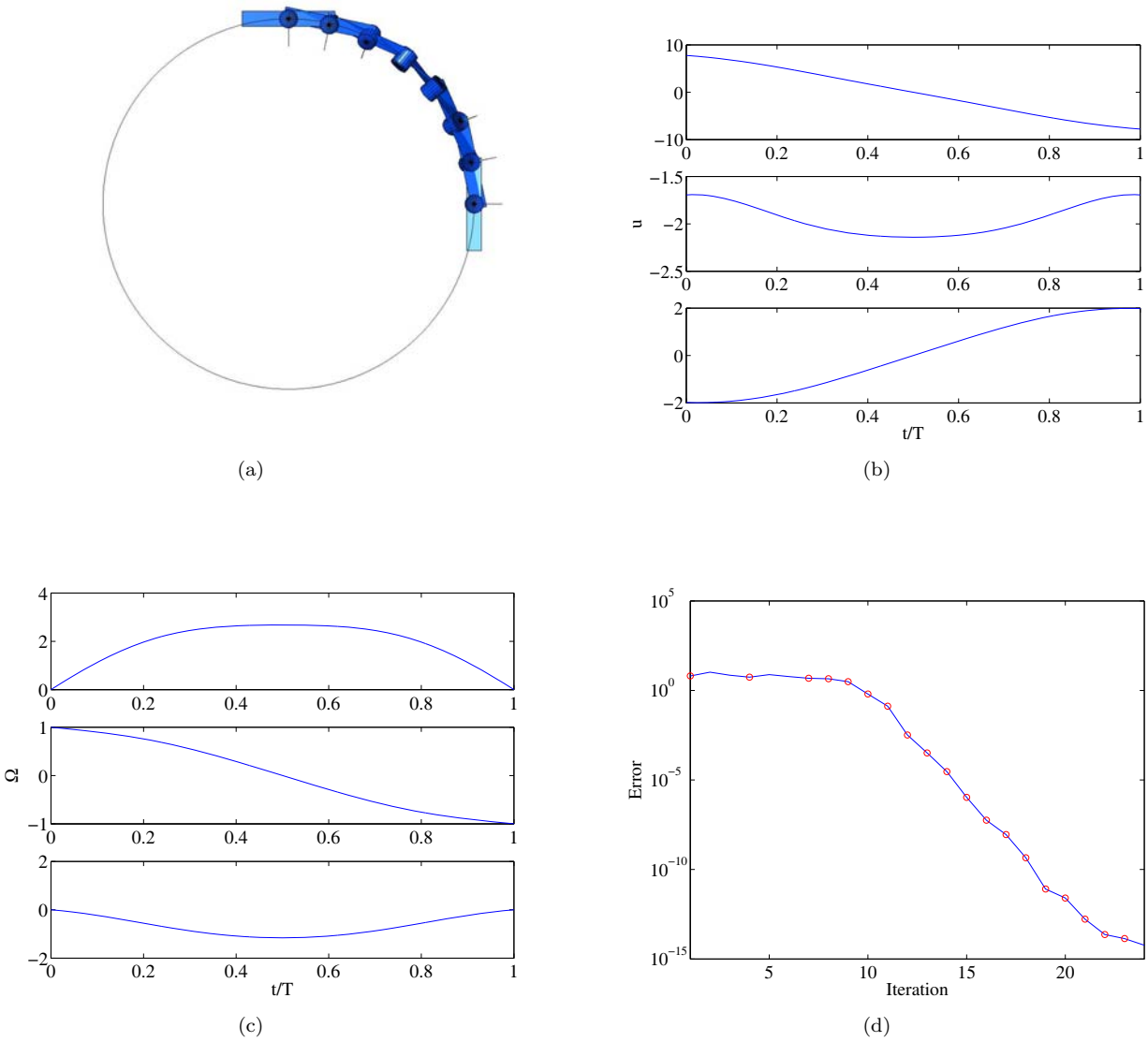


FIGURE 2. Case (i): 3D pendulum rotation from a hanging equilibrium to an inverted equilibrium



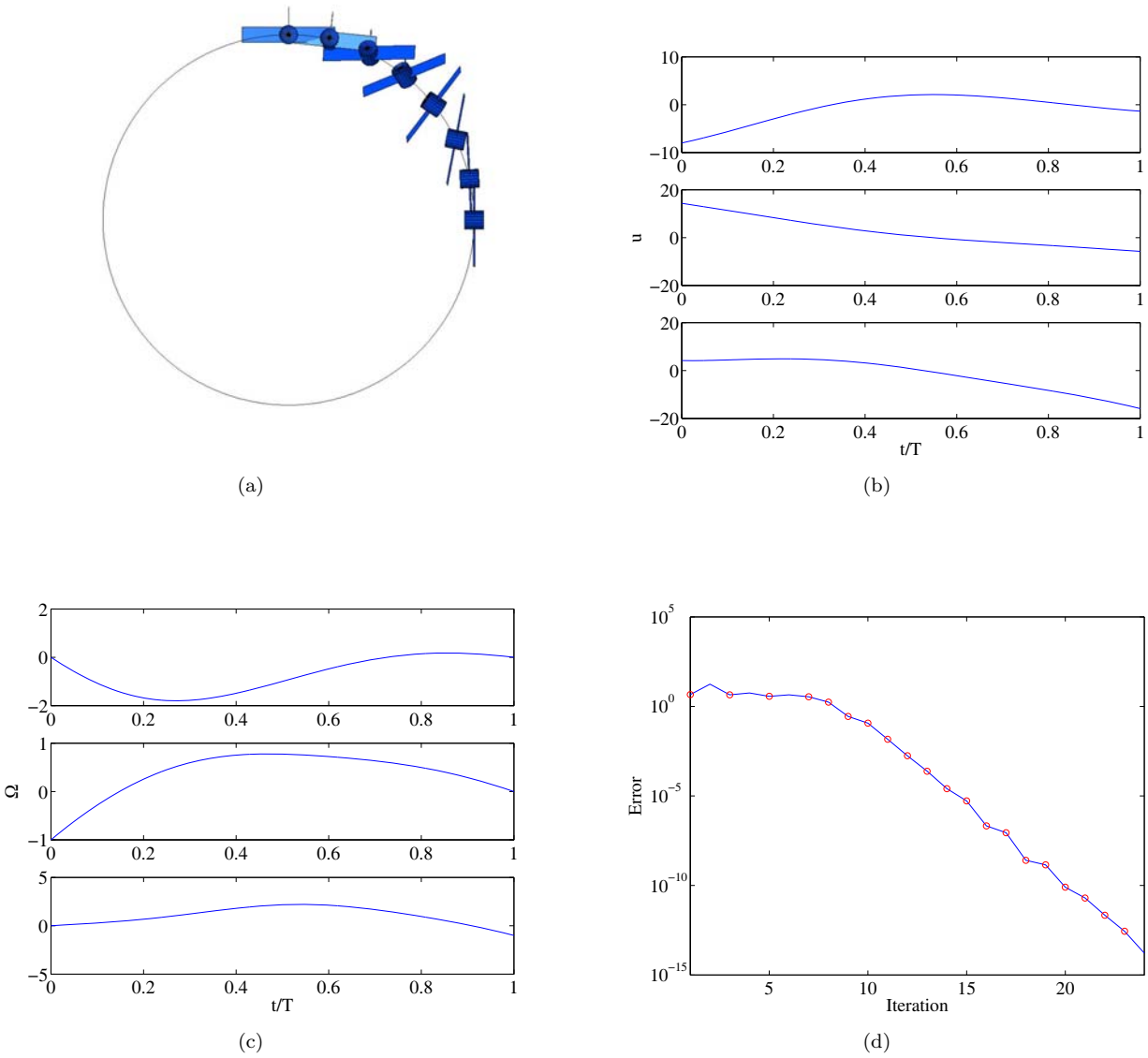
(a) (b) (c) (d)

FIGURE 3. Case (ii): 3D pendulum rotation from one hanging equilibrium to another hanging equilibrium



(a) (b) (c) (d)

FIGURE 4. Case (iii): Spacecraft rotation maneuver about the LVLH tangential axis



(a) (b) (c) (d)

FIGURE 5. Case (iv): Spacecraft rotation maneuver about the LVLH tangential and normal axes

DISCRETE CONTROL SYSTEMS

TAEYOUNG LEE, MELVIN LEOK, AND N. HARRIS MCCLAMROCH

Invited article for the Springer Encyclopedia of Complexity and System Science

Contents

Glossary and Notation

Discrete variational mechanics: A formulation of mechanics in discrete-time that is based on a discrete analogue of **Hamilton's principle**, which states that the system takes a trajectory for which the action integral is stationary.

Geometric integrator: A numerical method for obtaining numerical solutions of differential equations that preserves geometric properties of the continuous flow, such as symplecticity, momentum preservation, and the structure of the configuration space.

Lie group: A differentiable manifold with a group structure where the composition is differentiable. The corresponding **Lie algebra** is the tangent space to the Lie group based at the identity element.

Symplectic: A map is said to be symplectic if given any initial volume in phase space, the sum of the signed projected volumes onto each position-momentum subspace is invariant under the map. One consequence of symplecticity is that the map is volume-preserving as well.

1. Definition of the Subject

Discrete control systems, as considered here, refer to the control theory of discrete-time Lagrangian or Hamiltonian systems. These discrete-time models are based on a discrete variational principle, and are part of the broader field of geometric integration. Geometric integrators are numerical integration methods that preserve geometric properties of continuous systems, such as conservation of the symplectic form, momentum, and energy. They also guarantee that the discrete flow remains on the manifold on which the continuous system evolves, an important property in the case of rigid-body dynamics.

In nonlinear control, one typically relies on differential geometric and dynamical systems techniques to prove properties such as stability, controllability, and optimality. More generally, the geometric structure of such systems plays a critical role in the nonlinear analysis of the corresponding control problems. Despite the critical role of geometry and mechanics in the analysis of nonlinear control systems, nonlinear control algorithms have typically been implemented using numerical schemes that ignore the underlying geometry.

The field of discrete control systems aims to address this deficiency by restricting the approximation to the choice of a discrete-time model, and developing an associated control theory that does not introduce any

TL and ML have been supported in part by NSF Grant DMS-0504747 and DMS-0726263. TL and NHM have been supported in part by NSF Grant ECS-0244977 and CMS-0555797.

additional approximation. In particular, this involves the construction of a control theory for discrete-time models based on geometric integrators that yields numerical implementations of nonlinear and geometric control algorithms that preserve the crucial underlying geometric structure.

2. Introduction

The dynamics of Lagrangian and Hamiltonian systems have unique geometric properties; the Hamiltonian flow is symplectic, the total energy is conserved in the absence of non-conservative forces, and the momentum maps associated with symmetries of the system are preserved. Many interesting dynamics evolve on a non-Euclidean space. For example, the configuration space of a spherical pendulum is the two-sphere, and the configuration space of rigid body attitude dynamics has a Lie group structure, namely the special orthogonal group. These geometric features determine the qualitative behavior of the system, and serve as a basis for theoretical study.

Geometric numerical integrators are numerical integration algorithms that preserve structures of the continuous dynamics such as invariants, symplecticity, and the configuration manifold (see Hairer et al. [14]). The exact geometric properties of the discrete flow not only generate improved qualitative behavior, but also provide accurate and efficient numerical techniques. In this article, we view a geometric integrator as an intrinsically discrete dynamical system, instead of concentrating on the numerical approximation of a continuous trajectory.

Numerical integration methods that preserve the simplecticity of a Hamiltonian system have been studied (see Leimkuhler and Reich [28]; Sanz-Serna [36]). Coefficients of a Runge-Kutta method are carefully chosen to satisfy a simplecticity criterion and order conditions to obtain a symplectic Runge-Kutta method. However, it can be difficult to construct such integrators, and it is not guaranteed that other invariants of the system, such as a momentum map, are preserved. Alternatively, variational integrators are constructed by discretizing Hamilton's principle, rather than discretizing the continuous Euler-Lagrange equation (see Marsden and West [31]; Moser and Veselov [34]). The resulting integrators have the desirable property that they are symplectic and momentum preserving, and they exhibit good energy behavior for exponentially long times (see Benettin and Giorgilli [2]). Lie group methods are numerical integrators that preserve the Lie group structure of the configuration space (see Iserles et al. [18]). Recently, these two approaches have been unified to obtain Lie group variational integrators that preserve the geometric properties of the dynamics as well as the Lie group structure of the configuration space without the use of local charts, reprojection, or constraints (see Lee et al. [26]; Leok [29]; Marsden et al. [32]).

Optimal control problems involve finding a control input such that a certain optimality objective is achieved under prescribed constraints. An optimal control problem that minimizes a performance index is described by a set of differential equations, which can be derived using Pontryagin's maximum principle. Discrete optimal control problems involve finding a control input for a discrete dynamic system such that an optimality objective is achieved with prescribed constraints. Optimality conditions are derived from the discrete equations of motion, described by a set of discrete equations. This approach is in contrast to traditional techniques where a discretization appears at the last stage to solve the optimality condition numerically. Discrete mechanics and optimal control approaches determine optimal control inputs and trajectories more accurately with less computational load (see Junge et al. [19]). Combined with an indirect optimization technique, they are substantially more efficient (see Hussein et al. [17]; Lee et al. [23, 25]).

The geometric approach to mechanics can provide a theoretical basis for innovative control methodologies in geometric control theory. For example, these techniques allow the attitude of satellites to be controlled using changes in its shape, as opposed to chemical propulsion. While the geometric structure of mechanical systems plays a critical role in the construction of geometric control algorithms, these algorithms have typically been implemented using numerical schemes that ignore the underlying geometry. By applying geometric control algorithms to discrete mechanics that preserve geometric properties, we obtain an exact numerical implementation of the geometric control theory. In particular, the method of controlled Lagrangian systems is based on the idea of adopting a feedback control to realize a modification of either the potential energy or the kinetic energy, referred to as potential shaping or kinetic shaping, respectively. These ideas are applied to construct a real-time digital feedback controller that stabilizes the inverted equilibrium of the cart-pendulum (see Bloch et al. [10, 11]).

In this article, we will survey discrete Lagrangian and Hamiltonian mechanics, and their applications to optimal control and feedback control theory.

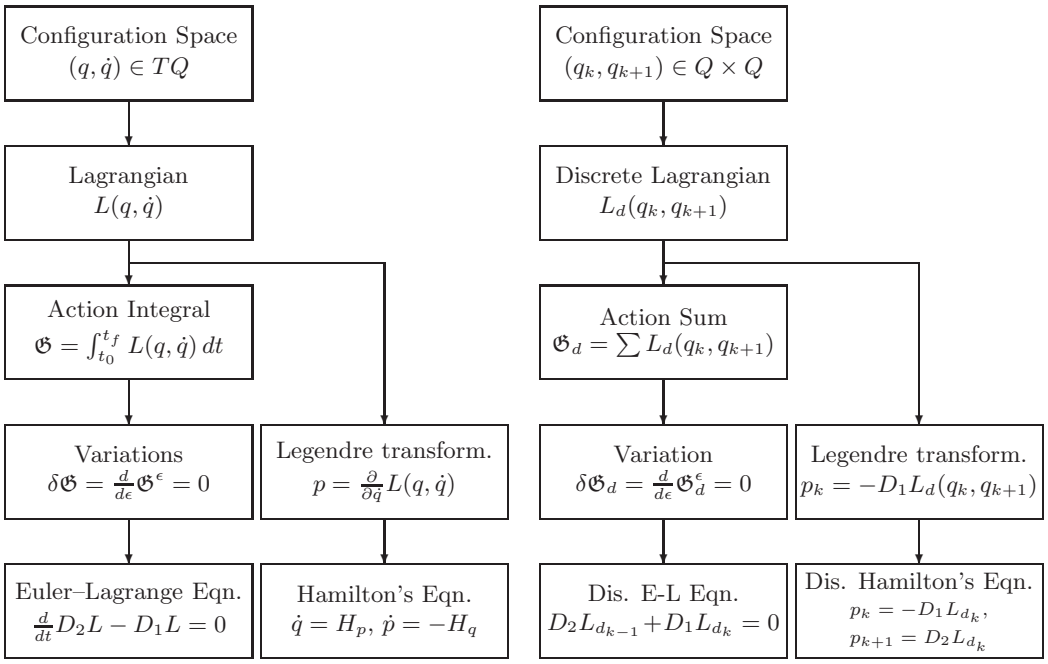


FIGURE 1. Procedures to derive continuous and discrete equations of motion

3. Discrete Lagrangian and Hamiltonian Mechanics

Mechanics studies the dynamics of physical bodies acting under forces and potential fields. In Lagrangian mechanics, the trajectory of the object is derived by finding the path that extremizes the integral of a Lagrangian over time, called the action integral. In many classical problems, the Lagrangian is chosen as the difference between kinetic energy and potential energy. The Legendre transformation provides an alternative description of mechanical systems, referred to as Hamiltonian mechanics.

Discrete Lagrangian and Hamiltonian mechanics has been developed by reformulating the theorems and the procedures of Lagrangian and Hamiltonian mechanics in a discrete time setting (see, for example, Marsden and West [31]). Therefore, discrete mechanics has a parallel structure with the mechanics described in continuous time, as summarized in Figure 1 for Lagrangian mechanics. In this section, we describe discrete Lagrangian mechanics in more detail, and we derive discrete Euler-Lagrange equations for several mechanical systems.

Consider a mechanical system on a configuration space Q, which is the space of possible positions. The Lagrangian depends on the position and velocity, which are elements of the tangent bundle to Q, denoted by TQ. Let $L : TQ \to \mathbb{R}$ be the Lagrangian of the system. The discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$ is an approximation to the exact discrete Lagrangian,

$$L_d^{\text{exact}}(q_0, q_1) = \int_0^h L(q_{01}(t), \dot{q}_{01}(t))dt, \tag{1}$$

where $q_{01}(0) = q_0$, $q_{01}(h) = q_1$, and $q_{01}(t)$ satisfies the Euler–Lagrange equation in the time interval $(0, h)$. A discrete action sum $\mathfrak{G}_d : Q^{N+1} \to \mathbb{R}$, analogous to the action integral, is given by

$$\mathfrak{G}_d(q_0, q_1, \ldots, q_N) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}). \tag{2}$$

The discrete Hamilton's principle states that

$$\delta\mathfrak{G}_d = 0$$

for any δq_k, which yields the ***discrete Euler–Lagrange (DEL)*** equation,

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

This yields a discrete Lagrangian flow map $(q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$. The discrete Legendre transformation, which from a pair of positions (q_0, q_1) gives a position-momentum pair $(q_0, p_0) = (q_0, -D_1 L_d(q_0, q_1))$ provides a discrete Hamiltonian flow map in terms of momenta.

The discrete equations of motion, referred to as variational integrators, inherit the geometric properties of the continuous system. Many interesting Lagrangian and Hamiltonian systems, such as rigid bodies evolve on a Lie group. Lie group variational integrators preserve the nonlinear structure of the Lie group configurations as well as geometric properties of the continuous dynamics (see Marsden et al. [32] and Leok [29]). The basic idea for all Lie group methods is to express the update map for the group elements in terms of the group operation,

$$g_1 = g_0 f_0, \tag{4}$$

where $g_0, g_1 \in G$ are configuration variables in a Lie group G, and $f_0 \in G$ is the discrete update represented by a right group operation on g_0. Since the group element is updated by a group operation, the group structure is preserved automatically without need of parameterizations, constraints, or re-projection. In the Lie group variational integrator, the expression for the flow map is obtained from the discrete variational principle on a Lie group, the same procedure presented in Figure 1. But, the infinitesimal variation of a Lie group element must be carefully expressed to respect the structure of the Lie group. For example, it can be expressed in terms of the exponential map as

$$\delta g = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} g \exp \epsilon\eta = g\eta,$$

for a Lie algebra element $\eta \in \mathfrak{g}$. This approach has been applied to the rotation group SO(3) and to the special Euclidean group SE(3) for dynamics of rigid bodies (see Lee et al. [24], Lee et al. [26], and Lee et al. [27]). Generalizations to arbitrary Lie groups gives the generalized ***discrete Euler–Poincaré (DEP)*** equation,

$$T_e^* L_{f_0} \cdot D_2 L_d(g_0, f_0) - \mathrm{Ad}_{f_0}^* \cdot (T_e^* L_{f_1} \cdot D_2 L_d(g_1, f_1)) + T_e^* L_{g_1} \cdot D_1 L_d(g_1, f_1) = 0, \tag{5}$$

for a discrete Lagrangian on a Lie group, $L_d : G \times G \to \mathbb{R}$. Here $L_f : G \to G$ denotes the left translation map given by $L_f g = fg$ for $f, g \in G$, $T_g L_f : T_g G \to T_{fg} G$ is the tangential map for the left translation, and $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is the adjoint map. A dual map is denoted by a superscript * (see Marsden and Ratiu [30] for detailed definitions).

We illustrate the properties of discrete mechanics using several mechanical systems, namely a mass-spring system, a planar pendulum, a spherical pendulum, and a rigid body.

Example 1 (**Mass-spring System**)**.**
Consider a mass-spring system, defined by a rigid body that moves along a straight frictionless slot, and is attached to a linear spring.

Continuous equation of motion: The configuration space is $Q = \mathbb{R}$, and the Lagrangian $L : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is given by

$$L(q, \dot{q}) = \frac{1}{2} m \dot{q}^2 - \frac{1}{2} \kappa q^2, \tag{6}$$

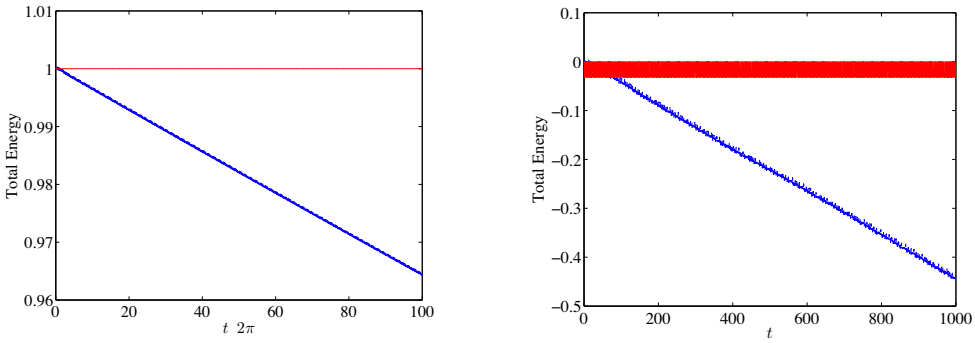
where $q \in \mathbb{R}$ is the displacement of the body measured from the point where the spring exerts no force. The mass of the body and the spring constant are denoted by $m, \kappa \in \mathbb{R}$, respectively. The Euler-Lagrange equation yields the continuous equation of motion.

$$m\ddot{q} + \kappa q = 0. \tag{7}$$

Discrete equation of motion: Let $h > 0$ be a discrete time step, and a subscript k denotes the k-th discrete variable at $t = kh$. The discrete Lagrangian $L_d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is an approximation of the integral of the continuous Lagrangian (6) along the solution of (7) over a time step. Here, we choose the following discrete Lagrangian.

$$L_d(q_k, q_{k+1}) = hL\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}\right) = \frac{1}{2h} m (q_{k+1} - q_k)^2 - \frac{h\kappa}{8} (q_k + q_{k+1})^2. \tag{8}$$

Direct application of the discrete Euler-Lagrange equation to this discrete Lagrangian yields the discrete equations of motion. We develop the discrete equation of motion using the discrete Hamilton's principle



(a) Mass-spring system (b) Planar pendulum

FIGURE 2. Computed total energy (RK45: blue, dotted, VI: red, solid)

to illustrate the principles more explicitly. Let $\mathfrak{G}_d : \mathbb{R}^{N+1} \to \mathbb{R}$ be the discrete action sum defined as $\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1})$, which approximates the action integral. The infinitesimal variation of the action sum can be written as

$$\delta\mathfrak{G}_d = \sum_{k=0}^{N-1} \delta q_{k+1} \left\{ \frac{m}{h}(q_{k+1} - q_k) - \frac{h\kappa}{4}(q_k + q_{k+1}) \right\} + \delta q_k \left\{ -\frac{m}{h}(q_{k+1} - q_k) - \frac{h\kappa}{4}(q_k + q_{k+1}) \right\}.$$

Since $\delta q_0 = \delta q_N = 0$, the summation index can be rewritten as

$$\delta\mathfrak{G}_d = \sum_{k=1}^{N-1} \delta q_k \left\{ -\frac{m}{h}(q_{k+1} - 2q_k + q_{k-1}) - \frac{h\kappa}{4}(q_{k+1} + 2q_k + q_{k-1}) \right\}.$$

From discrete Hamilton's principle, $\delta\mathfrak{G}_d = 0$ for any δq_k. Thus, the discrete equation of motion is given by

$$\frac{m}{h}(q_{k+1} - 2q_k + q_{k-1}) + \frac{h\kappa}{4}(q_{k+1} + 2q_k + q_{k-1}) = 0. \tag{9}$$

For a given (q_{k-1}, q_k), we solve the above equation to obtain q_{k+1}. This yields a discrete flow map $(q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$, and this process is repeated. The discrete Legendre transformation provides the discrete equation of motion in terms of the velocity as

$$\left(1 + \frac{h^2\kappa}{4m}\right) q_{k+1} = h\dot{q}_k + \left(1 - \frac{h^2\kappa}{4m}\right) q_k, \tag{10}$$

$$\dot{q}_{k+1} = \dot{q}_k - \frac{h\kappa}{2m} q_k - \frac{h\kappa}{2m} q_{k+1}. \tag{11}$$

For a given $(q_k, \dot{q}_k)$, we compute q_{k+1} and $\dot{q}_{k+1}$ by (10) and (11), respectively. This yields a discrete flow map $(q_k, \dot{q}_k) \mapsto (q_{k+1}, \dot{q}_{k+1})$. It can be shown that this variational integrator has second-order accuracy, which follows from the fact that the discrete action sum is a second-order approximation of the action integral.

Numerical example: We compare computational properties of the discrete equations of motion given by (10) and (11) with a 4(5)-th order variable step size Runge-Kutta method. We choose $m = 1\,\mathrm{kg}$, $\kappa = 1\,\mathrm{kg/s^2}$ so that the natural frequency is $1\,\mathrm{rad/s}$. The initial conditions are $q_0 = \sqrt{2}\,\mathrm{m}$, $\dot{q}_0 = 0$, and the total energy is $E = 1\,\mathrm{Nm}$. The simulation time is $200\pi\,\mathrm{sec}$, and the step-size $h = 0.035$ of the discrete equations of motion is chosen such that the CPU times are the same for both methods. Figure 2(a) shows the computed total energy. The variational integrator preserves the total energy well. The mean variation is $2.7327 \times 10^{-13}\,\mathrm{Nm}$. But, there is a notable dissipation of the computed total energy for the Runge-Kutta method

Example 2 (Planar Pendulum).
A planar pendulum is a mass particle connected to a frictionless, one degree-of-freedom pivot by a rigid massless link under a uniform gravitational potential. The configuration space is the one-sphere $\mathbb{S}^1 = \{q \in \mathbb{R}^2 \mid \|q\| = 1\}$. While it is common to parameterize the one-sphere by an angle, we develop parameter-free equations of motion in the special orthogonal group SO(2), which is a group of 2×2 orthogonal matrices

with determinant 1, i.e. $\mathrm{SO}(2) = \left\{R \in \mathbb{R}^{2\times 2} \,|\, R^T R = I_{2\times 2}, \det[R] = 1\right\}$. $\mathrm{SO}(2)$ is diffeomorphic to the one-sphere. It is also possible to develop global equations of motion on the one-sphere directly, as shown in the next example, but here we focus on the special orthogonal group in order to illustrate the key steps to develop a Lie group variational integrator.

We first exploit the basic structures of the Lie group $\mathrm{SO}(2)$. Define a hat map $\hat{\cdot}$, which maps a scalar Ω to a 2×2 skew-symmetric matrix $\hat{\Omega}$ as

$$\hat{\Omega} = \begin{bmatrix} 0 & -\Omega \\ \Omega & 0 \end{bmatrix}.$$

The set of 2×2 skew-symmetric matrices forms the Lie algebra $\mathfrak{so}(2)$. Using the hat map, we identify $\mathfrak{so}(2)$ with $\mathbb{R}$. An inner product on $\mathfrak{so}(2)$ can be induced from the inner product on $\mathbb{R}$ as $\langle \hat{\Omega}_1, \hat{\Omega}_2 \rangle = \frac{1}{2}\mathrm{tr}[\hat{\Omega}_1^T \hat{\Omega}_2] = \Omega_1 \Omega_2$ for any $\Omega_1, \Omega_2 \in \mathbb{R}$. The matrix exponential is a local diffeomorphism from $\mathfrak{so}(2)$ to $\mathrm{SO}(2)$ given by

$$\exp \hat{\Omega} = \begin{bmatrix} \cos\Omega & -\sin\Omega \\ \sin\Omega & \cos\Omega \end{bmatrix}.$$

The kinematics equation for $R \in \mathrm{SO}(2)$ can be written in terms of a Lie algebra element as

$$\dot{R} = R\hat{\Omega}. \tag{12}$$

Continuous equations of motion: The Lagrangian for a planar pendulum $L : \mathrm{SO}(2) \times \mathfrak{so}(2) \to \mathbb{R}$ can be written as

$$L(R, \hat{\Omega}) = \frac{1}{2} m l^2 \Omega^2 + mgle_2^T R e_2 = \frac{1}{2} m l^2 \langle \hat{\Omega}, \hat{\Omega} \rangle + mgle_2^T R e_2, \tag{13}$$

where the constant $g \in \mathbb{R}$ is the gravitational acceleration. The mass and the length of the pendulum are denoted by $m, l \in \mathbb{R}$, respectively. The second expression is used to define a discrete Lagrangian later. We choose the bases of the inertial frame and the body-fixed frame such that the unit vector along the gravity direction in the inertial frame, and the unit vector along the pendulum axis in the body-fixed frame are represented by the same vector $e_2 = [0; 1] \in \mathbb{R}^2$. Thus, for example, the hanging attitude is represented by $R = I_{2\times 2}$. Here, the rotation matrix $R \in \mathrm{SO}(2)$ represents the linear transformation from a representation of a vector in the body-fixed frame to the inertial frame.

Since the special orthogonal group is not a linear vector space, the expression for the variation should be carefully chosen. The infinitesimal variation of a rotation matrix $R \in \mathrm{SO}(2)$ can be written in terms of its Lie algebra element as

$$\delta R = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R \exp \epsilon\hat{\eta} = R\hat{\eta} \exp \epsilon\hat{\eta}\bigg|_{\epsilon=0} = R\hat{\eta}, \tag{14}$$

where $\eta \in \mathbb{R}$ so that $\hat{\eta} \in \mathfrak{so}(2)$. The infinitesimal variation of the angular velocity is induced from this expression and (12) as

$$\delta\hat{\Omega} = \delta R^T \dot{R} + R^T \delta\dot{R} = (R\hat{\eta})^T \dot{R} + R^T(\dot{R}\hat{\eta} + R\dot{\hat{\eta}}) = -\hat{\eta}\hat{\Omega} + \hat{\Omega}\hat{\eta} + \dot{\hat{\eta}} = \dot{\hat{\eta}}, \tag{15}$$

where we used the equality of mixed partials to compute $\delta\dot{R}$ as $\frac{d}{dt}(\delta R)$.

Define the action integral to be $\mathfrak{G} = \int_0^T L(R, \hat{\Omega})\, dt$. The infinitesimal variation of the action integral is obtained by using (14) and (15). Hamilton's principle yields the following continuous equations of motion.

$$\dot{\Omega} + \frac{g}{l} e_2^T R e_1 = 0, \tag{16}$$

$$\dot{R} = R\hat{\Omega}. \tag{17}$$

If we parameterize the rotation matrix as $R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, these equations are equivalent to

$$\ddot{\theta} + \frac{g}{l}\sin\theta = 0. \tag{18}$$

Discrete equations of motion: We develop a Lie group variational integrator on $\mathrm{SO}(2)$. Similar to (4), define $F_k \in \mathrm{SO}(2)$ such that

$$R_{k+1} = R_k F_k. \tag{19}$$

Thus, $F_k = R_k^T R_{k+1}$ represents the relative update between two integration steps. If we find $F_k \in \mathrm{SO}(2)$, the orthogonal structure is preserved through (19) since multiplication of orthogonal matrices is also orthogonal. This is a key idea of Lie group variational integrators.

Define the discrete Lagrangian $L_d : \mathrm{SO}(2) \times \mathrm{SO}(2) \to \mathbb{R}$ to be

$$\begin{aligned} L_d(R_k, F_k) &= \frac{1}{2h} ml^2 \langle F_k - I_{2\times 2}, F_k - I_{2\times 2} \rangle + \frac{h}{2} mgle_2^T R_k e_2 + \frac{h}{2} mgle_2^T R_{k+1} e_2, \\ &= \frac{1}{2h} ml^2 \mathrm{tr}[I_{2\times 2} - F_k] + \frac{h}{2} mgle_2^T R_k e_2 + \frac{h}{2} mgle_2^T R_{k+1} e_2, \end{aligned} \tag{20}$$

which is obtained by an approximation $h\hat{\Omega}_k \simeq R_k^T(R_{k+1} - R_k) = F_k - I_{2\times 2}$, applied to the continuous Lagrangian given by (13).

As for the continuous time case, expressions for the infinitesimal variations should be carefully chosen. The infinitesimal variation of a rotation matrix is the same as (14), namely

$$\delta R_k = R_k \hat{\eta}_k, \tag{21}$$

for $\eta_k \in \mathbb{R}$, and the constrained variation of F_k is obtained from (19) as

$$\delta F_k = \delta R_k^T R_{k+1} + R_k^T \delta R_{k+1} = -\hat{\eta}_k F_k + F_k \hat{\eta}_{k+1} = F_k(\hat{\eta}_{k+1} - F_k^T \hat{\eta}_k F_k) = F_k(\hat{\eta}_{k+1} - \hat{\eta}_k), \tag{22}$$

where we use the fact that $F\hat{\eta}F^T = \hat{\eta}$ for any $F \in \mathrm{SO}(2)$ and $\hat{\eta} \in \mathfrak{so}(2)$.

Define an action sum $\mathfrak{G}_d : \mathrm{SO}(2)^{N+1} \to \mathbb{R}$ as $\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(R_k, F_k)$. Using (21) and (22), the variation of the action sum is written as

$$\delta \mathfrak{G}_d = \sum_{k=1}^{N-1} \left\langle \frac{1}{2h} ml^2 (F_{k-1} - F_{k-1}^T) - \frac{1}{2h}(F_k - F_k^T) - hmgle_2^T \widehat{R_k e_1},\ \hat{\eta}_k \right\rangle.$$

From the discrete Hamilton's principle, $\delta \mathfrak{G}_d = 0$ for any $\hat{\eta}_k$. Thus, we obtain the Lie group variational integrator on $\mathrm{SO}(2)$ as

$$(F_k - F_k^T) - (F_{k+1} - F_{k+1}^T) - \frac{2h^2 g}{l} e_2^T \widehat{R_{k+1} e_1} = 0, \tag{23}$$

$$R_{k+1} = R_k F_k. \tag{24}$$

For a given (R_k, F_k) and $R_{k+1} = R_k F_k$, (23) is solved to find F_{k+1}. This yields a discrete map $(R_k, F_k) \mapsto (R_{k+1}, F_{k+1})$. If we parameterize the rotation matrices R and F with θ and $\Delta\theta$ and if we assume that $\Delta\theta \ll 1$, these equations are equivalent to

$$\frac{1}{h}(\theta_{k+1} - 2\theta_k + \theta_k) + \frac{hg}{l} \sin\theta_k = 0.$$

The discrete version of the Legendre transformation provides the discrete Hamiltonian map as follows.

$$F_k - F_k^T = 2h\hat{\Omega} - \frac{h^2 g}{l} e_2^T \widehat{R_k e_1}, \tag{25}$$

$$R_{k+1} = R_k F_k, \tag{26}$$

$$\Omega_{k+1} = \Omega_k - \frac{hg}{2l} e_2^T R_k e_1 - \frac{hg}{2l} e_2^T R_{k+1} e_1. \tag{27}$$

For a given (R_k, Ω_k), we solve (25) to obtain F_k. Using this, (R_{k+1}, Ω_{k+1}) is obtained from (26) and (27). This yields a discrete map $(R_k, \Omega_k) \mapsto (R_{k+1}, \Omega_{k+1})$.

Numerical example: We compare the computational properties of the discrete equations of motion given by (25)–(27) with a 4(5)-th order variable step size Runge-Kutta method. We choose $m = 1\,\mathrm{kg}$, $l = 9.81\,\mathrm{m}$. The initial conditions are $\theta_0 = \pi/2\,\mathrm{rad}$, $\Omega = 0$, and the total energy is $E = 0\,\mathrm{Nm}$. The simulation time is $1000\,\mathrm{sec}$, and the step-size $h = 0.03$ of the discrete equations of motion is chosen such that the CPU times are identical. Figure 2(b) shows the computed total energy for both methods. The variational integrator preserves the total energy well. There is no drift in the computed total energy, and the mean variation is $1.0835 \times 10^{-2}\,\mathrm{Nm}$. But, there is a notable dissipation of the computed total energy for the Runge-Kutta method. Note that the computed total energy would further decrease as the simulation time increases.

Example 3 (**Spherical Pendulum**).
A spherical pendulum is a mass particle connected to a frictionless, two degree-of-freedom pivot by a rigid massless link. The mass particle acts under a uniform gravitational potential. The configuration space is the two-sphere $\mathbb{S}^2 = \left\{q \in \mathbb{R}^3 \mid \|q\| = 1\right\}$. It is common to parameterize the two-sphere by two angles, but this description of the spherical pendulum has a singularity. Any trajectory near the singularity encounters numerical ill-conditioning. Furthermore, this leads to complicated expressions involving trigonometric functions.

Here we develop equations of motion for a spherical pendulum using the global structure of the two-sphere without parameterization. In the previous example, we develop equations of motion for a planar pendulum using the fact that the one-sphere $\mathbb{S}^1$ is diffeomorphic to the special orthogonal group SO(2). But, the two-sphere is not diffeomorphic to a Lie group. Instead, there exists a natural Lie group action on the two-sphere. That is the 3-dimensional special orthogonal group SO(3), a group of 3×3 orthogonal matrices with determinant 1, i.e. $\mathrm{SO}(3) = \left\{R \in \mathbb{R}^{3\times 3} \mid R^T R = I_{3\times 3}, \det[R] = 1\right\}$. The special orthogonal group SO(3) acts on the two-sphere in a transitive way; for any $q_1, q_2 \in \mathbb{S}^2$, there exists a $R \in \mathrm{SO}(3)$ such that $q_2 = Rq_1$. Thus, the discrete update for the two-sphere can be expressed in terms of a rotation matrix as (19). This is a key idea to develop a discrete equations of motion for a spherical pendulum.

Continuous equations of motion: Let $q \in \mathbb{S}^2$ be a unit vector from the pivot point to the point mass. The Lagrangian for a spherical pendulum can be written as

$$L(q, \dot{q}) = \frac{1}{2} m l^2 \dot{q} \cdot \dot{q} + mgle_3 \cdot q, \tag{28}$$

where the gravity direction is assumed to be $e_3 = [0; 0; 1] \in \mathbb{R}^3$. The mass and the length of the pendulum are denoted by $m, l \in \mathbb{R}$, respectively. The infinitesimal variation of the unit vector q can be written in terms of the vector cross product as

$$\delta q = \xi \times q, \tag{29}$$

where $\xi \in \mathbb{R}^3$ is constrained to be orthogonal to the unit vector, i.e. $\xi \cdot q = 0$. Using this expression for the infinitesimal variation, Hamilton's principle yields the following continuous equations of motion.

$$\ddot{q} + (\dot{q} \cdot \dot{q})q + \frac{g}{l}(q \times (q \times e_3)) = 0. \tag{30}$$

Since $\dot{q} = \omega \times q$ for some angular velocity $\omega \in \mathbb{R}^3$ satisfying $\omega \cdot q = 0$, this can also be equivalently written as

$$\dot{\omega} - \frac{g}{l} q \times e_3 = 0, \tag{31}$$

$$\dot{q} = \omega \times q. \tag{32}$$

These are global equations of motion for a spherical pendulum; these are much simpler than the equations expressed in term of angles, and they have no singularity.

Discrete equations of motion: We develop a variational integrator for the spherical pendulum defined on $\mathbb{S}^2$. Since the special orthogonal group acts on the two-sphere transitively, we can define the discrete update map for the unit vector as

$$q_{k+1} = F_k q_k \tag{33}$$

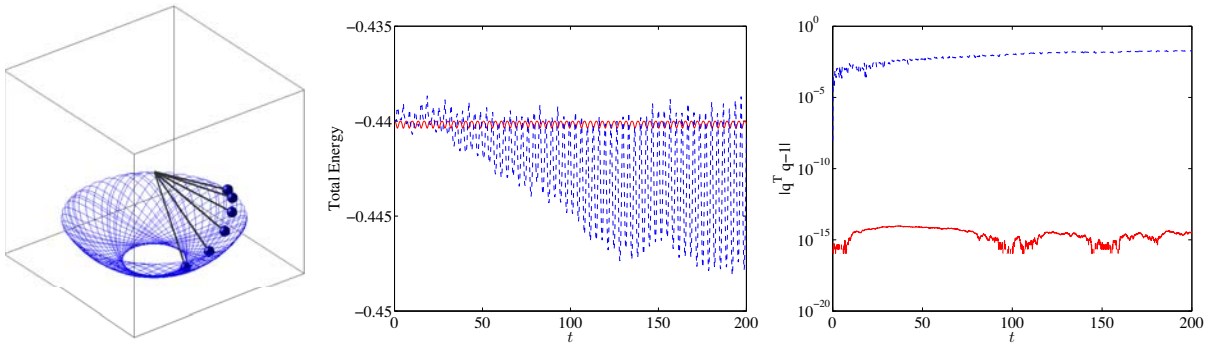
for $F_k \in \mathrm{SO}(3)$. The rotation matrix F_k is not uniquely defined by this condition, since $\exp(\lambda \hat{q}_k)$, which corresponds to a rotation about the q_k direction fixes the vector q_k. Consequently, if F_k satisfies (33), then $F_k \exp(\lambda \hat{q}_k)$ does as well. We avoid this ambiguity by requiring that F_k does not rotate about q_k, which can be achieved by letting $F_k = \exp(\hat{f}_k)$, where $f_k \cdot q_k = 0$.

Define a discrete Lagrangian $L_d : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}$ to be

$$L_d(q_k, q_{k+1}) = \frac{1}{2h} m l^2 (q_{k+1} - q_k) \cdot (q_{k+1} - q_k) + \frac{h}{2} mgle_3 \cdot q_k + \frac{h}{2} mgle_3 \cdot q_{k+1}.$$

The variation of q_k is the same as (29), namely

$$\delta q_k = \xi_k \times q_k \tag{34}$$



(a) Trajectory of pendulum (b) Total energy (c) Unit length error $|q_k \cdot q_k - 1|$

FIGURE 3. Numerical simulation of a spherical pendulum (RK45: blue, dotted, VI: red, solid)

for $\xi_k \in \mathbb{R}^3$ with a constraint $\xi_k \cdot q_k = 0$. Using this discrete Lagrangian and the expression for the variation, discrete Hamilton's principle yields the following discrete equations of motion for a spherical pendulum.

$$q_{k+1} = \left(h\omega_k + \frac{h^2 g}{2l} q_k \times e_3\right) \times q_k + \left(1 - \left\|h\omega_k + \frac{h^2 g}{2l} q_k \times e_3\right\|^2\right)^{1/2} q_k, \tag{35}$$

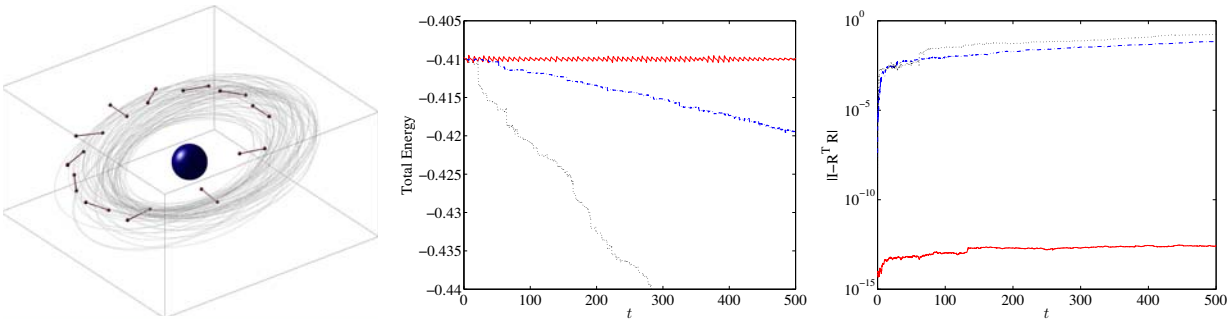
$$\omega_{k+1} = \omega_k + \frac{hg}{2l} q_k \times e_3 + \frac{hg}{2l} q_{k+1} \times e_3, \tag{36}$$

Since an explicit solution for $F_k \in \mathrm{SO}(3)$ can be obtained in this case, the rotation matrix F_k does not appear in the equations of motion. This variational integrator on $\mathbb{S}^2$ exactly preserves the unit length of q_k, the constraint $q_k \cdot \omega_k = 0$, and the third component of the angular velocity $\omega_k \cdot e_3$ which is conserved since gravity exerts no moment along the gravity direction e_3.

Numerical example: We compare the computational properties of the discrete equations of motion given by (35) and (36) with a 4(5)-th order variable step size Runge-Kutta method for (31) and (32). We choose $m = 1\,\mathrm{kg}$, $l = 9.81\,\mathrm{m}$. The initial conditions are $q_0 = [\frac{\sqrt{3}}{2}, 0, \frac{1}{2}]$, $\omega_0 = 0.1[\sqrt{3}, 0, 3]\,\mathrm{rad/sec}$, and the total energy is $E = -0.44\,\mathrm{Nm}$. The simulation time is 200 sec, and the step-size $h = 0.05$ of the discrete equations of motion is chosen such that the CPU times are identical. Figure 3 shows the computed total energy and the unit length errors. The variational integrator preserves the total energy and the structure of the two-sphere well. The mean total energy deviation is $1.5460 \times 10^{-4}\,\mathrm{Nm}$, and the mean unit length error is 3.2476×10^{-15}. But, there is a notable dissipation of the computed total energy for the Runge-Kutta method. The Runge-Kutta method also fails to preserve the structure of the two-sphere. The mean unit length error is 1.0164×10^{-2}.

Example 4 (Rigid Body in a Potential Field).
Consider a rigid body under a potential field that is dependent on the position and the attitude of the body. The configuration space is the special Euclidean group, which is a semi-direct product of the special orthogonal group and Euclidean space, i.e. $\mathrm{SE}(3) = \mathrm{SO}(3) \circledS \mathbb{R}^3$.

Continuous equations of motion: The equations of motion for a rigid body can be developed either from Hamilton's principle (see Lee et al. [26]) in a similar way as Example 2, or directly from the generalized discrete Euler–Poincaré equation given at (5). Here, we summarize the results. Let $m \in \mathbb{R}$ and $J \in \mathbb{R}^{3\times3}$ be the mass and the moment of inertia matrix of a rigid body. For $(R, x) \in \mathrm{SE}(3)$, the linear transformation from the body-fixed frame to the inertial frame is denoted by the rotation matrix $R \in \mathrm{SO}(3)$, and the position of the mass center in the inertial frame is denoted by a vector $x \in \mathbb{R}^3$. The vectors $\Omega, v \in \mathbb{R}^3$ are the angular velocity in the body-fixed frame, and the translational velocity in the inertial frame, respectively. Suppose that the rigid body acts under a configuration-dependent potential $U : \mathrm{SE}(3) \to \mathbb{R}$. The continuous



(a) Trajectory of rigid body (b) Total energy (c) Rotation matrix error $\|I - R^T R\|$

FIGURE 4. Numerical simulation of a dumbbell rigid body (LGVI: red, solid, RK45 with rotation matrices: blue, dash-dotted, RK45 with quaternions: black, dotted)

equations of motion for the rigid body can be written as

$$\dot{R} = R\hat{\Omega}, \tag{37}$$

$$\dot{x} = v, \tag{38}$$

$$J\dot{\Omega} + \Omega \times J\Omega = M, \tag{39}$$

$$m\dot{v} = -\frac{\partial U}{\partial x}, \tag{40}$$

where the hat map $\hat{\cdot}$ is an isomorphism from $\mathbb{R}^3$ to 3×3 skew-symmetric matrices $\mathfrak{so}(3)$, defined such that $\hat{x}y = x \times y$ for any $x, y \in \mathbb{R}^3$. The moment due to the potential $M \in \mathbb{R}^3$ is obtained by the following relationship.

$$\hat{M} = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}. \tag{41}$$

The matrix $\frac{\partial U}{\partial R} \in \mathbb{R}^{3\times 3}$ is defined such that $[\frac{\partial U}{\partial R}]_{ij} = \frac{\partial U}{\partial [R]_{ij}}$ for $i, j \in \{1, 2, 3\}$, where the i, j-th element of a matrix is denoted by $[\cdot]_{ij}$.

Discrete equations of motion: The corresponding discrete equations of motion are given by

$$h\widehat{J\Omega_k} + \frac{h^2}{2}\hat{M}_k = F_k J_d - J_d F_k^T, \tag{42}$$

$$R_{k+1} = R_k F_k, \tag{43}$$

$$x_{k+1} = x_k + hv_k - \frac{h^2}{2m}\frac{\partial U_k}{\partial x_k}, \tag{44}$$

$$J\Omega_{k+1} = F_k^T J\Omega_k + \frac{h}{2}F_k^T M_k + \frac{h}{2}M_{k+1}, \tag{45}$$

$$mv_{k+1} = mv_k - \frac{h}{2}\frac{\partial U_k}{\partial x_k} - \frac{h}{2}\frac{\partial U_{k+1}}{\partial x_{k+1}}, \tag{46}$$

where $J_d \in \mathbb{R}^{3\times 3}$ is a non-standard moment of inertia matrix defined as $J_d = \frac{1}{2}\mathrm{tr}[J]I_{3\times 3} - J$. For a given $(R_k, x_k, \Omega_k, v_k)$, we solve the implicit equation (42) to find $F_k \in \mathrm{SO}(3)$. Then, the configuration at the next step R_{k+1}, x_{k+1} is obtained by (43) and (44), and the moment and force $M_{k+1}, \frac{\partial U_{k+1}}{\partial x_{k+1}}$ can be computed. Velocities Ω_{k+1}, v_{k+1} are obtained from (45) and (46). This defines a discrete flow map, $(R_k, x_k, \Omega_k, v_k) \mapsto (R_{k+1}, x_{k+1}, \Omega_{k+1}, v_{k+1})$, and this process can be repeated. This Lie group variational integrator on SE(3) can be generalized to multiple rigid bodies acting under their mutual gravitational potential (see Lee et al. [26]).

Numerical example: We compare the computational properties of the discrete equations of motion given by (42)–(46) with a 4(5)-th order variable step size Runge-Kutta method for (37)–(40). In addition, we compute the attitude dynamics using quaternions on the unit three-sphere $\mathbb{S}^3$. The attitude kinematics

equation (37) is rewritten in terms of quaternions, and the corresponding equations are integrated by the same Runge-Kutta method.

We choose a dumbbell spacecraft, that is two spheres connected by a rigid massless rod, acting under a central gravitational potential. The resulting system is a restricted full two body problem. The dumbbell spacecraft model has an analytic expression for the gravitational potential, resulting in a nontrivial coupling between the attitude dynamics and the orbital dynamics.

As shown in Figure 4(a), the initial conditions are chosen such that the resulting motion is a near-circular orbit combined with a rotational motion. Figure 4(b) and Figure 4(c) show the computed total energy and the orthogonality errors of the rotation matrix. The Lie group variational integrator preserves the total energy and the Lie group structure of SO(3). The mean total energy deviation is 2.5983×10^{-4}, and the mean orthogonality error is 1.8553×10^{-13}. But, there is a notable dissipation of the computed total energy and the orthogonality error for the Runge-Kutta method. The mean orthogonality errors for the Runge-Kutta method are 0.0287 and 0.0753, respectively, using kinematics equation with rotation matrices, and using the kinematics equation with quaternions. Thus, the attitude of the rigid body is not accurately computed for Runge-Kutta methods. It is interesting to see that the Runge-Kutta method with quaternions, which is generally assumed to have better computational properties than the kinematics equation with rotation matrices, has larger total energy error and orthogonality error. Since the unit length of the quaternion vector is not preserved in the numerical computations, orthogonality errors arise when converted to a rotation matrix. This suggests that it is critical to preserve the structure of SO(3) in order to study the global characteristics of the rigid body dynamics.

The importance of simultaneously preserving the symplectic structure and the Lie group structure of the configuration space in rigid body dynamics can be observed numerically. Lie group variational integrators, which preserve both of these properties, are compared to methods that only preserve one, or neither, of these properties (see Lee et al. [27]). It is shown that the Lie group variational integrator exhibits greater numerical accuracy and efficiency. Due to these computational advantages, the Lie group variational integrator has been used to study the dynamics of the binary near-Earth asteroid 66391 (1999 KW_4) in joint work between the University of Michigan and the Jet Propulsion Laboratory, NASA (see Scheeres et al. [37]).

4. Optimal Control of Discrete Lagrangian and Hamiltonian Systems

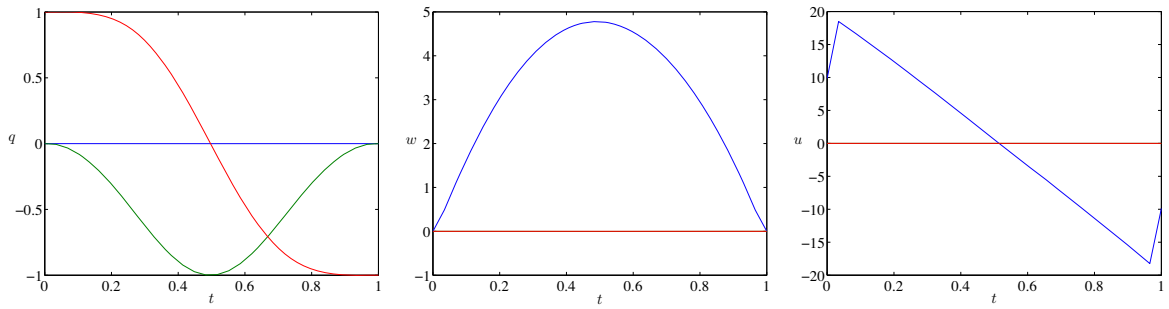
Optimal control problems involve finding a control input such that a certain optimality objective is achieved under prescribed constraints. An optimal control problem that minimizes a performance index is described by a set of differential equations, which can be derived using Pontryagin's maximum principle. The equations of motion for a system are constrained by Lagrange multipliers, and necessary conditions for optimality is obtained by the calculus of variations. The solution for the corresponding two point boundary value problem provides the optimal control input. Alternatively, a sub-optimal control law is obtained by approximating the control input history with finite data points.

Discrete optimal control problems involve finding a control input for a given system described by discrete Lagrangian and Hamiltonian mechanics. The control inputs are parameterized by their values at each discrete time step, and the discrete equations of motion are derived from the discrete Lagrange-d'Alembert principle (Kane et al. [21]),

$$\delta \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) = \sum_{k=0}^{N-1} [F_d^-(q_k, q_{k+1}) \cdot \delta q_k + F_d^+(q_k, q_{k+1}) \cdot \delta q_{k+1}],$$

which modifies the discrete Hamilton's principle by taking into account the virtual work of the external forces. Discrete optimal control is in contrast to traditional techniques such as collocation, wherein the continuous equations of motion are imposed as constraints at a set of collocation points, since this approach induces constraints on the configuration at each discrete timestep.

Any optimal control algorithm can be applied to the discrete Lagrangian or Hamiltonian system. For an indirect method, our approach to a discrete optimal control problem can be considered as a multiple stage variational problem. The discrete equations of motion are derived by the discrete variational principle. The corresponding variational integrators are imposed as dynamic constraints to be satisfied by using Lagrange multipliers, and necessary conditions for optimality, expressed as discrete equations on multipliers, are obtained from a variational principle. For a direct method, control inputs can be optimized by using



(a) Pendulum position q (b) Angular velocity ω (c) Control moment u

FIGURE 5. Optimal control of a spherical pendulum

parameter optimization tools such as a sequential quadratic programming. The discrete optimal control can be characterized by discretizing the optimal control problem from the problem formulation stage.

This method has substantial computational advantages when used to find an optimal control law. As discussed in the previous section, the discrete dynamics are more faithful to the continuous equations of motion, and consequently more accurate solutions to the optimal control problem are obtained. The external control inputs break the Lagrangian and Hamiltonian system structure. For example, the total energy is not conserved for a controlled mechanical system. But, the computational superiority of the discrete mechanics still holds for controlled systems. It has been shown that the discrete dynamics is more reliable even for controlled system as it computes the energy dissipation rate of controlled systems more accurately (see Marsden and West [31]). For example, this feature is extremely important in computing accurate optimal trajectories for long term spacecraft attitude maneuvers using low energy control inputs.

The discrete dynamics does not only provide an accurate optimal control input, but also enables us to find it efficiently. For the indirect optimal control approach, optimal solutions are usually sensitive to a small variation of multipliers. This causes difficulties, such as numerical ill-conditioning, when solving the necessary conditions for optimality expressed as a two point boundary value problem. Sensitivity derivatives along the discrete necessary conditions do not have numerical dissipation introduced by conventional numerical integration schemes. Thus, they are numerically more robust, and the necessary conditions can be solved computationally efficiently. For the direct optimal control approach, optimal control inputs can be obtained by using a larger discrete step size, which requires less computational load.

We illustrate the basic properties of the discrete optimal control using optimal control problems for the spherical pendulum and the rigid body model presented in the previous section.

Example 5 (Optimal Control of a Spherical Pendulum).
We study an optimal control problem for the spherical pendulum described in Example 3. We assume that an external control moment $u \in \mathbb{R}^3$ acts on the pendulum. Control inputs are parameterized by their values at each time step, and the discrete equations of motion are modified to include the effects of the external control inputs by using the discrete Lagrange-d'Alembert principle. Since the component of the control moment that is parallel to the direction along the pendulum has no effect, we parameterize the control input as $u_k = q_k \times w_k$ for $w_k \in \mathbb{R}^3$.

The objective is to transfer the pendulum from a given initial configuration (q_0, ω_0) to a desired configuration (q^d, ω^d) during a fixed maneuver time N, while minimizing the square of the weighted l_2 norm of the control moments.

$$\min_{w_k} \mathcal{J} = \sum_{k=0}^{N} \frac{h}{2} u_k^T u_k = \sum_{k=0}^{N} \frac{h}{2} (q_k \times w_k)^T (q_k \times w_k).$$

We solve this optimal control problem by using a direct numerical optimization method. The terminal boundary condition is imposed as an equality constraint, and the $3(N+1)$ control input parameters $\{w_k\}_{k=0}^{N}$ are numerically optimized using sequential quadratic programming. This method is referred to as a DMOC (Discrete Mechanics and Optimal Control) approach (see Junge et al. [19]).

Figure 5 shows a optimal solution transferring the spherical pendulum from a hanging configuration given by $(q_0, \omega_0) = (e_3, 0_{3\times 1})$ to an inverted configuration $(q^d, \omega^d) = (-e_3, 0_{3\times 1})$ during 1 second. The time step size is $h = 0.033$. Experiments have shown that the DMOC approach can compute optimal solutions using larger step sizes than typical Runge-Kutta methods, and consequently, it requires less computational load. In this case, using a second-order accurate Runge-Kutta method, the same optimization code fails while giving error messages of inaccurate and singular gradient computations. It is presumed that the unit length errors of the Runge-Kutta method, shown in Example 3, cause numerical instabilities for the finite-difference gradient computations required for the sequential quadratic programming algorithm.

Example 6 (**Optimal Control of a Rigid Body in a Potential Field**)**.**
We study an optimal control problem of a rigid body using a dumbbell spacecraft model described in Example 4 (see Lee et al. [25] for detail). We assume that external control forces $u^f \in \mathbb{R}^3$, and control moment $u^m \in \mathbb{R}^3$ act on the dumbbell spacecraft. Control inputs are parameterized by their values at each time step, and the Lie group variational integrators are modified to include the effects of the external control inputs by using discrete Lagrange-d'Alembert principle.

The objective is to transfer the dumbbell from a given initial configuration $(R_0, x_0, \Omega_0, v_0)$ to a desired configuration $(R^d, x^d, \Omega^d, v^d)$ during a fixed maneuver time N, while minimizing the square of the l_2 norm of the control inputs.

$$\min_{u_{k+1}} \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2}(u^f_{k+1})^T W_f u^f_{k+1} + \frac{h}{2}(u^m_{k+1})^T W_m u^m_{k+1},$$

where $W_f, W_m \in \mathbb{R}^{3\times 3}$ are symmetric positive definite matrices. Here we use a modified version of the discrete equations of motion with first order accuracy, as it yields a compact form for the necessary conditions.

Necessary conditions for optimality: We solve this optimal control problem by using an indirect optimization method, where necessary conditions for optimality are derived using variational arguments, and a solution of the corresponding two-point boundary value problem provides the optimal control. This approach is common in the optimal control literature; here the optimal control problem is discretized at the problem formulation level using the Lie group variational integrator presented in Section 3.

$$\begin{aligned}\mathcal{J}_a = \sum_{k=0}^{N-1} &\frac{h}{2}(u^f_{k+1})^T W^f u^f_{k+1} + \frac{h}{2}(u^m_{k+1})^T W^m u^m_{k+1} \\ &+ \lambda^{1,T}_k \left\{-x_{k+1} + x_k + h v_k\right\} + \lambda^{2,T}_k \left\{-m v_{k+1} + m v_k - h\frac{\partial U_{k+1}}{\partial x_{k+1}} + h u^f_{k+1}\right\} \\ &+ \lambda^{3,T}_k \left(\operatorname{logm}(F_k - R_k^T R_{k+1})\right)^\vee + \lambda^{4,T}_k \left\{-J\Omega_{k+1} + F_k^T J\Omega_k + h\left(M_{k+1} + u^m_{k+1}\right)\right\},\end{aligned}$$

where $\lambda^1_k, \lambda^2_k, \lambda^3_k, \lambda^4_k \in \mathbb{R}^3$ are Lagrange multipliers. The matrix logarithm is denoted by $\operatorname{logm} : \mathrm{SO}(3) \to \mathfrak{so}(3)$ and the vee map $\vee : \mathfrak{so}(3) \to \mathbb{R}^3$ is the inverse of the hat map introduced in Example 4. The logarithm form of (43) is used, and the constraint (42) is considered implicitly using constrained variations. Using similar expressions for the variation of the rotation matrix and the angular velocity given in (14) and (15), the infinitesimal variation can be written as

$$\delta\mathcal{J}_a = \sum_{k=1}^{N-1} h\delta u^{f,T}_k \left\{W_f u^f_k + \lambda^2_{k-1}\right\} + h\delta u^{m,T}_k \left\{W_m u^m_k + \lambda^4_{k-1}\right\} + z^T_k \left\{-\lambda_{k-1} + A^T_k \lambda_k\right\},$$

where $\lambda_k = [\lambda^1_k; \lambda^2_k; \lambda^3_k; \lambda^4_k] \in \mathbb{R}^{12}$, and $z_k \in \mathbb{R}^{12}$ represents the infinitesimal variation of $(R_k, x_k, \Omega_k, v_k)$, given by $z_k = [\operatorname{logm}(R_k^T \delta R_k)^\vee; \delta x_k, \delta\Omega_k, \delta v_k]$. The matrix $A_k \in \mathbb{R}^{12\times 12}$ is defined in terms of $(R_k, x_k, \Omega_k, v_k)$. Thus, necessary conditions for optimality are given by

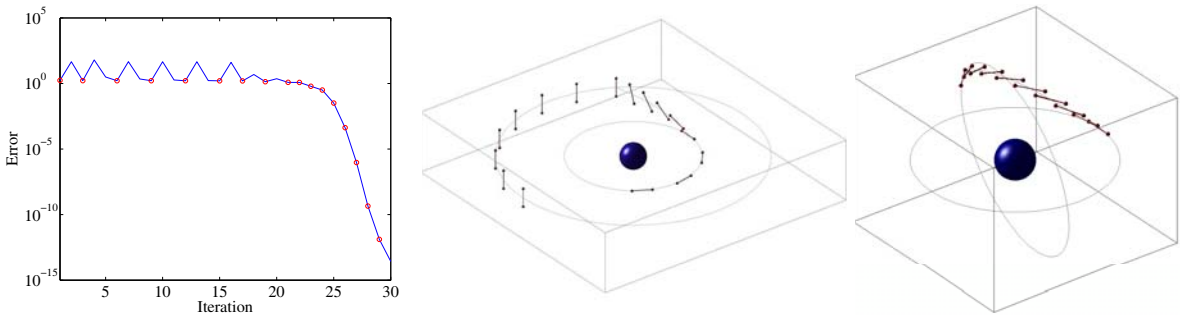
$$u^f_{k+1} = -W_f^{-1}\lambda^2_k, \tag{47}$$

$$u^m_{k+1} = -W_m^{-1}\lambda^4_k, \tag{48}$$

$$\lambda_k = A^T_{k+1}\lambda_{k+1} \tag{49}$$

together with the discrete equations of motion and the boundary conditions.

Computational approach: Necessary conditions for optimality are expressed in terms of a two point boundary problem. The problem is to find the optimal discrete flow, multipliers, and control inputs to satisfy the



(a) Convergence rate (b) Orbital radius change (c) Inclination angle change

FIGURE 6. Optimal control of a rigid body

equations of motion, optimality conditions, multiplier equations, and boundary conditions simultaneously. We use a neighboring extremal method (see Bryson and Ho [12]). A nominal solution satisfying all of the necessary conditions except the boundary conditions is chosen. The unspecified initial multiplier is updated by successive linearization so as to satisfy the specified terminal boundary conditions in the limit. This is also referred to as the shooting method. The main advantage of the neighboring extremal method is that the number of iteration variables is small.

The difficulty is that the extremal solutions are sensitive to small changes in the unspecified initial multiplier values. The nonlinearities also make it hard to construct an accurate estimate of sensitivity, thereby resulting in numerical ill-conditioning. Therefore, it is important to compute the sensitivities accurately to apply the neighboring extremal method. Here the optimality conditions (47) and (48) are substituted into the equations of motion and the multiplier equations, which are linearized to obtain sensitivity derivatives of an optimal solution with respect to boundary conditions. Using this sensitivity, an initial guess of the unspecified initial conditions is iterated to satisfy the specified terminal conditions in the limit. Any type of Newton iteration can be applied. We use a line search with backtracking algorithm, referred to as Newton-Armijo iteration (see Kelley [22]).

Figure 6 shows optimized maneuvers, where a dumbbell spacecraft on a reference circular orbit is transferred to another circular orbit with a different orbital radius and inclination angle. Figure 6(a) shows the violation of the terminal boundary condition according to the number of iterations on a logarithmic scale. Red circles denote outer iterations in the Newton-Armijo iteration to compute the sensitivity derivatives. The error in satisfaction of the terminal boundary condition converges quickly to machine precision after the solution is close to the local minimum at around the 20th iteration. These convergence results are consistent with the quadratic convergence rates expected of Newton methods with accurately computed gradients.

The neighboring extremal method, also referred to as the shooting method, is numerically efficient in the sense that the number of optimization parameters is minimized. But, in general, this approach may be prone to numerical ill-conditioning (see Betts [3]). A small change in the initial multiplier can cause highly nonlinear behavior of the terminal attitude and angular momentum. It is difficult to compute the gradient for Newton iterations accurately, and the numerical error may not converge. However, the numerical examples presented in this article show excellent numerical convergence properties. The dynamics of a rigid body arises from Hamiltonian mechanics, which have neutral stability, and its adjoint system is also neutrally stable. The proposed Lie group variational integrator and the discrete multiplier equations, obtained from variations expressed in the Lie algebra, preserve the neutral stability property numerically. Therefore the sensitivity derivatives are computed accurately.

5. Controlled Lagrangian Method for Discrete Lagrangian Systems

The method of controlled Lagrangians is a procedure for constructing feedback controllers for the stabilization of relative equilibria. It relies on choosing a parametrized family of controlled Lagrangians whose corresponding Euler–Lagrange flows are equivalent to the closed loop behavior of a Lagrangian system with external control forces. The condition that these two systems are equivalent results in matching conditions.

Since the controlled system can now be viewed as a Lagrangian system with a modified Lagrangian, the global stability of the controlled system can be determined directly using Lyapunov stability analysis.

This approach originated in Bloch et al. [5] and was then developed in Auckly et al. [1]; Bloch et al. [6, 7, 8, 9]; Hamberg [15, 16]. A similar approach for Hamiltonian controlled systems was introduced and further studied in the work of Blankenstein, Ortega, van der Schaft, Maschke, Spong, and their collaborators (see, for example, Maschke et al. [33]; Ortega et al. [35] and related references). The two methods were shown to be equivalent in Chang et al. [13] and a nonholonomic version was developed in Zenkov et al. [40, 41], and Bloch [4].

Since the method of controlled Lagrangians relies on relating the closed-loop dynamics of a controlled system with the Euler–Lagrange flow associated with a modified Lagrangian, it is natural to discretize this approach through the use of variational integrators. In Bloch et al. [10, 11], a discrete theory of controlled Lagrangians was developed for variational integrators, and applied to the feedback stabilization of the unstable inverted equilibrium of the pendulum on a cart.

The pendulum on a cart is an example of an underactuated control problem, which has two degrees of freedom, given by the pendulum angle and the cart position. Only the cart position has control forces acting on it, and the stabilization of the pendulum has to be achieved indirectly through the coupling between the pendulum and the cart. The controlled Lagrangian is obtained by modifying the kinetic energy term, a process referred to as kinetic shaping. Similarly, it is possible to modify the potential energy term using potential shaping.

Since the pendulum on a cart model involves both a planar pendulum, and a cart that translates in one-dimension, the configuration space is a cylinder, $\mathbb{S}^1 \times \mathbb{R}$.

Continuous kinetic shaping: The Lagrangian has the form kinetic minus potential energy

$$L(q,\dot{q}) = \tfrac{1}{2}\left[\alpha\dot{\theta}^2 + 2\beta(\theta)\dot{\theta}\dot{s} + \gamma\dot{s}^2\right] - V(q), \tag{50}$$

and the corresponding controlled Euler–Lagrange dynamics is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = 0, \tag{51}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{s}} = u, \tag{52}$$

where u is the control input.

Since the potential energy is translation invariant, *i.e.,* $V(q) = V(\theta)$, the *relative equilibria* $\theta = \theta_e$, $\dot{s} = \text{const}$ are unstable and given by non-degenerate critical points of $V(\theta)$. To stabilize the relative equilibria $\theta = \theta_e$, $\dot{s} = \text{const}$ with respect to θ, kinetic shaping is used. The controlled Lagrangian in this case is defined by

$$L^{\tau,\sigma}(q,\dot{q}) = L(\theta,\dot{\theta},\dot{s} + \tau(\theta)\dot{\theta}) + \tfrac{1}{2}\sigma\gamma(\tau(\theta)\dot{\theta})^2, \tag{53}$$

where $\tau(\theta) = \kappa\beta(\theta)$. This velocity shift corresponds to a new choice of the horizontal space (see Bloch et al. [8] for details). The dynamics is just the Euler–Lagrange dynamics for controlled Lagrangian (53),

$$\frac{d}{dt}\frac{\partial L^{\tau,\sigma}}{\partial \dot{\theta}} - \frac{\partial L^{\tau,\sigma}}{\partial \theta} = 0, \tag{54}$$

$$\frac{d}{dt}\frac{\partial L^{\tau,\sigma}}{\partial \dot{s}} = 0. \tag{55}$$

The Lagrangian (53) satisfies the simplified matching conditions of Bloch et al. [9] when the kinetic energy metric coefficient γ in (50) is constant.

Setting $u = -d\big(\gamma\tau(\theta)\dot{\theta}\big)/dt$ defines the control input, makes equations (52) and (55) identical, and results in controlled momentum conservation by dynamics (51) and (52). Setting $\sigma = -1/\gamma\kappa$ makes equations (51) and (54) identical when restricted to a level set of the controlled momentum.

Discrete kinetic shaping: Here, we adopt the following notation:

$$q_{k+1/2} = \frac{q_k + q_{k+1}}{2}, \quad \Delta q_k = q_{k+1} - q_k, \quad q_k = (\theta_k, s_k).$$

Then, a second-order accurate discrete Lagrangian is given by,

$$L_d(q_k, q_{k+1}) = hL(q_{k+1/2}, \Delta q_k/h).$$

The discrete dynamics is governed by the equations

$$\frac{\partial L_d(q_k, q_{k+1})}{\partial \theta_k} + \frac{\partial L_d(q_{k-1}, q_k)}{\partial \theta_k} = 0, \tag{56}$$

$$\frac{\partial L_d(q_k, q_{k+1})}{\partial s_k} + \frac{\partial L_d(q_{k-1}, q_k)}{\partial s_k} = -u_k, \tag{57}$$

where u_k is the control input. Similarly, the discrete controlled Lagrangian is,

$$L_d^{\tau,\sigma}(q_k, q_{k+1}) = hL^{\tau,\sigma}(q_{k+1/2}, \Delta q_k/h),$$

with discrete dynamics given by,

$$\frac{\partial L_d^{\tau,\sigma}(q_k, q_{k+1})}{\partial \theta_k} + \frac{\partial L_d^{\tau,\sigma}(q_{k-1}, q_k)}{\partial \theta_k} = 0, \tag{58}$$

$$\frac{\partial L_d^{\tau,\sigma}(q_k, q_{k+1})}{\partial s_k} + \frac{\partial L_d^{\tau,\sigma}(q_{k-1}, q_k)}{\partial s_k} = 0. \tag{59}$$

Equation (59) is equivalent to the *discrete controlled momentum conservation:*

$$p_k = \mu,$$

where

$$p_k = -\frac{\partial L_d^{\tau,\sigma}(q_k, q_{k+1})}{\partial s_k} = \frac{(1+\gamma\kappa)\beta(\theta_{k+1/2})\Delta\theta_k + \gamma\Delta s_k}{h}.$$

Setting

$$u_k = -\frac{\gamma\Delta\theta_k\tau(\theta_{k+1/2}) - \gamma\Delta\theta_{k-1}\tau(\theta_{k-1/2})}{h}$$

makes equations (57) and (59) identical and allows one to represent the discrete momentum equation (57) as the discrete momentum conservation law $p_k = p$.

The condition that (56-57) are equivalent to (58-59) yield the discrete matching conditions. The dynamics determined by equations (56-57) restricted to the momentum level $p_k = p$ is equivalent to the dynamics of equations (58-59) restricted to the momentum level $p_k = \mu$ if and only if the matching conditions

$$\sigma = -\frac{1}{\gamma\kappa}, \qquad \mu = \frac{p}{1+\gamma\kappa},$$

hold.

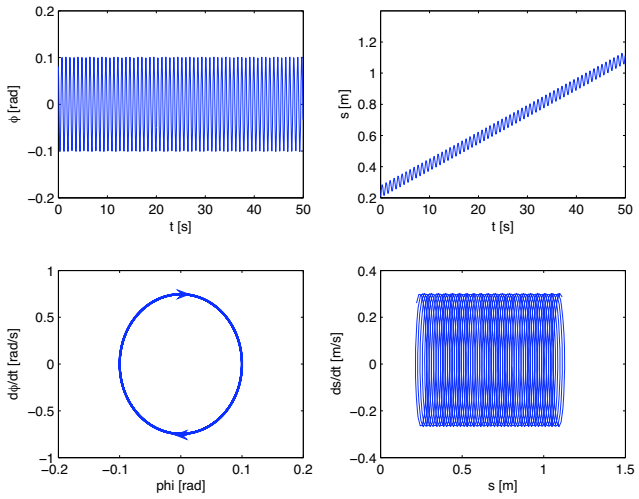


FIGURE 7. Discrete controlled dynamics with kinetic shaping and without dissipation. The discrete controlled system stabilizes the θ motion about the equilibrium, but the s dynamics is not stabilized; since there is no dissipation, the oscillations are sustained.

Numerical example: Simulating the behavior of the discrete controlled Lagrangian system involves viewing equations (58-59) as an implict update map $(q_{k-2}, q_{k-1}) \mapsto (q_{k-1}, q_k)$. This presupposes that the initial conditions are given in the form (q_0, q_1); however it is generally preferable to specify the initial conditions as $(q_0, \dot{q}_0)$. This is achieved by solving the boundary condition,

$$\frac{\partial L}{\partial \dot{q}}(q_0, \dot{q}_0) + D_1 L_d(q_0, q_1) + F_d^-(q_0, q_1) = 0,$$

for q_1. Once the initial conditions are expressed in the form (q_0, q_1), the discrete evolution can be obtained using the implicit update map.

We first consider the case of kinetic shaping on a level surface, when κ is twice the critical value, and without dissipation. Here, $h = 0.05\,\text{sec}$, $m = 0.14\,\text{kg}$, $M = 0.44\,\text{kg}$, and $l = 0.215\,\text{m}$. As shown in Figure 7, the θ dynamics is stabilized, but since there is no dissipation, the oscillations are sustained. The s dynamics exhibits both a drift and oscillations, as potential shaping is necessary to stabilize the translational dynamics.

6. Future Directions

Discrete Receding Horizon Optimal Control: The existing work on discrete optimal control has been primarily focused on constructing the optimal trajectory in an open loop sense. In practice, model uncertainty and actuation errors necessitate the use of feedback control, and it would be interesting to extend the existing work on optimal control of discrete systems to the feedback setting by adopting a receding horizon approach.

Discrete State Estimation: In feedback control, one typically assumes complete knowledge regarding the state of the system, an assumption that is often unrealistic in practice. The general problem of state estimation in the context of discrete mechanics would rely on good numerical methods for quantifying the propagation of uncertainty by solving the Liouville equation, which describes the evolution of a phase space distribution function advected by a prescribed vector field. In the setting of Hamiltonian systems, the solution of the Liouville equation can be solved by the method of characteristics (Scheeres et al. [38]). This implies that a collocational approach (Xiu [39]) combined with Lie group variational integrators, and interpolation based on noncommutative harmonic analysis on Lie groups could yield an efficient means of propagating uncertainty, and serve as the basis of a discrete state estimation algorithm.

Forced Symplectic-Energy-Momentum Variational Integrators: One of the motivations for studying the control of Lagrangian systems using the method of controlled Lagrangians is that the method provides a natural candidate Lyapunov function to study the global stability properties of the controlled system. In the discrete theory, this approach is complicated by the fact that the energy of a discrete Lagrangian system is not exactly conserved, but rather oscillates in a bounded fashion.

This can be addressed by considering the symplectic-energy-momentum (Kane et al. [20]) analogue of the discrete Lagrange-d'Alembert principle,

$$\delta \sum_{k=0}^{N_1} L_d(q_k, q_{k+1}, h_k) = \sum_{k=0}^{N-1} [F_d^-(q_k, q_{k+1}, h_k) \cdot \delta q_k + F_d^+(q_k, q_{k+1}, h_k) \cdot \delta q_{k+1}],$$

where the timestep h_k is allowed to vary, and is chosen to satisfy the variational principle. The variations in h_k yield an Euler–Lagrange equation that reduces to the conservation of discrete energy in the absence of external forces. By developing a theory of controlled Lagrangians around a geometric integrator based on the symplectic-energy-momentum version of the Lagrange-d'Alembert principle, one would potentially be able to use Lyapunov techniques to study the global stability of the resulting numerical control algorithms.

Primary Literature

[1] D. Auckly, L. Kapitanski, and W. White. Control of nonlinear underactuated systems. *Commun. Pure Appl. Math.*, 53:354–369, 2000.

[2] G. Benettin and A. Giorgilli. On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms. *J. Stat. Phys.*, 74:1117–1143, 1994.

[3] J. T. Betts. *Practical Methods for Optimal Control Using Nonlinear Programming*. SIAM, 2001.

[4] A. M. Bloch. *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Appl. Math.* Springer-Verlag, 2003.

[5] A. M. Bloch, N. Leonard, and J. E. Marsden. Matching and stabilization using controlled Lagrangians. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2356–2361, 1997.

[6] A. M. Bloch, N. Leonard, and J. E. Marsden. Matching and stabilization by the method of controlled Lagrangians. In *Proceedings of the IEEE Conference on Decision and Control*, pages 1446–1451, 1998.
[7] A. M. Bloch, N. Leonard, and J. E. Marsden. Potential shaping and the method of controlled Lagrangians. In *Proceedings of the IEEE Conference on Decision and Control*, pages 1652–1657, 1999.
[8] A. M. Bloch, N. E. Leonard, and J. E. Marsden. Controlled Lagrangians and the stabilization of mechanical systems I: The first matching theorem. *IEEE Trans. on Systems and Control*, 45:2253–2270, 2000.
[9] A. M. Bloch, D. E. Chang, N. E. Leonard, and J. E. Marsden. Controlled Lagrangians and the stabilization of mechanical systems II: Potential shaping. *IEEE Trans. on Autom. Contr.*, 46:1556–1571, 2001.
[10] A. M. Bloch, M. Leok, J. E. Marsden, and D. V. Zenkov. Controlled Lagrangians and stabilization of the discrete cart-pendulum system. In *Proceedings of the IEEE Conference on Decision and Control*, pages 6579–6584, 2005.
[11] A. M. Bloch, M. Leok, J. E. Marsden, and D. V. Zenkov. Controlled Lagrangians and potential shaping for stabilization of discrete mechanical systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 3333–3338, 2006.
[12] A. E. Bryson and Y. Ho. *Applied Optimal Control*. Hemisphere Publishing Corporation, 1975.
[13] D-E. Chang, A. M. Bloch, N. E. Leonard, J. E. Marsden, and C. Woolsey. The equivalence of controlled Lagrangian and controlled Hamiltonian systems. *Control and the Calculus of Variations (special issue dedicated to J. L. Lions)*, 8:393–422, 2002.
[14] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, second edition, 2006.
[15] J. Hamberg. General matching conditions in the theory of controlled Lagrangians. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2519–2523, 1999.
[16] J. Hamberg. Controlled Lagrangians, symmetries and conditions for strong matching. In *Lagrangian and Hamiltonian Methods for Nonlinear Control*. Elsevier, 2000.
[17] I. I. Hussein, M. Leok, A. K. Sanyal, and A.M. Bloch. A discrete variational integrator for optimal control problems in $SO(3)$. In *Proceedings of the IEEE Conference on Decision and Control*, pages 6636–6641, 2006.
[18] A. Iserles, H. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. In *Acta Numerica*, volume 9, pages 215–365. Cambridge University Press, 2000.
[19] O. Junge, J. E. Marsden, and S. Ober-Blöbaum. Discrete mechanics and optimal control. In *IFAC Congress*, Praha, 2005.
[20] C. Kane, J. E. Marsden, and M. Ortiz. Symplectic-energy-momentum preserving variational integrators. *J. Math. Phys.*, 40(7):3353–3371, 1999.
[21] C. Kane, J. E. Marsden, M. Ortiz, and M. West. Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *Int. J. Numer. Meth. Eng.*, 49(10):1295–1325, 2000.
[22] C. T. Kelley. *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.
[23] T. Lee, M. Leok, and N. H. McClamroch. Attitude maneuvers of a rigid spacecraft in a circular orbit. In *Proceedings of the American Control Conference*, pages 1742–1747, 2005.
[24] T. Lee, M. Leok, and N. H. McClamroch. A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum. In *Proceedings of the IEEE Conference on Control Applications*, pages 962–967, 2005.
[25] T. Lee, M. Leok, and N. H. McClamroch. Optimal control of a rigid body using geometrically exact computations on SE(3). In *Proceedings of the IEEE Conference on Decision and Control*, pages 2710–2715, 2006.
[26] T. Lee, M. Leok, and N. H. McClamroch. Lie group variational integrators for the full body problem. *Computer Methods in Applied Mechanics and Engineering*, 196:2907–2924, May 2007.
[27] T. Lee, M. Leok, and N. H. McClamroch. Lie group variational integrators for the full body problem in orbital mechanics. *Celestial Mechanics and Dynamical Astronomy*, 98(2):121–144, June 2007.
[28] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 2004.
[29] M. Leok. *Foundations of Computational Geometric Mechanics*. PhD thesis, California Instittute of Technology, 2004.

[30] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, second edition, 1999.
[31] J. E. Marsden and M. West. Discrete mechanics and variational integrators. In *Acta Numerica*, volume 10, pages 317–514. Cambridge University Press, 2001.
[32] J. E. Marsden, S. Pekarsky, and S. Shkoller. Discrete Euler–Poincaré and Lie–Poisson equations. *Nonlinearity*, 12(6):1647–1662, 1999.
[33] B. Maschke, R. Ortega, and A. van der Schaft. Energy-based Lyapunov functions for forced Hamiltonian systems with dissipation. *IEEE Trans. on Autom. Contr.*, 45:1498–1502, 2001.
[34] J. Moser and A. P. Veselov. Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Communications in Mathematical Physics*, 139:217–243, 1991.
[35] R. Ortega, M. W. Spong, F. Gómez-Estern, and G. Blankenstein. Stabilization of a class of underactuated mechanical systems via interconnection and damping assignment. *IEEE Trans. on Autom. Contr.*, 47:1218–1233, 2002.
[36] J. M. Sanz-Serna. Symplectic integrators for Hamiltonian problems: an overview. In *Acta Numerica*, volume 1, pages 243–286. Cambridge University Press, 1992.
[37] D. J. Scheeres, E. G. Fahnestock, S. J. Ostro, J. L. Margot, L. A. M. Benner, S. B. Broschart, J. Bellerose, J. D. Giorgini, M. C. Nolan, C. Magri, P. Pravec, P. Scheirich, R. Rose, R. F. Jurgens, E. M. De Jong, and S. Suzuki. Dynamical configuration of binary near-Earth asteroid (66391) 1999 KW4. *Science*, 314: 1280–1283, 2006.
[38] D. J. Scheeres, F.-Y. Hsiao, R. S. Park, B. F. Villac, and J. M. Maruskin. Fundamental limits on spacecraft orbit uncertainty and distribution propagation. *Journal of the Astronautical Sciences*, 54: 505–523, 2006.
[39] D. Xiu. Efficient collocational approach for parametric uncertainty analysis. *Comm. Comput. Phys.*, 2: 293–309, 2007.
[40] D. V. Zenkov, A. M. Bloch, N. E. Leonard, and J. E. Marsden. Matching and stabilization of low-dimensional nonholonomic systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 1289–1295, 2000.
[41] D. V. Zenkov, A. M. Bloch, N. E. Leonard, and J. E. Marsden. Flat nonholonomic matching. In *Proceedings of the American Control Conference*, pages 2812–2817, 2002.

Books and Reviews

[1] A. M. Bloch. *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Appl. Math.* Springer-Verlag, 2003.
[2] F. Bullo and A. D. Lewis. *Geometric control of mechanical systems*, volume 49 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 2005.
[3] E. Hairer, C. Lubich and G. Wanner. *Geometric Numerical Integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, second edition, 2006.
[4] A. Iserles, H. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. In *Acta Numerica*, volume 9, pages 215–365. Cambridge University Press, 2000.
[5] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 2004.
[6] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, second edition, 1999.
[7] J. E. Marsden and M. West. Discrete mechanics and variational integrators. In *Acta Numerica*, volume 10, pages 317–514. Cambridge University Press, 2001.
[8] J. M. Sanz-Serna. Symplectic integrators for Hamiltonian problems: an overview. In *Acta Numerica*, volume 1, pages 243–286. Cambridge University Press, 1992.

Department of Aerospace Engineering, University of Michigan, Ann Arbor, Michigan, USA

Department of Mathematics, Purdue University, West Lafayette, Indiana, USA

Department of Aerospace Engineering, University of Michigan, Ann Arbor, Michigan, USA

This paper was awarded in the I International Competition (1992/93) "First Step to Nobel Prize in Physics" and published in the competition proceedings (*Acta Phys. Pol. A* **85** Supplement, S-27 (1994)).
The paper is reproduced here due to kind agreement of the Editorial Board of "Acta Physica Polonica A".

ESTIMATING THE ATTRACTOR DIMENSION OF THE EQUATORIAL WEATHER SYSTEM

M. Leok B.T.

Blk 106 Spottiswoode Park Road, #05-140, Singapore 0208, Singapore

The correlation dimension and limit capacity serve theoretically as lower and upper bounds, respectively, of the fractal dimension of attractors of dynamic systems. In this paper we show that estimates of the correlation dimension grow rapidly with increasing noise level in the time-series, while estimates of the limit capacity remain relatively unaffected. It is therefore proposed that the limit capacity can be used in studies of noisy data, despite its heavier computational requirements. An analysis of Singapore wind data with the limit capacity estimate revealed a surprisingly low dimension (≈ 2.5). It is suggested that further studies can be made with comprehensive equatorial weather data.
PACS numbers: 47.52.+j, 47.53.+n, 92.60.Wc

1. Introduction

Estimations of the fractal dimension of meteorological time-series play a fundamental role in the development of dynamic models of meteorological phenomena. These dimensional estimates provide bounds for the number of independent variables necessary to model the system, and assist one in determining the appropriateness of a model.

There have been numerous efforts to determine the dimension of weather and climatic attractors [1–3], however the equatorial weather system has yet to be subjected to a thorough analysis. Previous dimensional estimates have been based upon the use of the correlation dimension, but the reliability of this measure is questionable as it is unstable when perturbed by noise in the observational time-series.

In this study, the stability of the estimates of the limit capacity and the correlation dimension is investigated by considering the Lorenz attractor perturbed by various levels of noise.

The limit capacity dimension of a 22-year equatorial wind time-series, with a sampling frequency of once a day is then estimated.

2. Basic concepts

The phase portrait of a dynamical system can be reconstructed from the observation of a single variable by the method of delays as proposed by Takens [4]. Takens proved that almost all smooth d-dimensional manifolds can be embedded in a $(m = 2d + 1)$-dimensional space while preserving geometrical invariants. The observational time series x_1,x_2,x_3, ...,x_N is represented by the set of vectors $\vec{x}_t = \{x_t, x_{t+\tau}, \ldots, x_{t+(m-1)\tau}\}$ (where $t = 1, 2, \ldots, \ N - (m-1)\tau$) in the reconstructed phase space.

The delay τ is usually chosen for which the autocorrelation function,

$$C(\tau) = \frac{1}{N} \sum_{i=1}^{N-\tau} (x_i - \overline{x})(x_{i+\tau} - \overline{x}) \tag{1}$$

(where $\overline{x}$ is the arithmetic mean) is close to zero, thus minimizing the statistical dependence among the coordinates of the vectors.

In practice, one does not know *a priori* the dimension of the dynamical system, and the embedding dimension which is necessary for the phase space reconstruction. Thus, the dimensional estimate is computed for increasing embedding dimensions until the dimensional estimate stabilizes.

2.1. Correlation dimension

The correlation dimension theoretically provides a lower bound to the fractal dimension of an attractor, thus it satisfies the following inequality:

$$d_{\mathrm{c}} \leq d_{\mathrm{f}}. \tag{2}$$

The correlation dimension is obtained by considering the cumulative correlation function, defined by Grassberger and Procaccia [5] as

$$C(r) = \lim_{n\to\infty} \frac{1}{N^2} \sum_{\substack{k,j=1 \\ k\neq j}} \theta(r - \| \vec{x}_k - \vec{x}_j \|), \tag{3}$$

where θ is the Heaviside function, such that $\theta(x) = 0$ if $x \leq 0$ and $\theta(x) = 1$ if $x > 0$. The Euclidean norm utilized in (3) is defined as

$$\| \vec{x}_k - \vec{x}_j \| = \sqrt{\sum_{i=1}^{m} (x_{k+(i-1)\tau} - x_{j-(i-1)\tau})^2}. \tag{4}$$

The cumulative correlation function is related to the correlation dimension by the power law

$$C(r) \sim r^{d_{\mathrm{c}}} \qquad (r \to 0). \tag{5}$$

This results in the correlation dimension d_{c} being defined as

$$d_{\mathrm{c}} = \lim_{r\to 0} \frac{\ln[C(r)]}{\ln[r]}. \tag{6}$$

2.2. Limit capacity

As opposed to the correlation dimension, the limit capacity, d_l serves as an upper bound to the fractal dimension of an attractor. Considering this property, and the inequality (2), we obtain

$$d_{\mathrm{c}} \leq d_{\mathrm{f}} \leq d_l. \tag{7}$$

Let $N(\varepsilon)$ be *the minimum number of spheres of radius ε necessary to cover all points of the attractor*. Takens [4] defined the limit capacity d_l as follows:

$$N(\varepsilon) \sim \varepsilon^{-d_l} \qquad (\varepsilon \to 0). \tag{8}$$

From the relation (8), we obtain the limit capacity d_l as

$$d_l = \lim_{\varepsilon\to 0} \frac{\ln[N(\varepsilon)]}{\ln[1/\varepsilon]}. \tag{9}$$

2.3. Methods of estimating dimensions

The reconstruction of an attractor by the method of delays as described in Sec. 2 is performed on the observational time-series. The embedding dimension m which is used in the attractor reconstruction begins from the value 2 and proceeds to successively higher dimensions until the dimensional estimate stabilizes. For the correlation dimension, $C(r)$ vs. r is plotted on a log–log graph, and the gradient of the region of the graph which exhibits scaling behaviour yields an estimate of the correlation dimension. Similarly, the limit capacity is obtained by plotting $N(\varepsilon)$ vs. $1/\varepsilon$ on a log–log graph, with the gradient providing an estimate of the limit capacity.

3. Effect of noise on dimensional estimates

We consider the effect of noise on the correlation dimension and limit capacity estimates by computing the dimensional estimates of the discretely sampled time-series obtained from the x component of the Lorenz attractor, which is described by the following set of ordinary differential equations:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma(y - x), \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = rx - y - xz, \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = xy - bz \tag{10}$$

with the following parameters: $\sigma = 10$, $r = 28$, $b = 8/3$. The Lorenz attractor is shown in Fig. 1.

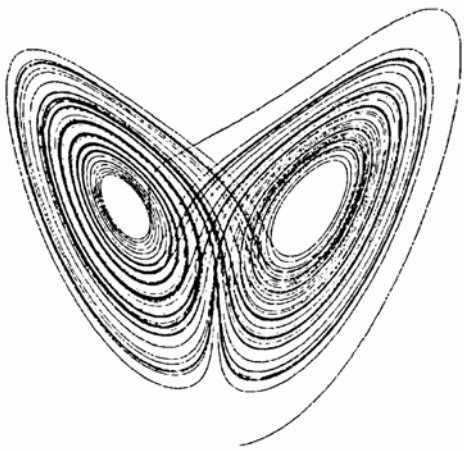


Fig. 1. Lorenz attractor.

The Lorenz equations were numerically integrated with a fourth-order Runge–Kutta method [6], with $\Delta t = 0.005$ and $\tau = 0.6$ (which was the time lag at which the autocorrelation function for the discrete time-series of the Lorenz attractor vanished). A discrete time-series with 2048 data points of the x component was then generated and subjected to dimensional estimation. Each data point of the time-series was perturbed by the introduction of noise with a maximum magnitude which is 1% and 10% of the maximum magnitude of the x component. The noise was generated by a pseudo-random multiplicative linear congruential generator proposed by Park and Miller in [7].

As can be seen from Fig. 2, the dimensional estimates of the Lorenz attractor without noise converge to a finite value with 2.06 for the correlation dimension and 2.31 for the limit capacity at an embedding dimension $m = 2d + 1$, thus satisfying inequality (7). The attractor with 1% noise as shown in Fig. 3, also behaves similarly with a correlation dimension of 2.183 and a limit capacity of 2.342. However, when the time-series is perturbed by 10% noise as shown in Fig. 4, inequality (7) is *not* satisfied, as the correlation dimension of 4.678 is *greater* than the limit capacity of 2.332.

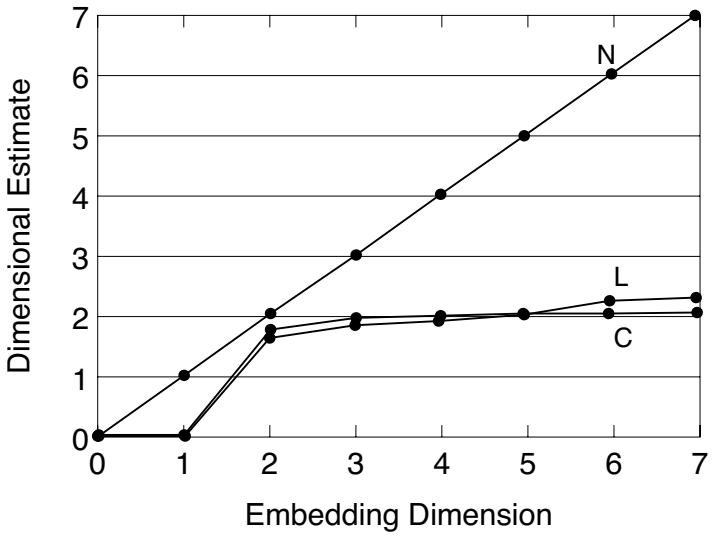


Fig. 2. Dimension estimate of Lorenz attractor: N — Gaussian noise, C — correlation dimension, L — limit capacity.

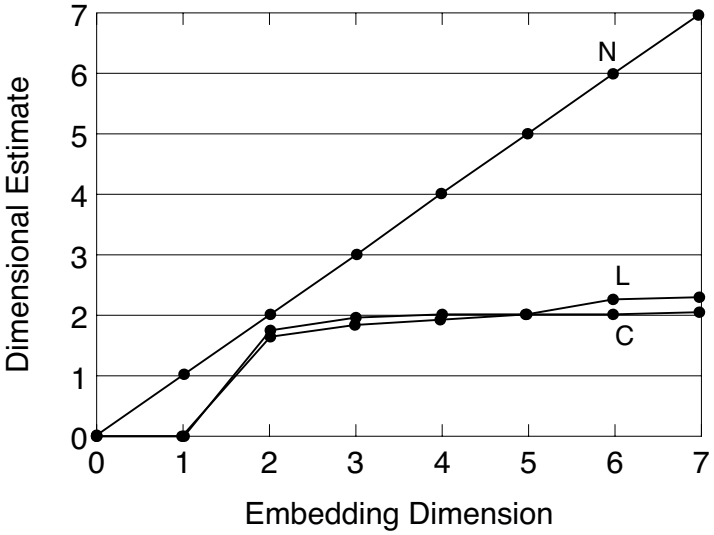


Fig. 3. Dimensional estimate of Lorenz attractor with 1% noise: N — Gaussian noise, C — correlation dimension, L — limit capacity.

3.1. Analysis of dimensional estimates

The correlation dimension is extremely sensitive to noise, as illustrated by the rate of divergence of the correlation dimension from the dimensional estimate of the noiseless attractor upon the addition of noise. This is significant, since the increase in the correlation dimension due to noise would potentially render the lower bound to the fractal dimension provided by the correlation dimension invalid, as in the case of the Lorenz attractor with 10% noise. Thus it would seem that the correlation dimension would tend to measure the dimension of the noise as opposed to the underlying dynamics which are of interest. A possible explanation for this is that for a noisy attractor, the form of the power law relation (5) is

$$\begin{aligned} C_m(r) &\sim r^m, \qquad r < r_{\text{noise}}, \\ C_m(r) &\sim r^{d_c}, \qquad r > r_{\text{noise}}. \end{aligned} \tag{11}$$

Given that the scaling region for the correlation dimension occurs in the region ($r \to 0$), it follows that the dimensional estimate would tend to be higher than that inherent in the dynamics of the system.

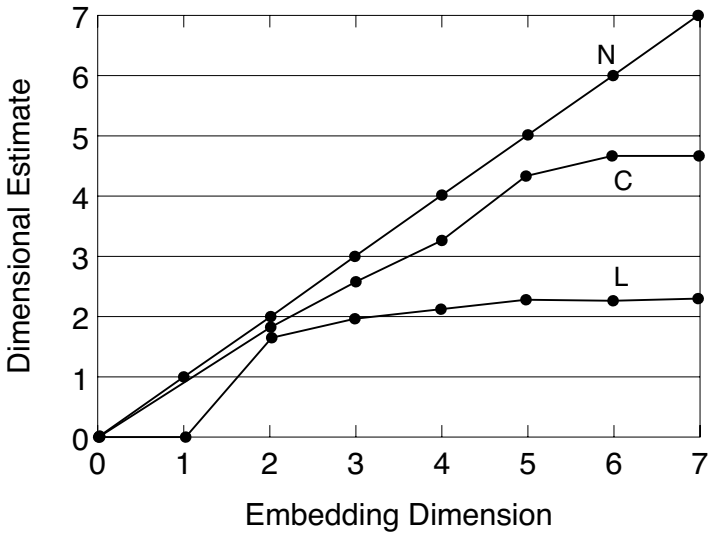


Fig. 4. Dimensional estimate of Lorenz attractor with 10% noise: N — Gaussian noise, C — correlation dimension, L — limit capacity.

The analogous situation does not occur in the case of the limit capacity as shown from the small deviation from the dimensional estimate obtained for the noiseless attractor. An intuitive explanation is as follows. In the region of small r, the limit capacity estimate suffers from the effect of saturation due to high embedding dimensions and a finite time-series, and hence it is generally necessary to obtain the limit capacity estimate at a moderate r region, where the effect of noise is less dominant.

4. Dimensional estimate of wind attractor

The time-series analysed are the x and y components of the daily wind velocity observations over Singapore. Missing data were linearly interpolated. A Fourier transform was performed upon each of the two time-series.

As shown in Fig. 5 and Fig. 6, there is a conspicuous peak in the yearly cycle range. The long term cycles are of such high magnitude that they have to be filtered off to allow for the analysis of the dynamics of the weather for periods of less than 140 days.

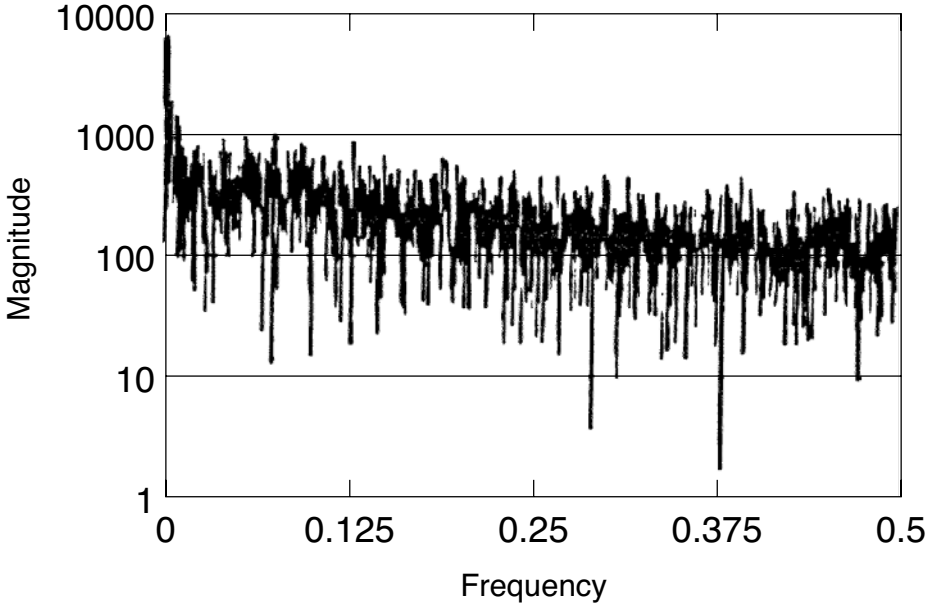


Fig. 5. FFT of X-winds.

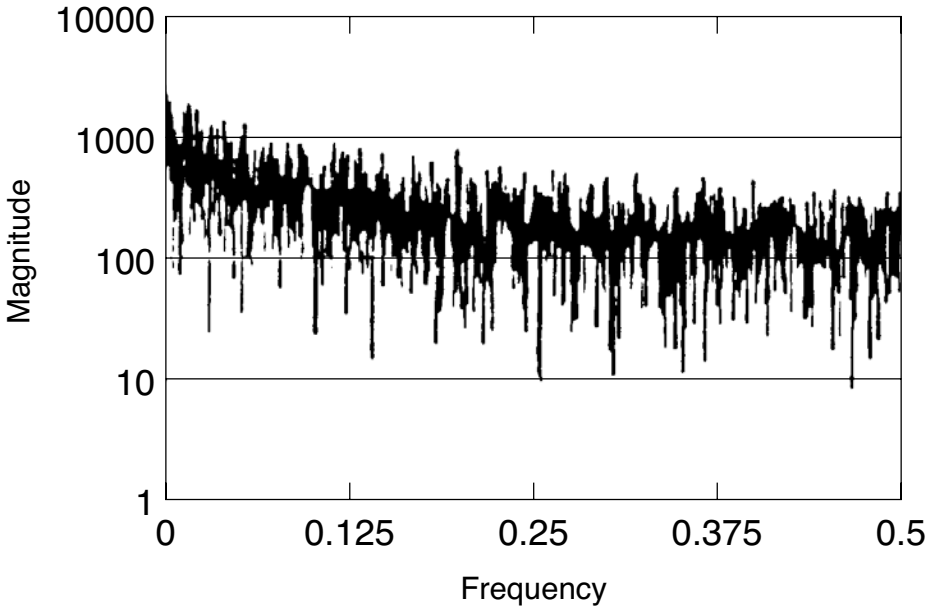


Fig. 6. FFT of Y-winds.

The autocorrelation function is plotted for this filtered data and is shown in Fig. 7 and Fig. 8. The value of τ for which the autocorrelation function vanishes is 4 days for the *X-winds* time-series and 8 days for the *Y-winds* time-series.

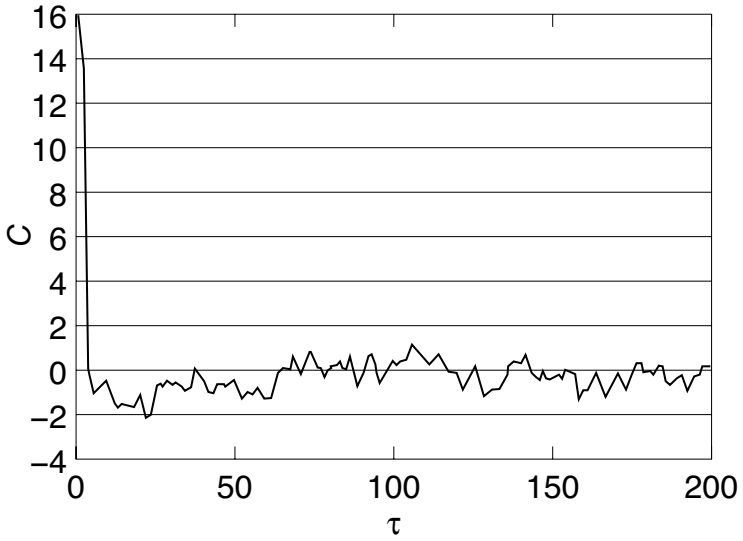


Fig. 7. Autocorrelation of X-winds.

The limit capacity dimensional estimate was then performed upon the two time-series for embedding dimensions from 2 to 10. The dimensional estimate was obtained by plotting the points $(1/\varepsilon, N(\varepsilon))$ on a log–log graph. The region satisfying the power law relation would be represented by a straight line on the log–log graph. A least-squares fit is then performed to provide the limit capacity dimensional estimate. As seen from Fig. 9, the estimated dimension deviates rapidly from the linear behaviour which would have been expected if the series consisted of Gaussian noise, thus implying that the dynamics are of a deterministic and finite-dimensional nature. It is however essential to realize that the choice of range for the scaling region from which the limit capacity dimensional estimate was derived is somewhat subjective, and a different range could result in a slightly differing dimensional estimate.

As seen from Table, the limit capacity converges to a fractional dimension between two and three, 2.508 for the *X-winds* time-series, and 2.313 for the *Y-winds* time-series.

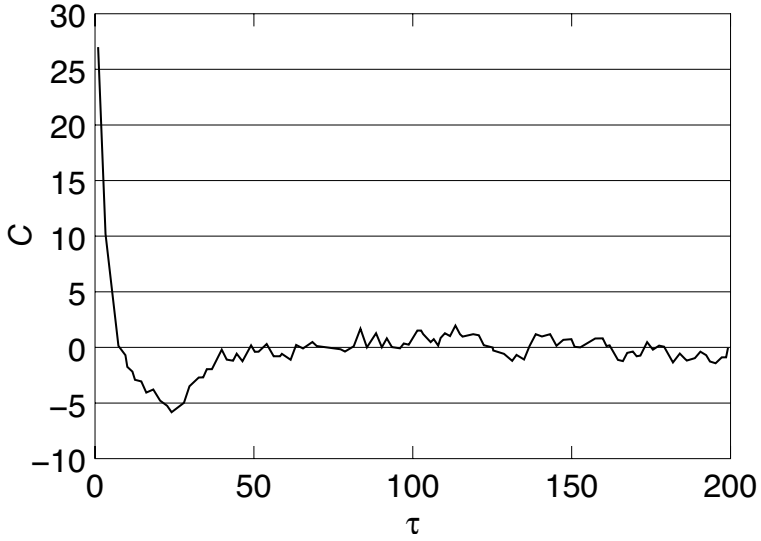


Fig. 8. Autocorrelation of Y-winds.

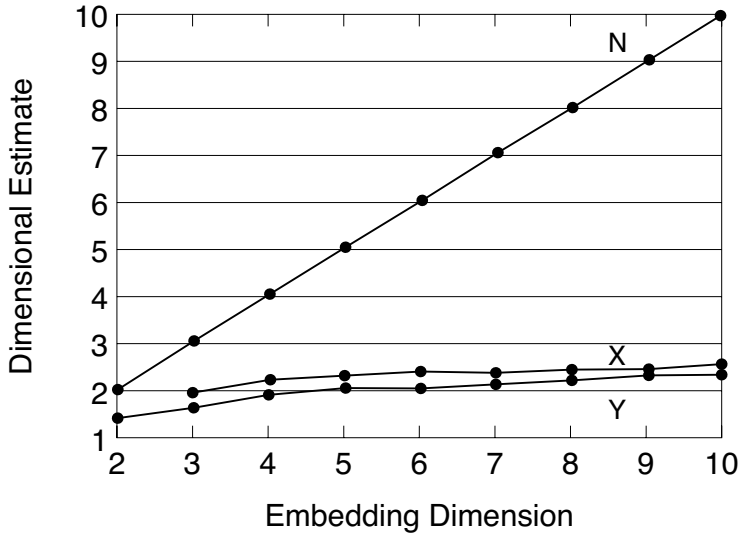


Fig. 9. Limit capacity estimate of wind data.

TABLE. Limit capacity estimate of wind data time-series.

Embedding	Gaussian	X-winds	Y-winds
2	2	1.537	1.392
3	3	1.939	1.634
4	4	2.190	1.901
5	5	2.271	2.017
6	6	2.350	2.025
7	7	2.343	2.109
8	8	2.402	2.193
9	9	2.430	2.294
10	10	2.508	2.313

This fractional dimension is indicative of the *sensitive dependence on initial conditions* which characterizes the dynamics of the weather system.

The correlation dimension estimates were also computed, and they exhibited behaviour similar to the dimensional estimates for the Lorenz attractor with 10% noise, and *did not* satisfy inequality (7).

5. Conclusion

The dimensional estimates obtained from the correlation dimension should be treated with caution as the dimensional estimates are not robust when the time-series is perturbed by noise. The limit capacity in comparison is less affected by noise and is thus recommended for dimensional estimates of noisy dynamical systems.

The low fractal dimensionality obtained for the equatorial weather attractor suggests that the inherent dynamics could be modelled by a low-dimensional deterministic system. The dimensional estimate for the Singapore wind data is lower than that reported for temperate systems. This is surprising as the equatorial weather systems are usually less organized than the temperate weather systems. A comprehensive study with regional data on a range of weather parameters may provide a deeper insight into the dynamics.

Acknowledgments

The author wishes to thank the Meteorological Service of Singapore for providing the upper air data used in this study. He would also like to express his gratitude to Mdm Alice Heng Wang Cheng, Mr Chang Ee Chien, Miss Ng Hwee Lang, Miss Koh Siew Lee, and especially to his mentor, Associate Professor Lim Hock, Physics Department, National University of Singapore, for his invaluable help and guidance.

References

[1] K. Fraedrich, J. Atmos. Sci. **43**, 419 (1986)

[2] C. Nicolis, G. Nicolis, Nature **311**, 529 (1984)

[3] P. Grassberger, Nature **323**, 609 (1986)

[4] F. Takens, in: *Lecture Notes in Mathematics 898*, Eds. D. Rand, B.S. Young, Springer-Verlag, Berlin 1981, p. 366

[5] P. Grassberger, I. Procaccia, Physica D **9**, 189 (1986)

[6] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in Pascal: The Art of Scientific Computing*, Cambridge University Press, New York 1989, p. 599

[7] S.K. Park, K.W. Miller, Commun. ACM **31**, 1192 (1988)

Proceedings of the
2005 IEEE Conference on Control Applications
Toronto, Canada, August 28-31, 2005

A Lie Group Variational Integrator for the Attitude Dynamics of a Rigid Body with Applications to the 3D Pendulum

Taeyoung Lee*, N. Harris McClamroch†
Department of Aerospace Engineering
University of Michigan, Ann Arbor, MI 48109
{tylee@, nhm@engin.}umich.edu

Melvin Leok*
Department of Mathematics
University of Michigan, Ann Arbor, MI 48109
mleok@umich.edu

***Abstract*— A numerical integrator is derived for a class of models that describe the attitude dynamics of a rigid body in the presence of an attitude dependent potential. The numerical integrator is obtained from a discrete variational principle, and exhibits excellent geometric conservation properties. In particular, by performing computations at the level of the Lie algebra, and updating the solution using the matrix exponential, the attitude automatically evolves on the rotation group embedded in the space of matrices. The geometric conservation properties of the numerical integrator imply long time numerical stability. We apply this variational integrator to the uncontrolled 3D pendulum, that is a rigid asymmetric body supported at a frictionless pivot acting under the influence of uniform gravity. Interesting dynamics of the 3D pendulum are exposed.**

I. Introduction

The dynamics of a rigid body has fundamental invariance properties. In the absence of nonconservative forces on the system, the Hamiltonian is preserved, and a consequence of Noether's theorem is that symmetries in the Lagrangian result in the conservation of the associated momentum map [1]. Furthermore, the configuration space of the rigid body is a Lie group. Conservation laws are very important for studying dynamics, and the structure of the configuration space is a crucial geometric feature.

General numerical integration methods, including the popular Runge–Kutta schemes, typically preserve neither first integrals nor the characteristics of the configuration space. Numerical methods that preserve these geometric properties have been studied recently [2].

Numerical methods that conserve energy and momentum for N-body systems were introduced in LaBudde and Greenspan [3], and in Simo, Tarnow, and Wong [4]. Lewis and Simo [5] derived an exact energy-momentum conserving algorithm for nonlinear rotational dynamics. In this method, the integration step is selected to maintain constant angular momentum, and energy conservation is enforced by a momentum-preserving projection onto the manifold of constant energy.

Alternatively, variational integrators provide a systematic method of constructing structure-preserving integrators based on discrete mechanics. The variational approach approximates the continuous dynamics by discretizing *Hamilton's principle*. The discrete dynamics obtained from the discrete *Hamilton's principle* preserves the symplectic form, satisfies a discrete version of Noether's theorem [6], and exhibits good energy properties [7].

This discrete variational mechanics was introduced in Moser and Veselov [8], where they developed a numerical integrator for a free rigid body by imposing an orthogonality constraint. Wendlandt and Marsden [9] derived a more general discrete *Euler-Lagrange equation*, and applied it to the rigid body problem using quaternions. However, these methods do not consider the geometry of the configuration space directly, nor do they include the effect of an attitude-dependent potential. Marsden, Pekarsky and Shkoller [10] developed a variational integrator for the reduced dynamics of a mechanical system on a Lie group by introducing the discrete *Euler-Poincaré equation*. Bobenko and Suris [11] presented related work in the context of the Lagrange top, which only works for axially symmetric bodies.

In this paper, a numerical integrator is derived for the attitude dynamics of a rigid body under the influence of an attitude-dependent potential. It is obtained from a discrete variational approach, and therefore exactly preserves the momentum and symplectic form, while exhibiting good energy behavior over exponentially long times. This method explicitly respects the geometry of the configuration space $\mathrm{SO}(3)$. Furthermore, it is shown that this integrator is equivalent to the integrator presented by Marsden, Pekarsky and Shkoller [10] and by Moser and Veselov [8] in the absence of a potential.

The idea of using the Lie group structure and the exponential map to numerically compute rigid body dynamics arises in the work of Simo, Tarnow, and Wong [4], and in more recent work by Krysl [12]. It can also be found in the abstract setting of symplectic methods on manifolds obtained from generating functions, as in the work of Leimkuhler and Patrick [13], and in the general setting of Lie group methods, as in Iserles, Munthe-Kaas, Nørsett, and Zanna [14]. However, the method presented in this paper represents the first time that the Lie group approach is explicitly adopted in the context of variational integrators.

This Lie group variational integrator is used to study the dynamics of an uncontrolled 3D pendulum, that is a rigid asymmetric body supported at a frictionless pivot acting under the influence of uniform gravity. The pendulum models have been studied extensively in [15], [16], [17].

*This research has been supported in part by a grant from the Rackham Graduate School, University of Michigan.

†This research has been supported in part by NSF under grant ECS-0244977.

II. Continuous Time Model

A. Lagrangian

In this paper, the attitude dynamics of a rigid body is considered. It is assumed that a point rigidly attached to the body is fixed in an inertial frame. We use $R \in \mathrm{SO}(3)$ to denote a rotation matrix from the body fixed frame to the inertial frame, and $\rho \in \mathbb{R}^3$ is a vector from the pivot to the center of mass of the body expressed in the body fixed frame. The potential $U : \mathrm{SO}(3) \mapsto \mathbb{R}$ depends on the attitude of the rigid body.

Let $\tilde{\rho} \in \mathbb{R}^3$ be a vector from the pivot to a mass element of the body expressed in the body fixed frame, and $\Omega \in \mathbb{R}^3$ be the angular velocity of the body in the body fixed frame. The mass element has a velocity $\Omega \times \tilde{\rho}$. Then, the Lagrangian of the 3D pendulum can be expressed as follows,

$$L(R,\Omega) = \frac{1}{2}\int_{Body} \|S(\Omega)\tilde{\rho}\|^2 dm - U(R), \tag{1}$$

$$= \frac{1}{2}\int_{Body} \|S(\tilde{\rho})\Omega\|^2 dm - U(R), \tag{2}$$

where $S(\cdot) : \mathbb{R}^3 \mapsto \mathbb{R}^{3\times 3}$ is a skew mapping such that $S(x)y = x \times y$ for $x, y \in \mathbb{R}^3$. Using the relationship $\|x\|^2 = x^T x = \mathrm{tr}\big[xx^T\big]$, (1) can be expressed as

$$L(R,\Omega) = \frac{1}{2}\int_{Body} \mathrm{tr}\big[S(\Omega)\tilde{\rho}\tilde{\rho}^T S(\Omega)^T\big]\, dm - U(R),$$

$$\triangleq \frac{1}{2}\mathrm{tr}\big[S(\Omega)J_d S(\Omega)^T\big] - U(R), \tag{3}$$

where $J_d = \int_{Body} \tilde{\rho}\tilde{\rho}^T dm \in \mathbb{R}^{3\times 3}$. Alternatively, (2) can be expressed as

$$L(R,\Omega) = \frac{1}{2}\Omega^T\left(\int_{Body} S(\tilde{\rho})^T S(\tilde{\rho}) dm\right)\Omega - U(R),$$

$$\triangleq \frac{1}{2}\Omega^T J\Omega - U(R),$$

where $J = \int_{Body} S(\rho)^T S(\rho) dm \in \mathbb{R}^{3\times 3}$. This provides two different but equivalent expressions for the kinetic energy term in the Lagrangian, where J is the standard moment of inertia matrix, and J_d is a nonstandard moment of inertia matrix. The relationship between J and J_d is given by $J = \mathrm{tr}[J_d]\, I_{3\times 3} - J_d$. Furthermore, it can be shown that

$$S(J\Omega) = S(\Omega)J_d + J_d S(\Omega). \tag{4}$$

B. Equations of Motion

The equations of motion can easily be obtained using Newtonian mechanics. However, in this paper, we derive them from *Hamilton's principle* to more clearly illustrate the analogy between the continuous and the discrete equations of motion.

Define the action integral using (3) as

$$\mathfrak{G} = \int_{t_0}^{t_f} \frac{1}{2}\mathrm{tr}\big[S(\Omega)J_d S(\Omega)^T\big] - U(R)\, dt. \tag{5}$$

Hamilton's principle states that the variation of the action integral is zero, i.e. $\delta\mathfrak{G} = 0$. Because the configuration space is the rotation group $\mathrm{SO}(3)$, variation should be consistent with the rotation group. The varied rotation matrix $R^\epsilon \in \mathrm{SO}(3)$ can be expressed as

$$R^\epsilon = Re^{\epsilon\eta}, \tag{6}$$

where $\epsilon \in \mathbb{R}$, $\eta \in \mathfrak{so}(3)$ denotes a variation in the Lie algebra of skew symmetric matrices vanishing at t_0 and t_f. Hence,

$$\dot{R}^\epsilon = \dot{R}e^{\epsilon\eta} + \epsilon R e^{\epsilon\eta}\dot{\eta}.$$

Using the kinematic relationship $\dot{R} = RS(\Omega)$, the variation of the angular velocity can be expressed as

$$S(\Omega^\epsilon) = R^{\epsilon T}\dot{R}^\epsilon = e^{-\epsilon\eta}S(\Omega)e^{\epsilon\eta} + \epsilon\dot{\eta},$$

$$= S(\Omega) + \epsilon\left[\dot{\eta} + S(\Omega)\eta - \eta S(\Omega)\right] + \mathcal{O}(\epsilon^2). \tag{7}$$

Substituting (6), (7) into (5), and taking variations yields

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0}\mathfrak{G}^\epsilon = \int_{t_0}^{t_f} -\frac{1}{2}\mathrm{tr}[\dot{\eta}\left\{J_d S(\Omega) + S(\Omega)J_d\right\}]$$

$$+\frac{1}{2}\mathrm{tr}[\eta S(\Omega)\left\{J_d S(\Omega) + S(\Omega)J_d\right\}]$$

$$-\frac{1}{2}\mathrm{tr}[\eta\left\{J_d S(\Omega) + S(\Omega)J_d\right\}S(\Omega)]$$

$$+\mathrm{tr}\left[\eta R^T\frac{\partial U}{\partial R}\right]\, dt, \tag{8}$$

where the following relationships are used.

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} U(R^\epsilon) = \sum_{i,j=1}^{3}\frac{\partial U}{\partial [R]_{ij}}\left.\frac{\partial\left[Re^{\epsilon\eta}\right]_{ij}}{\partial\epsilon}\right|_{\epsilon=0},$$

$$= \sum_{i,j=1}^{3}\frac{\partial U}{\partial [R]_{ij}}\left[R\eta\right]_{ij},$$

$$= -\mathrm{tr}\left[\eta R^T\frac{\partial U}{\partial R}\right], \tag{9}$$

where $[A]_{ij}$ denotes the i,jth element of a matrix A, and $\frac{\partial U}{\partial R} \in \mathbb{R}^{3\times 3}$ is defined such that $\left[\frac{\partial U}{\partial R}\right]_{ij} = \frac{\partial U(R)}{\partial [R]_{ij}}$.

Using (4) and the properties $S(x \times y) = S(x)S(y) - S(y)S(x)$ for $x, y \in \mathbb{R}^3$, (8) can be expressed as

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0}\mathfrak{G}^\epsilon = \int_{t_0}^{t_f}\frac{1}{2}\mathrm{tr}[-\dot{\eta}S(J\Omega) + \eta S(\Omega\times J\Omega)]$$

$$+\mathrm{tr}\left[\eta R^T\frac{\partial U}{\partial R}\right]\, dt,$$

$$= -\frac{1}{2}\mathrm{tr}[\eta S(J\Omega)]\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f}\frac{1}{2}\mathrm{tr}\left[\eta S(J\dot{\Omega})\right]\, dt$$

$$+\int_{t_0}^{t_f}\frac{1}{2}\mathrm{tr}[\eta S(\Omega\times J\Omega)] + \mathrm{tr}\left[\eta R^T\frac{\partial U}{\partial R}\right] dt,$$

$$= \int_{t_0}^{t_f}\frac{1}{2}\mathrm{tr}\left[\eta\left\{S(J\dot{\Omega} + \Omega\times J\Omega) + 2R^T\frac{\partial U}{\partial R}\right\}\right] dt.$$

From *Hamilton's principle*, the above equation should be zero for all variations $\eta \in \mathfrak{so}(3)$. Because η is skew-symmetric, the equation in the braces should be symmetric. Then, we obtain the following *equations of motion in Lagrangian form*

$$S(J\dot{\Omega} + \Omega\times J\Omega) = \frac{\partial U}{\partial R}^T R - R^T\frac{\partial U}{\partial R},$$

or equivalently,

$$J\dot{\Omega} + \Omega \times J\Omega = M, \tag{10}$$

where $M \in \mathbb{R}^3$ is determined by $S(M) = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}$. More explicitly, it can be shown that the torque due to the attitude-dependent potential is given by

$$M = r_1 \times v_{r_1} + r_2 \times v_{r_2} + r_3 \times v_{r_3}, \tag{11}$$

where r_i, $v_{r_i} \in \mathbb{R}^{1\times 3}$ are the ith row vectors of R and $\frac{\partial U}{\partial R}$, respectively. The attitude kinematic equation is

$$\dot{R} = RS(\Omega). \tag{12}$$

Denote by $\Pi \in \mathbb{R}^3$, the angular momentum of the pendulum expressed in the body fixed frame. Then, (10) can be expressed as *equations of motion in Hamiltonian form*:

$$\dot{\Pi} + \Omega \times \Pi = M. \tag{13}$$

Define $P = \mathbb{F}L(R, \Omega) \in \mathbb{R}^{3\times 3}$, where $\mathbb{F}L$ denotes a fiber derivative, the derivative of L along the fiber in the direction $\delta\Omega$ [1]. Then, we have

$$\begin{aligned}\mathrm{tr}\big[S(\delta\Omega)^T P\big] &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} L(R, \Omega + \epsilon\delta\Omega),\\ &= \frac{1}{2}\mathrm{tr}\big[S(\delta\Omega)J_d S(\Omega)^T + S(\Omega)J_d S(\delta\Omega)^T\big],\\ &= \frac{1}{2}\mathrm{tr}\big[S(\delta\Omega)^T S(J\Omega)\big],\end{aligned}$$

where (4) is used. Therefore,

$$\mathrm{tr}\left[S(\delta\Omega)^T\left\{P - \frac{1}{2}S(J\Omega)\right\}\right] = 0.$$

Because $S(\delta\Omega)$ is skew symmetric, the expression in braces should be symmetric. Then, we have

$$P - P^T = S(\Pi). \tag{14}$$

III. Variational Integrator

Variational integrators are obtained by discretizing *Hamilton's principle* rather than the equations of motion.

A. Discrete Equations of Motion

Consider the fixed step size $h > 0$. Let $R_k \in \mathrm{SO}(3)$ denote the attitude of the rigid body at time $t = kh + t_0$. Define $F_k \in \mathrm{SO}(3)$ such that $R_{k+1} = R_k F_k$, i.e. $F_k = R_k^T R_{k+1}$. Using the kinematic relationship $\dot{R} = RS(\Omega)$, $S(\Omega_k)$ can be approximated as

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

Consider the discrete Lagrangian L_d,

$$\begin{aligned}L_d(R_k, F_k) &\simeq \frac{h}{2}L(R_k, \Omega_k) + \frac{h}{2}L(R_{k+1}, \Omega_k),\\ &= \frac{1}{2h}\mathrm{tr}\Big[\left(F_k - I_{3\times 3}\right) J_d \left(F_k - I_{3\times 3}\right)^T\Big]\\ &\quad - \frac{h}{2}U(R_k) - \frac{h}{2}U(R_{k+1}),\end{aligned}$$

where we used (3). Simplifying, we obtain

$$L_d(R_k, F_k) = \frac{1}{h}\mathrm{tr}[\left(I_{3\times 3} - F_k\right) J_d] - \frac{h}{2}U(R_k) - \frac{h}{2}U(R_{k+1}).$$

The discrete Lagrangian L_d approximates $\int_{t_k}^{t_{k+1}} L(R, \Omega)\, dt$, and the discrete action sum $\mathfrak{G}_d$ is given by

$$\mathfrak{G}_d = \sum_{k=0}^{N-1} \frac{1}{h}\mathrm{tr}[\left(I_{3\times 3} - F_k\right) J_d] - \frac{h}{2}U(R_k) - \frac{h}{2}U(R_{k+1}), \tag{15}$$

where $N \in \mathbb{R}$ is defined such that $t_f = Nh + t_0$.

Consider admissible variations of R_k and F_k. As in (6), the variation R_k^ϵ can be expressed using the exponential

$$R_k^\epsilon = R_k e^{\epsilon\eta_k}, \tag{16}$$

where $\epsilon \in \mathbb{R}$ and the variation $\eta_k \in \mathfrak{so}(3)$. From $F_k = R_k^T R_{k+1}$, the variation F_k^ϵ can be expressed as

$$\begin{aligned}F_k^\epsilon &= R_k^{\epsilon T} R_{k+1}^\epsilon,\\ &= e^{-\epsilon\eta_k} F_k e^{\epsilon\eta_{k+1}}.\end{aligned} \tag{17}$$

Substituting (16), (17) into the discrete action sum (15), we obtain

$$\begin{aligned}\mathfrak{G}_d^\epsilon = &\sum_{k=0}^{N-1} \frac{1}{h}\mathrm{tr}\big[\left(I_{3\times 3} - e^{-\epsilon\eta_k} F_k e^{\epsilon\eta_{k+1}}\right) J_d\big]\\ &- \sum_{k=0}^{N-1}\left\{\frac{h}{2}U(R_k e^{\epsilon\eta_k}) + \frac{h}{2}U(R_{k+1}e^{\epsilon\eta_{k+1}})\right\}.\end{aligned}$$

Differentiating the above equation,

$$\begin{aligned}\frac{d}{d\epsilon}\Big|_{\epsilon=0}\mathfrak{G}_d^\epsilon = &\sum_{k=0}^{N-1}\frac{1}{h}\mathrm{tr}[(\eta_k F_k - F_k\eta_{k+1})\, J_d]\\ &+ \sum_{k=0}^{N-1}\frac{h}{2}\mathrm{tr}\left[\eta_k R_k^T \frac{\partial U_k}{\partial R_k} + \eta_{k+1} R_{k+1}^T \frac{\partial U_{k+1}}{\partial R_{k+1}}\right],\end{aligned}$$

where (9) is used, and $U_k = U(R_k)$. Using the fact that $\eta_0 = \eta_N = 0$, we can write the summation as

$$\begin{aligned}\frac{d}{d\epsilon}\Big|_{\epsilon=0}\mathfrak{G}_d^\epsilon &= \sum_{k=1}^{N-1}\frac{1}{h}\mathrm{tr}[\eta_k F_k J_d] - \sum_{k=1}^{N-1}\frac{1}{h}\mathrm{tr}[\eta_k J_d F_{k-1}]\\ &\quad + \sum_{k=1}^{N-1} h\mathrm{tr}\left[\eta_k R_k^T \frac{\partial U_k}{\partial R_k}\right],\\ &= \sum_{k=1}^{N-1}\mathrm{tr}\left[\eta_k\left\{\frac{1}{h}\left(F_k J_d - J_d F_{k-1}\right) + hR_k^T\frac{\partial U_k}{\partial R_k}\right\}\right].\end{aligned}$$

From *Hamilton's principle*, the above equation should be zero for all variations $\eta_k \in \mathfrak{so}(3)$. Because η_k is skew-symmetric, the equation in the braces should be symmetric. Then, we obtain the following *discrete equations of motion in Lagrangian form*.

$$\begin{aligned}&\frac{1}{h}\left(F_{k+1}J_d - J_d F_k - J_d F_{k+1}^T + F_k^T J_d\right)\\ &\quad = h\left(\frac{\partial U_{k+1}}{\partial R_{k+1}}^T R_{k+1} - R_{k+1}^T\frac{\partial U_{k+1}}{\partial R_{k+1}}\right),\end{aligned} \tag{18}$$

$$R_{k+1} = R_k F_k. \tag{19}$$

For given R_0, F_0 and $R_1 = R_0F_0$, we can obtain F_1 by solving the implicit equation (18), and R_2 is computed using (19). This yields a map $(R_0, F_0, R_1) \mapsto (R_1, F_1, R_2)$, and the process can be repeated to time march the solution.

The Hamiltonian form of the variational integrator is given in [6] by

$$P_k = -\mathbf{D}_{R_k} L_d(R_k, F_k), \tag{20}$$

$$P_{k+1} = \mathbf{D}_{R_{k+1}} L_d(R_k, F_k), \tag{21}$$

where $\mathbf{D}_{R_k} L_d$ is a partial derivative of L_d along R_k. Then,

$$\begin{aligned}\mathrm{tr}\big[\eta_k \mathbf{D}_{R_k} L_d(R_k, F_k)^T\big] &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L_d(R_k^\epsilon, R_k^{\epsilon T} R_{k+1}),\\ &= \mathrm{tr}\bigg[\eta_k \left\{\frac{1}{h} F_k J_d + \frac{h}{2} R_k^T \frac{\partial U_k}{\partial R_k}\right\}\bigg].\end{aligned}$$

Because η_k is skew symmetric, we have

$$-\mathbf{D}_{R_k} L_d + \mathbf{D}_{R_k} L_d{}^T = \frac{1}{h}\left(F_k J_d - J_d F_k^T\right) - \frac{h}{2} S(M_k).$$

where $M_k \in \mathbb{R}^3$ is defined such that $S(M_k) = \frac{\partial U_k}{\partial R_k}{}^T R_k - R_k^T \frac{\partial U_k}{\partial R_k}$ or (11) can be used. Substituting (14), (20) into the above equation, we obtain

$$S(\Pi_k) = \frac{1}{h}\left(F_k J_d - J_d F_k^T\right) - \frac{h}{2} S(M_k). \tag{22}$$

Similarly, we can derive the following equation using (21).

$$S(\Pi_{k+1}) = \frac{1}{h} F_k^T \left(F_k J_d - J_d F_k^T\right) F_k + \frac{h}{2} S(M_{k+1}). \tag{23}$$

Using (22) and (23), we can express the *discrete equations of motion in Hamiltonian form* as

$$\Pi_{k+1} = F_k^T \Pi_k + \frac{h}{2} F_k^T M_k + \frac{h}{2} M_{k+1}, \tag{24}$$

$$hS(\Pi_k + \frac{h}{2} M_k) = F_k J_d - J_d F_k^T, \tag{25}$$

$$R_{k+1} = R_k F_k. \tag{26}$$

Given R_0 and Π_0, we can obtain F_0 implicitly by solving (25), R_1 is updated by (26), and the angular momentum Π_1 is updated by (24). This yields a map $(R_0, \Pi_0) \mapsto (R_1, \Pi_1)$, and the process can be repeated to time march the solution.

B. Properties of the Variational Integrator

Variational integrators exhibit a discrete analogue of Noether's theorem [6], and symmetries of the discrete Lagrangian result in the conservation of the corresponding momentum maps. Our choice of discrete Lagrangian is such that it inherits the symmetries of the continuous Lagrangian. Therefore, all the conserved momenta in the continuous dynamics are preserved in the discrete dynamics.

During each integration step, $F_k \in \mathrm{SO}(3)$ is obtained by solving (25) and R_k is updated using (26). Since $\mathrm{SO}(3)$ is closed under matrix multiplication, the attitude matrix R_k remains in $\mathrm{SO}(3)$. We will make this more explicit in the following discussion by expressing F_k as the exponential of an element of the Lie algebra $\mathfrak{so}(3)$.

C. Computational Approach

The structure of the discrete equations of motion, (24), (25), and (26), suggests a computational approach. The only implicit equation is (25). Given $\Pi_k, M_k \in \mathbb{R}^3$, $F_k \in \mathrm{SO}(3)$ can be obtained by solving (25).

We now consider an iterative approach to solve (25) using the fact that $F_k \in \mathrm{SO}(3)$ can be expressed as an exponential of $S(f_k) \in \mathfrak{so}(3)$ where f_k is a vector in $\mathbb{R}^3$. Using the Rodrigues' formula,

$$\begin{aligned}F_k &= e^{S(f_k)},\\ &= I_{3\times 3} + c_1(f_k) S(f_k) + c_2(f_k) S(f_k)^2,\end{aligned} \tag{27}$$

where $c_1(f) = \frac{\sin\|f\|}{\|f\|}, c_2(f) = \frac{1-\cos\|f\|}{\|f\|^2}$. Substituting (27) into (25), we obtain the following equation after some algebra,

$$hS(\Pi_k + \frac{h}{2} M_k) = c_1(f_k) S(Jf_k) + c_2(f_k) S(f_k \times Jf_k).$$

Thus, (25) is converted into the equivalent vector equation for $f_k \in \mathbb{R}^3$,

$$h\Pi_k + \frac{h^2}{2} M_k = \frac{\sin\|f_k\|}{\|f_k\|} Jf_k + \frac{1-\cos\|f_k\|}{\|f_k\|^2} f_k \times Jf_k.$$

We use the *Newton iteration method* to solve the above equation.

D. Relation to Other Methods

The above discrete equations of motion can be viewed as generalizations of related work by Marsden, Pekarsky and Shkoller [10], and Moser and Veselov [8], to include the effect of an attitude-dependent potential, while respecting the geometry of $\mathrm{SO}(3)$.

Indeed, these three approaches are equivalent in the absence of a potential. When $U = 0$, it follows that $M_k = M_{k+1} = 0$. Then (24) and (25) reduces to,

$$F_{k+1} J_d - J_d F_k - J_d F_{k+1}^T + F_k^T J_d = 0, \tag{28}$$

which is equivalent to equation (5.4) in [10]. Furthermore, (26) is simply the reconstruction equation for a left-invariant system in discrete reduction, so we recover the discrete *Euler-Poincaré* method of Marsden, Pekarsky and Shkoller [10], which they show to be equivalent to Moser and Veselov's method [8] for a specific choice of discrete Lagrangian.

IV. Variational Integration of a 3D pendulum

We apply the variational integrator introduced in this paper to study the dynamics of the 3D pendulum, a rigid asymmetric body supported at a frictionless pivot acting under the influence of uniform gravity [16], [17]. The gravity potential, acting in the vertical or e_3 direction, and its derivative can be expressed as

$$\begin{aligned}U &= -mge_3^T R\rho,\\ \frac{\partial U}{\partial R} &= -mge_3\rho^T.\end{aligned} \tag{29}$$

A. Continuous Time Model

Using (29), we obtain

$$S(M) = mg\left(R^T e_3 \rho^T - \rho e_3^T R\right), \\ = mgS(\rho \times R^T e_3). \tag{30}$$

Then, from (13), the continuous equation of the motion is

$$\dot{\Pi} + \Omega \times \Pi = mg\rho \times R^T e_3, \tag{31}$$

which is also known as the *Euler-Poincaré equation*.

The Hamiltonian

$$H = \frac{1}{2}\Omega^T J \Omega - mge_3^T R\rho$$

is a first integral. Conservation of energy is easily verified by direct calculation using (12) and (31).

Equation (31) can be written as

$$\frac{d}{dt} R\Pi = mgR\rho \times e_3. \tag{32}$$

If we take the dot product of the above equation with e_3, the right side vanishes, which shows conservation of the vertical component of the angular momentum $e_3^T R\Pi$.

B. Variational Integrator

Substituting (30) into (24), we obtain as the discrete equation of motion

$$\Pi_{k+1} = F_k^T \Pi_k + \frac{h}{2} mg F_k^T \left(\rho \times R_k^T e_3\right) + \frac{h}{2} mg\rho \times R_{k+1}^T e_3. \tag{33}$$

Since the discrete Lagrangian is invariant under the left action of rotations about e_3, the discrete analogue of Noether's theorem [6] states that the spatial angular momentum about the vertical axis is preserved. We can also verify this directly. Substituting (26) into (33), we get

$$\Pi_{k+1} = R_{k+1}^T R_k \Pi_k \\ + \frac{h}{2} mg R_{k+1}^T R_k \left(\rho \times R_k^T e_3\right) + \frac{h}{2} mg\rho \times R_{k+1}^T e_3.$$

Multiplying both side by R_{k+1} and using property of the rotation matrix, the above equation is transformed into

$$R_{k+1}\Pi_{k+1} - R_k \Pi_k = \frac{h}{2} mg\, R_k \rho \times e_3 + \frac{h}{2} mg\, R_{k+1}\rho \times e_3. \tag{34}$$

If we take the dot product of the above equation with e_3, the right side is zero, which demonstrates conservation of $e_3^T R_k \Pi_k$. It is interesting to note that (34) can be viewed as a simple discretization of (32).

C. Numerical Simulations

The properties of a 3D pendulum are

$$J = \mathrm{diag}\,[1,\ 2.8,\ 2]\,\mathrm{kg\,m^2},\ \mathrm{m} = 1\mathrm{kg},\ \rho = [0, 0, 1]\,\mathrm{m}.$$

Two initial conditions are considered.

(i) Small perturbation from a hanging equilibrium

$$\Omega_0 = [0.5, -0.5, 0.4]\,\mathrm{rad/s},\quad R_0 = I_{3\times 3}.$$

TABLE I

STANDARD DEVIATIONS OF THE CONSERVED QUANTITIES

Method	Case	H (J)	$e_3^T R\Pi$ (Nms)	$\|I - R^T R\|$
Variational Integrator	(i)	1.74×10^{-7}	4.16×10^{-13}	3.96×10^{-14}
	(ii)	1.83×10^{-7}	3.51×10^{-12}	3.33×10^{-12}
Runge Kutta	(i)	6.59×10^{-4}	9.50×10^{-5}	6.17×10^{-5}
	(ii)	9.92×10^{-3}	5.91×10^{-3}	1.83×10^{-3}

(ii) Small perturbation from a inverted equilibrium

$$\Omega_0 = [0.5, -0.5, 0.4]\,\mathrm{rad/s},\quad R_0 = \mathrm{diag}\,[-1, 1, -1]\,.$$

Simulation results for case (i) are presented in Fig. 1, where the left figure shows time histories of the angular velocity, and the right figure shows the variation of the the total energy, the angular momentum around the vertical axis, and the SO(3) error, $\left\|I_{3\times 3} - R^T R\right\|$. The step size, h is 0.001. A simulation obtained from the variable step size Runge-Kutta method (*Matlab* ode45 function) is indicated as a comparison.

The angular velocity shows a regular motion of the pendulum. We can see that the variational integrator preserves the first integrals and the structure of the configuration space very well. Note that the conserved quantities diverge from the their true values when the Runge-Kutta method is used. The errors in the Runge-Kutta simulation increase as the simulation time increases, but the errors in the variational integrator do not. Standard deviations of the conserved quantities are presented at Table I for each method.

Simulation results for case (ii) are shown in Fig. 2. It shows that the time histories of the angular velocity are irregular, while the total energy and the angular momentum around the vertical axis remain nearly constant. (A simple animation which shows chaotic motion of the 3D pendulum can be found at `http://www.umich.edu/~tylee`.)

Note that the Runge-Kutta method does not preserve the conserved quantities, and it also fails to show the correct response of the angular velocity.

V. CONCLUSION

A numerical integrator is derived for the attitude dynamics of a rigid body under a potential that depends on the attitude of the body. This integrator is obtained from a discrete variational principle, and exhibits the characteristic symplectic and momentum preservation properties, and good energy behavior characteristic of variational integrators.

The discrete equations of motion are expressed in terms of the angular velocity vector and a rotation matrix that represents the attitude globally. This is particularly significant, since it represents the first variational integrator that adopts the approach of Lie group integrators [14] to express the updated solution in terms of the exponential of a Lie algebra element, thereby yielding an algorithm that automatically evolves on the rotation group while embedded in the space of matrices, without the need for reprojection techniques or constraints. Higher-order analogues of such Lie group variational integrators are described in Leok [18].

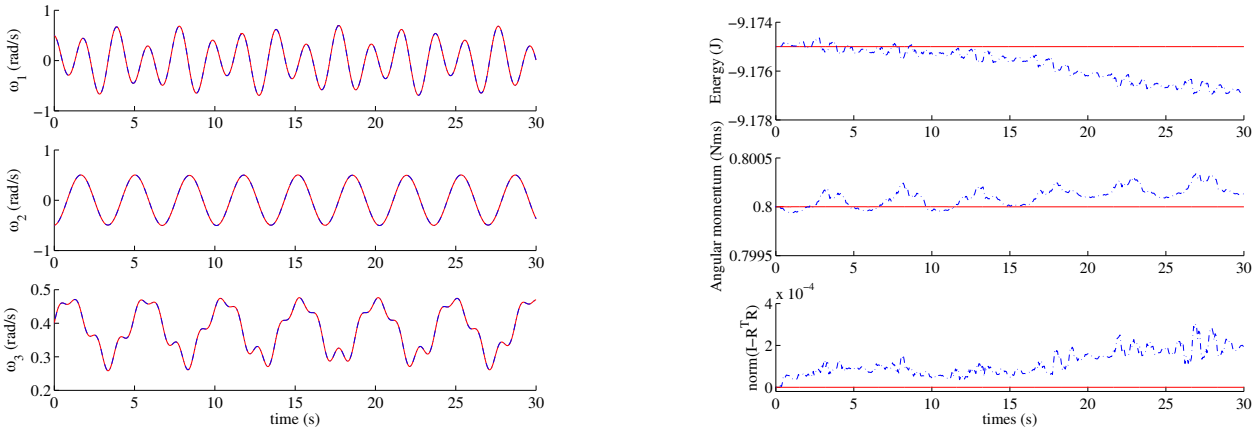


Fig. 1. Case (i), Small perturbation from the hanging equilibrium. (Variational integrator ———, Runge-Kutta —·—·—)

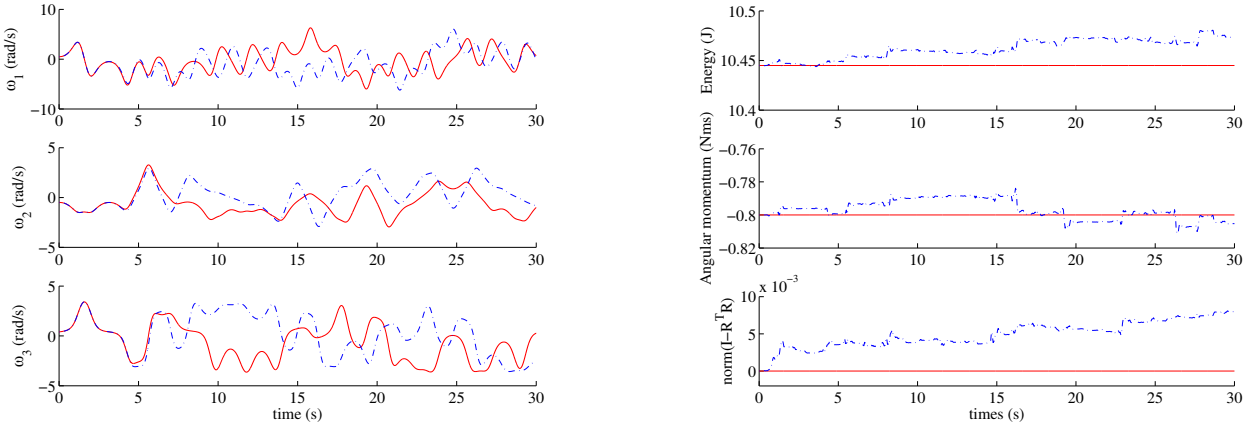


Fig. 2. Case (ii), Small perturbation from the inverted equilibrium. (Variational integrator ———, Runge-Kutta —·—·—)

The resulting Lie group variational integrator is used to study the dynamics of the 3D pendulum, thereby exposing some interesting dynamics. Small perturbations from a stable equilibrium are shown to lead to solution regularity; perturbations from an unstable equilibrium may lead to irregular or chaotic solutions.

REFERENCES

[1] R. Abraham and J. E. Marsden, *Foundations of Mechanics*, 2nd ed. Benjamin/Cummings Publishing Company, 1978.

[2] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*. Springer, 2000.

[3] R. A. LaBudde and D. Greenspan, "Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion. i. motion of a single particle," *Numerische Mathematik*, vol. 25, no. 4, pp. 323–346, 1976.

[4] J. C. Simo, N. Tarnow, and K. K. Wong, "Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics," *Computer Methods in Applied Mechanics and Engineering*, vol. 100, pp. 63–116, 1992.

[5] D. Lewis and J. Simo, "Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups," *Journal of Nonlinear Science*, vol. 4, pp. 253–299, 1994.

[6] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[7] E. Hairer, "Backward analysis of numerical integrators and symplectic methods," *Ann. Numer. Math.*, vol. 1, no. 1-4, pp. 107–132, 1994, scientific computation and differential equations (Auckland, 1993).

[8] J. Moser and A. P. Veselov, "Discrete versions of some classical integrable systems and factorization of matrix polynomials," *Communications in Mathematical Physics*, vol. 139, pp. 217–243, 1991.

[9] J. M. Wendlandt and J. E. Marsden, "Mechanical integrator derived from a discrete variational principle," *Physica D*, vol. 106, pp. 223–246, 1997.

[10] J. E. Marsden, S. Pekarsky, and S. Shkoller, "Discrete Euler-Poincaré and Lie-Poisson equation," *Nonlinearity*, vol. 12, pp. 1647–1662, 1999.

[11] A. I. Bobenko and Y. B. Suris, "Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top," *Communications in Mathematical Physics*, vol. 204, pp. 147–188, 1999.

[12] P. Krysl, "Explicit momentum-conserving integrator for dynamics of rigid bodies approximating the midpoint Lie algorithm," *Int. J. Numer. Meth. Eng.*, 2004, to appear.

[13] B. Leimkuhler and G. W. Patrick, "A symplectic integrator for Riemannian manifolds," *J. Nonlinear Sci.*, vol. 6, no. 4, pp. 367–384, 1996.

[14] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, "Lie-group methods," *Acta Numerica*, vol. 9, pp. 215–365, 2000.

[15] F. Furuta, "Control of pendulum: From super mechano-system to human adaptive mechatronics," in *Proceedings of 42nd IEEE Conference on Decision and Control*, Dec. 2003, pp. 1498–1507.

[16] S. Cho, J. Shen, and N. H. McClamroch, "Mathematical models for the Triaxial Attitude Control Testbed," *Mathematical and Computer Modeling of Dynamical Systems*, vol. 9, no. 2, pp. 165–192, 2003.

[17] J. Shen, A. K. Sanyal, N. A. Chaturvedi, D. Bernstein, and N. H. McClamroch, "Dynamics and control of a 3D pendulum," in *Proceedings of 43rd IEEE Conference on Decision and Control*, Dec. 2004, pp. 323–328.

[18] M. Leok, "Generalized Galerkin variational integrators: Lie group, multiscale, and pseudospectral methods," 2004, (in preparation).

Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Controlled Lagrangians and Stabilization of the Discrete Cart-Pendulum System

Anthony M. Bloch, Melvin Leok
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109
{abloch, mleok}@umich.edu

Jerrold E. Marsden
Control and Dynamical Systems
California Institute of Technology
Pasadena, CA 91125
marsden@cds.caltech.edu

Dmitry V. Zenkov
Department of Mathematics
North Carolina State University
Raleigh, NC 27695
dvzenkov@unity.ncsu.edu

***Abstract*— Matching techniques are developed for discrete mechanical systems with symmetry. We describe new phenomena that arise in the controlled Lagrangian approach for mechanical systems in the discrete context. In particular, one needs to either make an appropriate selection of momentum levels or introduce a new parameter into the controlled Lagrangian to complete the matching procedure. We also discuss digital and model predictive control.**

I. Introduction

The method of controlled Lagrangians for stabilization of relative equilibria (steady state motions) originated in Bloch, Leonard, and Marsden [3] and was then developed in Auckly [1], Bloch, Leonard, and Marsden [4], [5], [6], Bloch, Chang, Leonard, and Marsden [7], and Hamberg [10], [11]. A similar approach for Hamiltonian controlled systems was introduced and further studied in the work of Blankenstein, Ortega, van der Schaft, Maschke, Spong, and their collaborators (see, e.g., [18] and related references). The two methods were shown to be equivalent in [8] and a nonholonomic version was developed in [20], [21], and [2].

According to the method of controlled Lagrangians, the original controlled system is represented as a new, uncontrolled Lagrangian system for a suitable controlled Lagrangian. The energy associated with this controlled Lagrangian is designed to be positive or negative definite at the (relative) equilibrium to be stabilized. The time-invariant feedback control law is obtained from the equivalence requirement for the new and old systems of equations of motion. If asymptotic stabilization is desired, dissipation-emulating terms are added to the control input.

In this paper the method of controlled Lagrangians is applied to the discrete cart-pendulum system. This study is motivated by the importance of structure-preserving algorithms for numerical simulation of controlled systems. In particular, as the closed loop dynamics of a controlled Lagrangian system is itself Lagrangian, it is natural to adopt a variational discretization that exhibits good long-time numerical stability.

We carry out the matching procedure explicitly for the discrete cart-pendulum system and prove that we can asymptotically stabilize the upward vertical position of the pendulum. The theoretical analysis is validated by simulating the discrete cart-pendulum system with control, and when dissipation is added, the inverted pendulum configuration is asymptotically stabilized, as predicted.

We then use the discrete controlled dynamics to construct a real-time model predictive controller with piecewise constant control inputs. This serves to illustrate how discrete mechanics can be naturally applied to yield digital controllers for mechanical systems.

The paper is organized as follows: In Sections II and III we review discrete mechanics and the method of controlled Lagrangians for stabilization of (relative) equilibria of mechanical systems. The discrete version of the method of controlled Lagrangians is discussed in Section IV. The theory is illustrated with the discrete cart-pendulum system in Section V. Simulations and the construction of the digital controller are presented in Sections VI and VII.

In a future publication we intend to treat discrete systems with nonabelian symmetries as well as systems with nonholonomic constraints.

II. An Overview of Discrete Mechanics

A discrete analogue of Lagrangian mechanics can be obtained by considering a discretization of Hamilton's principle; this approach underlies the construction of variational integrators. See Marsden and West [17] and references therein for a more detailed discussion of discrete mechanics.

A key notion is that of the *discrete Lagrangian*, which is a map $L^d : Q \times Q \to \mathbb{R}$ that approximates the action integral along an exact solution of the Euler–Lagrange equations joining q_k and q_{k+1},

$$L^d(q_k, q_{k+1}) \approx \operatorname*{ext}_{q \in \mathcal{C}([0,h],Q)} \int_0^h L(q, \dot{q})\, dt, \qquad (1)$$

where $\mathcal{C}([0,h],Q)$ is the space of curves $q : [0,h] \to Q$ with $q(0) = q_k$, $q(h) = q_{k+1}$ and ext denotes extremum.

In the discrete setting, the action integral of Lagrangian mechanics is replaced by an action sum

$$S^d = \sum_{k=0}^{N-1} L^d(q_k, q_{k+1}),$$

where $q_k \in Q$. The equations are obtained by the discrete Hamilton's principle which extremizes the discrete action given fixed endpoints q_0 and q_N. Taking the extremum over $q_1, \ldots, q_{N-1}$ gives the *discrete Euler–Lagrange equations*

$$D_1 L^d(q_k, q_{k+1}) + D_2 L^d(q_{k-1}, q_k) = 0,$$

0-7803-9568-9/05/$20.00 ©2005 IEEE

for $k = 1, \ldots, N-1$. This implicitly defines the update map $\Phi : Q \times Q \to Q \times Q$, where $\Phi(q_{k-1}, q_k) = (q_k, q_{k+1})$ and $Q \times Q$ replaces the phase space TQ of Lagrangian mechanics.

In the rest of this paper, we will adopt the notations

$$q_{k+1/2} = \frac{q_k + q_{k+1}}{2}, \quad \Delta q_k = q_{k+1} - q_k.$$

This allows us to express a second-order accurate discrete Lagrangian as

$$L^d(q_{k,k+1}) = hL\left(q_{k+1/2}, \Delta q_k/h\right). \tag{2}$$

More generally, higher-order discrete Lagrangians can be obtained by using higher-order polynomial interpolation and numerical quadrature schemes. This yields the following approximation to (1):

$$L^d(q_k, q_{k+1}) = \operatorname*{ext}_{q \in \mathcal{C}^s([0,h],Q)} h \sum_{i=1}^{s} b_i L(q(c_i h), \dot{q}(c_i h)), \tag{3}$$

where c_i are a set of quadrature points, b_i are the associated maximal order weights, and $\mathcal{C}^s([0,h],Q) = \{q \in \mathcal{C}([0,h],Q) \mid q \text{ is a polynomial of degree } s\}$. The discrete Lagrangian (2) arises from this general formulation by using linear interpolation and the midpoint rule.

Since we are concerned with control, we need to consider the effect of external forces on Lagrangian systems. In the context of discrete mechanics, this is addressed by introducing the *discrete Lagrange–d'Alembert principle* (see, Kane, Marsden, Ortiz, West [14]), which states that

$$\delta \sum_{k=0}^{n-1} L^d(q_k, q_{k+1}) + \sum_{k=0}^{n-1} F^d(q_k, q_{k+1}) \cdot (\delta q_k, \delta q_{k+1}) = 0$$

for all variations $\boldsymbol{\delta q}$ of $\boldsymbol{q}$ that vanish at the endpoints. Here, $\boldsymbol{q}$ denotes the vector of positions $(q_0, q_1, \ldots, q_N)$, and $\boldsymbol{\delta q} = (\delta q_0, \delta q_1, \ldots, \delta q_N)$, where $\delta q_k \in T_{q_k}\mathcal{C}(Q)$. The discrete one-form F^d on $Q \times Q$ approximates the impulse integral between the points q_k and q_{k+1}, just as the discrete Lagrangian L^d approximates the action integral. We define the one-forms F^d_+ and F^d_- on $Q \times Q$ and the maps $F^d_1, F^d_2 : Q \times Q \to T^*Q$ by the relations

$$\begin{aligned} F^d_+(q_0, q_1) \cdot (\delta q_0, \delta q_1) &= F^d_2(q_0, q_1) \cdot \delta q_1 \\ &:= F^d(q_0, q_1) \cdot (0, \delta q_1), \\ F^d_-(q_0, q_1) \cdot (\delta q_0, \delta q_1) &= F^d_1(q_0, q_1) \cdot \delta q_0 \\ &:= F^d(q_0, q_1) \cdot (\delta q_0, 0). \end{aligned}$$

The discrete Lagrange–d'Alembert principle may then be rewritten as

$$\begin{aligned} &\delta \sum_{k=0}^{n-1} L^d(q_k, q_{k+1}) \\ &+ \sum_{k=0}^{n-1} \left[F^d_1(q_k, q_{k+1}) \cdot \delta q_k + F^d_2(q_k, q_{k+1}) \cdot \delta q_{k+1}\right] = 0 \end{aligned}$$

for all variations $\boldsymbol{\delta q}$ of $\boldsymbol{q}$ that vanish at the endpoints. This is equivalent to the *forced discrete Euler–Lagrange equations*

$$\begin{aligned} &D_1 L^d(q_k, q_{k+1}) + D_2 L^d(q_{k-1}, q_k) \\ &\qquad + F^d_1(q_k, q_{k+1}) + F^d_2(q_{k-1}, q_k) = 0. \end{aligned}$$

III. Matching and Controlled Lagrangians

In the controlled Lagrangian approach one considers a mechanical system with an uncontrolled (free) Lagrangian equal to kinetic energy minus potential energy. In the simplest setting we modify the kinetic energy to produce a new controlled Lagrangian which describes the dynamics of the controlled closed-loop system. The method may be extended in various ways including the incorporation of potential shaping.

Suppose our system has configuration space Q and a Lie group G acts freely and properly on Q. It is useful to keep in mind the case in which $Q = S \times G$ with G acting only on the second factor by the left group multiplication.

For example, for the inverted planar pendulum on a cart, $Q = S^1 \times \mathbb{R}$ with $G = \mathbb{R}$, the group of reals under addition (corresponding to translations of the cart).

Our goal is to control the variables lying in the *shape space* Q/G using controls that act directly on the variables lying in G.[1] The controlled Lagrangian is constructed to be G-invariant, thus providing modified or *controlled* conservation laws. In this paper we assume that G is an abelian group.

The key modification of the Lagrangian involves changing the kinetic energy metric $g(\cdot, \cdot)$. The tangent space to Q can be split into a sum of horizontal and vertical parts defined as follows: For each tangent vector v_q to Q at a point $q \in Q$, we can write a unique decomposition $v_q = \operatorname{Hor} v_q + \operatorname{Ver} v_q$, such that the vertical part is tangent to the orbits of the G-action and the horizontal part is metric-orthogonal to the vertical space, *i.e.*, it is uniquely defined by the identity

$$g(v_q, w_q) = g(\operatorname{Hor} v_q, \operatorname{Hor} w_q) + g(\operatorname{Ver} v_q, \operatorname{Ver} w_q) \tag{4}$$

with v_q and w_q arbitrary tangent vectors to Q at the point $q \in Q$. This choice of horizontal space coincides with that given by the *mechanical connection*; see, for example, Marsden [1992].

For the kinetic energy of our controlled Lagrangian, we use a modified version of the right-hand side of equation (4). The potential energy remains unchanged. The modification consists of three ingredients:

1) a new choice of horizontal space, denoted Hor_τ,
2) a change $g \to g_\sigma$ of the metric on horizontal vectors,
3) a change $g \to g_\rho$ of the metric on vertical vectors.

Let ξ_Q denote the infinitesimal generator corresponding to $\xi \in \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra of G (see Marsden [1992] or Marsden and Ratiu [1994]). Thus, for each $\xi \in \mathfrak{g}$, ξ_Q is a vector field on the configuration manifold Q and its value at a point $q \in Q$ is denoted $\xi_Q(q)$.

Definition 1: Let τ be a Lie-algebra-valued horizontal one-form on Q; that is, a one-form that annihilates vertical vectors. The ***τ-horizontal space*** *at $q \in Q$ consists of tangent vectors to Q at q of the form $\operatorname{Hor}_\tau v_q = \operatorname{Hor} v_q - [\tau(v)]_Q(q)$, which also defines $v_q \mapsto \operatorname{Hor}_\tau(v_q)$, the* ***$\tau$-horizontal projection****. The* ***τ-vertical projection operator*** *is defined by $\operatorname{Ver}_\tau(v_q) := \operatorname{Ver}(v_q) + [\tau(v)]_Q(q)$.*

[1] The shape space is S in the case $Q = S \times G$.

*Definition 2: Given g_σ, g_ρ and τ, the **controlled Lagrangian** equals a modified kinetic minus the given potential energy:*

$$L_{\tau,\sigma,\rho}(v) = \frac{1}{2}[g_\sigma(\mathrm{Hor}_\tau v_q, \mathrm{Hor}_\tau v_q) + g_\rho(\mathrm{Ver}_\tau v_q, \mathrm{Ver}_\tau v_q)] - V(q).$$

The equations corresponding to this Lagrangian will be our closed-loop equations. The new terms appearing in those equations corresponding to the directly controlled variables are interpreted as control inputs. The modifications to the Lagrangian are chosen so that no new terms appear in the equations corresponding to the variables that are not directly controlled. We refer to this process as *matching*.

Once the control law is derived using the controlled Lagrangian, the closed-loop stability of an equilibrium can be determined by energy methods, using any available freedom in the choice of τ, g_σ and g_ρ.

Under some reasonable assumptions on the metric g_σ, $L_{\tau,\sigma,\rho}(v)$ has the following useful structure.

Theorem 3: Assume that $g = g_\sigma$ on Hor *and* Hor *and* Ver *are orthogonal for g_σ. Then*

$$L_{\tau,\sigma,\rho}(v) = L(v + \tau(v)_Q) + \frac{1}{2}g_\sigma(\tau(v)_Q, \tau(v)_Q) + \frac{1}{2}\varpi(v),$$

where $v \in T_qQ$ and $\varpi(v) = (g_\rho - g)(\mathrm{Ver}_\tau(v), \mathrm{Ver}_\tau(v))$.

A useful example treated in earlier papers in the smooth setting is the *pendulum on a cart*. Let s denote the position of the cart on the s-axis and let θ denote the angle of the pendulum with the upright vertical, as in Figure 1. The

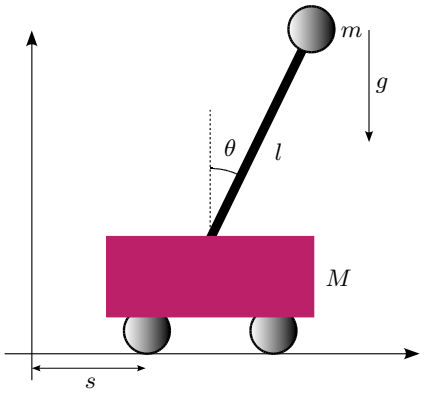


Fig. 1. The pendulum on a cart

configuration space for this system is $Q = S \times G = S^1 \times \mathbb{R}$, with the first factor being the pendulum angle θ and the second factor being the cart position s. The velocity phase space, TQ, has coordinates $(\theta, s, \dot\theta, \dot s)$. The length of the pendulum is l, the mass of the pendulum is m and that of the cart is M.

The symmetry group G of the pendulum-cart system is that of translation in the s variable, so $G = \mathbb{R}$. We do not destroy this symmetry when doing stabilization in θ.

For notational convenience, write the Lagrangian as

$$L(\theta, s, \dot\theta, \dot s) = \frac{1}{2}(\alpha\dot\theta^2 + 2\beta\cos\theta\dot s\dot\theta + \gamma\dot s^2) - U(\theta), \quad (5)$$

where $\alpha = ml^2, \beta = ml, \gamma = M + m$ are constants and $U(\theta) = -mgl\cos\theta$ is the potential energy. Note that $\alpha\gamma - \beta^2 > 0$.

IV. Discrete Matching

In discretizing the method of controlled Lagrangians, it is natural to combine the results of Theorem 3 with formula (3). To simplify the exposition in the remainder of the paper, we will restrict ourselves to the second-order controlled discrete Lagrangian that is defined by

$$L^d_{\tau,\sigma,\rho}(q_k, q_{k+1}) = h\Big[L(q_{k+\frac{1}{2}}, \Delta q_k/h + \tau(\Delta q_k/h)_Q) + \frac{1}{2}g_\sigma(\tau(\Delta q_k/h)_Q, \tau(\Delta q_k/h)_Q) + \frac{1}{2}\varpi(\Delta q_k/h)\Big]. \quad (6)$$

An alternative approach for parameterizing the space of controlled Lagrangians is through the use of discrete connections (see Leok, Marsden, and Weinstein [15]).

V. Stabilization of the Discrete Cart-Pendulum System

A. The Discrete Cart-Pendulum

For the discrete cart-pendulum system, $q_k = (\theta_k, s_k)$. According to (2), the discrete Lagrangian for the cart-pendulum system is

$$L^d(q_k, q_{k+1}) = hL(\theta_{k+\frac{1}{2}}, \Delta\theta_k/h, \Delta s_k/h),$$

where L is Lagrangian (5). The discrete dynamics is governed by the equations

$$\frac{\partial L^d(q_k, q_{k+1})}{\partial\theta_k} + \frac{\partial L^d(q_{k-1}, q_k)}{\partial\theta_k} = 0, \quad (7)$$

$$\frac{\partial L^d(q_k, q_{k+1})}{\partial s_k} + \frac{\partial L^d(q_{k-1}, q_k)}{\partial s_k} = u_k, \quad (8)$$

where u_k is the control input. Straightforward calculation shows that the relative equilibria $\theta_k = 0, s_k = \text{const}$ of the discrete cart-pendulum are unstable.

B. Discrete Matching Conditions

Using (6), we define the discrete controlled Lagrangian $L^d_{\tau,\sigma}(q_k, q_{k+1})$ for the cart-pendulum system by

$$hL\big(\theta_{k+\frac{1}{2}}, \Delta\theta_k/h, \Delta s_k/h + \tau(\theta_{k+\frac{1}{2}})\Delta\theta_k/h\big) + \frac{h}{2}\sigma\gamma\big(\tau(\theta_{k+\frac{1}{2}})\Delta\theta_k/h\big)^2, \quad (9)$$

where L is Lagrangian of the continuous-time cart-pendulum system (5).

The dynamics associated with (9) is

$$\frac{\partial L^d_{\tau,\sigma}(q_k, q_{k+1})}{\partial\theta_k} + \frac{\partial L^d_{\tau,\sigma}(q_{k-1}, q_k)}{\partial\theta_k} = 0, \quad (10)$$

$$\frac{\partial L^d_{\tau,\sigma}(q_k, q_{k+1})}{\partial s_k} + \frac{\partial L^d_{\tau,\sigma}(q_{k-1}, q_k)}{\partial s_k} = 0. \quad (11)$$

Equation (11) is equivalent to the *discrete controlled momentum conservation:*

$$p_k = \mu, \quad (12)$$

where

$$p_k = -\frac{\beta\Delta\theta_k \cos\theta_{k+1/2} + \gamma\Delta s_k + \gamma\Delta\theta_k\tau(\theta_{k+1/2})}{h}. \quad (13)$$

Setting

$$u_k = \frac{\gamma\Delta\theta_k\tau(\theta_{k+1/2}) - \gamma\Delta\theta_{k-1}\tau(\theta_{k-1/2})}{h} \quad (14)$$

matches equations (8) and (11) and allows one to represent the discrete momentum equation (8) as the discrete momentum conservation law

$$p_k = p. \quad (15)$$

Define the functions $F(\theta)$ and $G_\mu(\theta)$ by the formulae $F(\theta) = \beta\gamma\cos\theta + \gamma^2\sigma\tau(\theta)$ and $G_\mu(\theta) = h^2\beta(\mu - p)\cos\theta + h^2\gamma\mu\tau(\theta)$.

Theorem 4: The dynamics (7) *and* (8) *restricted to the momentum level $p_k = p$ is equivalent to the dynamics* (10) *and* (11) *restricted to the momentum level $p_k = \mu$ if and only if the matching condition*

$$\begin{aligned}\left[\Delta\theta_{k-1}^2\tau(\theta_{k-\frac{1}{2}})F'(\theta_{k-\frac{1}{2}}) - \Delta\theta_k^2\tau(\theta_{k+\frac{1}{2}})F'(\theta_{k+\frac{1}{2}})\right] \\ + 2\left[\Delta\theta_{k-1}\tau(\theta_{k-\frac{1}{2}})F(\theta_{k-\frac{1}{2}}) + \Delta\theta_k\tau(\theta_{k+\frac{1}{2}})F(\theta_{k+\frac{1}{2}})\right] \\ + \left[\Delta\theta_{k-1}G'_\mu(\theta_{k-\frac{1}{2}}) + \Delta\theta_k G'_\mu(\theta_{k+\frac{1}{2}})\right] \\ + 2\left[G_\mu(\theta_{k-\frac{1}{2}}) - G_\mu(\theta_{k+\frac{1}{2}})\right] = 0 \quad (16)\end{aligned}$$

holds.

Proof: Solve equations (12) and (15) for Δs_k and substitute the solutions in equations (7) and (10), respectively. This process is a simple version of discrete reduction [12]. A computation shows that the equations obtained this way are equivalent if and only if (16) is satisfied. ∎

Corollary 5: The matching condition (16) *is satisfied if*

$$\tau(\theta) = \kappa\cos\theta, \quad \sigma = -\frac{\beta}{\gamma\kappa}, \quad \mu = \frac{\beta}{\beta + \gamma\kappa}\, p. \quad (17)$$

Proof: Equation (17) implies $F(\theta) = G(\theta) = 0$, and therefore each term in the left-hand side of equation (16) vanishes. Note that the momentum levels p and μ *are not* the same. ∎

We now briefly discuss an alternative matching procedure. Define the discrete controlled Lagrangian $\Lambda^d_{\tau,\sigma,\lambda}(q_k, q_{k+1})$ by the formula

$$\begin{aligned}hL\left(\theta_{k+\frac{1}{2}}, \Delta\theta_k/h, \Delta s_k/h + \tau(\theta_{k+\frac{1}{2}})\Delta\theta_k/h\right) \\ + \frac{h}{2}\sigma\gamma\left(\tau(\theta_{k+\frac{1}{2}})\Delta\theta_k/h\right)^2 + h\lambda\tau(\theta_{k+\frac{1}{2}})\Delta\theta_k.\end{aligned}$$

The dynamics associated with this Lagrangian is

$$\frac{\partial\Lambda^d_{\tau,\sigma,\lambda}(q_k, q_{k+1})}{\partial\theta_k} + \frac{\partial\Lambda^d_{\tau,\sigma,\lambda}(q_{k-1}, q_k)}{\partial\theta_k} = 0, \quad (18)$$

$$\frac{\partial\Lambda^d_{\tau,\sigma}(q_k, q_{k+1})}{\partial s_k} + \frac{\partial\Lambda^d_{\tau,\sigma}(q_{k-1}, q_k)}{\partial s_k} = 0. \quad (19)$$

As before, the discrete controlled momentum is given by formula (13), and equation (19) is equivalent to the discrete momentum conservation (15).

Theorem 6: The dynamics (7) *and* (8) *restricted to the momentum level $p_k = p$ is equivalent to the dynamics* (18) *and* (19) *restricted to the same momentum level if and only if the matching condition*

$$\begin{aligned}\left[\Delta\theta_{k-1}^2\tau(\theta_{k-\frac{1}{2}})F'(\theta_{k-\frac{1}{2}}) - \Delta\theta_k^2\tau(\theta_{k+\frac{1}{2}})F'(\theta_{k+\frac{1}{2}})\right] \\ + 2\left[\Delta\theta_{k-1}\tau(\theta_{k-\frac{1}{2}})F(\theta_{k-\frac{1}{2}}) + \Delta\theta_k\tau(\theta_{k+\frac{1}{2}})F(\theta_{k+\frac{1}{2}})\right] \\ + h^2\gamma(p+\lambda)\left[2\tau(\theta_{k-\frac{1}{2}}) - 2\tau(\theta_{k+\frac{1}{2}})\right. \\ \left. + \Delta\theta_{k-1}\tau'(\theta_{k-\frac{1}{2}}) + \Delta\theta_k\tau'(\theta_{k+\frac{1}{2}})\right] = 0 \quad (20)\end{aligned}$$

holds.

Note that in this case we add an extra term to the controlled Lagrangian which eliminates the need for adjusting the momentum level. The proof of Theorem 6 is similar to that of Theorem 4. Further details will be given in a forthcoming publication.

Corollary 7: The matching condition (20) *is satisfied for the cart-pendulum system if*

$$\tau(\theta) = \kappa\cos\theta, \quad \sigma = -\frac{\beta}{\gamma\kappa}, \quad \lambda = -p.$$

Remark. The quantities $\tau(\theta)$ and σ obtained in Corollaries 5 and 7 are identical to those resulting from the continuous-time matching procedure for the cart-pendulum system.

C. Stabilization of the Discrete Cart-Pendulum System

Here we study the stability properties of the relative equilibria $\theta_k = 0$, $s_k = \text{const}$ of the discrete cart-pendulum system.

Theorem 8: The relative equilibria $\theta_k = 0$, $s_k = const$ of equations (7) *and* (8)*, with u_k defined by* (14)*, are spectrally stable if*

$$\kappa > \frac{\alpha\gamma - \beta^2}{\beta\gamma}. \quad (21)$$

Proof: Recall that $U''(0) = -C$, where $C > 0$. The linearization of the reduced dynamics (7) and (8) at $\theta = 0$ is computed to be

$$\begin{aligned}\frac{\alpha\gamma - \beta^2 - \beta\gamma\kappa}{h^2\gamma}\left(\Delta\theta_{k-1} - \Delta\theta_k\right) \\ + \frac{C}{4}\left(\theta_{k-1} + 2\theta_k + \theta_{k+1}\right) = 0. \quad (22)\end{aligned}$$

Observe that the value of p does not affect the linearized dynamics.

The linearized dynamics preserves the quadratic approximation of the discrete energy

$$\frac{\alpha\gamma - \beta^2 - \beta\gamma\kappa}{2h^2\gamma}\Delta\theta_k^2 - \frac{C\theta_{k+1/2}^2}{2}. \quad (23)$$

The equilibrium $\theta_k = 0$ of (22) is stable if and only if the function (23) is negative-definite at $\theta_k = \theta_{k+1} = 0$. The latter requirement is equivalent to condition (21). ∎

Remark. Stability condition (21) is identical to the stability condition of the continuous-time cart-pendulum system.

The spectrum of the linear map $(\theta_{k-1}, \theta_k) \mapsto (\theta_k, \theta_{k+1})$ defined by (22) belongs to the unit circle. Spectral stability in this situation is not sufficient to conclude nonlinear stability.

We now modify the control input (14) by adding the *discrete dissipation-emulating term*

$$-\frac{D(\Delta\theta_{k-1} + \Delta\theta_k)}{h}$$

in order to achieve the asymptotic stabilization of the upward position of the pendulum. In the above, D is a positive constant. The discrete momentum conservation law becomes

$$p_k = p - \frac{D\,\theta_{k+1/2}}{h}.$$

Straightforward calculation shows that the spectrum of the matrix of the linear map $(\theta_{k-1}, \theta_k) \mapsto (\theta_k, \theta_{k+1})$ defined by the reduced discrete dynamics belongs to the open unit disc. This implies that the equilibrium $\theta = 0$ is asymptotically stable.

VI. Simulations

Simulating the discrete behavior of the controlled Lagrangian system is simply a matter of viewing the forced discrete Euler–Lagrange equation as an implicit update map $\Phi : (q_{k-2}, q_{k-1}) \mapsto (q_{k-1}, q_k)$. This presupposes that the initial conditions are given in the form (q_0, q_1), however it might be preferable to specify the initial conditions as $(q_0, \dot{q}_0)$ instead. In this situation, one solves the boundary condition

$$\frac{\partial L}{\partial \dot{q}}(q_0, \dot{q}_0) + D_1 L^d(q_0, q_1) + F_1^d(q_0, q_1) = 0$$

for q_1, which can then be time marched using the implicit update map Φ. This procedure is described in Algorithm 1.

Algorithm 1 Variational Integrator ($q_0, \dot{q}_0, T_f, h$)

$p_0 \leftarrow \frac{\partial L}{\partial \dot{q}}(q_0, \dot{q}_0)$
$q_1 \leftarrow$ **solve** $p_0 + D_1 L^d(q_0, q_1) + F_1^d(q_0, q_1) = 0$
for $k = 2$ to T_f/h **do**
  $q_k \leftarrow$ **solve** $D_2 L^d(q_{k-2}, q_{k-1}) + D_1 L^d(q_{k-1}, q_k)$
    $+F_2^d(q_{k-2}, q_{k-1}) + F_1^d(q_{k-1}, q_k) = 0$
end for

When κ is twice the critical value, and dissipation is added, the θ dynamics is asymptotically stabilized, as predicted. This is illustrated in Figure 2.

VII. Model Predictive Controller

We now explore the use of the forced discrete Euler–Lagrange equations as a model for use in the context of a real-time model predictive controller, with piecewise constant control forces. Algorithm 2 describes the procedure in detail. The digital controller uses the position information it senses for $t = -2h, -h$ to estimate the positions at $t = 0, h, 2h$ during the time interval $t = [-h, 0]$. This allows

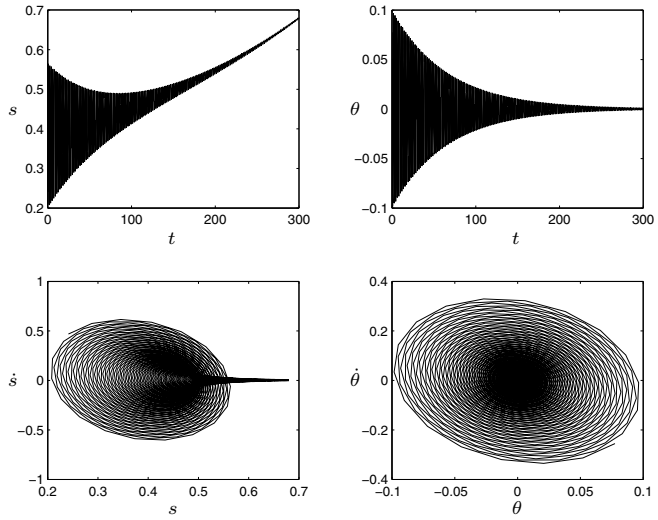


Fig. 2. Controlled dynamics with dissipation

Algorithm 2 Digital Controller ($q(\,\cdot\,), T_f, h$)

$q_0 \leftarrow$ **sense** $q(0)$
$q_1 \leftarrow$ **sense** $q(h)$
$\bar{q}_2 \leftarrow$ **solve** $D_2 L^d(q_0, q_1) + D_1 L^d(q_1, \bar{q}_2) = 0$
$\bar{q}_3 \leftarrow$ **solve** $D_2 L^d(q_1, \bar{q}_2) + D_1 L^d(\bar{q}_2, \bar{q}_3) + F_1^d(\bar{q}_2, \bar{q}_3) = 0$
$\bar{q}_4 \leftarrow$ **solve** $D_2 L^d(\bar{q}_2, \bar{q}_3) + D_1 L^d(\bar{q}_3, \bar{q}_4)$
    $+F_2^d(\bar{q}_2, \bar{q}_3) + F_1^d(\bar{q}_3, \bar{q}_4) = 0$
$u_{2+1/2} \leftarrow \frac{\gamma\kappa}{2h^2}\left[(\bar{\theta}_3 - \theta_1)\cos(\bar{\theta}_2) - (\bar{\theta}_4 - \bar{\theta}_2)\cos(\bar{\theta}_3)\right]$
actuate $u = u_{2+1/2}$ for $t \in [2h, 3h]$
$q_2 \leftarrow$ **sense** $q(2h)$
$\bar{q}_3 \leftarrow$ **solve** $D_2 L^d(q_1, q_2) + D_1 L^d(q_2, \bar{q}_3) + F_1^d(q_2, \bar{q}_3) = 0$
$\bar{q}_4 \leftarrow$ **solve** $D_2 L^d(q_2, \bar{q}_3) + D_1 L^d(\bar{q}_3, \bar{q}_4)$
    $+F_2^d(q_2, \bar{q}_3) + F_1^d(\bar{q}_3, \bar{q}_4) = 0$
$\bar{q}_5 \leftarrow$ **solve** $D_2 L^d(\bar{q}_3, \bar{q}_4) + D_1 L^d(\bar{q}_4, \bar{q}_5)$
    $+F_2^d(\bar{q}_3, \bar{q}_4) + F_1^d(\bar{q}_4, \bar{q}_5) = 0$
$u_{3+1/2} \leftarrow \frac{\gamma\kappa}{2h^2}\left[(\bar{\theta}_4 - \theta_2)\cos(\bar{\theta}_3) - (\bar{\theta}_5 - \bar{\theta}_3)\cos(\bar{\theta}_4)\right]$
actuate $u = u_{3+1/2}$ for $t \in [3h, 4h]$
for $k = 4$ to $(T_f/h - 1)$ **do**
  $q_{k-1} \leftarrow$ **sense** $q((k-1)h)$
  $\bar{q}_k \leftarrow$ **solve** $D_2 L^d(q_{k-2}, q_{k-1}) + D_1 L^d(q_{k-1}, \bar{q}_k)$
    $+F_2^d(q_{k-2}, q_{k-1}) + F_1^d(q_{k-1}, \bar{q}_k) = 0$
  $\bar{q}_{k+1} \leftarrow$ **solve** $D_2 L^d(q_{k-1}, \bar{q}_k) + D_1 L^d(\bar{q}_k, \bar{q}_{k+1})$
    $+F_2^d(q_{k-1}, \bar{q}_k) + F_1^d(\bar{q}_k, \bar{q}_{k+1}) = 0$
  $\bar{q}_{k+2} \leftarrow$ **solve** $D_2 L^d(\bar{q}_k, \bar{q}_{k+1}) + D_1 L^d(\bar{q}_{k+1}, \bar{q}_{k+2})$
    $+F_2^d(\bar{q}_k, \bar{q}_{k+1}) + F_1^d(\bar{q}_{k+1}, \bar{q}_{k+2}) = 0$
  $u_{k+1/2} \leftarrow \frac{\gamma\kappa}{2h^2}[(\bar{\theta}_{k+1} - \theta_{k-1})\cos(\bar{\theta}_k)$
    $-(\bar{\theta}_{k+2} - \bar{\theta}_k)\cos(\bar{\theta}_{k+1})]$
  actuate $u = u_{k+1/2}$ for $t \in [kh, (k+1)h]$
end for

it to compute a symmetric finite difference approximation to the control force at $t = h/2$ using the approximation

$$u(h/2) = -\left.\frac{d}{dt}\gamma\tau(\theta)\dot{\theta}\right|_{t=h/2}$$
$$\approx \frac{\gamma\kappa}{2h^2}[(\bar{\theta}_1 - \theta_{-1})\cos(\bar{\theta}_0) - (\bar{\theta}_2 - \bar{\theta}_0)\cos(\bar{\theta}_1)],$$

where the overbar indicates that the position variable is being estimated by the numerical model. This control is then applied as a constant control input for the time interval $[0, h]$. If the three forward solves can be computed within the time interval h, then this algorithm can be implemented in real-time.

The initialization of the discrete controller is somewhat involved, since the system is unforced during the time interval $[0, 2h]$ while the controller senses the initial states, and computes the appropriate control forces. As a consequence, we initially have to solve a combination of the discrete Euler–Lagrange equations and the forced discrete Euler–Lagrange equations to estimate the evolution of the system, until the feedback actuation comes fully online.

The numerical simulation of the digital controller is shown in Figure 3. We see that the system is asymptotically stabilized in the θ variable. The use of a piecewise constant control introduces dissipation-like effects, which are reduced as the time-step is decreased. The nature of this behavior will be studied in a forthcoming publication.

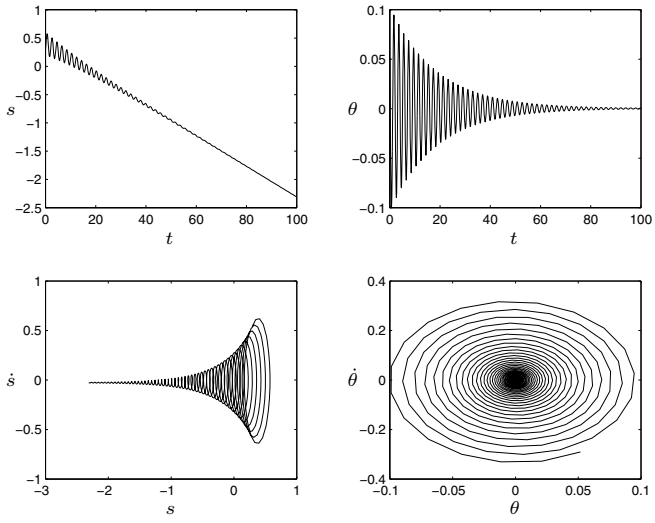


Fig. 3. Real-time piecewise constant model predictive controller

VIII. Conclusions

In this paper we have introduced the method of controlled Lagrangians for discrete systems and have shown that it leads to an effective numerical implementation for stabilization in the case of the discrete cart-pendulum model. The method in this paper is related to other discrete methods in control that have a long history; recent papers that use discrete mechanics in the context of optimal control and celestial navigation are [9], [13], and [19]. The full theory of discrete controlled Lagrangians will be developed in a forthcoming paper.

IX. Acknowledgments

The authors would like to thank the reviewers for helpful remarks. The research of AMB was supported by NSF grants DMS-0103895, DMS-030583, and CMS-0408542. The research of ML was partially supported by NSF grant DMS-0504747 and a University of Michigan Rackham faculty grant. The research of JEM was partially supported by ONR MURI Contract 00000916. The research of DVZ was partially supported by NSF grant DMS-0306017. AMB and DVZ would also like to thank the CRM, Barcelona, where part of this research was carried out.

References

[1] Auckly, D., L. Kapitanski, & W. White, Control of Nonlinear Underactuated Systems, *Commun. Pure Appl. Math.* **53**, 2000, 354–369.

[2] Bloch, A. M., *Nonholonomic Mechanics and Control*, Interdisciplinary Appl. Math. **24**, Springer-Verlag, 2003.

[3] Bloch, A. M., N. Leonard, & J. E. Marsden, Stabilization of Mechanical Systems Using Controlled Lagrangians, *Proc. CDC* **36**, 1997, 2356–2361.

[4] Bloch, A. M., N. Leonard, & J. E. Marsden, Matching and Stabilization by the Method of Controlled Lagrangians, *Proc. CDC* **37**, 1998, 1446–1451.

[5] Bloch, A. M., N. Leonard, & J. E. Marsden, Potential Shaping and the Method of Controlled Lagrangians, *Proc. CDC* **38**, 1999, 1652–1657.

[6] Bloch, A. M., N. E. Leonard, & J. E. Marsden, Controlled Lagrangians and the Stabilization of Mechanical Systems I: The First Matching Theorem, *IEEE Trans. on Systems and Control* **45**, 2000, 2253–2270.

[7] Bloch, A. M., D-E. Chang, N. E. Leonard, & J. E. Marsden, Controlled Lagrangians and the Stabilization of Mechanical Systems II: Potential Shaping, *Trans. IEEE on Autom. Contr.* **46**, 2001, 1556–1571.

[8] Chang, D-E., A. M. Bloch, N. E. Leonard, J. E. Marsden, & C. Woolsey, The Equivalence of Controlled Lagrangian and Controlled Hamiltonian Systems, *Control and the Calculus of Variations (special issue dedicated to J. L. Lions)* **8**, 2002, 393–422.

[9] Guibout, V. & A. M. Bloch, A Discrete Maximum Principle for Solving Optimal Control Problems, *Proc. CDC* **43**, 2004, 1806–1811.

[10] Hamberg, J., General Matching Conditions in the Theory of Controlled Lagrangians, *Proc. CDC* **38**, 1999, 2519–2523.

[11] Hamberg, J., Controlled Lagrangians, Symmetries and Conditions for Strong Matching, *In: Lagrangian and Hamiltonian Methods for Nonlinear Control,* Elsevier, 2000.

[12] Jalnapurkar, S. M., M. Leok, J. E. Marsden & M. West, Discrete Routh Reduction, 2005, arXiv:math.NA/0508330.

[13] Junge, O., J. Marsden, & S. Ober-Blbaum, Discrete Mechanics and Optimal Control, *Proc. of the 16th IFAC World Congress*, 2005 (to appear).

[14] Kane, C., J. E. Marsden, M. Ortiz, & M. West, Variational Integrators and the Newmark Algorithm for Conservative and Dissipative Mechanical Systems, *Int. J. Numer. Math. Eng.* **49**, 2000, 1295–1325.

[15] Leok, M., J. E. Marsden, & A. D. Weinstein, A Discrete Theory of Connections on Principal Bundles, 2005, arXiv:math.DG/0508338.

[16] Marsden, J. E., *Lectures on Mechanics,* London Mathematical Society Lecture Note Series **174**, Cambridge University Press, 1992.

[17] Marsden, J. E. & M. West, Discrete Mechanics and Variational Integrators, *Acta Numerica* **10**, 2001, 357–514.

[18] Maschke, B., R. Ortega, & A. van der Schaft, Energy-Based Lyapunov Functions for Forced Hamiltonian Systems with Dissipation, *IEEE Trans. Automat. Control* **45**, 2001, 1498–1502.

[19] Sanyal, A., J. Shen, N. H. McClamroch, & A. M. Bloch, Stability and Stabilization of Relative Equilibria of the Dumbbell Satellite in Central Gravity, 2005, *Journal of the American Institute of Aeronautics and Astronautics,* (to appear).

[20] Zenkov, D. V., A. M. Bloch, N. E. Leonard, & J. E. Marsden, Matching and Stabilization of Low-Dimensional Nonholonomic Systems, *Proc. CDC* **39**, 2000, 1289–1295.

[21] Zenkov, D. V., A. M. Bloch, & J. E. Marsden, Flat Nonholonomic Matching, *Proc. ACC*, 2002, 2812–2817.

Attitude Maneuvers of a Rigid Spacecraft in a Circular Orbit

Taeyoung Lee*†, N. Harris McClamroch†
Department of Aerospace Engineering
University of Michigan, Ann Arbor, MI 48109
{tylee, nhm}@umich.edu

Melvin Leok*
Department of Mathematics
University of Michigan, Ann Arbor, MI 48109
mleok@umich.edu

***Abstract*— A global model is presented that can be used to study attitude maneuvers of a rigid spacecraft in a circular orbit about a large central body. The model includes gravity gradient effects that arise from the non-uniform gravity field and characterizes the spacecraft attitude with respect to the uniformly rotating local vertical local horizontal coordinate frame. An accurate computational approach for solving a nonlinear boundary value problem is proposed, assuming that control torque impulses can be applied at initiation and at termination of the maneuver. If the terminal attitude condition is relaxed, then an accurate computational approach for solving the minimal impulse optimal control problem is presented. Since the attitude is represented by a rotation matrix, this approach avoids any singularity or ambiguity arising from other attitude representations such as Euler angles or quaternions.**

I. Introduction

The attitude dynamics of an uncontrolled rigid spacecraft in a circular orbit about a large central body, including gravity gradient effects, have been extensively studied; see [1], [2]. There are 24 distinct relative equilibria for which the principal axes are exactly aligned with the local vertical local horizontal (LVLH) axes, and the spacecraft angular velocity is identical to the orbital angular velocity of the LVLH coordinate frame. Linear rotational equations of motion that describe small perturbations from any relative equilibrium solutions are well known. Linear attitude control of a rigid spacecraft in a circular orbit, including linear gravity gradient effects, has also been addressed in [2]. However, linear controllers have the limitation that they are only applicable to small attitude change maneuvers.

The emphasis in this paper is on large angle attitude maneuvers of a rigid spacecraft in a circular orbit about a large central body, including gravity gradient effects. A nonlinear, globally defined model is introduced. This model describes the attitude of the spacecraft, relative to the uniformly rotating LVLH coordinate frame, by a rotation matrix. The model includes gravity gradient terms that reflect the rotation of the LVLH frame, and terms that reflect control input torques. In the problems studied in this paper, independent impulsive control torques can be applied about each principal axis. Gravity gradient moments are significant in Earth orbits for orbit altitudes between $400\,\mathrm{km}$ and $40,000\,\mathrm{km}$ and for attitude maneuver times that are not small compared with the orbital period.

*This research has been supported in part by NSF under grant DMS-0504747, and by a grant from the Rackham Graduate School, University of Michigan.

†This research has been supported in part by NSF under grant ECS-0244977.

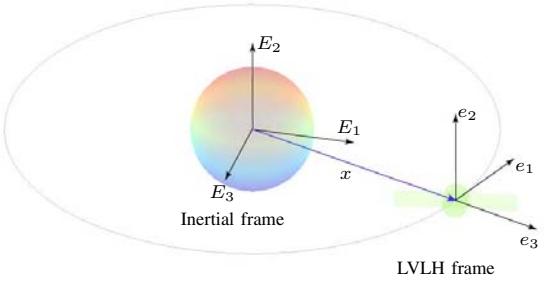


Fig. 1. Coordinate frames

We study open loop attitude maneuvers that can be accomplished by using two impulsive torque controls, one occurring at the initial time and one occurring at the terminal time of the maneuver. The attitude motion in between the initial time and the final time is uncontrolled. Two classes of attitude maneuver problems are studied. In Section III, a terminal attitude is specified so that there is a unique impulse sequence satisfying the boundary conditions. This problem can be solved computationally use a root finding algorithm. In Section IV, the specified terminal attitude condition is relaxed such that there are many impulse sequences that satisfy the boundary conditions. In this case we seek the minimum total impulse sequence that satisfies the boundary conditions. This problem can be solved computationally using a constrained minimization algorithm.

For both classes of attitude maneuver problems, our contribution is to demonstrate how sensitivity derivatives, used in the computational algorithms, can be determined effectively and accurately such that they satisfy the global geometry of the problem. Since the attitude is represented by a rotation matrix in the special orthogonal group $\mathrm{SO}(3)$, and the sensitivity derivatives are expressed in terms of the Lie algebra $\mathfrak{so}(3)$, this approach completely avoids singularities or ambiguities that arise from other representations of the rotation group, such as Euler angles and quaternions.

II. Rigid spacecraft models in a circular orbit

We assume that a rigid spacecraft is on a circular orbit with a constant orbital angular velocity $\omega_0 \in \mathbb{R}$. In this section, the continuous equations of motion and a geometric numerical integrator, referred to as Lie group variational integrator, for the attitude maneuver of a spacecraft in a circular orbit are given following the results of [3]. We identify $\mathrm{TSO}(3) \simeq$

$\mathrm{SO}(3)\times\mathfrak{so}(3)$ by left translation, and we identify $\mathfrak{so}(3) \simeq \mathbb{R}^3$ by an isomorphism $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$.

Continuous equations of motion: We define three rotation matrices in $\mathrm{SO}(3)$;

R^{bi} : from the body fixed frame to the inertial frame,

R^{li} : from the LVLH frame to the inertial frame,

R^{bl} : from the body fixed frame to the LVLH frame,

where the inertial frame and the LVLH frame are illustrated in Fig. 1. Thus, $R^{bl} = {R^{li}}^T R^{bi}$.

The on-orbit spacecraft equations of motion are given by

$$\dot{\Pi} + \Omega \times \Pi = M^g, \tag{1}$$

$$\dot{R}^{bi} = R^{bi} S(\Omega), \tag{2}$$

$$\dot{R}^{li} = R^{li} S(\omega_0 e_2), \tag{3}$$

$$\dot{R}^{bl} = R^{bl} S(\Omega - \omega_0 {R^{bl}}^T e_2), \tag{4}$$

where $\Pi, \Omega \in \mathbb{R}^3$ are the angular momentum and the angular velocity of the spacecraft expressed in the body fixed frame, respectively, and $M^g \in \mathbb{R}^3$ is the gravity gradient moment. The isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)$ is defined such that $S(x)y = x \times y$ for any $x, y \in \mathbb{R}^3$. Since the orbital angular velocity ω_0 is constant, the solution of (3) is given by $R^{li}(t) = R^{li}(0)e^{S(\omega_0 e_2)t}$.

Gravity gradient moment: The gravity gradient moment is derived in [2]. We present an alternative way to obtain the gravity gradient moment directly using the gravity potential;

$$U = -\int_{\mathcal{B}} \frac{GM}{\|x + R^{bi}\rho\|}\, dm,$$

where $x \in \mathbb{R}^3$ is the position of the spacecraft in the inertial frame, and $\rho \in \mathbb{R}^3$ is a vector from the center of mass of the spacecraft to a mass element in the body fixed frame. G is the gravitational constant and M is the mass of the Earth.

From [3], the gravity gradient moment M^g can be determined by using the following relationship;

$$S(M^g) = \frac{\partial U}{\partial R^{bi}}^T R^{bi} - {R^{bi}}^T \frac{\partial U}{\partial R^{bi}}. \tag{5}$$

We derive a closed form for M^g from (5), by assuming that the spacecraft is on a circular orbit so that the norm of x is constant, and the size of the spacecraft is much smaller than the size of the orbit.

As shown in Fig. 1, the coordinate of the spacecraft position in the LVLH frame is $r_0 e_3$, where $r_0 \in \mathbb{R}$ is the radius of the circular orbit, and $e_3 = [0,\, 0,\, 1]^T$. Therefore, the position of the spacecraft in the inertial frame is given by $x = r_0 R^{li} e_3$. Using this expression,

$$\begin{aligned}\frac{\partial U}{\partial R^{bi}} &= \int_{\mathcal{B}} \frac{GM\, r_0 R^{li}\, e_3 \rho^T}{\|\, r_0 e_3 + R^{bl}\rho\,\|^3}\, dm, \\ &= \frac{GM}{r_0} \int_{\mathcal{B}} \frac{\left(R^{li} e_3 \hat{\rho}^T\right) \frac{\|\rho\|}{r_0}}{\left[1 + 2\left(e_3^T R^{bl} \hat{\rho}\right) \frac{\|\rho\|}{r_0} + \frac{\|\rho\|^2}{r_0^2}\right]^{\frac{3}{2}}}\, dm,\end{aligned}$$

where $\hat{\rho} = \frac{\rho}{\|\rho\|} \in \mathbb{R}^3$ is the unit vector along the direction of ρ. Assume that the size of the spacecraft is significantly smaller than the size of the orbit, i.e. $\frac{\|\rho\|}{r_0} \ll 1$. Using a Taylor series expansion, we obtain the 2nd order approximation.

$$\frac{\partial U}{\partial R^{bi}} = \frac{GM}{r_0} \int_{\mathcal{B}} R^{li} e_3 \hat{\rho}^T \left\{ \frac{\|\rho\|}{r_0} - 3 e_3^T R^{bl} \hat{\rho} \frac{\|\rho\|^2}{r_0^2} \right\} dm.$$

Since the body fixed frame is located at the mass center of the spacecraft, $\int_{\mathcal{B}} \rho\ dm = 0$. Therefore, the first term in the above equation vanishes. Because $e_3^T R^{bl} \rho$ is a scalar quantity, we can rewrite the above equation as

$$\frac{\partial U}{\partial R^{bi}} = -3\omega_0^2 R^{li} e_3 e_3^T R^{bl} \left(\frac{1}{2} \mathrm{tr}[J]\, I_{3\times 3} - J \right), \tag{6}$$

where $\omega_0 = \sqrt{\frac{GM}{r_0^3}} \in \mathbb{R}$ is the orbital angular velocity, and $J \in \mathbb{R}^{3\times 3}$ is the moment of inertia matrix of the spacecraft. Substituting (6) into (5), and using the property $S(x \times y) = yx^T - xy^T$ for $x, y \in \mathbb{R}^3$, we obtain an expression for the gravity gradient moment as follows.

$$M^g = 3\omega_0^2 {R^{bl}}^T e_3 \times J {R^{bl}}^T e_3. \tag{7}$$

Discrete equations of motion: In the continuous equations of motion, the structure of (2), (3), and (4) ensures that R^{bi}, R^{li}, and R^{bl} evolve on the special orthogonal group, $\mathrm{SO}(3)$. However, general numerical integration methods, including the popular Runge-Kutta methods, do not preserve the orthogonality property of this group. For example, if we integrate (4) by a typical Runge-Kutta scheme, the quantity ${R^{bl}}^T R^{bl}$ inevitably drifts from the identity matrix as the simulation time increases.

The rotation matrix is commonly parameterized by Euler angles or quaternions. These attitude kinematics equations can be numerically integrated and are used to recompute the rotation matrix. However, Euler angles are not global expressions of the attitude since they have associated singularities. The analytical expressions for sensitivities are hard to develop since many trigonometric terms are encountered. Quaternions have no singularity, but quaternions must lie on the three sphere $\mathbb{S}^3$. General numerical integration methods do not preserve the unit length of a quaternion. Therefore, quaternions have the same numerical drift problem as rotation matrices. Furthermore, quaternions, which are diffeomorphic to $\mathrm{SU}(2)$, double covers $\mathrm{SO}(3)$. So there are inevitable ambiguities in expressing the attitude.

These cause significant inaccuracies in numerical simulations based on quaternions and Euler angles. In particular, the gravity gradient moment, as given in (7), depends on R^{bl} directly, and consequently, errors in computing R^{bl} cause errors in the gravity gradient moment. These effects are more pronounced when the simulation time is large.

Lie group variational integrators preserve the orthonormal structure of $\mathrm{SO}(3)$ without any reprojection or parametrization. They also conserve the momentum map, and the symplectic property of rigid body dynamics. In addition, the total energy is well-behaved, as it only oscillates in a bounded fashion about its true value. So, Lie group variational integrators are geometrically exact. Using the results

given in [3], a Lie group variational integrator for the attitude dynamics of a spacecraft in a circular orbit are given by

$$\Pi_{k+1} = F_k^T \Pi_k + \frac{h}{2} F_k^T M_k^g + \frac{h}{2} M_{k+1}^g, \tag{8}$$

$$hS(\Pi_k + \frac{h}{2} M_k^g) = F_k J_d - J_d F_k^T, \tag{9}$$

$$R_{k+1}^{bi} = R_k^{bi} F_k, \tag{10}$$

$$R_{k+1}^{li} = R_k^{li} e^{S(\omega_0 e_2)h}, \tag{11}$$

$$R_{k+1}^{bl} = e^{-S(\omega_0 e_2)h} R_k^{bl} F_k, \tag{12}$$

where the subscript k denotes variables at the kth time step, and $h \in \mathbb{R}$ is the integration step size. The matrix $J_d \in \mathbb{R}^{3\times 3}$ is a nonstandard moment of inertia matrix defined by $J_d = \frac{1}{2}\mathrm{tr}[J]\, I_{3\times 3} - J$.

The matrix $F_k = R_k^{bi^T} R_{k+1}^{bi} \in \mathrm{SO}(3)$ is the relative attitude between integration steps, and it is obtained by solving (9). Since F_k and $e^{S(\omega_0 e_2)h}$ are in $\mathrm{SO}(3)$, and $\mathrm{SO}(3)$ is closed under matrix multiplication, R_k^{bi}, R_k^{li}, and R_k^{bl} evolve in $\mathrm{SO}(3)$ automatically for all k according to (10), (11), and (12). The actual computation of F_k is done in the Lie algebra $\mathfrak{so}(3)$ of dimension 3, and the rotation matrices are updated by multiplication with the exponential of a skew-symmetric matrix. So, Lie group variational integrators are numerically efficient, and there is no excessive computational burden in updating the 9 elements of the rotation matrix. The properties of these discrete equations of motion are discussed more explicitly in [3] and [4].

Since ω_0 is constant, the solution of (11) is given by

$$R_k^{li} = R_0^{li} e^{S(\omega_0 e_2)kh}. \tag{13}$$

III. Spacecraft attitude maneuvers

A. Problem formulation

A two point boundary value problem is formulated for spacecraft rest-to-rest maneuver between two given attitudes for a fixed maneuver time. This is a Lambert type boundary value problem on $\mathrm{SO}(3)$. The initial attitude and the desired terminal attitude are expressed as rotation matrices with respect to the LVLH frame, namely R_0^{bl}, $R_{N_d}^{bl} \in \mathrm{SO}(3)$. Two impulsive control torques are applied at the initial time and the terminal time. We assume that the control torques are purely impulsive, which means that each impulse changes the angular velocity of the spacecraft instantaneously, but it does not have any effect on the attitude of the spacecraft at that instant. The rotational motion of the spacecraft between the initial time and the terminal time is uncontrolled. This assumption is reasonable when the operating time of the spacecraft control moment actuator is much smaller than the fixed maneuver time.

We define this boundary value problem directly on $\mathrm{SO}(3)$ instead of using parameterizations such as Euler angles and quaternions. This approach allows us to define the attitude of the spacecraft globally without singularities and ambiguities. We use the discrete equations of motion given in (8)–(12) for the problem formulation and for the following analysis.

We transform this two point boundary value problem into a nonlinear root finding problem. For a given initial angular momentum Π_0 and an initial attitude of the spacecraft R_0^{bl}, the terminal angular momentum Π_N and the terminal attitude R_N^{bl} are determined by the discrete equations of motion. By choosing the initial angular momentum so that the terminal attitude of the spacecraft is equal to the desired attitude, i.e. $R_N^{bl} = R_{N_d}^{bl}$, we obtain the initial impulse and the terminal impulse. Thus, the nonlinear boundary value problem for the spacecraft attitude maneuver is formulated as

$$\begin{gathered}\text{given} : R_0^{bl}, R_{N_d}^{bl}, N \\ \text{find} : \Pi_0 \\ \text{such that } R_N^{bl} = R_{N_d}^{bl} \text{ subject to (8)–(12),}\end{gathered}$$

where $N \in \mathbb{N}$ is the number of integration steps determined by $N = \frac{T}{h}$ for the fixed maneuver time T and the fixed integration step size h.

B. Computational approach

We solve a sequence of linear boundary value problems whose solutions converge to the solution of the nonlinear boundary value problem.

Linearization: The equations of motion are linearized about a given trajectory, and they are expressed in terms of the Lie algebra $\mathfrak{so}(3)$. Consider small perturbations from a given trajectory denoted by Π_k^ϵ, $R_k^{bi,\epsilon}$, $R_k^{bl,\epsilon}$, F_k^ϵ:

$$\Pi_k^\epsilon = \Pi_k + \epsilon \delta\Pi_k, \tag{14}$$

$$R_k^{bi,\epsilon} = R_k^{bi} + \epsilon \delta R_k^{bi} + \mathcal{O}(\epsilon^2), \tag{15}$$

$$R_k^{bl,\epsilon} = R_k^{bl} + \epsilon \delta R_k^{bl} + \mathcal{O}(\epsilon^2), \tag{16}$$

$$F_k^\epsilon = F_k + \epsilon \delta F_k + \mathcal{O}(\epsilon^2), \tag{17}$$

where $\epsilon \in \mathbb{R}$. Since the orbital angular velocity is constant, $\delta R_k^{li} = 0$.

The infinitesimal variation of the angular momentum $\delta\Pi_k$ can be expressed in $\mathbb{R}^3$. The variation of the rotation matrix $R_k^{bi,\epsilon} \in \mathrm{SO}(3)$ can be expressed as

$$R_k^{bi,\epsilon} = R_k^{bi} e^{\epsilon S(\zeta_k)},$$

where $\zeta_k \in \mathbb{R}^3$ and $S(\zeta_k) \in \mathfrak{so}(3)$ is a skew-symmetric matrix. Since the map $S(\cdot)$ is an isomorphism between $\mathfrak{so}(3)$ and $\mathbb{R}^3$, ζ_k is well defined. Then, the infinitesimal variation δR_k^{bi} is given by

$$\delta R_k^{bi} = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_k^{bi,\epsilon} = R_k^{bi} S(\zeta_k). \tag{18}$$

δF_k is obtained from definition (10), and (18) as

$$\begin{aligned}\delta F_k &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_k^{bi,\epsilon^T} R_{k+1}^{bi,\epsilon}, \\ &= -S(\zeta_k) F_k + F_k S(\zeta_{k+1}).\end{aligned} \tag{19}$$

Since $\delta R_k^{li} = 0$, δR_k^{bl} is given by

$$\delta R_k^{bl} = R_k^{li^T} \delta R_k^{bi} = R_k^{bl} S(\zeta_k). \tag{20}$$

In summary, equations (18), (19) and (20) describe the variations of rotation matrices in $\mathrm{SO}(3)$.

Substituting (14), (16), and (17) into (8), and (9), and ignoring higher-order terms, the linearized discrete equations of motion are given by

$$\begin{aligned}\delta\Pi_{k+1} &= \delta F_k^T \Pi_k + F_k^T \delta\Pi_k \\ &\quad + \frac{h}{2}\delta F_k^T M_k^g + \frac{h}{2} F_k^T \delta M_k^g + \frac{h}{2}\delta M_{k+1}^g, \end{aligned} \tag{21}$$

$$h S(\delta\Pi_k + \frac{h}{2}\delta M_k^g) = \delta F_k J_d - J_d \delta F_k^T, \tag{22}$$

where

$$\delta M_k^g = 3\omega_0^2 \left[\delta R_k^{bl^T} e_3 \times J R_k^{bl^T} e_3 + R_k^{bl^T} e_3 \times J \delta R_k^{bl^T} e_3 \right]. \tag{23}$$

These equations are not in standard form since (22) is an implicit equation in δF_k. By using (18), (19) and (20), we will obtain an explicit solution of (22), and rewrite the above equations in standard form.

Substituting (20) into (23), and using the property $S(x)y = -S(y)x$ for all $x, y \in \mathbb{R}^3$, δM_k^g can be written as

$$\begin{aligned}\delta M_k^g &= 3\omega_0^2 \Big[- S(J R_k^{bl^T} e_3) S(R_k^{bl^T} e_3) \\ &\qquad + S(R_k^{bl^T} e_3) J S(R_k^{bl^T} e_3) \Big] \zeta_k, \\ &= \mathcal{M}_k \zeta_k, \end{aligned} \tag{24}$$

where $\mathcal{M}_k \in \mathbb{R}^{3\times 3}$.

Using (19) and the property $S(Rx) = R S(x) R^T$ for $x \in \mathbb{R}^3$, $R \in \mathrm{SO}(3)$, the right hand side of (22) is written as

$$\begin{aligned}&\delta F_k J_d - J_d \delta F_k^T \\ &= - \left\{ S(\zeta_k) F_k J_d + J_d F_k^T S(\zeta_k) \right\} \\ &\quad + \left\{ S(F_k \zeta_{k+1}) F_k J_d + J_d F_k^T S(F_k \zeta_{k+1}) \right\}. \end{aligned} \tag{25}$$

Substituting (24) and (25) into (22), and using the property $S(x)A + A^T S(x) = S(\{\mathrm{tr}[A]\, I_{3\times 3} - A\}\, x)$ for $x \in \mathbb{R}^3$, $A \in \mathbb{R}^{3\times 3}$, (22) can be transformed into an equivalent vector form;

$$\begin{aligned} h\delta\Pi_k + \frac{h^2}{2}\mathcal{M}_k \zeta_k &= - \left\{ \mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d \right\} \zeta_k \\ &\quad + \left\{ \mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d \right\} F_k \zeta_{k+1}. \end{aligned}$$

Multiplying both sides by $F_k^T \left\{ \mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d \right\}^{-1}$ and rearranging, we obtain

$$\begin{aligned}\zeta_{k+1} &= F_k^T \left[\frac{h^2}{2} \left\{ \mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d \right\}^{-1} \mathcal{M}_k + I_{3\times 3} \right] \zeta_k \\ &\quad + h F_k^T \left\{ \mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d \right\}^{-1} \delta\Pi_k, \\ &\triangleq \mathcal{A}_k \zeta_k + \mathcal{B}_k \delta\Pi_k, \end{aligned} \tag{26}$$

where $\mathcal{A}_k, \mathcal{B}_k \in \mathbb{R}^{3\times 3}$. Substituting (26) into (19), δF_k is obtained as

$$\begin{aligned}\delta F_k &= -\eta_k F_k + F_k \eta_{k+1}, \\ &= -S(\zeta_k) F_k + F_k S(\zeta_{k+1}), \\ &= -S(\zeta_k) F_k + F_k S(\mathcal{A}_k \zeta_k + \mathcal{B}_k \delta\Pi_k). \end{aligned} \tag{27}$$

Equation (27) is an explicit solution of (22).

Now we rewrite (21) in the standard form of linear discrete equations of motion. Substituting (24), (27) into (21), and rearranging, we obtain

$$\begin{aligned}\delta\Pi_{k+1} &= \Big[F_k^T S(\Pi_k + \frac{h}{2} M_k^g) \left\{ -I_{3\times 3} + F_k \mathcal{A}_k \right\} \\ &\qquad + \frac{h}{2} F_k^T \mathcal{M}_k + \frac{h}{2} \mathcal{M}_{k+1} \mathcal{A}_k \Big] \zeta_k \\ &\quad + \left[S(F_k^T \left\{ \Pi_k + \frac{h}{2} M_k^g \right\}) \mathcal{B}_k + F_k^T + \frac{h}{2} \mathcal{M}_{k+1} \mathcal{B}_k \right] \delta\Pi_k, \\ &\triangleq \mathcal{C}_k \zeta_k + \mathcal{D}_k \delta\Pi_k, \end{aligned} \tag{28}$$

where $\mathcal{C}_k, \mathcal{D}_k \in \mathbb{R}^{3\times 3}$.

In summary, (26) and (28) are linear discrete equations equivalent to (21), (22) and (23), and they can be written as

$$\begin{aligned}\begin{bmatrix} \zeta_{k+1} \\ \delta\Pi_{k+1} \end{bmatrix} &= \begin{bmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{bmatrix} \begin{bmatrix} \zeta_k \\ \delta\Pi_k \end{bmatrix}, \\ &\triangleq A_k \begin{bmatrix} \zeta_k \\ \delta\Pi_k \end{bmatrix}, \end{aligned} \tag{29}$$

where $A_k \in \mathbb{R}^{6\times 6}$. Equation (29) is the linear discrete equation for perturbations of the attitude dynamics of a spacecraft in a circular orbit, expressed in terms of $\mathbb{R}^3 \simeq \mathfrak{so}(3)$. The important feature of (29) is that it is linearized in such a way that it respects the geometry of the special orthogonal group $\mathrm{SO}(3)$.

Linear boundary value problem: The solution of (29) is given by

$$\begin{aligned}\begin{bmatrix} \zeta_N \\ \delta\Pi_N \end{bmatrix} &= \left(\prod_{k=0}^{N-1} A_k \right) \begin{bmatrix} \zeta_0 \\ \delta\Pi_0 \end{bmatrix}, \\ &\triangleq \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} \zeta_0 \\ \delta\Pi_0 \end{bmatrix}, \end{aligned} \tag{30}$$

where $\Phi_{ij} \in \mathbb{R}^{3\times 3}$ for $i, j = 1, 2$.

For the given boundary value problem, $\zeta_0 = 0$ since the initial attitude of the spacecraft is given and fixed, and $\delta\Pi_N$ is free since the terminal angular momentum is compensated by the terminal impulse. Then, we obtain

$$\zeta_N = \Phi_{12} \delta\Pi_0.$$

This equation provides Φ_{12}, the sensitivity derivative of the terminal attitude with respect to a change in the initial angular momentum. It states that for a given trajectory, if we update the initial angular velocity by $\delta\Pi_0$, then the terminal attitude is changed from R_N^{bl} to $R_N^{bl} e^{S(\zeta_N)} = R_N^{bl} e^{S(\Phi_{12}\delta\Pi_0)}$.

We choose the change of the initial angular momentum so that the updated terminal attitude is equal to the desired terminal attitude; $R_N^{bl} e^{S(\Phi_{12}\delta\Pi_0)} = R_{N_d}^{bl}$, or equivalently,

$$\delta\Pi_0 = \Phi_{12}^{-1} S^{-1} \Big(\mathrm{logm} \Big(R_N^{bl^T} R_{N_d}^{bl} \Big) \Big), \tag{31}$$

where $S^{-1}(\cdot) : \mathfrak{so}(3) \mapsto \mathbb{R}^3$ is the inverse mapping of $S(\cdot)$, and logm denotes the matrix logarithm. Equation (31) provides a solution of the linear boundary value problem for the attitude dynamics of a spacecraft in a circular orbit, assuming that Φ_{12} is invertible.

Nonlinear boundary value problem: The linear boundary value problem is solved successively so that its solution converges to the solution of the nonlinear boundary value problem. A numerical algorithm is summarized as follows.

1: Set $\text{Error} = 2\epsilon_S$.
2: Guess an initial condition $\Pi_0^{(0)}$.
3: Set $i = 0$.
4: **while** $\text{Error} > \epsilon_S$.
5: Find $\Pi_k^{(i)}, R_k^{bl(i)}$ using $\Pi_0^{(i)}$ and (8), (9), (12).
6: Compute the error;
$\zeta_N^{(i)} = S^{-1}\Big(\text{logm}\Big({R_N^{bl(i)}}^T R_{N_d}^{bl}\Big)\Big)$, $\text{Error} = \Big\|\zeta_N^{(i)}\Big\|$.
7: Update the initial condition; $\Pi_0^{(i+1)} = \Pi_0^{(i)} + c\Phi_{12}^{-1}\zeta_N^{(i)}$.
8: Set $i = i + 1$.
9: **end while**

Here the superscript (i) denotes the ith iteration, and $\epsilon_S, c \in \mathbb{R}$ are a stopping criterion and a scaling factor, respectively.

This computational approach utilizes an exact and efficient method to compute sensitivity derivatives in the special orthogonal group $\text{SO}(3)$. The sensitivity derivatives are then used to solve the two point boundary value problem.

C. Numerical example

Three spacecraft rest-to-rest maneuvers between relative equilibrium attitudes are considered. The resulting motions are highly nonlinear, large angle maneuvers.

The mass, length and time dimensions are normalized. The moment of inertia of the spacecraft and simulation parameters are chosen as $J = \text{diag}\,[1,\, 2.8,\, 2]$, $\epsilon_s = 10^{-14}$, $c = 0.1$, $h = 0.001$. Each maneuver is completed in a quarter of the orbit, $T = \frac{\pi}{2}$. The boundary conditions and the corresponding computed impulsive control are as follows.

(i) Rotational maneuver about the LVLH axis e_1:

$$R_0^{bl} = I_{3\times 3}, \quad R_{N_d}^{bl} = \text{diag}\,[1, -1, -1]\,,$$
$$\Pi_0 = [2.116,\, 1.531,\, -1.782]^T\,,$$
$$-\Pi_N = [-2.116,\, 1.531,\, 1.782]^T\,.$$

(ii) Rotational maneuver about the LVLH axes e_1 and e_2:

$$R_0^{bl} = \text{diag}\,[1, -1, -1]\,, \quad R_{N_d}^{bl} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix},$$
$$\Pi_0 = [-1.323,\, 1.798,\, 0.932]^T\,,$$
$$-\Pi_N = [0.397,\, -1.586,\, -1.310]^T\,.$$

(iii) Rotational maneuver about the LVLH axes e_2 and e_3:

$$R_0^{bl} = \text{diag}\,[1, -1, -1]\,, \quad R_{N_d}^{bl} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
$$\Pi_0 = [1.047,\, 0.437,\, 2.800]^T\,,$$
$$-\Pi_N = [-1.416,\, -1.761,\, 1.159]^T\,.$$

Fig. 2 shows the attitude maneuver of the spacecraft, and the angular velocity response for each case. (Simple animations which show these spacecraft maneuvers can be found at `http://www.umich.edu/~tylee`.)

IV. Optimal spacecraft attitude maneuvers

A. Problem formulation

An optimization problem is formulated as a rest-to-rest maneuver of an axially-symmetric spacecraft from a given initial attitude to a given terminal reduced attitude for a fixed maneuver time. The initial attitude is expressed by a rotation matrix with respect to the LVLH frame, namely R_0^{bl}. The terminal desired attitude is given by the reduced attitude, $\Lambda_{N_d} = R_N^{bl} e_3 \in \mathbb{S}^2$. This reduced attitude represents the direction of the spacecraft axis of symmetry e_3 in the LVLH frame. Two impulsive control moments are applied at the initial time and the terminal time, and the maneuver of the spacecraft between the initial time and the terminal time is uncontrolled.

In the problem studied in section III, the initial parameter Π_0 is exactly prescribed by the constraint $R_{N_d}^{bl} = R_N^{bl}$, and the discrete dynamics. In this section, we relax the terminal constraint by only specifying it up to a rotation about the axis of symmetry of the spacecraft. Then, we can formulate an optimal attitude maneuver problem.

The performance index is the sum of the magnitudes of the initial impulse and the terminal impulse. Equivalently, one can minimize the change in the initial angular momentum and the change in the terminal angular momentum. Since the initial attitude and the terminal time are fixed, Π_N and Λ_N can be considered as functions of Π_0 through the discrete equations of motion. The optimization problem is equivalent to

$$\text{given} : R_0^{bl}, \Lambda_{N_d}, N,$$
$$\min_{\Pi_0} \mathcal{J} = \left\|\Pi_0 - \omega_0 R_0^T J e_2\right\| + \left\|\omega_0 R_N^T J e_2 - \Pi_N\right\|,$$
$$= \|H_0\| + \|H_N\|\,,$$
$$\text{such that } \mathcal{C} = \|\Lambda_N - \Lambda_{N_d}\|^2 = 0,$$
$$\text{subject to (8)–(12).}$$

B. Computational approach

This problem is optimized by the Sequential Quadratic Programming (SQP) method using analytical expressions for the sensitivity derivatives of the performance index and of the constraint equation.

The variation of the performance index is

$$\delta\mathcal{J} = \frac{H_0^T}{\|H_0\|}\delta\Pi_0 + \frac{H_N^T}{\|H_N\|}\left\{-\omega_0 S(\zeta_N) R_N^T J e_2 - \delta\Pi_N\right\}.$$

Since the initial attitude is given and fixed, the perturbation of the initial attitude ζ_0 is zero. Therefore $\delta\Pi_N = \Phi_{22}\delta\Pi_0$ and $\zeta_N = \Phi_{12}\delta\Pi_0$ from (30). Then, $\delta\mathcal{J}$ is given by

$$\delta\mathcal{J} = \left[\frac{H_0^T}{\|H_0\|} + \frac{H_N^T}{\|H_N\|}\left\{\omega_0 S(R_N^T J e_2)\Phi_{12} - \Phi_{22}\right\}\right]\delta\Pi_0. \tag{32}$$

Since Λ_{N_d} is fixed and $\Lambda_N \in \mathbb{S}^2$, the variation of the constraint can be written as

$$\delta\mathcal{C} = -2\Lambda_{N_d}^T\delta\Lambda_N = 2\Lambda_{N_d}^T R_N^{bl} S(e_3)\zeta_N,$$

where $\zeta_N = \Phi_{12}\delta\Pi_0$ from (30). Thus, $\delta\mathcal{C}$ is

$$\delta\mathcal{C} = \left[2\Lambda_d^T R_N^{bl} S(e_3)\Phi_{12}\right]\delta\Pi_0. \tag{33}$$

Equations (32) and (33) are analytical expressions for the sensitivity derivatives of the performance index and the constraint.

C. Numerical example

The mass, length and time dimensions are normalized. The moment of inertia of the spacecraft is chosen as $J = \mathrm{diag}\,[3,\, 3,\, 2]$, so that e_3 is the axis of symmetry of the spacecraft.

The desired maneuver is to rotate the axis of symmetry from the radial direction to the normal to the orbital plane during a quarter orbit. The boundary conditions are given by

$$R_0^{bl} = \mathrm{diag}\,[1,\, -1,\, -1]\,,\ \Lambda_{N_d} = [0,\, -1,\, 0]^T\,.$$

We use MATLAB's `fmincon` function as an optimization tool. The sensitivity derivatives of the performance index and the constraint are provided by (32) and (33). The initial guess of the initial angular momentum is chosen as $\Pi_0^{(0)} = J\,[1,\, 1,\, 0]^T$. The optimized performance index and the violation of constraints are $\mathcal{J} = 6.771$, $\mathcal{C} = 4.80 \times 10^{-14}$. The corresponding angular momenta and the terminal attitude are

$$\Pi_0 = [-2.915,\, -2.347,\, -2.734]^T\,,$$
$$-\Pi_N = [-0.343,\, 2.686,\, 2.734]^T\,,$$
$$R_N^{bl} = \begin{bmatrix} 0.633 & 0.733 & 0.000 \\ 0.000 & 0.000 & -1.000 \\ -0.773 & 0.633 & 0.0000 \end{bmatrix},$$

so that $\Lambda_N = R_N^{bl} e_3 = [0,\, 0,\, -1]^T = \Lambda_{N_d}$. Fig. 3 shows the optimal maneuver of the spacecraft.

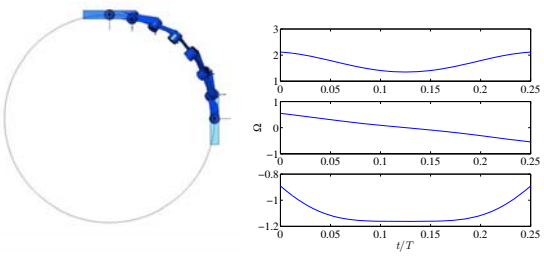
V. Conclusion

A global model for a rigid spacecraft in a circular orbit about a large central body is presented. This model includes gravity gradient effects that arise from the non-uniform gravity field.

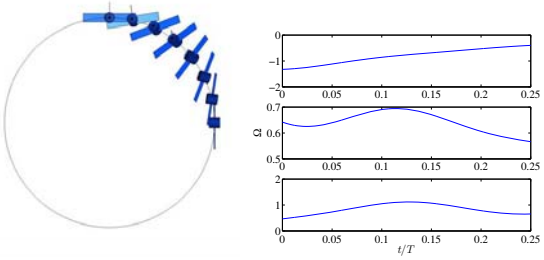
The sensitivity derivatives for attitude dynamics of a rigid body are derived while satisfying the global geometry of the problem. Accurate computational approaches for solving a nonlinear boundary value problem and the minimal impulse optimal control problem for spacecraft attitude maneuvers are studied using sensitivity derivatives.

The attitude dynamics are represented by a rotation matrix in the Lie group $\mathrm{SO}(3)$, and it is updated by Lie group variational integrators that preserve the structure of $\mathrm{SO}(3)$ as well as other geometric invariants of motion. The sensitivity derivatives are expressed in terms of the Lie algebra $\mathfrak{so}(3)$. This approach completely avoids the singularities and ambiguities associated with Euler angles or quaternions, and it leads to a geometrically exact and numerically efficient method for rigid body attitude dynamics problems.

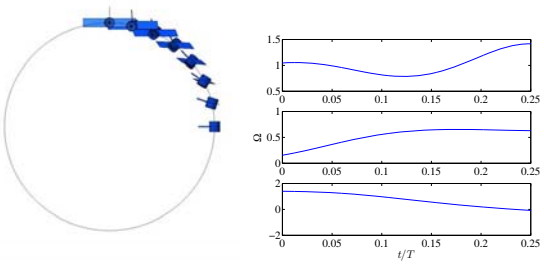
Although the development in this paper includes a gravity gradient moment and the rotation of the LVLH frame, the results presented reduce to the case of a free rigid body if $\omega_0 = 0$. That is, the computational approach suggested applies directly to attitude maneuvers of the free rigid body.



(a) Rotation about the LVLH axis e_1



(b) Rotation about the LVLH axes e_1 and e_2



(c) Rotation about the LVLH axes e_2 and e_3

Fig. 2. Spacecraft attitude maneuvers

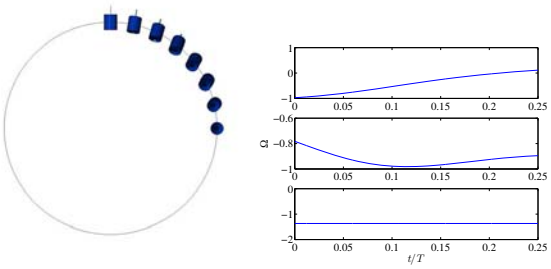


Fig. 3. Optimal spacecraft attitude maneuver

References

[1] P. C. Hughes, *Spacecraft attitude dynamcis*. John Wiley & Sons, 1986.

[2] B. Wie, *Space Vehicle Dynamics and Control*. AIAA, 1998.

[3] T. Lee, M. Leok, and N. H. McClamroch, "A Lie group variational integrator for the attitude dynamics of a rigid body with application to the 3D pendulum," in *Proceedings of the IEEE Conference on Control Application*, Toronto, Canada, Aug 2005, pp. 962–967.

[4] ——, "Lie group variational integrators for the Full Body problem," *Computer Methods in Applied Mechanics and Engineering*, submitted, Available: http://arxiv.org/abs/math.NA/0508365.

AIAA/AAS Astrodynamics Specialist Conference and Exhibit
21 - 24 August 2006, Keystone, Colorado

AIAA 2006-6289

Polyhedral Potential and Variational Integrator Computation of the Full Two Body Problem

Eugene G. Fahnestock*, Taeyoung Lee*, Melvin Leok†,
N. Harris McClamroch‡, and Daniel J. Scheeres§
University of Michigan, Ann Arbor, Michigan, USA

We present a combination of tools which allows for investigation of the coupled orbital and rotational dynamics of two rigid bodies with nearly arbitrary shape and mass distribution, under the influence of their mutual gravitational potential. Methods for calculating that mutual potential and resulting forces and moments for a polyhedral body representation are simple and efficient. Discrete equations of motion, referred to as the Lie Group Variational Integrator (LGVI), preserve the structure of the configuration space, SE(3), as well as the geometric features represented by the total energy and the total angular momentum. The synthesis of these approaches allows us to simulate the full two body problem accurately and efficiently. Simulation results are given for two octahedral rigid bodies for comparison with other integration methods and to show the qualities of the results thus obtained. A significant improvement is seen over other integration methods while correctly capturing the interesting effects of strong orbit and attitude dynamics coupling, in multiple scenarios.

Nomenclature

E	Total energy of system
F	Update matrix, $\in$ SO(3), for relative attitude rotation matrix R
F_2	Update matrix, $\in$ SO(3), for rotation matrix R_2
G	Universal gravitational constant
J_1	Standard moment of inertia matrix of $\mathcal{B}_1$ expressed in its own frame
J_2	Standard moment of inertia matrix of $\mathcal{B}_2$ expressed in its own frame
J_R	Standard moment of inertia matrix of $\mathcal{B}_1$ expressed in frame fixed to $\mathcal{B}_2$
J_{dR}	Non-standard moment of inertia matrix of $\mathcal{B}_1$ expressed in frame fixed to $\mathcal{B}_2$
J_{d_1}	Non-standard moment of inertia matrix of $\mathcal{B}_1$ expressed in its own frame
J_{d_2}	Non-standard moment of inertia matrix of $\mathcal{B}_2$ expressed in its own frame
KE	Kinetic energy of system
L	Reduced Lagrangian for system
M	Moment due to mutual gravitational potential
N	Number of time-steps of length h to go from the initial time to the final time in simulation
PE	Potential energy of system
R	Relative attitude rotation matrix, $\in$ SO(3), mapping from the frame fixed to $\mathcal{B}_1$ to the frame fixed to $\mathcal{B}_2$
R_1	Rotation matrix, $\in$ SO(3), mapping from the frame fixed to $\mathcal{B}_1$ to the inertial reference frame
R_2	Rotation matrix, $\in$ SO(3), mapping from the frame fixed to $\mathcal{B}_2$ to the inertial reference frame
$S(.)$	Denotes the cross-product operation matrix.

*Graduate Student, Department of Aerospace Engineering, University of Michigan, 2008 FXB Building, 1320 Beal Avenue, Ann Arbor, MI 48109, AIAA Student Member.
†Assistant Professor, Department of Mathematics, University of Michigan.
‡Professor, Department of Aerospace Engineering, University of Michigan, AIAA Senior Member.
§Associate Professor, Department of Aerospace Engineering, University of Michigan, AIAA Associate Fellow.

T Jacobian determinant of the matrix containing coordinates of non-centroid vertices of a simplex
U Mutual gravitational potential
$\hat{U}_{\#}$ The #'th Legendre series term of increasing order within mutual gravitational potential equation
V Relative velocity vector of $\mathcal{B}_1$ with respect to $\mathcal{B}_2$ expressed in the frame fixed to $\mathcal{B}_2$
V_1 Velocity vector of $\mathcal{B}_1$'s centroid expressed in the frame fixed to $\mathcal{B}_1$
V_2 Velocity vector of $\mathcal{B}_2$'s centroid expressed in the frame fixed to $\mathcal{B}_2$
X Relative position vector between body centroids expressed in the frame fixed to $\mathcal{B}_2$
X_1 Position vector of $\mathcal{B}_1$'s centroid expressed in the frame fixed to $\mathcal{B}_1$
X_2 Position vector of $\mathcal{B}_2$'s centroid expressed in the frame fixed to $\mathcal{B}_2$
$\mathcal{B}_1$, $\mathcal{B}_2$ Labels for the two polyhedral rigid bodies
Ω Angular velocity vector of $\mathcal{B}_1$ expressed in the frame fixed to $\mathcal{B}_2$
Ω_1 Angular velocity vector of $\mathcal{B}_1$ expressed in the frame fixed to $\mathcal{B}_1$
Ω_2 Angular velocity vector of $\mathcal{B}_2$ expressed in the frame fixed to $\mathcal{B}_2$
δ Rank 2 tensor defined by the Kronecker delta function
γ_T Total linear momentum of system expressed in inertial reference frame
π_T Total angular momentum of system expressed in inertial reference frame
ρ Density, kg/m^3
a Simplex in $\mathcal{B}_1$
b Simplex in $\mathcal{B}_2$
h Integration step size
m Scalar mass parameter for the system
m_1 Mass of body $\mathcal{B}_1$
m_2 Mass of body $\mathcal{B}_2$
r Scalar magnitude of the relative position vector between body centroids, $r = \|X\|$
t Simulation time
t_0 Time at start of simulation
t_f Time at end of simulation
v_1 Velocity vector of $\mathcal{B}_1$'s centroid expressed in the inertial reference frame
v_2 Velocity vector of $\mathcal{B}_2$'s centroid expressed in the inertial reference frame
x_1 Position vector of $\mathcal{B}_1$'s centroid expressed in the inertial reference frame
x_2 Position vector of $\mathcal{B}_2$'s centroid expressed in the inertial reference frame

D Rank 4 tensor, $\in \mathbb{R}^{3\times6\times3\times3}$, the partial derivative of **v** with respect to a rotation matrix
Q Tensor that is symmetric along every dimension, in which each element is a rational number, with form illustrated in Ref. 1
r Rank 2 tensor, $\in \mathbb{R}^{6\times6}$, dependent on relative attitude but not relative position
v Rank 2 tensor, $\in \mathbb{R}^{3\times6}$, containing coordinates of non-centroid vertices of both simplices in a simplex pairing, expressed in any desired coordinate frame
w Rank 1 tensor, $\in \mathbb{R}^6$, dependent on relative attitude and relative position

Subscripts

$_{\phi\ \theta}$ Tensor indices not eliminated by summation
$_a$ Denotes "for simplex a in $\mathcal{B}_1$"
$_b$ Denotes "for simplex b in $\mathcal{B}_2$"
$_{c1\ c2\ c3}$ Denotes the first, second, third column vectors of a matrix, respectively
$_{i\ j\ k\ p}$ Tensor indices eliminated by summation in left hand side of equations
$_n$ Second level subscript, denotes the value of variables at simulation time $t = nh + t_0$

Superscripts

$^{\phi\ \theta}$ Tensor indices not eliminated by summation
$^{i\ j\ k\ p}$ Tensor indices eliminated by summation in left hand side of equations
^{T} Denotes matrix transpose

I. Introduction

THE full two body problem studies the dynamics of two irregular rigid bodies interacting under a mutual potential. The full two body problem arises in numerous engineering and scientific fields. Our focus is on

the dynamics of two rigid bodies in space due to their mutual gravitational potential. This depends on both the relative position and the relative attitude of the bodies. Therefore, the translational orbit dynamics and the rotational attitude dynamics are coupled in the full two body problem. For example, the trajectory of a very large spacecraft around the Earth is affected by the attitude of the spacecraft, and the dynamics of a binary asteroid pair are characterized by the non-spherical mass distributions of the two bodies. Recently, interest in the full (two) body problem as applied to binary objects in space has increased, due to evidence that such binary systems are common among the overall asteroid population and form an especially large percentage of near-earth objects, with some estimates as high as 16 percent.[2–5] In addition, advances in radar and optical observation methods[6] have allowed for better modelling of the shapes and characteristics of bodies in binary systems, encouraging further study of their dynamics.

Some previous work of relevance includes Maciejewski's presentation of equations of motion of the full two body problem in inertial and relative coordinates.[7] He also discussed the existence of relative equilibria. Scheeres derived a stability condition for the full two body problem,[8] and he studied the planar stability of an ellipsoid-sphere model.[9] Spacecraft motion about binary asteroids has been discussed using the restricted three body model,[10,11] and the four body model.[12]

The mutual gravitational potential of two rigid celestial bodies has been expressed using spherical harmonics.[13,14] But, the harmonic expansion is not guaranteed to converge. Convergence is shown to be an unstable property of such spherical harmonic series.[15] By this we mean that an arbitrarily small change to the mass distribution may cause a previously convergent series to diverge. Another commonly used approach for evaluating the mutual gravitational potential is to fill each rigid body's volume with a distribution of point masses, fixed with respect to one another, the sum of which equals the respective body's total mass.[16,17] Although the mutual potential obtained for two rigid bodies using this approach converges to the true gravity field in the limit as the number of point masses becomes arbitrarily large, there are significant errors in the computation of gravitational forces from that mutual potential.[18]

These problems are avoided with the approach presented in Ref. 1 for calculating mutual gravitational potential based on polyhedral body models. This method converges at all points exterior to the two bodies. By representing the bodies as polyhedra, the flexibility of the formalism over more specialized representations such as spheres and ellipsoids is increased. A polyhedral rigid body model can directly include important features such as craters, caves, or deep clefts where contact binaries meet. The entire body does not have to be modelled uniformly at a high resolution. The errors in the mutual potential computation can be reduced to the level of error in each body's shape determination, and to the level of discretization chosen for that shape. Derivatives of this polyhedral mutual potential formulation are given in Ref. 19 and, with some corrections and improvement of notation, in Ref. 20. These derivatives determine the forces and torques exerted by the bodies on each other, for use in either inertial or relative equations of motion that describe the full body dynamics.

These full body dynamics arise from Lagrangian and Hamiltonian mechanics; they are characterized by symplectic, momentum and energy preserving properties. These geometric features determine the qualitative behavior of the full body dynamics, and they can serve as a basis for further theoretical study of the full body problem. The configuration space of the full body dynamics have a Lie group structure referred to as the Euclidean group, SE(3). However, general numerical integration methods, including the widely used Runge-Kutta schemes, neither preserve the Lie group structure nor these geometric properties.[21]

The variational approach[22] and Lie group methods[23] provide systematic methods of constructing structure preserving numerical integrators. The idea of the variational approach is to discretize Hamilton's principle rather than the continuous equations of motion.[22] The numerical integrator obtained from the discrete Hamilton's principle exhibits excellent energy properties, conserves first integrals, and preserves the symplectic structure. Lie group methods consist of numerical integrators that preserve the geometry of the configuration space by automatically remaining on the Lie group.[23] A Lie group method is explicitly adopted for the variational integrator in Ref. 24 and 25. This unified integrator, hereafter referred to as the Lie Group Variational Integrator (or LGVI for short), is symplectic and momentum preserving, and it exhibits good total energy behavior for exponentially long time periods. It also preserves the Euclidian Lie group structure without the use of local charts, reprojection, or constraints.

Numerical simulation of the full body problem involves two major problems; a large computational burden in computing mutual gravitational forces and moments, and inaccuracy of numerical integrators. The forces and moments must constantly be reevaluated with any position change or any orientation change, not only at each time step but at each sub-step involved in the differencing scheme behind any general

numerical integrator. Therefore such general numerical integrators amplify the computation cost for finding the forces and moments. The accuracy of such general integrators also rapidly degrades as the simulation time increases. They fail to preserve the conserved quantities such as total energy and angular momentum, which determine the qualitative behavior of the full body dynamics. Attitude errors tend to accumulate as a consequence of numerical errors, and this attitude degradation causes significant errors in the gravitational force and moment computation.

The unified treatment given in this paper of the polyhedral mutual potential formulation combined with the LGVI presents a solution to these two major dynamic simulation problems. Using polyhedral models, we can approximate irregular bodies to a specified accuracy, and we can control the computational burden to compute the mutual gravitational forces and moments by choosing the level of body discretization and the number of series terms employed in our formulation. The LGVI, as presented in this paper, preserves the conserved quantities of the full body dynamics as well as the orthogonal structure of the rotation matrices. The obtained simulation results exhibit good stability properties for the invariants of motion for exponentially long times. The computational load is further minimized since the LGVI requires one force and torque evaluation per integration step for second order accuracy.

Subsequent sections of this paper are organized as follows. Algorithms for the mutual gravitational forces and moments computation for polyhedral body models are presented in section II. Some description of the full two body problem and its mathematical properties and the continuous relative equations of motion are given in section III. The Lie Group Variational Integrator for computing the dynamics of the full body problem is given in section IV. Selected simulation results are given in section V for two octahedral rigid bodies in several scenarios, for comparison with other integration methods and to show the qualities of the results thus obtained. Conclusions drawn from the results for these scenarios about the accuracy and suitability of our methods, and about computational burden, are given in the last section.

II. Polyhedral Mutual Gravitational Potential, Forces, and Moments

In this section, we present computational algorithms for determining the mutual gravitational potential, forces, and moments given polyhedral models of each of the two rigid bodies. The algorithm to compute the mutual potential relatively efficiently with such a modelling approach is outlined in Ref. 1. The methods outlined in Ref. 19, and in Ref. 20 with some corrections and improvement of notation, can then be used to compute the gradients of the mutual potential (i.e. forces and moments). A more detailed description and derivation of materials in this section, including extensions to computing the forces and moments for inertial equations of motion, can be found in Ref. 20.

A. Context for Methodology

In Ref. 1, a uniform density or homogenous mass distribution within the entire volume of each rigid body is assumed. This may be a realistic assumption for binary asteroids based on some empirical data.[26] However, the underlying method *does* also allow for approximating arbitrary density variations, as required for greater asteroid modelling fidelity and for systems in which one or more bodies is a spacecraft. This is mentioned without proper explanation in Ref. 20, so we will be more specific here. Rather than representing each body with just a single polyhedron whose triangular faces form the body's outer surface, as in prior work, we can represent each body by multiple polyhedra partially or fully nested inside of one another. Each polyhedron has a closed surface consisting of triangular faces, and each triangular face is defined by three vertices. A tetrahedron is formed by these three vertices plus the body centroid. This tetrahedron is referred to hereafter as a simplex. We require each simplex to have a constant density, but different densities can be assigned to each simplex. This allows for density variation over the two angular spherical coordinates within each polyhedron, with resolution determined by the size of the faces. The possible nesting of multiple polyhedra allows for density variation over the radial spherical coordinate within the body volume as well. It should be noted though that to make practical use of the method herein, the coordinates with respect to the body centroid of the three non-centroid vertices of each simplex must be known. While this vertex coordinate information is known throughout for spacecraft bodies, and can be derived for exterior faces of natural bodies from shape data obtained via optical and radar observations or LIDAR surveying, it is difficult to specify for interior faces within natural bodies without specific knowledge of internal mass distribution.

The underlying principle of what follows is that the evaluation of the mutual potential's double volume

integral over both bodies is equivalent to a global sum of the results of evaluating that double volume integral over each possible pairing of simplices, one being drawn from each body. For each such pairing, the contribution to the mutual potential is given by a Legendre-series expansion. Successive terms of this expansion are linear combinations of other terms which factor into symmetric tensors of increasing rank that are independent of relative position and attitude, and other tensors that depend on the relative position and attitude. The gradients of the mutual potential according to this formulation make use of efficient tensor differentiation rules and the chain rule, resulting in a series expansion for the forces and moments that converges for all points exterior to all polyhedra representing the bodies, and hence for all points exterior to the bodies.

It is useful to note here that the number of terms kept in the Legendre-series expansion determines the order, in the inverse of the distance between body centroids, of the errors in the forces and the moments. When the rigid bodies have small separation distances, more terms are required to maintain the same error levels. A possibility for adaptivity during simulation is to adjust the number of terms used in the series expansion for the force and moment computations, depending on the body separation distance. Another possibility is to refine one or more of the polyhedra representing each body as separation distance decreases. However, this must be done carefully, in a manner which involves no change to the relevant conserved quantities and no discontinuity in the motion states.

It is also worth mentioning that one can consider our approach as a special case of a more general problem, not treated in this paper, in which the two bodies are no longer rigid but fragmented, i.e. they are so-called "rubble-piles". The rubble-pile model of asteroids is being supported by an increasing body of evidence from recent asteroid exploration missions.[27] A number of researchers have represented such rubble-pile asteroids as collections of nonintersecting spheres held together under self-gravity,[28] not to be confused with representing a rigid body by filling its volume with spheres fixed with respect to one another. One could just as well represent rubble-pile asteroids as collections of ellipsoids or polyhedra of arbitrary size and shape held together under self-gravity, rather than smooth spheres, as in Refs. 29 and 30 respectively. This allows for filling a body's volume more fully or at different porosities than are possible with spheres (even zero initial porosity with polyhedra). With polyhedra, different levels of rigidity within a body can then be considered simply by initial combination of smaller polyhedra with coincident or adjacent faces into larger polyhedra. The full two rigid body problem that is the subject of this paper can then be viewed from a different perspective as the limiting case of that process of combination within two separated bodies. This is similar to using a tree-code method,[31] but only using the root- or top-level cell (or group) within each body. In other cases employing multiple non-intersecting polyhedra within each body, the gravitational force and moment couples between every possible pair of polyhedra in the binary system can still be accurately evaluated using the methodology of this section, considered independently from the rest of this paper.

B. Mutual Gravitational Potential

We label the two polyhedral rigid bodies $\mathcal{B}_1$ and $\mathcal{B}_2$. Consistent with the earlier discussion and definition of the polyhedral modelling approach, let body $\mathcal{B}_1$ be divided into a set of simplices indexed by a and let body $\mathcal{B}_2$ be divided into a set of simplices indexed by b . Evaluating the double volume integrals over $\mathcal{B}_1$ and $\mathcal{B}_2$ is equivalent to the double summation over all a and over all b of the result of evaluating the double volume integrals over each simplex combination (a,b). This is shown in the following expression for the mutual potential.[1] Note that at this point, we make use of tensor notation and the Einstein convention of summation over repeated indices:

$$U = -G \sum_{a\in\mathcal{B}_1} \sum_{b\in\mathcal{B}_2} \rho_a T_a \rho_b T_b \left\{ \left[\frac{\mathbf{Q}}{r}\right] + \left[-\frac{\mathbf{Q}_i \mathbf{w}^i}{r^3}\right] + \left[-\frac{\mathbf{Q}_{ij}\mathbf{r}^{ij}}{2r^3} + \frac{3\mathbf{Q}_{ij}\mathbf{w}^i\mathbf{w}^j}{2r^5}\right] + \left[\frac{3\mathbf{Q}_{ijk}\mathbf{r}^{ij}\mathbf{w}^k}{2r^5} - \frac{5\mathbf{Q}_{ijk}\mathbf{w}^i\mathbf{w}^j\mathbf{w}^k}{2r^7}\right] + \dots \right\} \quad (1)$$

The scalars T_a and T_b and the tensors $\mathbf{Q}$ of increasing rank are all independent of both relative position between centroids and relative attitude between the bodies, so they can be computed before any dynamic simulation. However the vector $\mathbf{w}$ is dependent on relative position, and both the vector $\mathbf{w}$ and the matrix $\mathbf{r}$ are dependent on relative attitude through a matrix $\mathbf{v}$, where

$$\mathbf{w}^i = \mathbf{v}^i_j X^j \qquad , \qquad \mathbf{r}^{ij} = \mathbf{v}^i_p \mathbf{v}^j_p .$$

The **Q**'s are defined in Ref. 1, in which they are written out explicitly up to the third rank for illustration of their form. The series within the braces in Eq. (1) is infinite but sufficient accuracy seems to be obtained with just the first several terms in square brackets. We denote these bracketed terms as scalars $\hat{U}_0$, $\hat{U}_1$, $\hat{U}_2$, and so on.

C. Mutual Gravitational Forces and Moments

The gravitational force existing between the two bodies is given by the partial derivative of the mutual potential with respect to the relative position vector X. To obtain this we first derive simple tensor differentiation rules:[20]

$$\frac{\partial r}{\partial X_\theta} = \frac{X_\theta}{r} \qquad , \qquad \frac{\partial \mathbf{w}^i}{\partial X_\theta} = \mathbf{v}^i_\theta. \tag{2}$$

These are used in finding each successive $\partial\hat{U} \,/\, \partial X_\theta$ term. The first few of these terms are illustrated below:

$$\begin{aligned}
\frac{\partial \hat{U}_0}{\partial X_\theta} &= -\frac{\mathbf{Q} X_\theta}{r^3}, \\
\frac{\partial \hat{U}_1}{\partial X_\theta} &= \frac{3\mathbf{Q}_i X_\theta \mathbf{w}^i}{r^5} - \frac{\mathbf{Q}_i \mathbf{v}^i_\theta}{r^3}, \\
\frac{\partial \hat{U}_2}{\partial X_\theta} &= \frac{3\mathbf{Q}_{ij}\mathbf{r}^{ij} X_\theta}{2r^5} - \frac{15\mathbf{Q}_{ij} X_\theta \mathbf{w}^i \mathbf{w}^j}{2r^7} + \frac{3\mathbf{Q}_{ij}\mathbf{w}^i \mathbf{v}^j_\theta}{r^5}, \\
\frac{\partial \hat{U}_3}{\partial X_\theta} &= -\frac{15\mathbf{Q}_{ijk}\mathbf{r}^{ij} X_\theta \mathbf{w}^k}{2r^7} + \frac{3\mathbf{Q}_{ijk}\mathbf{r}^{ij}\mathbf{v}^k_\theta}{2r^5} + \frac{35\mathbf{Q}_{ijk} X_\theta \mathbf{w}^i \mathbf{w}^j \mathbf{w}^k}{2r^9} - \frac{15\mathbf{Q}_{ijk}\mathbf{w}^i \mathbf{w}^j \mathbf{v}^k_\theta}{2r^7}.
\end{aligned}$$

Such terms are used in the overall expression for the force vector that is in turn used in the relative equations of motion, namely (laying aside the tensor notation for the moment)

$$\frac{\partial U}{\partial X} = -G \sum_{a\in\mathcal{B}_1} \sum_{b\in\mathcal{B}_2} \rho_a T_a \rho_b T_b \left(\frac{\partial \hat{U}_0}{\partial X} + \frac{\partial \hat{U}_1}{\partial X} + \frac{\partial \hat{U}_2}{\partial X} + \cdots \right). \tag{3}$$

The torque or moment between the bodies due to their mutual gravitational interaction is given by an expression requiring the partial derivative of the mutual potential with respect to the relative attitude R . Evaluating this requires partial differentiation of one attitude rotation matrix, an element of SO(3), with respect to another such matrix. For this we use the basic rule

$$\frac{\partial R^{jk}}{\partial R^{\phi\theta}} = \delta^\phi_j \delta^k_\theta \tag{4}$$

inside of the expression for the tensor $\mathbf{D}$. The details of this tensor and an alternate rule are discussed in Ref. 20. It is in turn used in the two tensor differentiation rules:[20]

$$\frac{\partial \mathbf{w}^i}{\partial R^{\phi\theta}} = X^j \mathbf{D}^{\phi i}_{j\theta} \qquad , \qquad \frac{\partial \mathbf{r}^{ij}}{\partial R^{\phi\theta}} = 2\,\mathbf{v}^i_p \mathbf{D}^{\phi j}_{p\theta} \tag{5}$$

These are used in finding each successive $\partial\hat{U} \,/\, \partial R^{\phi\theta}$ term. There is no attitude dependence in $\hat{U}_0$, and then the next few partials are:

$$\begin{aligned}
\frac{\partial \hat{U}_1}{\partial R^{\phi\theta}} &= -\frac{\mathbf{Q}_i X^j \mathbf{D}^{\phi i}_{j\theta}}{r^3}, \\
\frac{\partial \hat{U}_2}{\partial R^{\phi\theta}} &= -\frac{\mathbf{Q}_{ij}\mathbf{v}^i_p \mathbf{D}^{\phi j}_{p\theta}}{r^3} + \frac{3\,\mathbf{Q}_{ij}\mathbf{w}^i X^p \mathbf{D}^{\phi j}_{p\theta}}{r^5}, \\
\frac{\partial \hat{U}_3}{\partial R^{\phi\theta}} &= \frac{3\,\mathbf{Q}_{ijk}}{2\,r^5}\left(2\mathbf{v}^i_p \mathbf{D}^{\phi j}_{p\theta}\mathbf{w}^k + \mathbf{r}^{ij} X^p \mathbf{D}^{\phi k}_{p\theta}\right) - \frac{15\,\mathbf{Q}_{ijk}\mathbf{w}^i \mathbf{w}^j X^p \mathbf{D}^{\phi k}_{p\theta}}{2\,r^7}.
\end{aligned}$$

Such terms are used in the overall expression

$$\frac{\partial U}{\partial R^{\phi\theta}} = -G \sum_{a\in\mathcal{B}_1} \sum_{b\in\mathcal{B}_2} \rho_a T_a \rho_b T_b \left(\frac{\partial \hat{U}_1}{\partial R^{\phi\theta}} + \frac{\partial \hat{U}_2}{\partial R^{\phi\theta}} + \frac{\partial \hat{U}_3}{\partial R^{\phi\theta}} + \cdots \right). \tag{6}$$

Using this within Eq. (17) below, we can evaluate the moment due to the mutual gravitational potential.

III. Full Two Body Dynamics and Continuous Equations of Motion

In this section, we describe the full two rigid body problem, and we present the continuous relative equations of motion (EOM) and the conserved quantities for the full two rigid body dynamics. Continuous EOM for the full two rigid body problem are given in Ref. 7, and they are formally derived in the context of Lagrangian mechanics in Ref. 25. A more detailed description and derivation of materials in this section, including derivation of inertial EOM in addition to relative EOM, can be found in the latter reference.

The physical configuration of the full two rigid body problem is described in terms of the position vector of each rigid body in an inertial reference frame, an element of $\mathbb{R}^3$, and the attitude of each rigid body with respect to that frame. This attitude is represented by a rotation matrix, which is a 3×3 orthogonal matrix with determinant $+1$, and hence is an element of the Lie group SO(3). The configuration space of each rigid body is therefore a Lie group referred to as the Euclidean Lie group, $\mathrm{SE}(3) = \mathbb{R}^3 \circledS \mathrm{SO}(3)$ (with $\circledS$ meaning the semi-direct product). The motion of the two rigid bodies depends only on the relative positions and the relative attitudes of the bodies. This is a consequence of the property that the mutual gravitational potential depends only on these relative variables. Thus, the Lagrangian of the two rigid bodies is invariant under a group action of SE(3), and it is natural to reduce the EOM for the full body problem by writing them in one of the body fixed frames. Without loss of generality, we present the EOM in the frame fixed to $\mathcal{B}_2$. Then the configuration space of the reduced system is SE(3).

We start with the relative variables

$$X = R_2^T (x_1 - x_2), \tag{7}$$

$$R = R_2^T R_1, \tag{8}$$

where $X \in \mathbb{R}^3$ is the relative position of $\mathcal{B}_1$ with respect to $\mathcal{B}_2$ expressed in the frame fixed to $\mathcal{B}_2$, and $R \in \mathrm{SO}(3)$ is the relative attitude of $\mathcal{B}_1$ with respect to $\mathcal{B}_2$. The corresponding linear and angular velocities are

$$V = R_2^T (v_1 - v_2), \tag{9}$$

$$\Omega = R\Omega_1, \tag{10}$$

where $V \in \mathbb{R}^3$ represents the relative velocity of $\mathcal{B}_1$ with respect to $\mathcal{B}_2$ expressed in the frame fixed to $\mathcal{B}_2$, and $\Omega \in \mathbb{R}^3$ is the angular velocity of $\mathcal{B}_1$ expressed in the frame fixed to $\mathcal{B}_2$. The standard and nonstandard moment of inertia matrices (see Ref. 25 for the distinction) of $\mathcal{B}_1$ can also be expressed in the frame fixed to $\mathcal{B}_2$ according to $J_R = RJ_1R^T$ and $J_{dR} = RJ_{d_1}R^T$. Note that J_R and J_{dR} are not constant matrices.

Assuming that the total linear momentum of the whole system is zero, the kinetic energy of the system is expressed in terms of the relative variables as

$$KE = \frac{1}{2} m \|V\|^2 + \frac{1}{2} \mathrm{tr}\big[S(\Omega) J_{dR} S(\Omega)^T\big] + \frac{1}{2} \mathrm{tr}\big[S(\Omega_2) J_{d_2} S(\Omega_2)^T\big], \tag{11}$$

where $m = \frac{m_1 m_2}{m_1 + m_2} \in \mathbb{R}$, and $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is an isomorphism between $\mathbb{R}^3$ and skew-symmetric matrices, defined such that $S(x)y = x \times y$ for any $x, y \in \mathbb{R}^3$. We also have from Eq. (1) the potential energy expressed in terms of the relative variables, that is to say $PE = U(X, R)$, recalling that in Eq. (1) r and $\mathbf{w}$ are functions of X and both $\mathbf{w}$ and $\mathbf{r}$ are functions of R. The following continuous EOM of the full two rigid body problem in relative coordinates can be obtained by taking variations of the reduced Lagrangian, $L = KE - PE$, as follows:

$$\dot{X} + \Omega_2 \times X = V, \tag{12}$$

$$\dot{R} = S(\Omega)R - S(\Omega_2)R, \tag{13}$$

$$m\dot{V} + m\Omega_2 \times V = -\frac{\partial U}{\partial X}, \tag{14}$$

$$(\dot{J_R\Omega}) + \Omega_2 \times J_R\Omega = -M, \tag{15}$$

$$J_2\dot{\Omega}_2 + \Omega_2 \times J_2\Omega_2 = X \times \frac{\partial U}{\partial X} + M, \tag{16}$$

where the vector $J_R\Omega = RJ_1\Omega_1 \in \mathbb{R}^3$ is the angular momentum of the first body expressed in the second body fixed frame. The moment due to the gravity potential, $M \in \mathbb{R}^3$, is determined by the following relationship

$$S(M) = \frac{\partial U}{\partial R}R^T - R\frac{\partial U}{\partial R}^T,$$

or more explicitly,

$$M = R_{c1} \times \left[\frac{\partial U}{\partial R}\right]_{c1} + R_{c2} \times \left[\frac{\partial U}{\partial R}\right]_{c2} + R_{c3} \times \left[\frac{\partial U}{\partial R}\right]_{c3}. \tag{17}$$

These equations are also given in different notation in Ref. 20. The total energy, the total linear momentum expressed in the inertial frame, and the total angular momentum expressed in the inertial frame are conserved quantities, given by

$$E = KE + PE \tag{18}$$

$$\gamma_T = m_1 v_1 + m_2 v_2, \tag{19}$$

$$\pi_T = x_1 \times m_1 v_1 + R_1 J_1 \Omega_1 + x_2 \times m_2 v_2 + R_2 J_2 \Omega_2. \tag{20}$$

IV. Discrete Equations of Motion: the Lie Group Variational Integrator

This section presents the LGVI, essentially a statement of discrete relative EOM in Hamiltonian form, as opposed to continuous relative EOM in Lagrangian form in the previous section. Again, a detailed description and derivation of materials in this section, including development of inertial discrete EOM in addition to these relative discrete EOM, can be found in Ref. 25. Here we simply state the results.

A. Discrete Equations of Motion

Since the dynamics of the full two rigid bodies arise from Lagrangian or Hamiltonian mechanics, they are characterized by symplectic, momentum and energy preserving properties. These geometric features determine the qualitative behavior of the dynamics, and they can serve as a basis for further theoretical study of the full two rigid body problem. Furthermore, the configuration space has a group structure denoted by SE(3).

However, general numerical integration methods, including the popular Runge-Kutta methods, fail to preserve these geometric characteristics. General integration methods are obtained by approximating continuous EOM by directly discretizing them with respect to time. With each integration step, the updates involve additive operations, so that the underlying Lie group structure is not necessarily preserved as time progresses. This is caused by the fact that the Euclidean Lie group is not closed under addition.

For example, if we use a Runge-Kutta method for numerical integration of (13), then the rotation matrices drift from the orthogonal rotation group, SO(3); the quantity R^TR drifts from the identity matrix. Then the attitudes of the rigid bodies cannot be determined accurately, resulting in significant errors in the gravitational force and moment computations that depend on the attitude, and consequently errors in the entire simulation. It is often proposed to parameterize (13) by Euler angles or unit quaternions. However, Euler angles are not global expressions of the attitude since they have associated singularities. Unit quaternions do not exhibit singularities, but are constrained to lie on the unit three-sphere $\mathbb{S}^3$, and general numerical integration methods do not preserve the unit length constraint. Therefore, quaternions lead to the same numerical drift problem. Re-normalizing the quaternion vector at each step tends to break the conservation properties. Furthermore, unit quaternions double cover SO(3), so that there are inevitable ambiguities in expressing the attitude.

In contrast, the Lie Group Variational Integrator has desirable properties such as symplecticity, momentum preservation, and good energy stability for exponentially long time periods. It also preserves the Euclidian Lie group structure without the use of local charts, reprojection, or constraints. The LGVI is obtained by discretizing Hamilton's principle; the velocity phase space of the continuous Lagrangian is replaced by discrete variables, and a discrete Lagrangian is chosen such that it approximates a segment of the action integral. Taking the variation of the resulting action sum, we obtain discrete EOM referred to as a variational integrator. Since the discrete variables are updated by Lie group operations, the group structure

is preserved. Here we present the resulting discrete EOM as follows; the detailed development can be found in Ref. 25.

$$X_{n+1} = F_{2_n}^T \left(X_n + hV_n - \frac{h^2}{2m} \frac{\partial U_n}{\partial X_n} \right), \tag{21}$$

$$hS \left(J_{R_n} \Omega_n - \frac{h}{2} M_n \right) = F_n J_{dR_n} - J_{dR_n} F_n^T, \tag{22}$$

$$hS \left(J_2 \Omega_{2_n} + \frac{h}{2} X_n \times \frac{\partial U_n}{\partial X_n} + \frac{h}{2} M_n \right) = F_{2_n} J_{d_2} - J_{d_2} F_{2_n}^T, \tag{23}$$

$$R_{n+1} = F_{2_n}^T F_n R_n, \tag{24}$$

$$V_{n+1} = F_{2_n}^T \left(V_n - \frac{h}{2m} \frac{\partial U_n}{\partial X_n} \right) - \frac{h}{2m} \frac{\partial U_{n+1}}{\partial X_{n+1}}, \tag{25}$$

$$J_{R_{n+1}} \Omega_{n+1} = F_{2_n}^T \left(J_{R_n} \Omega_n - \frac{h}{2} M_n \right) - \frac{h}{2} M_{n+1}, \tag{26}$$

$$J_2 \Omega_{2_{n+1}} = F_{2_n}^T \left(J_2 \Omega_{2_n} + \frac{h}{2} X_n \times \frac{\partial U_n}{\partial X_n} + \frac{h}{2} M_n \right) + \frac{h}{2} X_{n+1} \times \frac{\partial U_{n+1}}{\partial X_{n+1}} + \frac{h}{2} M_{n+1}. \tag{27}$$

To propagate these equations, we start with a set of initial states, $(X_0, V_0, R_0, \Omega_0, \Omega_{2_0})$, and perform one initial evaluation of the mutual potential gradients, obtaining $\partial U_0 / \partial X_0$ and M_0 with Eqs. (3) and (17). We then find X_1 from Eq. (21). Solving the implicit equations (22) and (23) yields the matrix-multiplication update matrices F_0 and F_{2_0} for the attitude rotation matrices, and R_1 follows from Eq. (24). After that, we use X_1 and R_1 in a new evaluation of the mutual potential gradients. We then compute V_1, Ω_1, and Ω_{2_1} from equations (25), (26) and (27), respectively. This yields a discrete map $(X_0, V_0, R_0, \Omega_0, \Omega_{2_0}) \mapsto (X_1, V_1, R_1, \Omega_1, \Omega_{2_1})$, and this process can be repeated for each time step. Note that only one new evaluation of the potential gradients is required per time step. The discrete trajectory in reduced variables can be used to reconstruct the inertial motion of the bodies. Either concurrently with that propagation or later after completion of it, through storing values, we can use the gradient $\partial U / \partial X$, the relative attitude R, and the update matrix F_2 with these equations:

$$x_{2_{n+1}} = x_{2_n} + h v_{2_n} + \frac{h^2}{2m_2} R_n \frac{\partial U_n}{\partial X_n}, \tag{28}$$

$$v_{2_{n+1}} = v_{2_n} + \frac{h}{2m_2} R_n \frac{\partial U_n}{\partial X_n} + \frac{h}{2m_2} R_{n+1} \frac{\partial U_{n+1}}{\partial X_{n+1}}, \tag{29}$$

$$R_{2_{n+1}} = R_{2_n} F_{2_n}. \tag{30}$$

In the discrete map defined by the LGVI above, the only implicit parts are Eqs. (22) and (23). These two equations have the same structure, which suggests a specific computational approach. Using the Rodrigues formula, we rewrite those equations as an equivalent vector equations, and we solve them numerically using Newton's iteration. Numerical simulations show that two or three iterations are sufficient to achieve a tolerance of $\epsilon = 10^{-15}$.

B. Properties of the Lie Group Variational Integrator

Since the LGVI is obtained by discretizing Hamilton's principle, it is symplectic and preserves the structure of the configuration space, SE(3), as well as the relevant geometric features of the full two rigid body problem dynamics represented by the conserved first integrals of total angular momentum and total energy. The total energy oscillates around its initial value with small bounds on a comparatively short timescale, but there is no tendency for the mean of the oscillation in the total energy to drift (increase or decrease) from the initial value for exponentially long time. In contrast, the total energy behavior with general numerical methods such as the Runge-Kutta schemes tends to drift dramatically over exponentially long time.

The LGVI preserves the group structure. By using the given computational approach, the matrices F_n and F_{2_n}, representing the change in relative attitude and attitude of $\mathcal{B}_2$ over a time step, are guaranteed to be rotation matrices. The group operation of the Lie group SO(3) is matrix multiplication. Hence rotation

matrices R_n and R_{2_n} are updated by the group operation in Eqs. (24) and (30), so that they evolve on SO(3) automatically without constraints or reprojection. Therefore, the orthogonal structure of the rotation matrices is preserved, and the attitude of each rigid body is determined accurately and globally without the need to use local charts (parameterizations) such as Euler angles or quaternions.

This geometrically exact numerical integration method yields a highly efficient and accurate computational algorithm, especially for the full two rigid body problem examined here. In the full two rigid body problem there is a large burden in computing the mutual gravitational force and moment for arbitrary bodies, so the number of force and moment evaluations should be minimized. We have seen that the LGVI requires only one such evaluation per integration step, the minimum number of evaluations consistent with the presented LGVI having second order accuracy (because it is a self-adjoint method). Within the LGVI, two implicit equations must be solved at each time step to determine the matrix-multiplication updates for R and R_2. However the LGVI is only weakly implicit in the sense that the iteration for each implicit equation is independent of the much more costly gravitational force and moment computation. The computational load to solve each implicit equation is comparatively negligible; only two or three iterations are required. Altogether, the entire method could be considered "almost explicit".

The LGVI is a fixed step size integrator, but all of the properties above are independent of the step size. Consequently, we can achieve the same level of accuracy while choosing a larger step size as compared to other numerical integrators of the same order. All of these features are revealed in the simulation results below.

V. Numerical Examples

A. Implementation Details

Here we describe our specific implementation of the combined computational methods outlined in this paper. Our codes consist of pre-processing scripts and post-processing scripts written in the MATLAB scripting language, and actual executables written in the C language.

Body models for a natural body or for a spacecraft are originally provided to us already in the form of ordered vertex and face lists. Each row of the vertex list contains three numbers for the coordinates of that vertex's position with respect to a specified reference frame, preferably with origin at the center of mass of the body and axes aligned with the body's principle axis. Each row of the face list contains the three row numbers for the rows of the vertex list corresponding to the three vertices that form the corners of that face. The row numbers reference the vertex information in order, moving counterclockwise around the vertices as viewed from outside of the body (faces are oriented with outward-pointing normal vectors according to the right hand rule). If enough information about density distribution for the body is known, so that a density number is also assigned to the simplex associated with each face, those numbers are present in an additional column in the ordered face list. Otherwise, all simplices may be given the same density value. In this paper we do not address the process of generating such formatted body model files from actual observation data for asteroids or from actual CAD software models for spacecraft. The face and vertex lists can be manually generated from scratch for simple arbitrary polyhedron shapes, as for the octahedra bodies used in the next section. In addition to the body models themselves, the initial conditions are needed, including the initial attitude of each body with respect to some common reference frame, the initial spin axis orientation and spin rate of each body in that frame, and the initial mutual orbital elements or equivalent relative translational motion parameters with respect to that frame.

The pre-processing scripts handle a number of preliminary computations. For each body, the centroid (center of mass) and the principal axis directions are found, and if not at the origin and along the elementary unit vectors, respectively, of the frame in which the vertex coordinates are given, the coordinates of every vertex are shifted and rotated so that this becomes the case. The moment of inertia contribution of each simplex in the new frame is found, and these are summed to obtain the total standard moment of inertia matrix. Other information such as the surface area, volume, and mean equivalent radius of each body is also found and reported. All vertex coordinates and other quantities can then optionally be nondimensionalized using user-entered length, mass, and time factors, for better numerical conditioning. Finally, five files are produced. The first two of these files, one for each body, contains a re-ordering of the elements of the position vectors, with respect to the centroid, for the vertices of the simplices. The other three files, respectively, contain the initial state vector $(X_0, mV_0, J_2\Omega_{2_0}, J_{R_0}\Omega_0, R_0, R_{2_0})$, other system physical data (densities, body volumes, J_1, J_2, m_1, m_2, m, nondimensionalized G, and nondimensionalization factors), and the integration

parameters (starting and stopping times and truncation error tolerance, if needed).

A first executable makes use of the Runge-Kutta-Fehlberg 7(8) integration method, hereafter referred to as RKF7(8) or just RKF, to propagate the system, after a starting calculation of the time-invariant $\mathbf{Q}$ tensors. It is noted that this one-time only calculation of the $\mathbf{Q}'s$ is analogous to finding the successively higher-order mass moments of a "normalized" simplex with vertices at $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$. It is also noted that for this high-order scheme, the EOM are evaluated thirteen times within each integration step, and an evaluation of the mutual gravitational potential force and moment is required each of those times. After this the state update is performed and the new state vector is written to an output file only if the truncation error is within the tolerance specified. Step size adjustment is performed with every step. With each state update the mutual potential itself, force, and moment, while not required for the propagation of the dynamics, are also evaluated again using the new state vector and written to another output file. This allows for checking the total energy conservation and the linear and angular accelerations. Hence fourteen force and moment evaluations are needed per integration step.

Another executable uses the LGVI. At the start of this, the initial state vector from the input file, $(X_{_0}, mV_{_0}, J_2\Omega_{2_0}, J_{R_0}\Omega_{_0}, R_{_0}, R_{2_0})$, is converted to the vector for the discrete mapping, $(X_{_0}, V_{_0}, R_{_0}, \Omega_{_0}, \Omega_{2_0})$. Again, rather than fourteen force and moment evaluations, this LGVI involves only one such evaluation per integration step, plus the quick solution of the implicit equations. Step size is fixed, but this does not present a significant problem for most scenarios observed in binary asteroids having mutual orbits with relatively low-eccentricity.

With the aim of completing simulations much faster than in a single-processor environment, a parallelized version of each executable above was also written in C with the addition of MPI. In using the methods presented in this paper, most of the computation time is associated with the evaluation of the potential gradients, and that involves performing the same operations for all of the different simplex combinations, followed by a global sum. This is well-suited for parallelization. The parallelized version of each executable is flexible in that any number of nodes or processors can be specified by the user. Then the process 0 assigns to each of the other processors the task of calculating the portion of the potential gradients double summations of Eqs. (3) and (6) that arises from pairing a single simplex in $\mathcal{B}_2$ with all simplices in $\mathcal{B}_1$ in succession. If the number of other processors specified by the user or found available on the cluster is less than the number of simplices in $\mathcal{B}_2$, this is done in rounds until the portion of the problem matching with every a is obtained. The parallelized version of each executable has been used on Myrinet clusters at the Center for Advanced Computing (CAC) at the University of Michigan and at the Supercomputing and Visualization Center at NASA's Jet Propulsion Laboratory. Though compiler and user environment differences produced markedly different capabilities in each cluster environment, eventually a further two orders of magnitude reduction in computation time over otherwise identical single-processor runs was achieved with both the VI and RKF7(8) schemes. It should be noted however, that for the simulations with results presented in the following section, the parallel computing capability was not needed for such small (in number of faces) body models, and was not utilized.

The MATLAB script post-processing of the output files generates all desired plots of various dynamic quantities, with optional capability for animation generation. All pre- and post-processing steps and scripts are identical regardless of whether the RKF7(8) or LGVI executable is used, and regardless of whether the single-processor or parallel version of each is used.

B. Simulation Results

Simulation results for two octahedral rigid bodies with eight faces and eight simplices each are given for a variety of scenarios. Octahedra are used rather than more complex shapes because they are the simplest polyhedral shapes that manifest the coupled dynamics behavior desired in all of the scenarios. For greater simplicity, the octahedra are made symmetric about all axes, although they are of different sizes. The extents data defining the locations of the corners of each octahedron are given in Table 1, as are various physical parameters of each octahedron including mass and moment of inertia properties. We present simulation results for four scenarios, and the initial condition for each scenario is given in Table 2.

Scenario 1 The first scenario presented here is that of short-duration simulation of the two octahedra starting from initial conditions matching with a medium eccentricity elliptical mutual orbit. Both the RKF7(8) and LGVI integrators are used, with the intent of making a direct comparison between the trajectories of the configuration variables that result from using each integrator over a short simulation duration.

Table 1. Properties of octahedral body models used in simulations

Property	$\mathcal{B}_2$		$\mathcal{B}_1$	
Surface area (m^2)	8.839		2.002	
Volume (m^3)	1.800		0.1561	
Equiv. radius (m)	0.7546		0.3340	
Mass (kg)	4500		390.3	
Density (kg/m^3)	2500		2500	
I_{xx} (kg-m^2)	1377.0		9.24	
I_{yy} (kg-m^2)	814.5		42.99	
I_{zz} (kg-m^2)	1462.5		44.32	
Extents (m)	min	max	min	max
body frame X	-1.0	1.0	-1.0	1.0
body frame Y	-1.5	1.5	-1/exp(1)	1/exp(1)
body frame Z	-0.9	0.9	$-1/\pi$	$1/\pi$

Table 2. Initial Conditions

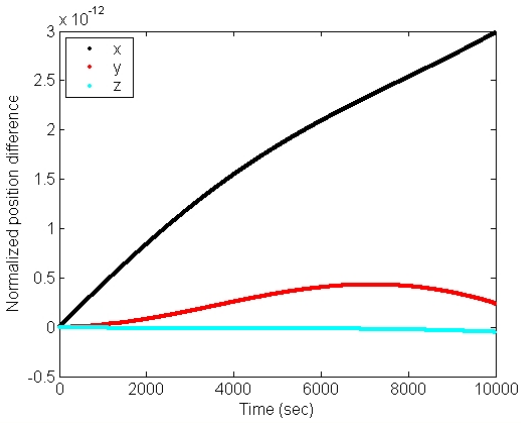
Scenario	Attitude* (deg)	Body spin† (rad/s)	Orbital elements (m,deg) OR state vector (m,m/s)
1	(100, 9.8, 175) (160, -5, 165)	(0, 0, 5.0×10^{-5}) (0, 0, 9.2×10^{-5})	$(a, e, i, \Omega, \omega, \nu)$=(4m, 0.3, 5°, 15°, 60°, 10°)
2	(180, 0, 30) (270, 0, 30)	(0, 0, 0.566) (0, 0, -0.566)	$X_0 = [0, 6, 0]^T$, $V_0 = [-0.006, 0, 0]^T$
3	(-22.6, 5, 180) (50.3, 5, -180)	(0, 0, 1.63×10^{-4}) (0, 0, 1.55×10^{-4})	$(a, e, i, \Omega, \omega, \nu)$=(52.9m, 0.942, 5°, 0°, 88.2°, -107.1°)
4	(-75, 30, 180) (-75, 30, 180)	(0.007, 0.007, 0.05) (-0.003, 0.002, 0.004)	$X_0 = [-0.5, 1.8, 1.1]^T$, $V_0 = [-0.3, -0.1, 0]^T$

* 3-1-3 Euler sequence for $\mathcal{B}_1$ (first line) and $\mathcal{B}_2$ (second line).

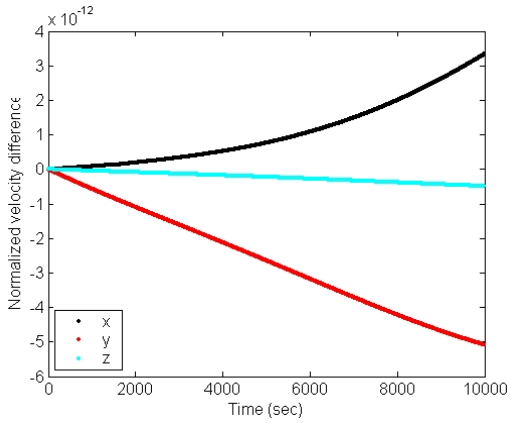
† Components of angular velocity of each body expressed in its own body-fixed frame for $\mathcal{B}_1$ (first line) and $\mathcal{B}_2$ (second line).

Figure 1 shows the difference between the output of the RKF7(8) and that of the LGVI in components of reconstructed inertial position, inertial velocity, and body-frame angular velocity vectors for $\mathcal{B}_2$, plus the difference in body attitude parameters for $\mathcal{B}_2$. The corresponding output difference plots for body $\mathcal{B}_1$ look very similar. The differences in vector components of Figure 1(a) are normalized by the system's semi-major axis ($a = 4.0$ m). The differences in vector components of Figure 1(b) are normalized by the equivalent circular velocity ($\sqrt{\mu/a} = 2.856 \times 10^{-4}$ m/s), and those of Figure 1(c) are normalized by the equivalent meanmotion ($\sqrt{\mu/a^3} = 7.141 \times 10^{-5}$ radians/s). To obtain the results compared here, the total number of mutual potential derivatives evaluations and actual running time using the RKF7(8) routine were 70014 and 494 seconds, respectively, while the number of such evaluations and actual running time using the LGVI were 70001 and 539 seconds. Therefore the computational effort and resources used were roughly the same in each case. All of these results show that the LGVI can be trusted to produce almost exactly the same trajectory as a standard RKF7(8) integration routine over short time scales. As the simulation duration increases the trajectories from the two integrators begin to diverge. The behavior of integrals of motion and appropriate error metrics must then be used to discern which trajectory is to be taken as the "truth".

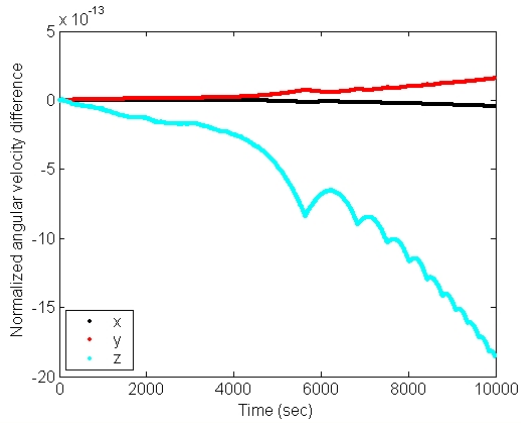
Scenario 2 Another scenario is that of propagation from an initial condition with the bodies aligned but possessing relatively large magnitude centroid velocity vectors that are antiparallel and perpendicular to the initial line between centroids. This scenario is simulated both with the LGVI at different step sizes and



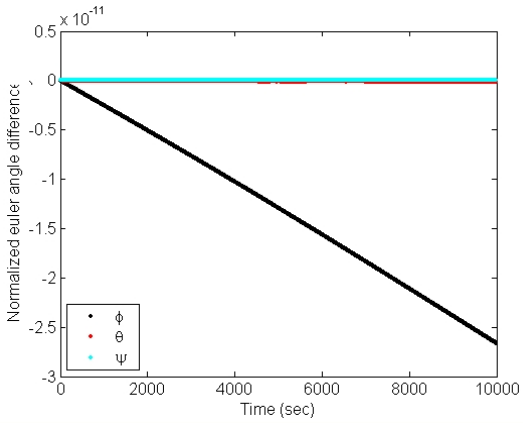
(a) Inertial position for $\mathcal{B}_2$



(b) Inertial velocity for $\mathcal{B}_2$



(c) Angular velocity for $\mathcal{B}_2$



(d) 3-1-3 Euler angles for $\mathcal{B}_2$

Figure 1. (Scenario 1) Difference between RKF7(8) and LGVI output.

with the variable step size RKF7(8) at different error tolerances. This allows for a comparison between the integrators of their performance, in terms of the total energy and total angular momentum integrals and the attitude error metric growth vs. computational burden. The results in Table 3 illustrate the general superiority of the LGVI approach over Runge-Kutta-type approaches.

Here we see that for any pair of simulations, one using the LGVI and the other using the RKF7(8) scheme, for which the total energy metric performs about the same, the computation time needed to complete the simulation using the LGVI is a fraction of that needed using the RKF7(8). Simultaneous with this improvement in run time, the total angular momentum and attitude error metrics still perform better in the LGVI run than in the RKF7(8) run by multiple orders of magnitude. Going in the other direction, as the step size for the LGVI is reduced so that the computational burden using it begins to approach that for any chosen run using the RKF7(8), all error metrics remain at the same level as or else orders of magnitude smaller than those for the chosen RKF7(8) run. For the LGVI, the round-off error accumulates when multiplying rotation matrices at (24). The rotation matrix error of the LGVI is caused only by the floating-point arithmetic operation, and it is increased as the number of integration steps is increased. A similar trend is observed in the total angular momentum error for the LGVI, because determination of the total angular momentum in the inertial frame from the states written to the output file makes use of the rotation matrices.

Scenario 3 The next scenario illustrates the ability of our methods to capture the interesting effects of coupling in a mutual orbit configuration that the Keplerian two-body approximation incorrectly predicts as being perpetual. Simulation with the LGVI yields the trajectory illustrated in Figure 2(a), which transitions from a highly elliptical orbit to a hyperbolic escape path. This is shown by the plots in Figure 2(b) of the

Table 3. (Scenario 2) Performance comparison between RKF7(8) and LGVI

Method	h^*	$N_u^\star$	$t_W^\diamond$	$\epsilon^\triangleleft$	$\mathrm{E}[\lvert\Delta\mathrm{TE}\rvert]^{\dagger\ddagger}$	$\mathrm{E}[\lVert\Delta\pi_T\rVert]^{\dagger\ddagger}$	$\mathrm{E}[\lVert I - R^T R\rVert]^{\dagger}$
RKF7(8)	0.236	2368912	23439	10^{-12}	3.901×10^{-12}	1.493×10^{-9}	1.151×10^{-7}
RKF7(8)	0.421	1331414	9102	10^{-10}	1.274×10^{-10}	2.630×10^{-7}	1.985×10^{-5}
RKF7(8)	0.749	747376	5252	10^{-8}	2.284×10^{-8}	4.620×10^{-5}	3.173×10^{-3}
LGVI	0.0169	2370000	13511	-	1.698×10^{-11}	5.167×10^{-10}	2.525×10^{-11}
LGVI	0.04	1000000	9920	-	1.928×10^{-11}	1.189×10^{-10}	2.120×10^{-11}
LGVI	0.08	500000	5127	-	9.879×10^{-11}	4.139×10^{-11}	2.004×10^{-12}
LGVI	0.4	100000	983	-	2.234×10^{-9}	6.266×10^{-12}	3.386×10^{-14}
LGVI	0.8	50000	431	-	9.326×10^{-9}	1.279×10^{-11}	6.352×10^{-14}
LGVI	1.0	40000	335	-	1.512×10^{-8}	3.991×10^{-12}	4.786×10^{-14}

* h is integration step size, in seconds, fixed for LGVI but averaged over the run's duration for RKF7(8)

⋆ N_u is the total number of calculations of the mutual potential derivatives made within the run

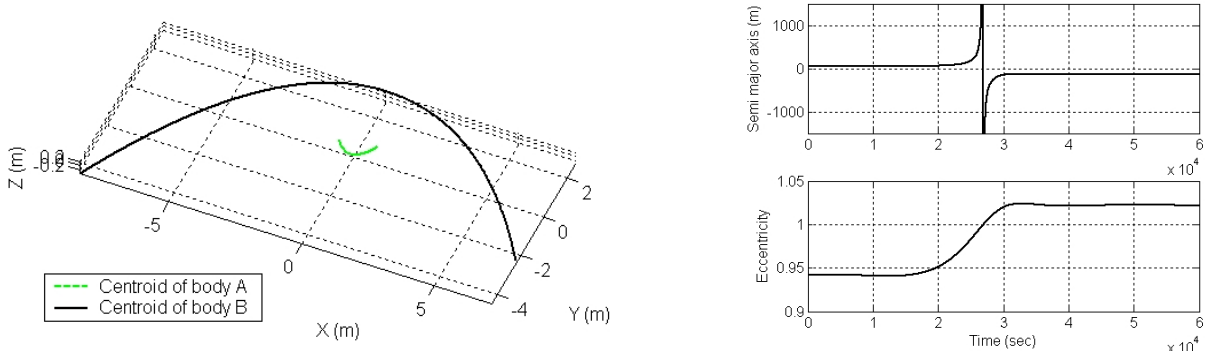
◇ t_W is the "wall-clock" time to complete each simulation run, in seconds

◁ ϵ is the error tolerance for the variable step size in RKF7(8)

‡ TE and π_T are total energy and the total angular momentum, respectively, while Δ refers to deviation from the initial value over simulation

† E[·] denotes mean

semi-major axis and eccentricity change during the close encounter, which occurs roughly midway through the run duration of 60,000 seconds. The initial conditions and body configurations are symmetric about the initial orbital plane, and as such the motion of the centroids should be restricted to the initial orbital plane. This is observed numerically, as the body centroids remain within 8.6×10^{-14} meters of the initial orbital plane throughout the simulation.



(a) Trajectory of binary octahedra system components

(b) Eccentricity and semi major axis

Figure 2. (Scenario 3) Disruption of the binary octahedra system.

Scenario 4 Finally, we examine a very long duration simulation starting from initial conditions that are stable in the sense that orbital trajectories are confined to separated and bounded regions but are highly unstable in the sense that body attitudes vary greatly and irregularly. The trajectory of binary octahedra system is shown in Figure 3. Note that this motion cannot be observed with the point mass assumption in the classical two body problem. This scenario is simulated both with LGVI at different step sizes and with RKF7(8) at different error tolerances, for 5×10^6 (sec) maneuver time. Table 4 illustrates the general superiority of the LGVI approach, similar to Scenario 2. In addition, we see that the simulation results of RKF7(8) with larger error tolerances, 10^{-13} and 10^{-10}, are completely unreliable since the rotation matrix error is increased to the unacceptable levels of 10^{-1} and 10^{1}.

The quantitative comparison is summarized in Figure 4, where the total computation time and the stan-

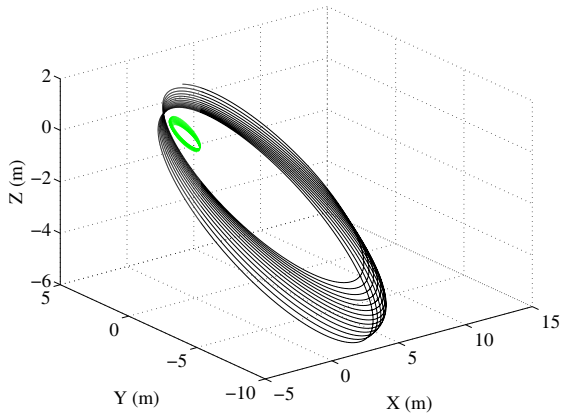


Figure 3. (Scenario 4) Trajectory of binary octahedra system

Table 4. (Scenario 4) Performance comparison between RKF7(8) and LGVI

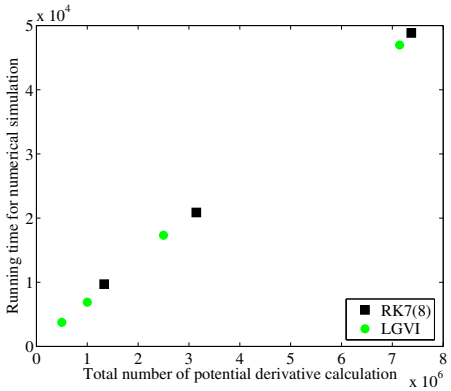
Method	h	N_u	t_W	ϵ	E[$\lvert\Delta$TE$\rvert$]	E[$\lVert\Delta\pi_T\rVert$]	E[$\lVert I - R^T R\rVert$]
RKF7(8)	8.81	7373106	48850	10^{-16}	6.198×10^{-8}	3.830×10^{-6}	6.089×10^{-5}
RKF7(8)	20.66	3145532	20873	10^{-13}	8.731×10^{-5}	7.427×10^{-3}	1.015×10^{-1}
RKF7(8)	48.74	1333475	9699	10^{-10}	1.959×10^{-1}	5.246×10^{0}	1.834×10^{1}
LGVI	0.7	7142858	46963	-	6.947×10^{-11}	2.190×10^{-12}	6.645×10^{-13}
LGVI	2	2500000	17320	-	5.687×10^{-10}	5.077×10^{-13}	3.663×10^{-13}
LGVI	5	1000000	6897	-	3.601×10^{-9}	8.447×10^{-13}	1.642×10^{-13}
LGVI	10	500000	3750	-	1.517×10^{-8}	2.567×10^{-13}	3.521×10^{-13}

See Table 3 for notations.

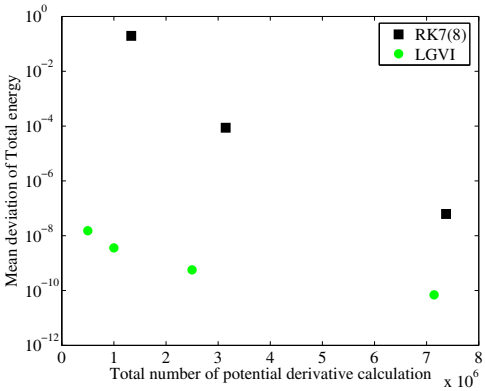
dard deviation of the total energy are shown over the number of the mutual potential derivatives calculations. In Figure 4(a), we see that the total computation time is directly proportional to the number of the mutual potential derivative calculations regardless of the integration scheme used. This justifies the statement that the LGVI is an "almost explicit" computational method; the computational load to solve the implicit equations is comparatively negligible. Considering that the mutual potential derivative computations are the major computational burden for the numerical simulation of full two rigid body dynamics, an efficient integration scheme with fewer mutual potential derivative calculations should exhibit smaller error measures. In Figure 4(b), the lower left portion represents higher computational efficiency, and the upper right portion represents lower efficiency. We see that LGVI is more efficient than RKF7(8). Figure 5 compares the total energy deviation history and the relative rotation matrix error history. For LGVI, the total energy varies within the level of 10^{-11}. The repeated peaks of the total energy deviation correspond to the perigees of the orbit shown in Figure 3, but but there is no tendency for the mean of the variation to drift for the entire simulation time. The rotation matrix error is slightly increasing, since the round off error in multiplying orthogornal matrices at (24) is accumulated. But the error is below 10^{-12} over the entire simulation's span. However, for RKF7(8), the variation of the total energy is linearly increasing over time, and the rotation matrix error is increasing to the level of 10^{-4}. We have seen that LGVI exhibits good total energy behavior for exponentially long time periods, and it also preserves the group structure well. The accuracy of RKF7(8) is vulnerable to a long time simulation of the full two rigid body dynamics.

VI. Conclusions

This paper presents a complementary combination of tools or algorithms that achieves superior performance in the computationally demanding simulation of the full two rigid bodies problem: a simple method

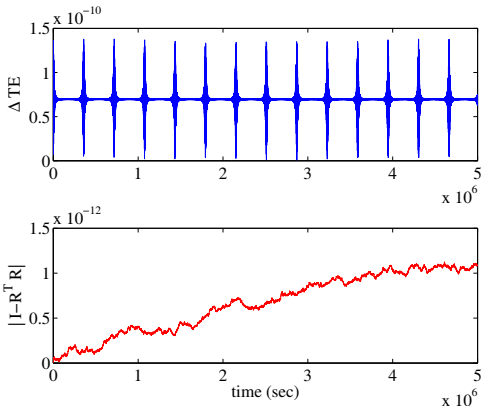


(a) Running time v.s. number mutual potential derivatives calculations

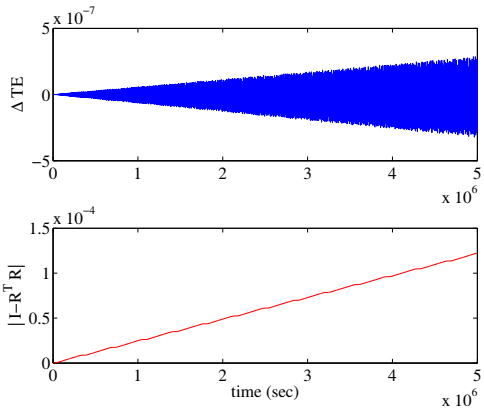


(b) Mean deviation of TE v.s. number of mutual potential derivatives calculations

Figure 4. (Scenario 4) Quantitative comparisons of running time and standard deviation of total energy over number of mutual potential derivatives calculations.



(a) LGVI



(b) RKF7(8)

Figure 5. (Scenario 4) Qualitative comparisons of behavior of deviation in total energy and rotation matrix error over time with similar computational load for LGVI and RKF7(8).

for efficient calculation of the mutual gravitational potential and its derivatives given versatile polyhedral models of the rigid bodies and a Lie group variational integrator consisting of discrete relative equations of motion that preserves the geometric features and structure of the configuration space. The use of these easily implemented methods together allows for simulation of the full two rigid bodies, capturing their fully coupled dynamics, that is both accurate and efficient. The results above show the accuracy maintained for energy, momenta, and attitude geometry constraints, even for long simulation run times. The presented mutual potential computations for polyhedral approximations are themselves efficient, but for bodies of relevant (i.e. large) model size, they still represent the bulk of the computational burden during simulation, regardless of the integration scheme used. Hence use of the Lie Group Variational Integrator, which requires minimal mutual potential computations, achieves an even greater combined efficiency. This yields maximal simulation durations for fixed computational resources.

Acknowledgments

The first author would like to acknowledge the support of the U. S. Air Force Office of Scientific Research (AFOSR) for the period in which this paper was written. The research of the third author was partially supported by NSF grant DMS-0504747 and a University of Michigan Rackham faculty grant.

References

[1]Werner, R. A. and Scheeres, D. J., "Mutual Potential of Homogenous Polyhedra," *Celestial Mechanics and Dynamical Astronomy*, Vol. 91, No. 3, March 2005, pp. 337–349.

[2]Bottke, W. F. and Melosh, H. J., "The Formation of Binary Asteroids and Doublet Craters," *Icarus*, Vol. 124, 1996, pp. 372391.

[3]Bottke, W. F. and Melosh, H. J., "The Formation of Asteroid Satellites and Doublet Craters by Planetary Tidal Forces," *Nature*, Vol. 381, 1996, pp. 5153.

[4]Margot, J. L., Nolan, M. C., Benner, L. A. M., Ostro, S. J., Jurgens, R. F., Giorgini, J. D., Slade, M. A., and Campbell, D. B., "Binary Asteroids in the Near-Earth Object Population," *Science*, Vol. 296, May 2002, pp. 1445–1448.

[5]Merline, W. J., Weidenschilling, S. J., Durda, D. D., Margot, J. L., Pravec, P., and Storrs, A. D., "Asteroids do have satellites," *Asteroids III*, edited by W. F. Bottke et al., Space Science Series, University of Arizona, Tuscon, AZ, 2002, pp. 289–312.

[6]Ostro, S. J., Hudson, R. S., Benner, L. A. M., Giorgini, J. D., Magri, C., Margot, J.-L., and Nolan, M. C., "Asteroid Radar Astronomy," *Asteroids III*, edited by W. F. Bottke et al., Space Science Series, University of Arizona, Tuscon, AZ, 2002, pp. 151–168.

[7]Maciejewski, A. J., "Reduction, Relative Equilibria and Potential in the Two Rigid Bodies Problem," *Celestial Mechanics and Dynamical Astronomy*, Vol. 63, No. 1, 1995, pp. 1–28.

[8]Scheeres, D. J., "Stability in the Full Two-Body Problem," *Celestial Mechanics and Dynamical Astronomy*, Vol. 83, 2002, pp. 155–169.

[9]Scheeres, D. J., "Stability of Relative equilibria in the Full Two-Body Problem," *New Trends in Astrodynamics Conference*, Jan. 2003.

[10]Scheeres, D. J. and Augenstein, S., "Spacecraft motion about binary asteroids," *Proc. AAS/AIAA Astrodynamics Specialist Conference*, Aug. 2003.

[11]Gabern, F., Koon, W. S., and Marsden, J. E., "Spacecraft dynamics near a binary asteroid," *Proceedings of the fifth international conference on dynamical systems and differential equations*, Jun 2004.

[12]Scheeres, D. J. and Bellerose, J., "The Restricted Hill Full 4-Body Problem: application to spacecraft motion about binary asteroids," *Dynamical Systems: An International Journal*, Vol. 20, No. 1, 2005, pp. 23–44.

[13]Borderies, N., "Mutual gravitational potential of N solid bodies," *Celestial Mechanics*, Vol. 18, No. 3, 1978, pp. 295–307.

[14]Braun, C. V., *The gravitational potential of two arbitrary, rotating bodies with applications to the Earth-Moon system*, Ph.D. thesis, University of Texas at Austin, 1991.

[15]Moritz, H., *Advanced Physical Geodesy*, Abacus Press, 1980.

[16]Geissler, P., Petit, J.-M., Durda, D. D., Greenberg, R., Bottke, W., Nolan, M., and Moore, J., "Erosion and Ejecta Reaccretion of 243 Ida and Its Moon," *Icarus*, Vol. 120, No. 1, 1996, pp. 140–157.

[17]Ashenberg, J., "Proposed Method for Modeling the Gravitational Interaction Between Finite Bodies," *Journal of Guidance, Control, and Dynamics*, Vol. 28, No. 4, 2005, pp. 768–774.

[18]Werner, R. A. and Scheeres, D. J., "Exterior Gravitation of a Polyhedron Derived and Compared with Harmonic and Mascon Gravitation Representations of Asteroid 4769 Castalia," *Celestial Mechanics and Dynamical Astronomy*, Vol. 65, No. 3, 1997, pp. 313–344.

[19]Fahnestock, E. G., Scheeres, D. J., McClamroch, N. H., and Werner, R. A., "Simulation and Analysis of Binary Asteroid Dynamics Using Mutual Potential and Potential Derivatives Formulation," *Proc. AAS/AIAA Astrodynamics Specialist Conference*, Lake Tahoe, CA, Aug. 2005.

[20]Fahnestock, E. G. and Scheeres, D. J., "Simulation of the Full Two Rigid Body Problem Using Polyhedral Mutual Potential and Potential Derivatives Approach," *Celestial Mechanics and Dynamical Astronomy*, submitted for publication.

[21]Hairer, E., Lubich, C., and Wanner, G., *Geometric Numerical Integration*, Springer, 2000.

[22]Marsden, J. E. and West, M., "Discrete mechanics and variational integrators," *Acta Numerica*, Vol. 10, 2001, pp. 357–514.

[23]Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P., and Zanna, A., "Lie-group methods," *Acta Numerica*, Vol. 9, 2000, pp. 215–365.

[24]Lee, T., Leok, M., and McClamroch, N. H., "A Lie group variational integrator for the attitude dynamics of a rigid body with application to the 3D pendulum," *Proceedings of the IEEE Conference on Control Application*, Toronto, Canada, Aug. 2005, pp. 962–967.

[25]Lee, T., Leok, M., and McClamroch, N. H., "Lie group variational integrators for the Full Body Problem," *Computer Methods in Applied Mechanics and Engineering*, submitted, Available: http://arxiv.org/abs/math.NA/0508365.

[26]Yeomans, D. K. et al., "Radio Science Results During the NEAR-Shoemaker Spacecraft Rendezvous with Eros," *Science*, Vol. 289, Sept. 2000, pp. 2085–2088.

[27]Fujiwara, A., Kawaguchi, J., Yeomans, D. K., Abe, M., Mukai, T., Okada, T., Saito, J., Yano, H., Yoshikawa, M., Scheeres, D. J., Barnouin-Jha, O., Cheng, A. F., Demura, H., Gaskell, R. W., Hirata, N., Ikeda, H., Kominato, T., Miyamoto, H., Nakamura, A. M., Nakamura, R., Sasaki, S., , and Uesugi, K., "The Rubble-Pile Asteroid Itokawa as Observed by Hayabusa," *Science*, accepted for publication.

[28]Richardson, D. C., William F. Bottke, J., and Love, S. G., "Tidal Distortion and Disruption of Earth-Crossing Asteroids," *Icarus*, Vol. 134, 1998, pp. 47–76.

[29]Roig, F., Duffard, R., Penteado, P., Lazzaro, D., and Kodama, T., "Interacting ellipsoids: a minimal model for the dynamics of rubble-pile bodies," *Icarus*, Vol. 165, No. 2, 2003, pp. 355–370.

[30]Korycansky, D. G., "Orbital Dynamics for Rigid Bodies," *Astrophysics and Space Science*, Vol. 291, 2004, pp. 57–74.

[31]Duan, Z.-H. and Krasny, R., "An Adaptive Treecode for Computing Nonbonded Potential Energy in Classical Molecular Systems," *Journal of Computational Chemistry*, Vol. 22, No. 2, 2001, pp. 184–195.

Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

Optimal Control of a Rigid Body using Geometrically Exact Computations on $\mathrm{SE}(3)$

Taeyoung Lee*†, N. Harris McClamroch†
Department of Aerospace Engineering
University of Michigan, Ann Arbor, MI 48109
{tylee, nhm}@umich.edu

Melvin Leok*
Department of Mathematics
Purdue University, West Lafayette, IN 47907
mleok@math.purdue.edu

***Abstract*— Optimal control problems are formulated and efficient computational procedures are proposed for combined orbital and rotational maneuvers of a rigid body in three dimensions. The rigid body is assumed to act under the influence of forces and moments that arise from a potential and from control forces and moments. The key features of this paper are its use of computational procedures that are guaranteed to preserve the geometry of the optimal solutions. The theoretical basis for the computational procedures is summarized, and examples of optimal spacecraft maneuvers are presented.**

I. Introduction

Discrete optimal control problems for translational and rotational dynamics of a rigid body under a potential are studied. Optimal control of a rigid body arises in numerous engineering and scientific fields. These problems provide both a theoretical challenge and a numerical challenge in the sense that the configuration space has a Lie group structure denoted by $\mathrm{SE}(3)$ that defines a fundamental constraint.

Optimal control problems on a Lie group have been studied in [1], [2]. These studies are based on the driftless kinematics of a Lie group. The dynamics are ignored, and it is assumed that elements in the corresponding Lie algebra are controlled directly.

General-purpose numerical integration methods, including the popular Runge–Kutta schemes, typically preserve neither the group structure of the configuration space nor geometric invariants of the dynamics. Geometric structure-preserving integrators, referred to as Lie group variational integrators [3], preserve the group structure without the use of local charts, reprojection, or constraints, and they have the desirable property that they are symplectic and momentum preserving, and they exhibit good energy behavior for an exponentially long time period.

This paper presents geometrically exact and numerically efficient computational approaches to solve optimal control problems of a rigid body on a Lie group, $\mathrm{SE}(3)$. The dynamics and the kinematics are discretized by a Lie group variational integrator, and discrete optimality conditions are constructed. Efficient numerical algorithms to solve the necessary condition are developed. This method provide a substantial advantage over current methods for optimal control on a Lie group in the sense that the dynamics of a rigid body as well as the kinematics equation are explicitly utilized, and the proposed computational approaches respect the group structure.

This paper is organized as follows. In Section II, a Lie group variational integrator is developed. Optimal control problems using impulsive controls are studied in Section III, and optimal control problems with smooth controls are studied in Section IV. Numerical results for a rigid dumbbell spacecraft are given in Section V.

II. Lie group variational integrator on $\mathrm{SE}(3)$

The configuration space for the translational and rotational motion of a rigid body is the special Euclidean group, $\mathrm{SE}(3) = \mathbb{R}^3 \circledS \mathrm{SO}(3)$. We identify the cotangent bundle $\mathrm{T}^*\mathrm{SE}(3)$ with $\mathrm{SE}(3) \times \mathfrak{se}(3)^*$ by left translation, and we identify $\mathfrak{se}(3)^*$ with $\mathbb{R}^6$ by an isomorphism between $\mathbb{R}^6$ and $\mathfrak{se}(3)$, and the standard inner product on $\mathbb{R}^6$. We denote the attitude, position, angular momentum, and linear momentum of the rigid body by $(R, x, \Pi, \gamma) \in \mathrm{T}^*\mathrm{SE}(3)$.

The continuous equations of motion are given by

$$\dot{x} = \frac{\gamma}{m}, \tag{1}$$

$$\dot{\gamma} = f + u^f, \tag{2}$$

$$\dot{R} = RS(\Omega), \tag{3}$$

$$\dot{\Pi} + \Omega \times \Pi = M + u^m, \tag{4}$$

where $\Omega \in \mathbb{R}^3$ is the angular velocity, and $u^f, u^m \in \mathbb{R}^3$ are the control force in the inertial frame and the control moment in the body fixed frame, respectively. The constant mass of the rigid body is $m \in \mathbb{R}$, and $J \in \mathbb{R}^{3\times 3}$ denotes the moment of inertia, i.e. $\Pi = J\Omega$. The map $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is an isomorphism between $\mathfrak{so}(3)$ and $\mathbb{R}^3$ defined by the condition $S(x)y = x \times y$ for all $x, y \in \mathbb{R}^3$.

We assume that the potential is dependent on the position and the attitude; $U(\cdot) : \mathrm{SE}(3) \mapsto \mathbb{R}$. The corresponding force and the moment due to the potential are given by

$$f = -\frac{\partial U}{\partial x}, \tag{5}$$

$$M = r_1 \times u_{r_1} + r_2 \times u_{r_2} + r_3 \times u_{r3}, \tag{6}$$

where $r_i, u_{r_i} \in \mathbb{R}^3$ are the ith row vector of R and $\frac{\partial U}{\partial R}$, respectively.

Since the dynamics of a rigid body has the structure of a Lagrangian or Hamiltonian system, they are characterized

*This research has been supported in part by NSF under grant DMS-0504747, and by a grant from the Rackham Graduate School, University of Michigan.

†This research has been supported in part by NSF under grant ECS-0244977.

by symplectic, momentum and energy preserving properties. These geometric features determine the qualitative behavior of the rigid body dynamics, and they can serve as a basis for theoretical study of rigid body dynamics.

In contrast, the most common numerical integration methods, including the widely used Runge-Kutta schemes, neither preserve the Lie group structure nor these geometric properties. In addition, standard Runge-Kutta methods fail to capture the energy dissipation of a controlled system accurately [4]. Additionally, if we integrate (3) by a typical Runge-Kutta scheme, the quantity $R^T R$ inevitably drifts from the identity matrix as the simulation time increases. It is often proposed to parameterize (3) by Euler angles or unit quaternions. However, Euler angles are not global expressions of the attitude since they have associated singularities. Unit quaternions do not exhibit singularities, but are constrained to lie on the unit three-sphere $\mathbb{S}^3$, and general numerical integration methods do not preserve the unit length constraint. Therefore, quaternions have the same numerical drift problem. Renormalizing the quaternion vector at each step tends to break other conservation properties. Furthermore, unit quaternions, which are diffeomorphic to $\mathrm{SU}(2)$, double cover $\mathrm{SO}(3)$. So there are inevitable ambiguities in expressing the attitude.

In [3], Lie group variational integrators are introduced by explicitly adapting Lie group methods [5] to the discrete variational principle [4]. They have the desirable property that they are symplectic and momentum preserving, and they exhibit good energy behavior for an exponentially long time period. They also preserve the Euclidian Lie group structure without the use of local charts, reprojection, or constraints. These geometrically exact numerical integration methods yield highly efficient and accurate computational algorithms for rigid body dynamics. They avoid singularities and ambiguities.

Using the results presented in [6], a Lie group variational integrator on $\mathrm{SE}(3)$ for equations (1)–(4) is given by

$$x_{k+1} = x_k + \frac{h}{m}\gamma_k + \frac{h^2}{2m}\left(f_k + u_k^f\right), \tag{7}$$

$$\gamma_{k+1} = \gamma_k + \frac{h}{2}\left(f_k + u_k^f\right) + \frac{h}{2}\left(f_{k+1} + u_{k+1}^f\right), \tag{8}$$

$$hS(\Pi_k + \frac{h}{2}\left(M_k + u_k^m\right)) = F_k J_d - J_d F_k^T, \tag{9}$$

$$R_{k+1} = R_k F_k, \tag{10}$$

$$\Pi_{k+1} = F_k^T \Pi_k + \frac{h}{2} F_k^T \left(M_k + u_k^m\right) + \frac{h}{2}\left(M_{k+1} + u_{k+1}^m\right), \tag{11}$$

where the subscript k denotes the kth discrete variables for a fixed integration step size $h \in \mathbb{R}$. $J_d \in \mathbb{R}^{3\times 3}$ is a nonstandard moment of inertia matrix defined by $J_d = \frac{1}{2}\mathrm{tr}[J]\, I_{3\times 3} - J$. $F_k \in \mathrm{SO}(3)$ is the relative attitude between adjacent integration steps.

For given $(R_k, x_k, \Pi_k, \gamma_k)$ and control inputs, (9) is solved to find F_k. Then (R_{k+1}, x_{k+1}) are obtained by (10),(7). Using (5),(6), (f_{k+1}, M_{k+1}) are computed, and they are used to find $(\Pi_{k+1}, \gamma_{k+1})$ by (11),(8). This yields a map $(R_k, x_k, \Pi_k, \gamma_k) \mapsto (R_{k+1}, x_{k+1}, \Pi_{k+1}, \gamma_{k+1})$, and this process is repeated. The only implicit part is (9). The actual computation of F_k is done in the Lie algebra $\mathfrak{so}(3)$ of dimension 3, and the rotation matrices are updated by multiplication. This approach is completely different from integration of the kinematics equation (3); there is no excessive computational burden. It can be shown that this integrator has second order accuracy. The properties of these discrete equations of motion are discussed in more detail in [3], [6].

III. Optimal Impulsive Control of a Rigid Body

We formulate an optimal impulsive control problem for a rigid body on $\mathrm{SE}(3)$, and we develop sensitivity derivatives. They are used in our computational method for solve optimal impulsive control problems.

A. Problem formulation

An optimal impulsive control problem is formulated as a maneuver of a rigid body from a given initial configuration $(R_0, x_0, \Pi_0, \gamma_0)$ to a desired configuration described by

$$\left\{(R_N, x_N, \Pi_N, \gamma_N) \in \mathrm{T}^*\mathrm{SE}(3) \middle| \mathcal{C}(R_N, x_N, \Pi_N, \gamma_N) = 0\right\},$$

where $\mathcal{C}(\cdot) : \mathrm{T}^*\mathrm{SE}(3) \mapsto \mathbb{R}^c$ during the given maneuver time N. Two impulsive control inputs are applied at the initial time and the terminal time. We assume that the control inputs are purely impulsive, which means that each impulse changes the momentum of the rigid body instantaneously, but it does not have any effect on the position and the attitude of the rigid body at that instant. The motion of the rigid body between the initial time and the terminal time is uncontrolled. i.e. $u_k^f = u_k^m = 0$. The performance index is the sum of the magnitudes of the initial impulse and the terminal impulse. It is equivalent to minimizing the sums of the initial momentum change and the terminal momentum change.

We transform this optimal impulsive control problem into a parameter optimization problem. Let (Π_0^+, γ_0^+) be the initial momentum after the initial impulsive control. Then, the terminal states are determined by the discrete equations of motion, and the momentum after the terminal impulsive control, (Π_N^+, γ_N^+), can be computed by the terminal constraint. Therefore, the performance index and the constraint are completely determined by (Π_0^+, γ_0^+). Thus, the optimal impulsive control on $\mathrm{SE}(3)$ is formulated as

$$\begin{gathered}
\text{given} : (R_0, x_0, \Pi_0, \gamma_0), N \\
\min_{\Pi_0^+, \gamma_0^+} \mathcal{J} = \left\|\Pi_0^+ - \Pi_0\right\| + \left\|\gamma_0^+ - \gamma_0\right\| \\
+ \left\|\Pi_N^+ - \Pi_N\right\| + \left\|\gamma_N^+ - \gamma_N\right\|, \\
\text{such that } \mathcal{C}(R_N, x_N, \Pi_N^+, \gamma_N^+) = 0, \\
\text{subject to discrete equations of motion (7)–(11).}
\end{gathered}$$

If the desired values for all of the terminal states are specified by the constraints, then there is no freedom for optimization. This problem degenerates to a two point boundary value problem on $\mathrm{SE}(3)$, which can be considered as an extension of the Lambert problem for the restricted two body problem. A similar optimal control problem for attitude dynamics of a rigid body on $\mathrm{SO}(3)$ is studied in [7].

B. Sensitivity derivatives

Variational model: The variation of $g_k = (R_k, x_k) \in \mathrm{SE}(3)$ can be expressed in terms of a Lie algebra element $\eta_k \in \mathfrak{se}(3)$ and the exponential map as $g_k^\epsilon = g_k \exp \epsilon\eta_k$. The corresponding infinitesimal variation is given by

$$\delta g_k = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} g_k \exp \epsilon\eta_k = \mathrm{T}_e L_{g_k} \cdot \eta_k.$$

Using homogeneous coordinates [8], the above equation is written in a matrix equation as

$$\begin{aligned}\begin{bmatrix}\delta R_k & \delta x_k \\ 0 & 0\end{bmatrix} &= \begin{bmatrix}R_k & x_k \\ 0 & 1\end{bmatrix}\begin{bmatrix}S(\zeta_k) & \chi_k \\ 0 & 0\end{bmatrix}, \\ &= \begin{bmatrix}R_k S(\zeta_k) & R_k\chi_k \\ 0 & 0\end{bmatrix}, \end{aligned} \tag{12}$$

where $\zeta_k, \chi_k \in \mathbb{R}^3$ so that $(S(\zeta_k), \chi_k) \in \mathfrak{se}(3)$. This gives an expression for the infinitesimal variation of a Lie group element in terms of its Lie algebra. Then, small perturbations from a given trajectory on $\mathrm{T}^*\mathrm{SE}(3)$ can be written as

$$x_k^\epsilon = x_k + \epsilon\delta x_k, \tag{13}$$
$$\gamma_k^\epsilon = \gamma_k + \epsilon\delta\gamma_k, \tag{14}$$
$$\Pi_k^\epsilon = \Pi_k + \epsilon\delta\Pi_k, \tag{15}$$
$$R_k^\epsilon = R_k + \epsilon R_k S(\zeta_k) + \mathcal{O}(\epsilon^2), \tag{16}$$

where $\delta x_k, \delta\gamma_k, \delta\Pi_k, \zeta_k$ are considered in $\mathbb{R}^3$.

We derive expressions for the constrained variation of F_k using (10) and (16). Since $F_k = R_k^T R_{k+1}$ by (10), the infinitesimal variation δF_k is given by

$$\delta F_k = \delta R_k^T R_{k+1} + R_k^T \delta R_{k+1} = -S(\zeta_k)F_k + F_k S(\zeta_{k+1}).$$

We can also express $\delta F_k = F_k S(\xi_k)$ for $\xi_k \in \mathbb{R}^3$, using (12). Using the property $S(R^T x) = R^T S(x) R$ for all $R \in \mathrm{SO}(3)$ and $x \in \mathbb{R}^3$, we obtain the constrained variation of F_k

$$\xi_k = -F_k^T \zeta_k + \zeta_{k+1}. \tag{17}$$

Linearized equations of motion: Substituting the variation model (13)–(16) and the constrained variation (17) into the equations of motion (7)–(11), and ignoring higher order terms, the linearized equation of motion can be written as

$$z_{k+1} = A_k z_k, \tag{18}$$

where $z_k = [\delta x_k; \delta\gamma_k; \zeta_k; \delta\Pi_k] \in \mathbb{R}^{12}$, and $A_k \in \mathbb{R}^{12\times 12}$ can be suitably defined. The solution of (18) is obtained as

$$z_N = \Phi z_0, \tag{19}$$

where $\Phi \in \mathbb{R}^{12\times 12}$ represents the sensitivity derivatives of the terminal state with respect to the initial state on $\mathrm{SE}(3)$.

C. Computational approach

We solve the optimal impulsive control problem by the Sequential Quadratic Programming (SQP) method using analytical expressions for the gradients of the performance index and the constraints. The exact computation of the gradients are crucial for efficient numerical optimization. For the given problem, $\delta x_0 = \zeta_0 = 0$ since the initial position and the initial attitude are fixed. Thus, (19) is written as

$$\begin{bmatrix}\delta x_N \\ \delta\gamma_N \\ \zeta_N \\ \delta\Pi_N\end{bmatrix} = \begin{bmatrix}\Phi^{12} & \Phi^{14} \\ \Phi^{22} & \Phi^{24} \\ \Phi^{32} & \Phi^{34} \\ \Phi^{42} & \Phi^{44}\end{bmatrix}\begin{bmatrix}\delta\gamma_0^+ \\ \delta\Pi_0^+\end{bmatrix}, \tag{20}$$

where $\Phi^{ij} \in \mathbb{R}^{3\times 3}$, $i, j \in (1, 2, 3, 4)$ are submatrices of Φ. The above equation represents the sensitivities of the terminal state with respect to the initial momentum (Π_0^+, γ_0^+). Therefore, we can obtain expressions for gradients of the performance index and the constraints, and any Newton type numerical approach can be applied.

IV. Optimal Control of a Rigid Body

We formulate an optimal control problem for a rigid body on $\mathrm{SE}(3)$ assuming that control forces and moments are applied during the maneuver. Necessary conditions for optimality are developed and computational approaches are presented to solve the corresponding two point boundary value problem.

A. Problem formulation

An optimal impulsive control problem is formulated as a maneuver of a rigid body from a given initial configuration $(R_0, x_0, \Pi_0, \gamma_0)$ to a desired configuration $(R_N^d, x_N^d, \Pi_N^d, \gamma_N^d)$ during the given maneuver time N. Control inputs are parameterized by their value at each time step. The performance index is the square of the weighted l_2 norm of the control inputs.

$$\begin{gathered}\text{given: } (x_0, \gamma_0, R_0, \Pi_0),\ (x_N^d, \gamma_N^d, R_N^d, \Pi_N^d),\ N, \\ \min_{u_{k+1}} \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2}(u_{k+1}^f)^T W_f u_{k+1}^f + \frac{h}{2}(u_{k+1}^m)^T W_m u_{k+1}^m, \\ \text{such that } (x_N, \gamma_N, R_N, \Pi_N) = (x_N^d, \gamma_N^d, R_N^d, \Pi_N^d), \\ \text{subject to discrete equations of motion (7)–(11),}\end{gathered}$$

where $W_f, W_m \in \mathbb{R}^{3\times 3}$ are symmetric positive definite matrices. Here we use a modified version of the discrete equations of motion with first order accuracy, because it yields a compact form for the necessary conditions, which are developed the following subsection. A similar optimal control problem for attitude dynamics on $\mathrm{SO}(3)$ is studied in [9].

B. Necessary conditions for optimality

Define an augmented performance index as

$$\begin{aligned}\mathcal{J}_a = \sum_{k=0}^{N-1} &\frac{h}{2}(u_{k+1}^f)^T W^f u_{k+1}^f + \frac{h}{2}(u_{k+1}^m)^T W^m u_{k+1}^m \\ &+ \lambda_k^{1,T}\left\{-x_{k+1} + x_k + \frac{h}{m}\gamma_k\right\} \\ &+ \lambda_k^{2,T}\left\{-\gamma_{k+1} + \gamma_k + hf_{k+1} + hu_{k+1}^f\right\} \\ &+ \lambda_k^{3,T} S^{-1}\big(\mathrm{logm}(F_k - R_k^T R_{k+1})\big) \\ &+ \lambda_k^{4,T}\left\{-\Pi_{k+1} + F_k^T\Pi_k + h\left(M_{k+1} + u_{k+1}^m\right)\right\},\end{aligned}$$

where $\lambda_k^i \in \mathbb{R}^3$ are Lagrange multipliers. The constraint (9) is considered implicitly using a constrained variation. Using the variational model (13)–(16), the constrained variation (17), and the fact that the variations vanish at $k = 0, N$, we obtain the infinitesimal variation of $\mathcal{J}_a$ as

$$\delta\mathcal{J}_a = \sum_{k=1}^{N-1} h\delta u_k^{f,T} \left\{ W_f u_k^f + \lambda_{k-1}^2 \right\} + h\delta u_k^{m,T} \left\{ W_m u_k^m + \lambda_{k-1}^4 \right\} + z_k^T \left\{ -\lambda_{k-1} + A_k^T \lambda_k \right\},$$

where $\lambda_k = [\lambda_k^1; \lambda_k^2; \lambda_k^3; \lambda_k^4] \in \mathbb{R}^{12}$, and $A_k \in \mathbb{R}^{12\times 12}$ is presented in (18).

Since $\delta\mathcal{J}_a = 0$ for all variations, we obtain necessary conditions for optimality as follows.

$$x_{k+1} = x_k + \frac{h}{m}\gamma_k, \tag{21}$$

$$\gamma_{k+1} = \gamma_k + hf_{k+1} + hu_{k+1}^f, \tag{22}$$

$$hS(\Pi_k) = F_k J_d - J_d F_k^T, \tag{23}$$

$$R_{k+1} = R_k F_k, \tag{24}$$

$$\Pi_{k+1} = F_k^T \Pi_k + hM_{k+1} + hu_{k+1}^m, \tag{25}$$

$$u_{k+1}^f = -W_f^{-1}\lambda_k^2, \tag{26}$$

$$u_{k+1}^m = -W_m^{-1}\lambda_k^4, \tag{27}$$

$$\lambda_k = A_{k+1}^T \lambda_{k+1}. \tag{28}$$

In the above equations, the only implicit part is (23). For a given initial condition $(R_0, x_0, \Pi_0, \gamma_0)$ and λ_0, we can find F_0 by solving (23). Then, R_1, x_1 is obtained by (24),(21), and the control input u_1^f, u_1^m is obtained by (26),(27). γ_1, Π_1 can be obtained by (22),(25). Now we compute $(R_1, x_1, \Pi_1, \gamma_1)$. We solve (23) to find F_1. Finally, λ_1 can be obtained by (28). This yields a map $\{(R_0, x_0, \Pi_0, \gamma_0), \lambda_0\} \mapsto \{(R_1, x_1, \Pi_1, \gamma_1), \lambda_1\}$, and this process can be repeated.

C. Computational Approach

The necessary conditions for optimality are expressed in terms of a two point boundary problem on $\mathrm{T}^*\mathrm{SE}(3)$ and its dual. This problem is to find the optimal discrete flow, multiplier, and control inputs to satisfy the equations of motion (21)–(25), optimality conditions (26),(27), multiplier equations (28), and boundary conditions simultaneously.

We use a neighboring extremal method [10]. A nominal solution satisfying all of the necessary conditions except the boundary conditions is chosen. The unspecified initial multiplier is updated by successive linearization so as to satisfy the specified terminal boundary conditions in the limit. This is also referred to as a shooting method. The main advantage of the neighboring extremal method is that the number of iteration variables is small. In other approaches, the initial guess of control input history or multiplier variables are iterated, so the number of optimization parameters are proportional to the number of discrete time steps.

The difficulty is that the extremal solutions are sensitive to small changes in the unspecified initial multiplier values. The nonlinearities also make it hard to construct an accurate estimate of sensitivity, and it may result in numerical ill-conditioning. Therefore, it is important to compute the sensitivities accurately to apply the neighboring extremal method.

Here the optimality conditions (26) and (27) are substituted into the equations of motion and the multiplier equations. The sensitivities of the specified terminal boundary conditions with respect to the unspecified initial multiplier conditions is obtained by a linear analysis.

Similar to (18), the linearized equations of motion can be written as

$$z_{k+1} = A_k z_k + \mathcal{A}^{12}\delta\lambda_k, \tag{29}$$

where $\mathcal{A}_k^{12} = -h\mathrm{diag}[0, W_f^{-1}, 0, W_m^{-1}] \in \mathbb{R}^{12\times 12}$. We can linearize the multiplier equations (28) to obtain

$$\delta\lambda_k = \mathcal{A}_{k+1}^{21} z_{k+1} + A_{k+1}^T \delta\lambda_{k+1}, \tag{30}$$

where $\mathcal{A}_{k+1}^{21} \in \mathbb{R}^{12\times 12}$ can be defined properly. The solution of the linear equations (29) and (30) can be obtained as

$$\begin{bmatrix} z_N \\ \delta\lambda_N \end{bmatrix} = \begin{bmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{bmatrix} \begin{bmatrix} z_0 \\ \delta\lambda_0 \end{bmatrix},$$

where $\Psi^{ij} \in \mathbb{R}^{12\times 12}$.

For the given two point boundary value problem $z_0 = 0$ since the initial condition is fixed, and λ_N is free. Thus,

$$z_N = \Psi_{12}\delta\lambda_0. \tag{31}$$

The matrix Ψ_{12} represents the sensitivity of the specified terminal boundary conditions with respect to the unspecified initial multipliers. Using this sensitivity, an initial guess of the unspecified initial conditions is iterated to satisfy the specified terminal conditions in the limit.

Any type of Newton iteration can be applied. We use a line search with backtracking algorithm, referred to as Newton-Armijo iteration in [11]. The procedure is summarized as follows.

```
Guess an initial multiplier λ_0.
Find x_k, γ_k, Π_k, R_k, λ_k using (21)–(28).
Compute the terminal B.C. error; Error = ‖z_N‖.
Set Error^t = Error,  i = 1.
while Error > ε_S.
    Find a line search direction; D = Ψ_12^{-1}.
    Set c = 1.
    while Error^t > (1 − 2αc)Error
        Choose a trial multiplier λ_0^t = λ_0 + cDz_N.
        Find x_k, γ_k, Π_k, R_k, λ_k using (21)–(28).
        Compute the error; Error^t = ‖z_N^t‖.
        Set c = c/10,  i = i + 1.
    end while
    Set λ_0 = λ_0^t, Error = Error^t. (accept the trial)
end while
```

Here i is the number of iterations, and $\epsilon_S, \alpha \in \mathbb{R}$ are a stopping criterion and a scaling factor, respectively. The outer loop finds a search direction by computing the sensitivity derivatives, and the inner loop performs a line search to find the largest step size $c \in \mathbb{R}$ along the search direction. The error in satisfaction of the terminal boundary condition is determined at each inner iteration.

V. Numerical Examples

A. Restricted Full Two Body Problem

We study a maneuver of a rigid spacecraft under a central gravity field. We assume that the mass of the spacecraft is negligible compared to the mass of a central body, and we consider a fixed frame attached to the central body as an inertial frame. The resulting model is a Restricted Full Two Body Problem (RF2BP).

The spacecraft is modeled as a dumbbell, which consists of two equal spheres and a massless rod. The gravitational potential is given by

$$U(x, R) = -\frac{GMm}{2} \sum_{q=1}^{2} \frac{1}{\|x + R\rho^q\|}, \qquad (32)$$

where $G \in \mathbb{R}$ is the gravitational constant, $M, m \in \mathbb{R}$ are the mass of the central body, and the mass of the dumbbell, respectively. The vector $\rho^q \in \mathbb{R}^3$ is the position of the qth sphere from the mass center of the dumbbell expressed in the body fixed frame ($q \in \{1, 2\}$). The mass, length, and time dimensions are normalized by the mass of the dumbbell, the radius of a reference circular orbit, and its orbital period.

B. Optimal Impulsive Control

We study an impulsive orbital transfer problem with an attitude change. Initially, the spacecraft is on a reference circular orbit. We consider two cases. In the first case, the spacecraft moves to a desired circular orbit and the desired values for all of the terminal state are specified. There is no freedom for optimization, and the resulting problem is a two point boundary value problem on $\mathrm{SE}(3)$. This maneuver can be considered as a generalization of Hohmann transfer [12]. The desired maneuver involves doubling the orbital radius in addition to a large angle attitude change.

In the second case, the terminal constraints are relaxed such that the spacecraft is allowed to transfer to any point on the desired orbit. The desired terminal orbit is described by its orbital radius $r_d \in \mathbb{R}$, and a directional vector $e_n \in \mathbb{S}^2$ normal to the orbital plane. Two constraints are imposed to locate the dumbbell in the desired orbital plane with the desired orbital radius, and one constraint is applied to align the dumbbell to the normal direction.

The gradients of the performance index and the constraints are obtained by using (20). We use Matlab `fmincon` function as an implementation of the SQP algorithm. Figures 1 and 2 show the spacecraft maneuver, and linear velocity and angular velocity responses, where red circles denote the velocities before the initial impulse and the velocities after the terminal impulse. Thus, differences between solid lines and red circles are proportional to the impulsive controls. (Simple animations which show these maneuvers of the spacecraft can be found at `http://www.umich.edu/~tylee`.) The error in satisfaction of the terminal boundary value of the first case is 4.77×10^{-15}. The performance index and the maximum violations of the constraints for the second case are 1.2305 and 3.88×10^{-15}, respectively.

C. Optimal Control

We study an optimal orbital transfer problem to increase the orbital inclination by $60\,\mathrm{deg}$, and an orbital capture problem to the reference circular orbit.

Figures 3 and 4 show the optimized spacecraft maneuver, control inputs history. For each case, the performance indices are 13.03 and 20.90, and the maximum violations of the constraint are 3.35×10^{-13} and 3.26×10^{-13}, respectively.

Figures 3.(b) and 4.(b) show the violation of the terminal boundary condition according to the number of iterations in a logarithmic scale. Red circles denote outer iterations in Newton-Armijo iteration to compute the sensitivity derivatives. For all cases, the initial guesses of the unspecified initial multiplier are arbitrarily chosen. The error in satisfaction of the terminal boundary condition converges quickly to machine precision after the solution is close to the local minimum at around 20th iteration. These convergence results are consistent with the quadratic convergence rates expected of Newton methods with accurately computed gradients.

The neighboring extremal method, also referred to as the shooting method, is numerically efficient in the sense that the number of optimization parameters is minimized. But, this approach may be prone to numerical ill-conditioning [13]. A small change in the initial multiplier can cause highly nonlinear behavior of the terminal attitude and angular momentum. It is difficult to compute the gradient for Newton iterations accurately, and the numerical error may not converge.

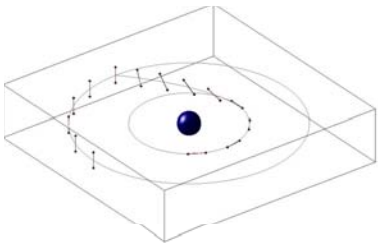
However, the numerical examples presented in this paper show excellent numerical convergence properties. This is because the proposed computational algorithms on $\mathrm{SE}(3)$ are geometrically exact and numerically accurate.

The dynamics of a rigid body arises from Hamiltonian mechanics, which have neutral stability, and its adjoint system is also neutrally stable. The proposed Lie group variational integrator and the discrete multiplier equations, obtained from variations expressed in the Lie algebra, preserve the neutral stability property numerically. Therefore the sensitivity derivatives are computed accurately.

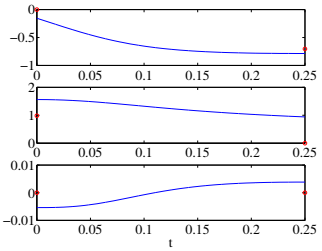
VI. Conclusions

Optimal control problems for combined orbital and rotational maneuvers of a rigid body are formulated and efficient computational procedures are proposed. The dynamics are discretized by a Lie group variational integrator, and sensitivity derivatives are developed by a linear analysis. Discrete necessary conditions for optimality are constructed, and the corresponding two point boundary value problem is solved efficiently.

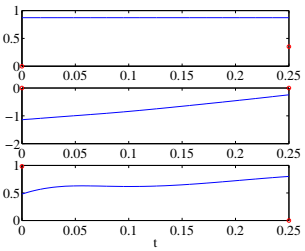
This approach is geometrically exact in the sense that the Lie group variational integrator preserves the group structure as well as the geometric invariant properties, and the sensitivity derivatives are expressed in terms of its Lie algebra. Since the configuration of a rigid body is defined globally using an element of $\mathrm{SE}(3)$, this approach completely avoids singularity or ambiguity arising from other representations such as Euler angles and quaternions. Numerical examples show the efficiency of the proposed computational approach.



(a) Spacecraft maneuver

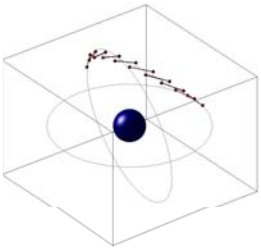


(b) Velocity v

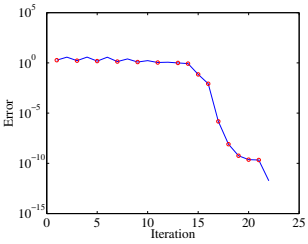


(c) Angular velocity Ω

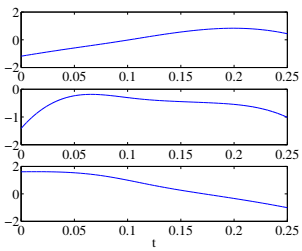
Fig. 1. TPBVP: Orbital radius change



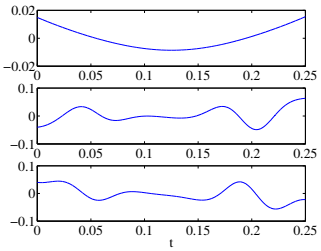
(a) Spacecraft maneuver



(b) Convergence rate

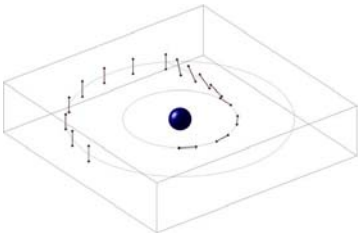


(c) Control force u^f

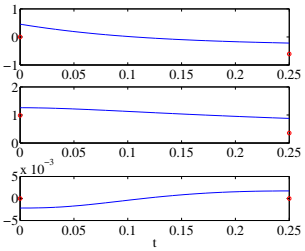


(d) Control moment u^m

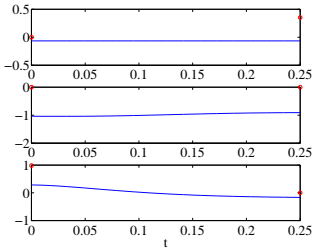
Fig. 3. Optimal control: Orbital inclination change



(a) Spacecraft maneuver

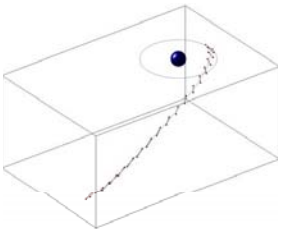


(b) Velocity v

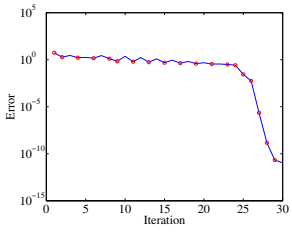


(c) Angular velocity Ω

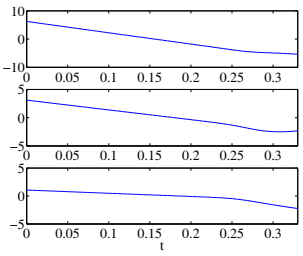
Fig. 2. Optimal impulsive control: Orbital radius change



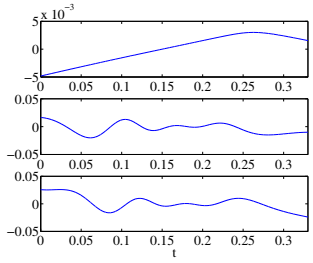
(a) Spacecraft maneuver



(b) Convergence rate



(c) Control force u^f



(d) Control moment u^m

Fig. 4. Optimal control: Orbital capture

REFERENCES

[1] K. Spindler, "Optimal control on Lie groups with applications to attitude control," *Mathematics of Control, Signals, and Systems*, vol. 11, pp. 197–219, 1998.

[2] S. Sastry, "Optimal control on Lie groups," in *Proceedings of the Third International Congress on Industrial and Applied Mathematics (ICIAM)*, 1995.

[3] T. Lee, M. Leok, and N. H. McClamroch, "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum," in *Proceedings of the IEEE Conference on Control Applications*, 2005, pp. 962–967.

[4] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[5] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, "Lie-group methods," *Acta Numerica*, vol. 9, pp. 215–365, 2000.

[6] T. Lee, M. Leok, and N. H. McClamroch, "Lie group variational integrators for the full body problem," *Computer Methods in Applied Mechanics and Engineering*, 2005, submitted. [Online]. Available: http://arxiv.org/math.NA/0508365

[7] ——, "Attitude maneuvers of a rigid spacecraft in a circular orbit," in *Proceedings of the American Control Conference*, 2006, pp. 1742–1747. [Online]. Available: http://arxiv.org/math.NA/0509299

[8] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1993.

[9] T. Lee, M. Leok, and N. H. McClamroch, "Optimal attitude control of a rigid body using geometrically exact computations on SO(3)," *Journal of Optimization Theory and Applications*, 2006, submitted. [Online]. Available: http://arxiv.org/math.OC/0601424

[10] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control*. Hemisphere Publishing Corporation, 1975.

[11] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.

[12] J. M. A. Danby, *Fundamentals of Celestial Mechanics*. Willmann Bell Inc., 1988.

[13] J. T. Betts, *Practical Methods for Optimal Control Using Nonlinear Programming*. SIAM, 2001.

Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

Deterministic Global Attitude Estimation

Taeyoung Lee*†, Amit Sanyal, Melvin Leok*, and N. Harris McClamroch†

Abstract— **A deterministic attitude estimation problem for a rigid body in an attitude dependent potential field with bounded measurement errors is studied. An attitude estimation scheme that does not use generalized coordinate representations of the attitude is presented here. Assuming that the initial attitude, angular velocity and measurement noise lie within given ellipsoidal bounds, an uncertainty ellipsoid that bounds the attitude and the angular velocity of the rigid body is obtained. The center of the uncertainty ellipsoid provides point estimates, and its size gives the accuracy of the estimates. The point estimates and the uncertainty ellipsoids are propagated using a Lie group variational integrator and its linearization, respectively. The estimation scheme is optimal in the sense that the attitude estimation error and the size of the uncertainty ellipsoid is minimized at each measurement instant, and it is global since the attitude is represented by a rotation matrix.**

I. Introduction

Attitude estimation is often a prerequisite for controlling aerospace and underwater vehicles, mobile robots, and other mechanical systems moving in space. The attitude determination problem for a rigid body from vector measurements was first posed in [1]. A sample of the literature in attitude estimation can be found in [2], [3], [4].

Most existing attitude estimation schemes use generalized coordinates to represent the attitude. As is well known, minimal coordinate representations of the rotation group, like Euler angles, Rodrigues parameters, and modified Rodrigues parameters, lead to geometric or kinematic singularities. Non-minimal coordinate representations, like quaternions used in the quaternion estimation (QUEST) algorithm and its several variants ([3], [5]), have their own associated problems. Besides the extra unit norm constraint one needs to impose on the quaternion, the quaternion representation, which is diffeomorphic to $\mathrm{SU}(2)$, double covers $\mathrm{SO}(3)$. As such, it has an inevitable ambiguity in expressing the attitude.

A stochastic state estimator requires probabilistic models for the state uncertainty and the noise. However, statistical properties of the uncertainty and the noise are often not available. We usually make statistical assumptions on disturbance and noise in order to make the estimation problem mathematically tractable. In many practical situations such idealized assumptions are not appropriate, and this may cause poor estimation performance [6].

An alternative deterministic approach is to specify bounds on the uncertainty and the measurement noise without an assumption on their distribution. Noise bounds are available in many cases, and deterministic estimation is robust to the noise distribution. An efficient but flexible way to describe the bounds is using ellipsoidal sets, referred to as uncertainty ellipsoids. The idea of the deterministic estimation process is based on set theory results developed in [7]; optimal deterministic estimation problems are studied in [8] and [9] using uncertainty ellipsoids.

In this paper, we study attitude estimation problems for the uncontrolled dynamics of a rigid body in an attitude-dependent potential field using uncertainty ellipsoids. The estimation scheme we present has the following important features: (1) the attitude is globally represented by a rotation matrix without using coordinates, (2) the deterministic estimator is distinguished from a Kalman or extended Kalman filter, (3) the measurement errors are assumed to be bounded but there is no restriction on their distribution, and (4) the estimates are optimal in the sense that the size of uncertainty is minimized at each estimation step.

This paper is organized as follows. The attitude determination problem from vector observations is introduced in Section II. The attitude estimation problem is formulated in Section III, and the attitude estimation scheme with angular velocity measurements is developed in Section IV. Numerical examples are presented in Section V.

II. Attitude Determination from vector observations

Attitude of a rigid body is defined as the orientation of a body fixed frame with respect to a reference frame. It is represented by a rotation matrix that is a 3×3 orthogonal matrix with determinant 1. Rotation matrices have a group structure denoted by $\mathrm{SO}(3)$. The group action of $\mathrm{SO}(3)$ on $\mathbb{R}^3$ transforms a vector represented in the body frame into the reference frame. In the attitude estimation problem, we measure directions in the body frame to fixed points with known directions in the reference frame. The directions in the body frame are transformed into the known reference directions by pre-multiplying by the rotation matrix defining the attitude of the rigid body. The rotation matrix can be estimated by minimizing an error between the transformed measured directions and the known reference directions.

We denote the ith known direction vector in the reference frame as $e^i \in \mathbb{S}^2$, and the corresponding vector represented

Taeyoung Lee, Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 tylee@umich.edu

Amit Sanyal, Mechanical and Aerospace Engineering, Arizona State University, Tempe, AZ 85287 sanyal@asu.edu

Melvin Leok, Mathematics, Purdue University, West Lafayette, IN 47907 mleok@math.purdue.edu

N. Harris McClamroch, Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 nhm@umich.edu

*This research has been supported in part by NSF under grant DMS-0504747, and by a grant from the Rackham Graduate School, University of Michigan.

†This research has been supported in part by NSF under grant ECS-0244977.

in the body frame as $b^i \in \mathbb{S}^2$. These direction vectors are normalized to have unit lengths. The e^i and b^i vectors are related by a rotation matrix $R \in \mathrm{SO}(3)$ that defines the attitude of the rigid body; $e^i = Rb^i$, for all $i \in \{1, 2, \cdots, m\}$, where m is the number of measurements.

We assume that b^i is measured by sensors in the body frame. Let the measured direction vector be $\tilde{b}^i \in \mathbb{S}^2$, which contains sensor errors, and denote the estimated rotation matrix by $\hat{R} \in \mathrm{SO}(3)$. The estimation error is given by $e^i - \hat{R}\tilde{b}^i$. The attitude determination problem consists of finding $\hat{R} \in \mathrm{SO}(3)$ such that the weighted 2 norm of those errors is minimized.

$$\min_{\hat{R}} \mathcal{J} = \frac{1}{2}\sum_{i=1}^{m} w_i (e^i - \hat{R}\tilde{b}^i)^T (e^i - \hat{R}\tilde{b}^i), \qquad (1)$$
$$\text{subject to } \hat{R} \in \mathrm{SO}(3),$$

where $E = \left[e^1, \cdots, e^m\right] \in \mathbb{R}^{3\times m}$, $\tilde{B} = \left[\tilde{b}^1, \cdots, \tilde{b}^m\right] \in \mathbb{R}^{3\times m}$, and $W = \mathrm{diag}\left[w^1, \cdots, w^m\right] \in \mathbb{R}^{m\times m}$ is a weighting factor for each measurement.

This problem is known as Wahba's problem [1]. The original solution of Wahba's problem is given in [10], and a solution expressed in terms of quaternions (QUEST) is presented in [11]. We use the solution expressed in terms of a rotation matrix without using generalized coordinates [12]. A necessary and sufficient condition for optimality of (1) is given by

$$\hat{R} = SL \in \mathrm{SO}(3), \quad S = S^T > 0 \qquad (2)$$

where $L = EW\tilde{B}^T \in \mathbb{R}^{3\times 3}$ is non-singular. The unique solution of (2) is obtained by QR factorization of $L = Q_q Q_r$

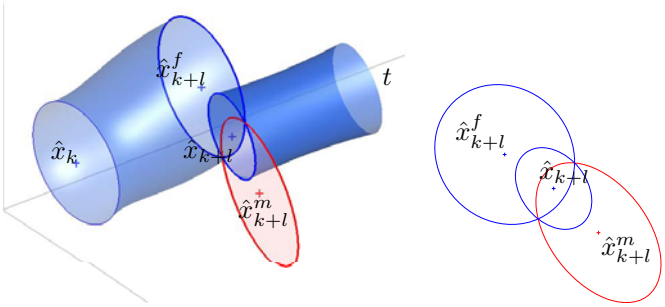
$$\hat{R} = \left(Q_q \sqrt{(Q_r Q_r^T)^{-1}} Q_q^T\right) L, \qquad (3)$$

where $Q_q \in \mathrm{SO}(3)$, $Q_r \in \mathbb{R}^{3\times 3}$ is an upper triangular matrix, and the symmetric positive definite (principal) square root is used. Equation (3) is the unique solution of Wahba's problem [12].

III. Attitude Estimation Problem Formulation

A. State bounding estimation

We use deterministic state bounding estimation using ellipsoidal sets, referred to as uncertainty ellipsoids, to describe state uncertainty and measurement noise. This deterministic estimation procedure has steps similar to those in the Kalman filter, and is illustrated in Fig. 1. The left figure shows time evolution of an uncertainty ellipsoid, and the right figure shows a cross section at a fixed measurement instant. At the kth time step, the state is bounded by an uncertainty ellipsoid centered at $\hat{x}_k$. This initial ellipsoid is propagated through time. Suppose that the state is measured next at the $(k+l)$th time step, when the predicted uncertainty ellipsoid is centered at $\hat{x}^f_{k+l}$. At this instant, the measurement uncertainty ellipsoid is centered at $\hat{x}^m_{k+l}$. The actual state then lies in the intersection of the two ellipsoids. In the estimation process, we find a new ellipsoid that contains this intersection, as shown in the right figure. The center of the new ellipsoid, $\hat{x}_{k+l}$ gives a point estimate of the state at time step $k+l$, and the magnitude of the new uncertainty ellipsoid measures the estimation accuracy. The deterministic estimates are optimal in the sense that the sizes of the ellipsoids are minimized.



(a) Propagation of uncertainty ellipsoid (b) Filtering procedure

Fig. 1. Uncertainty ellipsoids

B. Equations of motion

We consider estimation of the attitude dynamics of a rigid body in the presence of an attitude dependent potential, $U(\cdot) : \mathrm{SO}(3) \mapsto \mathbb{R}$, $R \in \mathrm{SO}(3)$. Systems that can be so modeled include a free rigid body, spacecraft on a circular orbit with gravity gradient effects [13], or a 3D pendulum [14]. The continuous equations of motion are

$$J\dot{\Omega} + \Omega \times J\Omega = M, \qquad (4)$$
$$\dot{R} = RS(\Omega), \qquad (5)$$

where $J \in \mathbb{R}^{3\times 3}$ is the moment of inertia matrix of the rigid body, $\Omega \in \mathbb{R}^3$ is the angular velocity of the body expressed in the body fixed frame, and $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is a skew mapping defined by $S(x)y = x \times y$ for all $x, y \in \mathbb{R}^3$. The vector $M \in \mathbb{R}^3$ is the moment due to the potential, determined by $S(M) = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}$, or more explicitly,

$$M = r_1 \times v_{r_1} + r_2 \times v_{r_2} + r_3 \times v_{r_3}, \qquad (6)$$

where $r_i, v_{r_i} \in \mathbb{R}^{1\times 3}$ are the ith row vectors of R and $\frac{\partial U}{\partial R}$, respectively.

General numerical integration methods like the popular Runge-Kutta schemes, typically preserve neither first integrals nor the characteristics of the configuration space, $\mathrm{SO}(3)$. In particular, the orthogonal structure of the rotation matrices is not preserved numerically. It is often proposed to parameterize (5) by Euler angles or quaternions instead of integrating (5) directly. However, Euler angles yield only local representations of the attitude and they have singularities. Unit quaternions do not exhibit singularities, but they have the manifold structure of the three sphere $\mathbb{S}^3$, and double cover $\mathrm{SO}(3)$. Consequently, the unit quaternion representing the attitude is inevitably ambiguous. In addition, general numerical integration methods do not preserve the unit length constraint. Therefore, quaternions have the same numerical drift problem as rotation matrices.

Lie group variational integrators preserve the group structure without the use of local charts, reprojection, or constraints, they are symplectic and momentum preserving, and they exhibit good energy behavior for an exponentially long time period. The following Lie group variational integrator for the attitude dynamics of a rigid body is presented in [14]:

$$hS(J\Omega_k + \frac{h}{2}M_k) = F_k J_d - J_d F_k^T, \tag{7}$$

$$R_{k+1} = R_k F_k, \tag{8}$$

$$J\Omega_{k+1} = F_k^T J\Omega_k + \frac{h}{2}F_k^T M_k + \frac{h}{2}M_{k+1}, \tag{9}$$

where $J_d \in \mathbb{R}^{3\times 3}$ is a nonstandard moment of inertia matrix defined by $J_d = \frac{1}{2}\mathrm{tr}[J]\, I_{3\times 3} - J$, and $F_k \in \mathrm{SO}(3)$ is the relative attitude over an integration step. The constant $h \in \mathbb{R}$ is the integration step size, and the subscript k denotes the kth integration step. This integrator yields a map $(R_k, \Omega_k) \mapsto (R_{k+1}, \Omega_{k+1})$ by solving (7) to obtain $F_k \in \mathrm{SO}(3)$ and substituting it into (8) and (9) to obtain R_{k+1} and Ω_{k+1}.

It preserves the orthogonal structure of $\mathrm{SO}(3)$ because the rotation matrix is updated by a product of two rotation matrices in (8). Since this integrator is obtained from a discrete variational principle, it is symplectic, momentum preserving, and has good energy behavior, properties that are characteristic of variational integrators.

C. Uncertainty Ellipsoid

An uncertainty ellipsoid in $\mathbb{R}^n$ is defined as

$$\mathcal{E}_{\mathbb{R}^n}(\hat{x}, P) = \left\{ x \in \mathbb{R}^n \;\middle|\; (x-\hat{x})^T P^{-1} (x-\hat{x}) \le 1 \right\}, \tag{10}$$

where $\hat{x} \in \mathbb{R}^n$, and $P \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. We call $\hat{x}$ the center of the uncertainty ellipsoid, and P is the uncertainty matrix that determines the size and the shape of the uncertainty ellipsoid. The size of an uncertainty ellipsoid is measured by $\mathrm{tr}[P]$ which is the sum of the squares of the semi principal axes of the ellipsoid.

The state evolves in the 6 dimensional tangent bundle, $\mathrm{TSO}(3)$. We identify $\mathrm{TSO}(3)$ with $\mathrm{SO}(3) \times \mathfrak{so}(3)$ by left trivialization, and we identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ by the isomorphism $S(\cdot)$. The uncertainty ellipsoid centered at $(\hat{R}, \hat{\Omega}) \in \mathrm{TSO}(3)$ is induced from an uncertainty ellipsoid in $\mathbb{R}^6$;

$$\mathcal{E}(\hat{R}, \hat{\Omega}, P) = \left\{ R \in \mathrm{SO}(3),\, \Omega \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} \zeta \\ \delta\Omega \end{bmatrix} \in \mathcal{E}_{\mathbb{R}^6}(0_6, P) \right\}, \tag{11}$$

where $S(\zeta) = \mathrm{logm}\left(\hat{R}^T R\right) \in \mathfrak{so}(3)$, $\delta\Omega = \Omega - \hat{\Omega} \in \mathbb{R}^3$, and $P \in \mathbb{R}^{6\times 6}$ is a symmetric positive definite matrix. An element $(R, \Omega) \in \mathcal{E}(\hat{R}, \hat{\Omega}, P)$ can be written as

$$R = \hat{R} e^{S(\zeta)}, \quad \Omega = \hat{\Omega} + \delta\Omega,$$

for some $x = [\zeta;\, \delta\Omega] \in \mathbb{R}^6$ satisfying $x^T P^{-1} x \le 1$.

D. Uncertainty model

We define the measurement error models for the direction vector and for the angular velocity. The measurement error is modeled by rotation of the measured direction;

$$\begin{aligned} b^i &= e^{S(\nu^i)} \tilde{b}^i, \\ &\simeq \tilde{b}^i + S(\nu^i)\tilde{b}^i, \end{aligned} \tag{12}$$

where $\nu^i \in \mathbb{R}^3$ is the sensor error, which represents the Euler axis of rotation vector from $\tilde{b}^i$ to b^i, and $\left\|\nu^i\right\|$ is the corresponding rotation angle in radians. The approximation is obtained by assuming that the measurement error is small.

The angular velocity measurement errors are modeled as

$$\Omega_k = \tilde{\Omega}_k + \upsilon_k, \tag{13}$$

where $\tilde{\Omega}_k \in \mathbb{R}^3$ is the measured angular velocity, and $\upsilon_k \in \mathbb{R}^3$ is an additive error.

We assume that the initial conditions and the sensor noise are bounded by prescribed uncertainty ellipsoids.

$$(R_0, \Omega_0) \in \mathcal{E}(\hat{R}_0, \hat{\Omega}_0, P_0), \tag{14}$$

$$\nu_k^i \in \mathcal{E}_{\mathbb{R}^3}(0, S_k^i), \tag{15}$$

$$\upsilon_k \in \mathcal{E}_{\mathbb{R}^3}(0, T_k), \tag{16}$$

where $P_0 \in \mathbb{R}^{6\times 6}$, $S_k^i, T_k \in \mathbb{R}^{3\times 3}$ are symmetric positive definite matrices that define the shape and the size of the uncertainty ellipsoids.

IV. Attitude Estimation with Attitude and Angular Velocity Measurements

In this section, we develop a deterministic estimator for the attitude and the angular velocity of a rigid body assuming that both attitude and angular velocity measurements are available. The estimator consists of three stages; flow update, measurement update, and filtering. The flow update predicts the uncertainty ellipsoid in the future. The measurement update obtains an uncertainty ellipsoid using new measurements and the sensor error model. Filtering obtains a new uncertainty ellipsoid compatible with the predicted and the measured uncertainty ellipsoids.

The subscript k denotes the kth discrete index, and the superscript i denotes ith directional sensor. The superscripts f and m denote the variables related to the flow update and the measurement update, respectively. $\tilde{\cdot}$ denotes a measured variable, and $\hat{\cdot}$ denotes an estimated variable.

A. Flow update

Suppose that the attitude and the angular momentum at the kth step lie in a given uncertainty ellipsoid

$$(R_k, \Omega_k) \in \mathcal{E}(\hat{R}_k, \hat{\Omega}_k, P_k),$$

and that new measurements are taken at $(k+l)$th time step. Flow update predicts the center and the uncertainty matrix that define the uncertainty ellipsoid at the $(k+l)$th step using the given uncertainty ellipsoid at the kth step. Since the attitude dynamics is nonlinear, the admissible boundary of the state at the $(k+l)$th step is not an ellipsoid in general.

We assume that the uncertainty ellipsoid at the kth step is sufficiently small that states in the uncertainty ellipsoid can be approximated using the linearized equations of motion.

Center: For the given center $(\hat{R}_k, \hat{\Omega}_k)$, the center of the uncertainty ellipsoid at step $(k+l)$ is $(\hat{R}^f_{k+l}, \hat{\Omega}^f_{k+l})$ obtained using the discrete equations of motion, (7), (8), and (9):

$$hS(J\hat{\Omega}_k + \frac{h}{2}\hat{M}_k) = \hat{F}_k J_d - J_d \hat{F}_k^T, \tag{17}$$

$$\hat{R}^f_{k+1} = \hat{R}_k \hat{F}_k, \tag{18}$$

$$J\hat{\Omega}^f_{k+1} = \hat{F}_k{}^T \hat{\Omega}_k + \frac{h}{2}\hat{F}_k{}^T \hat{M}_k + \frac{h}{2}\hat{M}_{k+1}. \tag{19}$$

This integrator yields a map $(\hat{R}_k, \hat{\Omega}_k) \mapsto (\hat{R}^f_{k+1}, \hat{\Omega}^f_{k+1})$, and this process is repeatedly applied to find the center at the $(k+l)$th step, $(\hat{R}^f_{k+l}, \hat{\Omega}^f_{k+l})$.

Uncertainty matrix: At the $(k+1)$th step, the state is represented by perturbations from the center $(\hat{R}^f_{k+1}, \hat{\Omega}^f_{k+1})$:

$$R_{k+1} = \hat{R}^f_{k+1} e^{S(\zeta^f_{k+1})},$$
$$\Omega_{k+1} = \hat{\Omega}^f_{k+1} + \delta\Omega^f_{k+1},$$

for some $\zeta^f_{k+1}, \delta\Omega^f_{k+1} \in \mathbb{R}^3$. The uncertainty matrix at the $(k+1)$th step is obtained by finding a bound on $\zeta^f_{k+1}, \delta\Omega^f_{k+1} \in \mathbb{R}^3$. Assume that the uncertainty ellipsoid at the kth step is sufficiently small. Then, $\zeta^f_{k+1}, \delta\Omega^f_{k+1}$ are represented by the following linear equations in [13]

$$x^f_{k+1} = A^f_k x_k,$$

where $x_k = [\zeta_k; \delta\Omega_k] \in \mathbb{R}^6$, and $A^f_k \in \mathbb{R}^{6\times 6}$ can be suitably defined. Since $(R_k, \Omega_k) \in \mathcal{E}(\hat{R}_k, \hat{\Omega}_k, P_k)$, $x_k \in \mathcal{E}_{\mathbb{R}^6}(0, P_k)$ by the definition of the uncertainty ellipsoid given in (11). Then we can show that $A^f_k x_k$ lies in

$$A^f_k x_k \in \mathcal{E}_{\mathbb{R}^6}\left(0, A^f_k P_k \left(A^f_k\right)^T\right).$$

Thus, the uncertainty matrix at the $(k+1)$th step is given by

$$P^f_{k+1} = A^f_k P_k \left(A^f_k\right)^T. \tag{20}$$

In summary, the uncertainty ellipsoid at the $(k+l)$th step is computed using (17), (18), (19), and (20) as:

$$(R_{k+l}, \Omega_{k+l}) \in \mathcal{E}(\hat{R}^f_{k+l}, \hat{\Omega}^f_{k+l}, P^f_{k+l}). \tag{21}$$

B. Measurement update

The measurement update finds an uncertainty ellipsoid in the state space using the measurements and sensor error models. The measured attitude and the angular velocity define the center of the measurement uncertainty ellipsoid, and the sensor error models give the uncertainty matrix.

Center: The center of the uncertainty ellipsoid, $(\hat{R}^m_{k+l}, \hat{\Omega}^m_{k+l})$ is obtained from measurements. Let the measured directions to the known points be $\tilde{B}_{k+l} = [\tilde{b}^1, \cdots, \tilde{b}^m] \in \mathbb{R}^{3\times m}$. Then, the attitude $\hat{R}^m_{k+l}$ satisfies the following necessary condition given in (2)

$$\left(\hat{R}^m_{k+l}\right)^T \tilde{L}_{k+l} - \tilde{L}^T_{k+l} \hat{R}^m_{k+l} = 0, \tag{22}$$

where $\tilde{L}_{k+l} = E_{k+l} W_{k+l} \tilde{B}^T_{k+l} \in \mathbb{R}^{3\times 3}$. The attitude matrix is given by a QR factorization of $\tilde{L}_{k+l}$ as in (3)

$$\hat{R}^m_{k+l} = \left(Q_q \sqrt{(Q_r Q_r^T)^{-1}} Q_q^T\right) \tilde{L}_{k+l}, \tag{23}$$

where $Q_q \in \mathrm{SO}(3)$ is an orthogonal matrix and $Q_r \in \mathbb{R}^{3\times 3}$ is a upper triangular matrix satisfying $\tilde{L}_{k+l} = Q_q Q_r$.

The angular velocity is measured directly by

$$\hat{\Omega}^m_{k+l} = \tilde{\Omega}_{k+l}. \tag{24}$$

Uncertainty matrix: We represent the actual state at the $(k+l)$th step as perturbations from the measured center:

$$R_{k+l} = \hat{R}^m_{k+l} e^{S(\zeta^m_{k+l})}, \tag{25}$$

$$\Omega_{k+l} = \hat{\Omega}^m_{k+l} + \delta\Omega^m_{k+l}, \tag{26}$$

for $\zeta^m_{k+l}, \delta\Omega^m_{k+l} \in \mathbb{R}^3$. The uncertainty matrix is obtained by finding a bound on $\zeta^m_{k+l}, \delta\Omega^m_{k+l}$.

We transform the uncertainties in measuring the body directions to known fixed points into uncertainties in the rotation matrix by (22). Using the error model in (12), the actual directions corresponding to B_{k+l} are given by

$$B_{k+l} = \tilde{B}_{k+l} + \delta\tilde{B}_{k+l}, \tag{27}$$

where $\delta B_{k+l} = \left[S(\nu^1)\tilde{b}^1, \cdots, S(\nu^m)\tilde{b}^m\right] \in \mathbb{R}^{3\times m}$.

The actual directions B_{k+l} and the actual attitude R_{k+l} at the $(k+l)$th step also satisfy (23);

$$R^T_{k+l} L_{k+l} - L^T_{k+l} R_{k+l} = 0, \tag{28}$$

where $L_{k+l} = E_{k+l} W_{k+l} B^T_{k+l} \in \mathbb{R}^{3\times 3}$. Substitute (25) and (27) into (28), and use $S(x)A + A^T S(x) = S(\{\mathrm{tr}[A]\, I_{3\times 3} - A\}\, x)$ for $A \in \mathbb{R}^{3\times 3}, x \in \mathbb{R}^3$, to get:

$$\left\{\mathrm{tr}\left[\left(\hat{R}^m_{k+l}\right)^T \tilde{L}_{k+l}\right] - \left(\hat{R}^m_{k+l}\right)^T \tilde{L}_{k+l}\right\} \zeta^m_{k+l}$$
$$= -\sum_{i=1}^m w_i \left\{\mathrm{tr}\left[\tilde{b}^i (e^i)^T \hat{R}^m_{k+l}\right] I_{3\times 3} - \tilde{b}^i (e^i)^T \hat{R}^m_{k+l}\right\} \nu^i.$$

We can rewrite the above equation as

$$\zeta^m_{k+l} = \sum_{i=1}^m \mathcal{A}^{m,i}_{k+l} \nu^i, \tag{29}$$

where $\mathcal{A}^{m,i}_{k+l} \in \mathbb{R}^{3\times 3}$ is defined appropriately.

The perturbation of the angular velocity $\delta\Omega^m_{k+l}$ is equal to the angular velocity measurement error υ_{k+l},

$$\delta\Omega^m_{k+l} = \upsilon_{k+l}. \tag{30}$$

Define $x^m_{k+l} = \left[\zeta^m_{k+l};\, \delta\Omega^m_{k+l}\right] \in \mathbb{R}^6$. Using (29) and (30),

$$x^m_{k+l} = H_1 \sum_{i=1}^m \mathcal{A}^{m,i}_{k+l} \nu^i_{k+l} + H_2 \upsilon_{k+l},$$

where $H_1 = [I_{3\times 3},\, 0_{3\times 3}]^T, H_2 = [0_{3\times 3},\, I_{3\times 3}]^T \in \mathbb{R}^{6\times 3}$. This expresses x^m_{k+l} as a linear combination of the sensor errors ν^i and υ. Using the measurement uncertainties (15)

and (16), we can show that the terms in the right hand side of the above equation are in the following uncertainty ellipsoids:

$$H_1\mathcal{A}^{m,i}_{k+l}\nu^i_{k+l} \in \mathcal{E}_{\mathbb{R}^6}\left(0, H_1\mathcal{A}^{m,i}_{k+l}S^i_{k+l}\left(\mathcal{A}^{m,i}_{k+l}\right)^T H_1^T\right),$$
$$H_2\upsilon_{k+l} \in \mathcal{E}_{\mathbb{R}^6}\left(0, H_2T_{k+l}H_2^T\right).$$

Thus, the uncertainty ellipsoid for x^m_{k+l} is obtained as the vector sum of the above uncertainty ellipsoids. The measurement update obtains a minimal ellipsoid that contains the vector sum of these uncertainty ellipsoids. Using expressions for such a minimal ellipsoid given in [8] and [9], we get:

$$P^m_{k+l} = \left\{\sum_{i=1}^m \sqrt{\mathrm{tr}\left[P^{m,i}_{k+l,R}\right]} + \sqrt{\mathrm{tr}\left[P^m_{k+l,\Omega}\right]}\right\} \times \left\{\sum_{i=1}^m \frac{P^{m,i}_{k+l,R}}{\sqrt{\mathrm{tr}\left[P^{m,i}_{k+l,R}\right]}} + \frac{P^m_{k+l,\Omega}}{\sqrt{\mathrm{tr}\left[P^m_{k+l,\Omega}\right]}}\right\}, \quad (31)$$

where

$$P^{m,i}_{k+l,R} = H_1\mathcal{A}^{m,i}_{k+l}S^i_{k+l}\left(\mathcal{A}^{m,i}_{k+l}\right)^T H_1^T,$$
$$P^m_{k+l,\Omega} = H_2T_{k+l}H_2^T.$$

In summary, the measured uncertainty ellipsoid at the $(k+l)$th step is defined by (23), (24), and (31);

$$(R_{k+l}, \Omega_{k+l}) \in \mathcal{E}(\hat{R}^m_{k+l}, \hat{\Omega}^m_{k+l}, P^m_{k+l}). \quad (32)$$

C. Filtering procedure

The filtering procedure obtains a new uncertainty ellipsoid compatible with both the predicted and the measured uncertainty ellipsoids. From (21) and (32), we know that:

$$(R_{k+l}, \Omega_{k+l}) \in \mathcal{E}(\hat{R}^f_{k+l}, \hat{\Omega}^f_{k+l}, P^f_{k+l}) \bigcap \mathcal{E}(\hat{R}^m_{k+l}, \hat{\Omega}^m_{k+l}, P^m_{k+l}).$$

The intersection of two ellipsoids is not generally an ellipsoid, and it is inefficient to describe an irregular subset in the multidimensional space numerically. We find a minimal uncertainty ellipsoid containing this intersection. We omit the subscript $(k+l)$ here for convenience.

The measurement uncertainty ellipsoid, $\mathcal{E}(\hat{R}^m, \hat{\Omega}^m, P^m)$, is identified by its center $(\hat{R}^m, \hat{\Omega}^m)$, and the uncertainty ellipsoid in $\mathbb{R}^6$:

$$(\zeta^m, \delta\Omega^m) \in \mathcal{E}_{\mathbb{R}^6}(0_{6\times1}, P^m), \quad (33)$$

where $S(\zeta^m) = \mathrm{logm}\left((\hat{R}^m)^T R\right) \in \mathfrak{so}(3)$, $\delta\Omega^m = \Omega - \hat{\Omega}^m \in \mathbb{R}^3$. Similarly, the flow uncertainty ellipsoid, $\mathcal{E}(\hat{R}^f, \hat{\Omega}^f, P^f)$, is identified by its center $(\hat{R}^f, \hat{\Omega}^f)$, and the uncertainty ellipsoid in $\mathbb{R}^6$:

$$(\zeta^f, \delta\Omega^f) \in \mathcal{E}_{\mathbb{R}^6}(0_{6\times1}, P^f), \quad (34)$$

where $S(\zeta^f) = \mathrm{logm}\left((\hat{R}^f)^T R\right) \in \mathfrak{so}(3)$, $\delta\Omega^f = \Omega - \hat{\Omega}^f \in \mathbb{R}^3$. An element $(R^f, \Omega^f) \in \mathcal{E}(\hat{R}^f, \hat{\Omega}^f, P^f)$ is given by

$$R^f = \hat{R}^f e^{S(\zeta^f)}, \quad (35)$$
$$\Omega^f = \hat{\Omega}^f + \delta\Omega^f. \quad (36)$$

Define $\hat{\zeta}^{mf}, \delta\hat{\Omega}^{mf} \in \mathbb{R}^3$ such that

$$\hat{R}^f = \hat{R}^m e^{S(\hat{\zeta}^{mf})}, \quad (37)$$
$$\hat{\Omega}^f = \hat{\Omega}^m + \delta\hat{\Omega}^{mf}. \quad (38)$$

Thus, $\hat{\zeta}^{mf}, \delta\hat{\Omega}^{mf}$ represent the difference between the centers of the two ellipsoids.

Substituting (37), (38) into (35), (36), we obtain

$$R^f = \hat{R}^m e^{S(\hat{\zeta}^{mf})}e^{S(\zeta^f)}, \simeq \hat{R}^m e^{S(\hat{\zeta}^{mf}+\zeta^f)}, \quad (39)$$
$$\Omega^f = \hat{\Omega}^m + \left(\delta\hat{\Omega}^{mf} + \delta\Omega^f\right), \quad (40)$$

where we assumed that $\hat{\zeta}^{mf}, \zeta^f$ are sufficiently small. Thus, the uncertainty ellipsoid obtained by the flow update, $\mathcal{E}(\hat{R}^f, \hat{\Omega}^f, P^f)$ is identified by the measured $(\hat{R}^m, \hat{\Omega}^m)$ and the following uncertainty ellipsoid in $\mathbb{R}^6$:

$$\mathcal{E}_{\mathbb{R}^6}(\hat{x}^{mf}, P^f), \quad (41)$$

where $\hat{x}^{mf} = \left[\hat{\zeta}^{mf}; \delta\hat{\Omega}^{mf}\right] \in \mathbb{R}^6$.

We seek a minimal ellipsoid that contains the intersection:

$$\mathcal{E}_{\mathbb{R}^6}(0_{6\times1}, P^m) \bigcap \mathcal{E}_{\mathbb{R}^6}(\hat{x}^{mf}, P^f) \subset \mathcal{E}_{\mathbb{R}^6}(\hat{x}, P), \quad (42)$$

where $\hat{x} = [\hat{\zeta}; \delta\hat{\Omega}] \in \mathbb{R}^6$. Using the expression for a minimal ellipsoid containing the intersection of two ellipsoids presented in [8], $\hat{x}$ and P are given by

$$\hat{x} = L\hat{x}^{mf},$$
$$P = \beta(q)(I - L)P^m,$$

where

$$\beta(q) = 1 + q - (\hat{x}^{mf})^T(P^m)^{-1}L\hat{x}^{mf},$$
$$L = P^m(P^m + q^{-1}P^f)^{-1}.$$

The constant q is chosen to minimize $\mathrm{tr}[P]$. We convert $\hat{x}$ to points in $\mathrm{TSO}(3)$ using the common center $(\hat{R}^m, \hat{\Omega}^m)$.

In summary, the uncertainty ellipsoid at $(k+l)$th step is

$$(R_{k+l}, \Omega_{k+l}) \in \mathcal{E}(\hat{R}_{k+l}, \hat{\Omega}_{k+l}, P_{k+l}), \quad (43)$$

where

$$\hat{R}_{k+l} = \hat{R}^m_{k+l}e^{S(\hat{\zeta})}, \quad \hat{\Omega}_{k+l} = \hat{\Omega}^m_{k+l} + \delta\hat{\Omega}, \quad P_{k+l} = P. \quad (44)$$

D. Properties of the estimator

The steps outlined above are repeated to get a dynamic filter. This attitude estimator has no singularities since the attitude is represented by a rotation matrix. Orthogonality of the rotation matrix is preserved as it is updated by the structure-preserving Lie group variational integrator. This estimator can be used for highly nonlinear large angle maneuvers of a rigid body. It is also robust to the distribution of the sensor noise since we only use ellipsoidal bounds on the noise. The measurements need not be periodic, the estimation is repeated whenever new measurements become available. We can also extend this attitude estimator to the case when angular velocity measurements are not available. The filtering step is modified to find an intersection of the non-degenerate predicted uncertainty ellipsoid and the degenerate measurement uncertainty ellipsoid.

V. Numerical Simulation

Numerical simulation results are presented for estimation of the attitude dynamics of an uncontrolled rigid spacecraft in a circular orbit about a large central body, including gravity gradient effects. The on orbit spacecraft model is given in [13].

The mass, length and time dimensions are normalized by the spacecraft mass, the maximum length of the spacecraft, and the orbital angular velocity, respectively. The inertia of the spacecraft is chosen as $J = \mathrm{diag}\,[1,\, 2.8,\, 2]$. The maneuver is an arbitrary large attitude change completed in a quarter of the orbit. The initial conditions are chosen as

$$R_0 = \mathrm{diag}[-1, -1, 1], \quad \Omega_0 = [2.316,\, 0.446,\, -0.591]\,\mathrm{rad/s},$$
$$\hat{R}_0 = I_{3\times 3}, \qquad\qquad \hat{\Omega}_0 = [2.116,\, 0.546,\, -0.891]\,\mathrm{rad/s}.$$

The corresponding initial estimation errors are $\|\zeta_0\| = 180\,\mathrm{deg}$, $\|\delta\Omega_0\| = 21.43\frac{\pi}{180}\,\mathrm{rad/s}$. Note that the actual initial attitude is opposite to the estimated initial attitude. The initial uncertainty matrix is given by

$$P_0 = 2\,\mathrm{diag}\left[\pi^2[1,\,1,\,1],\, \left(\frac{\pi}{6}\right)^2[1,\,1,\,1]\right],$$

so that $x_0^T P_0^{-1} x_0 = 0.7553 \leq 1$.

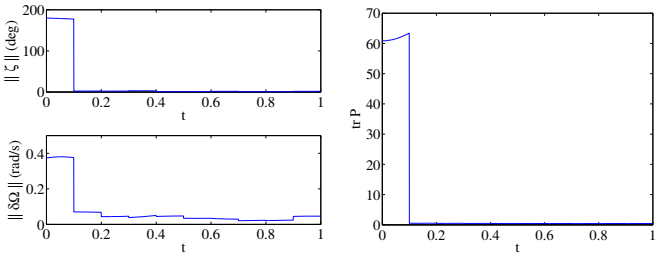
We assume that the measurements are available ten times per quarter orbit. The measurement noise is assumed to be normally distributed with uncertainty matrices given by

$$S_k^i = \left(7\frac{\pi}{180}\right)^2 I_{3\times 3}\,\mathrm{rad}^2, \quad T_k = \left(7\frac{\pi}{180}\right)^2 I_{3\times 3}\,\mathrm{rad}^2/\mathrm{s}^2.$$

We consider two cases. Fig. 2 shows simulation results when both the attitude and the angular velocity are measured. Fig. 3 shows simulation results when angular velocity measurements are not available. In each figure, the left plot shows the attitude and angular velocity estimation errors, and the right plot shows the size of the uncertainty ellipsoid. The estimation errors and the size of uncertainty decrease rapidly after the first measurement. When the angular velocity measurements are not available, the estimation error for the angular velocity converges relatively slowly as seen in Fig. 3.(a). For both cases, the terminal attitude error, and the terminal angular velocity error are less than $0.88\,\mathrm{deg}$, and $0.04\,\mathrm{rad/s}$, respectively.

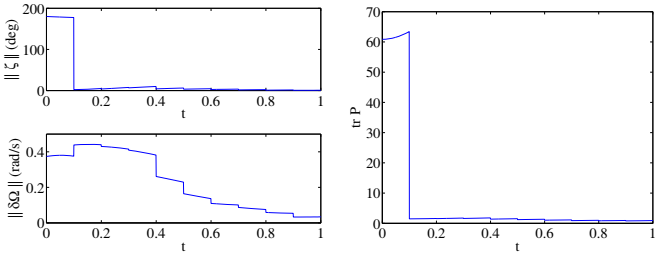
VI. Conclusion

A deterministic estimator for the attitude dynamics of a rigid body in a potential field with bounded measurement errors is presented. An uncertainty ellipsoid is obtained at each estimation step, and the dynamics is propagated using Lie group variational integrators. The center of the uncertainty ellipsoid is the point estimate, and its size determines the accuracy of the estimate. The estimation scheme is optimal in the sense that the size of the uncertainty is minimized at each estimation step. It is also global and robust to the distribution of measurement noise. This estimator can be extended to include the effects of process noise and to the case when only attitude measurements are available. These extensions are not described in this paper.



(a) Estimation error $\|\zeta_k\|$, $\|\delta\Omega_k\|$ (b) Size of uncertainty $\mathrm{tr}[P_k]$

Fig. 2. Estimation with attitude and angular velocity measurement



(a) Estimation error $\|\zeta_k\|$, $\|\delta\Omega_k\|$ (b) Size of uncertainty $\mathrm{tr}[P_k]$

Fig. 3. Estimation with attitude measurement only

References

[1] G. Wahba, "A least squares estimate of satellite attitude, Problem 65-1," *SIAM Review*, vol. 7, no. 5, p. 409, 1965.

[2] I. Y. Bar-Itzhack and Y. Oshman, "Attitude determination from vector observations; quaternion estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 21, no. 1, pp. 128–136, 1985.

[3] M. D. Shuster, "Kalman filtering of spacecraft attitude and the QUEST model," *Journal of the Astronautical Sciences*, vol. 38, no. 3, pp. 377–393, 1990.

[4] H. Rehbinder and X. Hu, "Drift-free attitude estimation for accelerated rigid bodies," *Automatica*, vol. 40, no. 4, pp. 653–659, 2004.

[5] M. L. Psiaki, "Attitude determination filtering via extended quaternion estimation," *AIAA Journal of Guidance, Control and Dynamics*, vol. 23, no. 2, pp. 206–214, 2000.

[6] Y. Theodor, U. Shaked, and C. E. de Souza, "A game theory approach to robust discrete-time H_∞-estimation," *IEEE Transactions on Signal Processing*, vol. 42, no. 6, pp. 1486–1495, 1994.

[7] F. C. Schweppe, "Recursive state estimation: Unknown but bounded errors and system inputs," *IEEE Transactions on Automatic Control*, vol. 13, no. 1, pp. 22–28, 1968.

[8] D. G. Maksarov and J. P. Norton, "State bounding with ellipsoidal set description of the uncertainty," *International Journal of Control*, vol. 65, no. 5, pp. 847–866, 1996.

[9] C. Durieu, E. Walter, and B. Polyak, "Multi-input multi-output ellipsoidal state bounding," *Journal of Optimization Theory and Applications*, vol. 111, no. 2, pp. 273–303, 2001.

[10] J. L. Farrell, J. C. Stuelpnagel, R. H. Wessner, J. R. Velman, and J. E. Brock, "A least squares estimate of satellite attitude, Solution 65-1," *SIAM Review*, vol. 8, no. 3, pp. 384–386, 1966.

[11] M. D. Shuster and S. D. Oh, "Three-axis attitude determination from vector observations," *Journal of Guidance Control and Dynamics*, vol. 4, no. 1, pp. 70–77, 1981.

[12] A. K. Sanyal, "Optimal attitude estimation and filtering without using local coordinates, Part I: Uncontrolled and deterministic attitude dynamics," in *Proceedings of the American Control Conference*, 2006, pp. 5734–5739.

[13] T. Lee, M. Leok, and N. H. McClamroch, "Attitude maneuvers of a rigid spacecraft in a circular orbit," in *Proceedings of the American Control Conference*, 2006, pp. 1742–1747.

[14] ——, "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum," in *Proceedings of the IEEE Conference on Control Applications*, 2005, pp. 962–967.

Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

Controlled Lagrangians and Potential Shaping for Stabilization of Discrete Mechanical Systems

Anthony M. Bloch
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109
abloch@umich.edu

Melvin Leok
Department of Mathematics
Purdue University
West Lafayette, IN 47907
mleok@math.purdue.edu

Jerrold E. Marsden
Control and Dynamical Systems
California Institute of Technology 107-81
Pasadena, CA 91125
marsden@cds.caltech.edu

Dmitry V. Zenkov
Department of Mathematics
North Carolina State University
Raleigh, NC 27695
dvzenkov@unity.ncsu.edu

Abstract— **The method of controlled Lagrangians for discrete mechanical systems is extended to include potential shaping in order to achieve complete state-space asymptotic stabilization. New terms in the controlled shape equation that are necessary for matching in the discrete context are introduced. The theory is illustrated with the problem of stabilization of the cart-pendulum system on an incline. We also discuss digital and model predictive control.**

I. Introduction

The method of controlled Lagrangians for stabilization of relative equilibria (steady state motions) originated in Bloch, Leonard, and Marsden [4] and was then developed in Auckly [1], Bloch, Leonard, and Marsden [5], [6], [7], Bloch, Chang, Leonard, and Marsden [8], and Hamberg [11], [12]. A similar approach for Hamiltonian controlled systems was introduced and further studied in the work of Blankenstein, Ortega, van der Schaft, Maschke, Spong, and their collaborators (see, e.g., [18] and related references). The two methods were shown to be equivalent in [9] and a nonholonomic version was developed in [20], [21], and [2].

According to the method of controlled Lagrangians, the original controlled system is represented as a new, uncontrolled Lagrangian system for a *controlled Lagrangian*, a modification of the original Lagrangian. The controlled Lagrangian is designed so that its associated energy has a maximum or minimum at the (relative) equilibrium to be stabilized. The time-invariant feedback control law is obtained by requiring that the new and old systems of equations of motion are equivalent. To obtain asymptotic stabilization, dissipation-emulating terms are added to the control input.

The method of controlled Lagrangians for discrete mechanical systems was introduced in Bloch, Leok, Marsden, and Zenkov [3]. In the present paper this formalism is further developed to include *potential shaping* which is used for complete state-space stabilization of equilibria. This study is motivated by the recent development of structure-preserving algorithms for numerical simulation of controlled systems. In particular, as the closed loop dynamics of a controlled Lagrangian system is itself Lagrangian, it is natural to adopt a variational discretization that exhibits good long-time numerical stability.

We carry out the matching procedure explicitly for discrete systems with two degrees of freedom and prove that we can asymptotically stabilize the equilibria of interest. The theoretical analysis is validated by simulating the discrete cart-pendulum system on an incline. When dissipation is added, the inverted pendulum configuration is asymptotically stabilized, as predicted. We then use the discrete controlled dynamics to construct a real-time model predictive controller with piecewise constant control inputs. This serves to illustrate how discrete mechanics can be naturally applied to yield digital controllers for mechanical systems.

The paper is organized as follows: In Sections II and III we review discrete mechanics and the method of controlled Lagrangians for stabilization of equilibria of mechanical systems. The discrete version of the potential shaping procedure and related stability analysis are discussed in Sections IV and V. The theory is illustrated with the discrete cart-pendulum system. Simulations and the construction of the digital controller are presented in Sections VI and VII.

In a future publication we intend to treat discrete systems with nonabelian symmetries as well as systems with nonholonomic constraints.

II. An Overview of Discrete Mechanics

A discrete analogue of Lagrangian mechanics can be obtained by considering a discretization of the Hamilton principle; this approach underlies the construction of variational integrators. See Marsden and West [17], and references therein, for a more detailed discussion of discrete mechanics.

A key notion is that of the *discrete Lagrangian*, which is a map $L^d : Q \times Q \to \mathbb{R}$ that approximates the action integral along an exact solution of the Euler–Lagrange equations

1-4244-0171-2/06/$20.00 ©2006 IEEE.

joining the configurations $q_k, q_{k+1} \in Q$,

$$L^d(q_k, q_{k+1}) \approx \operatorname*{ext}_{q \in \mathcal{C}([0,h],Q)} \int_0^h L(q, \dot{q})\, dt, \qquad (1)$$

where $\mathcal{C}([0,h],Q)$ is the space of curves $q : [0,h] \to Q$ with $q(0) = q_k$, $q(h) = q_{k+1}$, and ext denotes extremum.

In the discrete setting, the action integral of Lagrangian mechanics is replaced by an action sum

$$S^d(q_0, q_1, \ldots, q_N) = \sum_{k=0}^{N-1} L^d(q_k, q_{k+1}),$$

where $q_k \in Q$, $k = 0, 1, \ldots, N$, is a finite sequence in the configuration space. The equations are obtained by the discrete Hamilton principle, which extremizes the discrete action given fixed endpoints q_0 and q_N. Taking the extremum over $q_1, \ldots, q_{N-1}$ gives the *discrete Euler–Lagrange equations*

$$D_1 L^d(q_k, q_{k+1}) + D_2 L^d(q_{k-1}, q_k) = 0,$$

for $k = 1, \ldots, N-1$. This implicitly defines the update map $\Phi : Q \times Q \to Q \times Q$, where $\Phi(q_{k-1}, q_k) = (q_k, q_{k+1})$ and $Q \times Q$ replaces the phase space TQ of Lagrangian mechanics.

Since we are concerned with control, we need to consider the effect of external forces on Lagrangian systems. In the context of discrete mechanics, this is addressed by introducing the *discrete Lagrange–d'Alembert principle* (see Kane, Marsden, Ortiz, and West [14]), which states that

$$\delta \sum_{k=0}^{n-1} L^d\left(q_k, q_{k+1}\right) + \sum_{k=0}^{n-1} F^d\left(q_k, q_{k+1}\right) \cdot \left(\delta q_k, \delta q_{k+1}\right) = 0$$

for all variations $\boldsymbol{\delta q}$ of $\boldsymbol{q}$ that vanish at the endpoints. Here, $\boldsymbol{q}$ denotes the vector of positions $(q_0, q_1, \ldots, q_N)$, and $\boldsymbol{\delta q} = (\delta q_0, \delta q_1, \ldots, \delta q_N)$, where $\delta q_k \in T_{q_k} Q$. The discrete one-form F^d on $Q \times Q$ approximates the impulse integral between the points q_k and q_{k+1}, just as the discrete Lagrangian L^d approximates the action integral. We define the maps $F_1^d, F_2^d : Q \times Q \to T^*Q$ by the relations

$$F_2^d\left(q_0, q_1\right) \delta q_1 := F^d\left(q_0, q_1\right) \cdot \left(0, \delta q_1\right),$$
$$F_1^d\left(q_0, q_1\right) \delta q_0 := F^d\left(q_0, q_1\right) \cdot \left(\delta q_0, 0\right).$$

The discrete Lagrange–d'Alembert principle may then be rewritten as

$$\delta \sum_{k=0}^{n-1} L^d\left(q_k, q_{k+1}\right) + \sum_{k=0}^{n-1} \left[F_1^d\left(q_k, q_{k+1}\right) \delta q_k + F_2^d\left(q_k, q_{k+1}\right) \delta q_{k+1}\right] = 0$$

for all variations $\boldsymbol{\delta q}$ of $\boldsymbol{q}$ that vanish at the endpoints. This is equivalent to the *forced discrete Euler–Lagrange equations*

$$D_1 L^d\left(q_k, q_{k+1}\right) + D_2 L^d\left(q_{k-1}, q_k\right) + F_1^d\left(q_k, q_{k+1}\right) + F_2^d\left(q_{k-1}, q_k\right) = 0.$$

III. Matching and Controlled Lagrangians

In the controlled Lagrangian approach, one considers a mechanical system with an uncontrolled (free) Lagrangian equal to kinetic energy minus potential energy. In the simplest setting, we modify the kinetic energy to produce a new controlled Lagrangian which describes the dynamics of the controlled closed-loop system. The method is extended by the incorporation of potential shaping in [8].

Suppose our system has configuration space Q and a Lie group G acts freely and properly on Q. It is useful to keep in mind the case in which $Q = S \times G$ with G acting only on the second factor by the left group multiplication. For example, for the inverted planar pendulum on a cart, $Q = S^1 \times \mathbb{R}$ with $G = \mathbb{R}$, the group of reals under addition (corresponding to translations of the cart).

Our goal is to control the variables lying in the *shape space* Q/G using controls that act directly on the variables lying in G.[1] For kinetic shaping, the controlled Lagrangian is constructed to be G-invariant, thus providing modified or *controlled* conservation laws. In this paper, we assume that G is an abelian group.

The key modification of the Lagrangian involves changing the kinetic energy metric $g(\cdot, \cdot)$. The tangent space to Q can be split into a sum of horizontal and vertical parts defined as follows: For each tangent vector v_q to Q at a point $q \in Q$, we can write a unique decomposition $v_q = \operatorname{Hor} v_q + \operatorname{Ver} v_q$, such that the vertical part is tangent to the orbits of the G-action and the horizontal part is metric-orthogonal to the vertical space, *i.e.*, it is uniquely defined by the identity

$$g(v_q, w_q) = g(\operatorname{Hor} v_q, \operatorname{Hor} w_q) + g(\operatorname{Ver} v_q, \operatorname{Ver} w_q) \qquad (2)$$

with v_q and w_q arbitrary tangent vectors to Q at the point $q \in Q$. This choice of horizontal space coincides with that given by the *mechanical connection*; see, for example, Marsden [15].

For the kinetic energy of our controlled Lagrangian, we use a modified version of the right-hand side of equation (2). The potential energy remains unchanged. The modification consists of three ingredients:

1) a new choice of horizontal space, denoted Hor_τ,
2) a change $g \to g_\sigma$ of the metric on horizontal vectors,
3) a change $g \to g_\rho$ of the metric on vertical vectors.

Let ξ_Q denote the infinitesimal generator corresponding to $\xi \in \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra of G (see Marsden [15] or Marsden and Ratiu [16]). Thus, for each $\xi \in \mathfrak{g}$, ξ_Q is a vector field on the configuration manifold Q and its value at a point $q \in Q$ is denoted $\xi_Q(q)$.

Definition 1: Let τ be a Lie-algebra-valued horizontal one-form on Q; that is, a one-form that annihilates vertical vectors. The ***τ-horizontal space*** *at $q \in Q$ consists of tangent vectors to Q at q of the form $\operatorname{Hor}_\tau v_q := \operatorname{Hor} v_q - [\tau(v)]_Q(q)$, which also defines $v_q \mapsto \operatorname{Hor}_\tau v_q$, the* ***$\tau$-horizontal projection****. The* ***τ-vertical projection operator*** *is defined by $\operatorname{Ver}_\tau v_q := \operatorname{Ver} v_q + [\tau(v)]_Q(q)$.*

[1] The shape space is S in the case $Q = S \times G$.

*Definition 2: Given g_σ, g_ρ, and τ, the associated **controlled Lagrangian** $L_{\tau,\sigma,\rho}$ is given by a modified kinetic minus the given potential energy, namely*

$$L_{\tau,\sigma,\rho}(v_q) = \tfrac{1}{2}[g_\sigma(\mathrm{Hor}_\tau v_q, \mathrm{Hor}_\tau v_q) + g_\rho(\mathrm{Ver}_\tau v_q, \mathrm{Ver}_\tau v_q)] - V(q).$$

The equations corresponding to this Lagrangian will be our closed-loop equations. The new terms appearing in those equations corresponding to the directly controlled variables are interpreted as control inputs. The modifications to the Lagrangian are chosen so that no new terms appear in the equations corresponding to the variables that are not directly controlled. We refer to this process as *matching*.

Once the form of the control law is derived using the controlled Lagrangian, the closed-loop stability of an equilibrium can be determined by energy methods, using any available freedom in the choice of τ, g_σ and g_ρ.

We can extend the method of controlled Lagrangians to the class of Lagrangian mechanical systems with potential energy that may break symmetry, *i.e.*, we still have a symmetry group G for the kinetic energy of the system but we now have a potential energy $V(x^\alpha, \theta^a)$ that need not be G-invariant (see [8]). Further, we consider a modification to the potential energy that also breaks symmetry in the group variables. Let the potential energy for the controlled Lagrangian be defined as

$$V(x^\alpha, \theta^a) + V_\varepsilon(x^\alpha, \theta^a),$$

where V_ε is the modification—to be determined—that depends on a new real parameter ε.

For many systems it is sufficient to use the so-called *simplified matching conditions* [8]. For potential shaping in the setting where the simplified matching conditions hold we take $g_\rho = \rho g_{ab}$ where ρ is a scalar constant. The controlled Lagrangian takes the form

$$L_{\tau,\sigma,\rho,\varepsilon}(v) = L_{\tau,\sigma}(v) - V_\varepsilon(x^\alpha, \theta^a) + \tfrac{1}{2}(\rho-1)\times g_{ab}\left(\dot\theta^a + (g^{ac}g_{\alpha c} + \tau^a_\alpha)\dot x^\alpha\right)\left(\dot\theta^b + (g^{bd}g_{\beta d} + \tau^b_\beta)\dot x^\beta\right),$$

where

$$L_{\tau,\sigma}(v) = L(x^\alpha, \dot x^\beta, \theta^a, \dot\theta^a + \tau^a_\alpha \dot x^\alpha) + \tfrac{1}{2}\sigma g_{ab}\tau^a_\alpha\tau^b_\beta \dot x^\alpha \dot x^\beta.$$

This has sufficient generality to handle many examples of interest.

A basic example treated in earlier papers in the smooth setting is the *pendulum on a cart*. Let s denote the position of the cart on the s-axis, ϕ denote the angle of the pendulum with the upright vertical, and ψ denote the elevation angle of the incline, as in Figure 1. The configuration space for this system is $Q = S \times G = S^1 \times \mathbb{R}$, with the first factor being the pendulum angle ϕ and the second factor being the cart position s. The velocity phase space, TQ, has coordinates $(\phi, s, \dot\phi, \dot s)$. The length of the pendulum is l, the mass of the pendulum is m and that of the cart is M.

The symmetry group G of the kinetic energy of the pendulum-cart system is that of translation in the s variable, so $G = \mathbb{R}$.

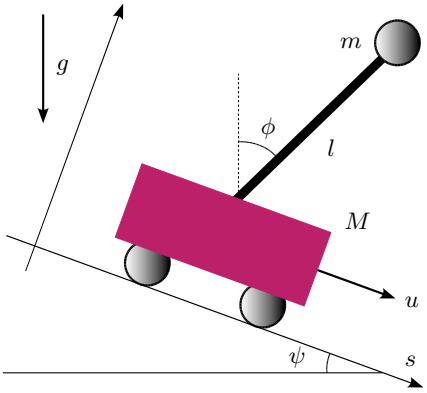


Fig. 1. The pendulum on a cart going down an inclined plane under gravity. The control force is in the direction s, the overall motion of the cart.

IV. Discrete Potential Shaping

For simplicity, we consider systems with one shape and one group degree of freedom. We further assume that the configuration space Q is the direct product of a one-dimensional shape space S and a one-dimensional Lie group G. The continuous-time Lagrangian $L : TQ \to \mathbb{R}$ and the form τ are

$$L(\phi, s, \dot\phi, \dot s) = \tfrac{1}{2}(\alpha\dot\phi^2 + 2\beta(\phi)\dot\phi\dot s + \gamma\dot s^2) - V_1(\phi) - V_2(s) \quad (3)$$

and

$$\tau = k(\phi)\dot\phi \quad \text{with} \quad k(\phi) = -\frac{\beta(\phi)}{\sigma\gamma}.$$

This Lagrangian (3) satisfies the simplified matching conditions of [8].

The continuous-time controlled Lagrangian $L_{\tau,\sigma,\rho,\varepsilon} : TQ \to \mathbb{R}$ becomes

$$L_{\tau,\sigma,\rho,\varepsilon}(\phi, s, \dot\phi, \dot s) = L(\phi, s, \dot\phi, \dot s + k(\phi)\dot\phi) + \tfrac{1}{2}\sigma\gamma(k(\phi)\dot\phi)^2 + \tfrac{1}{2}(\rho-1)\gamma(\dot s + (\sigma-1)k(\phi)\dot\phi)^2 + V_2(s) - V_\varepsilon(y), \quad (4)$$

where

$$y = s - \int_{\phi_e}^{\phi} \frac{1}{\gamma}\left(\frac{1}{\sigma} - \frac{\rho-1}{\rho}\right)\beta(z)\,dz,$$

the function $V_\varepsilon(y)$ is arbitrary, and (ϕ_e, s_e) is the equilibrium of interest. As in Bloch, Chang, Leonard, and Marsden [8], the kinetic energies in (3) and (4) are G-invariant.

For the cart-pendulum system, α, $\beta(\phi)$, γ, $V_1(\phi)$, and $V_2(s)$ are $\alpha = ml^2$, $\beta(\phi) = ml\cos(\phi - \psi)$, $\gamma = M + m$, $V_1(\phi) = -mgl\cos\phi$, and $V_2(s) = -\gamma g s \sin\psi$. Note that $\alpha\gamma - \beta^2(\phi) > 0$.

In discretizing the method of controlled Lagrangians, we combine formulae (1), (3), and (4). In the rest of this paper, we will adopt the notations

$$q_{k+1/2} = \frac{q_k + q_{k+1}}{2}, \quad \Delta q_k = q_{k+1} - q_k.$$

This allows us to construct a *second-order accurate* discrete Lagrangian and discrete controlled Lagrangian as

$$L^d(q_k, q_{k+1}) = hL(q_{k+1/2}, \Delta q_k/h),$$
$$L^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1}) = hL_{\tau,\sigma,\rho,\varepsilon}(q_{k+1/2}, \Delta q_k/h), \quad (5)$$

with $q_k = (\phi_k, s_k)$.

The discrete dynamics is governed by the equations

$$\frac{\partial L^d(q_k, q_{k+1})}{\partial \phi_k} + \frac{\partial L^d(q_{k-1}, q_k)}{\partial \phi_k} = 0, \tag{6}$$

$$\frac{\partial L^d(q_k, q_{k+1})}{\partial s_k} + \frac{\partial L^d(q_{k-1}, q_k)}{\partial s_k} = u_k, \tag{7}$$

where u_k is the control input.

The dynamics associated with (5) is amended by the term w_k in the discrete shape equation:

$$\frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})}{\partial \phi_k} + \frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1}, q_k)}{\partial \phi_k} = w_k, \tag{8}$$

$$\frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})}{\partial s_k} + \frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1}, q_k)}{\partial s_k} = 0. \tag{9}$$

This term w_k is important for matching systems (6), (7) and (8), (9). *The presence of the terms w_k represents an interesting (but manageable) departure from the continuous theory.* Let

$$J_k = \rho\gamma\big(\Delta s_k/h - (\sigma-1)k(\phi_{k+\frac{1}{2}})\Delta\phi_k/h\big).$$

The following statement is proved by a straightforward calculation:

Theorem 3: The dynamics (6)*,* (7) *is equivalent to the dynamics* (8)*,* (9) *if and only if u_k and w_k are given by*

$$\begin{aligned} u_k = &-\frac{h}{2}\left[V_2'(s_{k+\frac{1}{2}}) + V_2'(s_{k-\frac{1}{2}})\right] \\ &+ \frac{h}{2\rho}\left[V_\varepsilon'(s_{k+\frac{1}{2}}) + V_\varepsilon'(s_{k-\frac{1}{2}})\right] \\ &+ \frac{\gamma\Delta\phi_k k(\phi_{k+1/2}) - \gamma\Delta\phi_{k-1}k(\phi_{k-1/2})}{h}, \end{aligned} \tag{10}$$

and

$$\begin{aligned} w_k = -\Big(1-\sigma+\frac{\sigma}{\rho}\Big)\Big(&k(\phi_{k+\frac{1}{2}})\Big[-\gamma\rho J_k + \frac{h}{2}V_\varepsilon'(y_{k+\frac{1}{2}})\Big] \\ &+ k(\phi_{k-\frac{1}{2}})\Big[\gamma\rho J_{k-1} + \frac{h}{2}V_\varepsilon'(y_{k-\frac{1}{2}})\Big] \\ &- k'(\phi_{k+\frac{1}{2}})J_k\Delta\phi_k - k'(\phi_{k-\frac{1}{2}})J_{k-1}\Delta\phi_{k-1}\Big). \end{aligned}$$

Remark. The terms w_k vanish when $\beta(\phi) = \text{const}$ as they become proportional to the left-hand side of equation (9).

V. Stabilization of the Discrete Controlled System

The stability analysis in this paper is done by means of an analysis of the spectrum of the linearized discrete equations. We assume that the equilibrium to be stabilized is $(\phi_k, s_k) = (0,0)$.

Theorem 4: The equilibrium $(\phi_k, s_k) = (0,0)$ of equations (8) *and* (9) *is spectrally stable if*

$$-\frac{\beta^2(0)}{\alpha\gamma - \beta^2(0)} < \sigma < 0, \quad \rho < 0, \quad \textit{and} \quad V_\varepsilon''(0) < 0. \tag{11}$$

Proof: The linearized discrete equations are

$$\frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})}{\partial \phi_k} + \frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1}, q_k)}{\partial \phi_k} = 0, \tag{12}$$

$$\frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})}{\partial s_k} + \frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1}, q_k)}{\partial s_k} = 0, \tag{13}$$

where $\mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})$ is the quadratic approximation of $L^d_{\tau,\sigma,\rho,\varepsilon}$ at the equilibrium (*i.e.*, $\beta(\phi)$, $V_1(\phi)$, and $V_\varepsilon(y)$ in $L^d_{\tau,\sigma,\rho,\varepsilon}$ are replaced by $\beta(0)$, $\frac{1}{2}V_1''(0)\phi^2$, and $\frac{1}{2}V_\varepsilon''(0)y^2$, respectively). *Note the absence of the term w_k in equation* (12).

The linearized dynamics preserves the quadratic approximation of the discrete energy

$$\begin{aligned} &\frac{\alpha\gamma\sigma^2 - \beta(0)^2(\sigma-1)(\rho(\sigma-1)-\sigma)}{2\gamma\sigma^2 h}\Delta\phi_k^2 \\ &\quad + \frac{\beta(0)\rho(\sigma-1)}{\sigma h}\Delta\phi_k\Delta s_k + \frac{\gamma\rho}{2h}\Delta s_k^2 \\ &\quad + \frac{h}{2}V_1''(0)\phi_{k+\frac{1}{2}}^2 + \frac{h}{2}V_\varepsilon''(0)x_{k+\frac{1}{2}}^2, \end{aligned} \tag{14}$$

where

$$x = s + \left(\frac{\rho-1}{\rho} - \frac{1}{\sigma}\right)\frac{\beta(0)}{\gamma}\,\phi.$$

Since $V_1''(0)$ is negative, the equilibrium $(\phi_k, s_k) = (0,0)$ of equations (12) and (13) is stable if the function (14) is negative-definite. The latter requirement is equivalent to conditions (11). The spectrum of the linearized discrete dynamics in this case belongs to the unit circle. ■

Remarks. Spectral stability in this situation is not sufficient to conclude nonlinear stability. The stability conditions (11) are identical to the stability conditions of the corresponding continuous-time system.

Following [8], we now modify the control input (10) by adding the *discrete dissipation-emulating term*

$$-\frac{D(\Delta y_{k-1} + \Delta y_k)}{h} \tag{15}$$

in order to achieve the asymptotic stabilization of the equilibrium $(\phi_k, s_k) = (0,0)$. In the above, D is a constant. The linearized discrete dynamics becomes

$$\begin{aligned} &\frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})}{\partial \phi_k} + \frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1}, q_k)}{\partial \phi_k} \\ &\qquad = -\left(\frac{\rho-1}{\rho} - \frac{1}{\sigma}\right)\frac{\beta(0)}{\gamma}\,\frac{D(\Delta x_{k-1} + \Delta x_k)}{h}, \end{aligned} \tag{16}$$

$$\begin{aligned} &\frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})}{\partial s_k} + \frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1}, q_k)}{\partial s_k} \\ &\qquad = -\frac{D(\Delta x_{k-1} + \Delta x_k)}{h}. \end{aligned} \tag{17}$$

Theorem 5: The equilibrium $(\phi_k, s_k) = (0,0)$ of equations (16) *and* (17) *is asymptotically stable if conditions* (11) *are satisfied and D is positive.*

Proof: Multiplying equations (16) and (17) by $(\Delta\phi_{k-1} + \Delta\phi_k)/2$ and $(\Delta s_{k-1} + \Delta s_k)/2$, respectively, we obtain

$$E_{k,k+1} = E_{k-1,k} + \frac{Dh}{4}\left(\frac{\Delta x_{k-1}}{h} + \frac{\Delta x_k}{h}\right)^2,$$

where $E_{k,k+1}$ is the quadratic approximation of the discrete energy (14). Recall that $E_{k,k+1}$ is negative-definite. It is possible to show that, in some neighborhood of $(\phi_k, s_k) = (0,0)$, the quantity $\Delta x_{k-1} + \Delta x_k \not\equiv 0$ along a solution of

equations (16) and (17) unless this solution is the equilibrium $(\phi_k, s_k) = (0, 0)$. Therefore, $E_{k,k+1}$ increases along non-equilibrium solutions of (16) and (17). Since equations (16) and (17) are linear, this is only possible if the spectrum of (16) and (17) is inside the open unit disk, which implies asymptotic stability of the equilibrium of both linear system (16) and (17) and nonlinear system (6) and (7) with discrete dissipation-emulating term (15) added to u_k. ■

VI. Simulations

Simulating the discrete behavior of the controlled Lagrangian system involves viewing equations (6) and (9) as an implict update map $\Phi : (q_{k-2}, q_{k-1}) \mapsto (q_{k-1}, q_k)$. This presupposes that the initial conditions are given in the form (q_0, q_1); however it is generally preferable to specify the initial conditions as $(q_0, \dot{q}_0)$. This is achieved by solving the boundary condition

$$\frac{\partial L}{\partial \dot{q}}(q_0, \dot{q}_0) + D_1 L^d(q_0, q_1) + F_1^d(q_0, q_1) = 0$$

for q_1. Once the initial conditions are expressed in the form (q_0, q_1), the discrete evolution can be obtained using the implicit update map Φ.

In Figure 2, we present a MATLAB simulation of discrete controlled dynamics of the cart-pendulum system in the absence of dissipation.

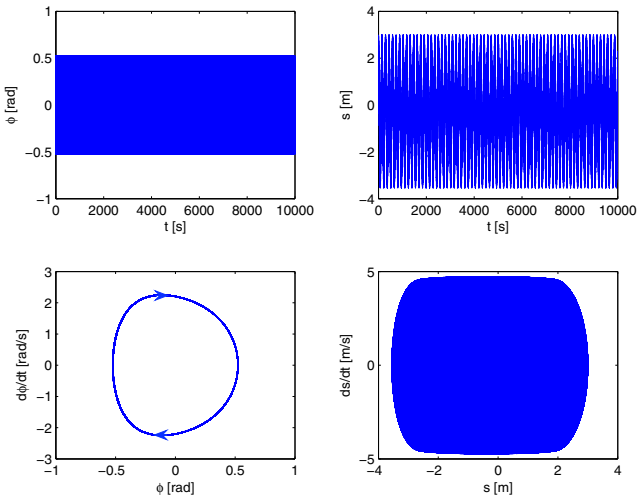


Fig. 2. Discrete controlled dynamics without dissipation. The discrete controlled system stabilizes the motion about the equilibrium; since there is no dissipation, the oscillations are sustained.

Here, $h = 0.05\,\text{sec}$, $m = 0.14\,\text{kg}$, $M = 0.44\,\text{kg}$, $l = 0.215\,\text{m}$, and $\psi = \frac{\pi}{9}$ radians. Our goal is to regulate the cart at $s = 0$ and the pendulum at $\phi = 0$. The control gains are chosen to be $\kappa = 20$, $\rho = -0.02$, and $\varepsilon = 0.00001$. It is worth noting that the discrete dynamics remain bounded near the desired equilibrium, and this behavior persists even for significantly longer simulation runs involving 10^6 time-steps. The exceptional stability of the discrete controlled trajectory can presumably be understood in terms of the bounded energy oscillations characteristic of symplectic and variational integrators.

When dissipation is added, we obtain an asymptotically stabilizing control law, as illustrated in Figure 3. This is consistent with the stability analysis of Section V.

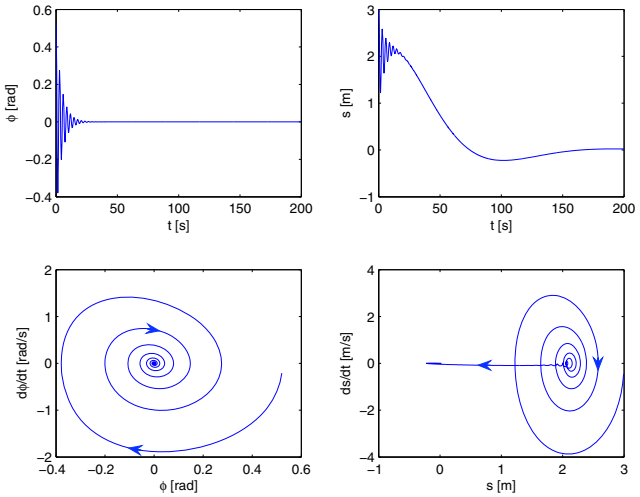


Fig. 3. Discrete controlled dynamics with dissipation. Here the oscillations die out and the cart converges to the desired point $s = 0$.

VII. Model Predictive Controller

We now explore the use of the forced discrete Euler–Lagrange equations as the model in a real-time model predictive controller, with piecewise constant control forces. Algorithm 1 below describes the details of the procedure.

Algorithm 1 Digital Controller ($q(\,\cdot\,), T_f, h$)

$q_0 \leftarrow$ **sense** $q(0)$
$q_1 \leftarrow$ **sense** $q(h)$
$\bar{q}_2 \leftarrow$ **solve** $D_2 L^d(q_0, q_1) + D_1 L^d(q_1, \bar{q}_2) = 0$
$\bar{q}_3 \leftarrow$ **solve** $D_2 L^d(q_1, \bar{q}_2) + D_1 L^d(\bar{q}_2, \bar{q}_3) + F_1^d(\bar{q}_2, \bar{q}_3) = 0$
$u_{2+1/2} \leftarrow u\left(\frac{\bar{q}_2+\bar{q}_3}{2}, \frac{\bar{q}_3-\bar{q}_2}{h}\right)$
actuate $u = u_{2+1/2}$ for $t \in [2h, 3h]$
$q_2 \leftarrow$ **sense** $q(2h)$
$\bar{q}_3 \leftarrow$ **solve** $D_2 L^d(q_1, q_2) + D_1 L^d(q_2, \bar{q}_3) + F_1^d(q_2, \bar{q}_3) = 0$
$\bar{q}_4 \leftarrow$ **solve** $D_2 L^d(q_2, \bar{q}_3) + D_1 L^d(\bar{q}_3, \bar{q}_4)$
$\quad + F_2^d(q_2, \bar{q}_3) + F_1^d(\bar{q}_3, \bar{q}_4) = 0$
$u_{3+1/2} \leftarrow u\left(\frac{\bar{q}_3+\bar{q}_4}{2}, \frac{\bar{q}_4-\bar{q}_3}{h}\right)$
actuate $u = u_{3+1/2}$ for $t \in [3h, 4h]$
for $k = 4$ to $(T_f/h - 1)$ **do**
  $q_{k-1} \leftarrow$ **sense** $q((k-1)h)$
  $\bar{q}_k \leftarrow$ **solve** $D_2 L^d(q_{k-2}, q_{k-1}) + D_1 L^d(q_{k-1}, \bar{q}_k)$
  $\quad + F_2^d(q_{k-2}, q_{k-1}) + F_1^d(q_{k-1}, \bar{q}_k) = 0$
  $\bar{q}_{k+1} \leftarrow$ **solve** $D_2 L^d(q_{k-1}, \bar{q}_k) + D_1 L^d(\bar{q}_k, \bar{q}_{k+1})$
  $\quad + F_2^d(q_{k-1}, \bar{q}_k) + F_1^d(\bar{q}_k, \bar{q}_{k+1}) = 0$
  $u_{k+1/2} \leftarrow u\left(\frac{\bar{q}_k+\bar{q}_{k+1}}{2}, \frac{\bar{q}_{k+1}-\bar{q}_k}{h}\right)$
  actuate $u = u_{k+1/2}$ for $t \in [kh, (k+1)h]$
end for

The digital controller uses the position information it senses for $t = -2h, -h$ to estimate the positions at $t = 0, h$ during the time interval $t = [-h, 0]$. This allows it to compute a symmetric finite difference approximation to the continuous control force $u(\phi, s, \dot{\phi}, \dot{s})$ at $t = h/2$ using the approximation

$$u_{1/2} = u\left(\frac{\bar{\phi}_0 + \bar{\phi}_1}{2}, \frac{\bar{s}_0 + \bar{s}_1}{2}, \frac{\bar{\phi}_1 - \bar{\phi}_0}{h}, \frac{\bar{s}_1 - \bar{s}_0}{h}\right),$$

where the overbar indicates that the position variable is being estimated by the numerical model. This control is then applied as a constant control input for the time interval $[0, h]$. This algorithm can be implemented in real-time if the two forward solves can be computed within the time interval h.

The initialization of the discrete controller is somewhat involved, since the system is unforced during the time interval $[0, 2h]$ while the controller senses the initial states, and computes the appropriate control forces.

The numerical simulation of the digital controller is shown in Figure 4. We see that the system is asymptotically stabilized in both the ϕ and s variables.

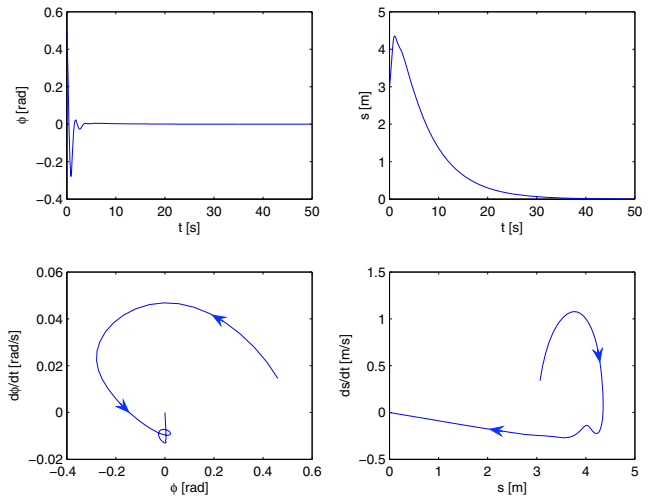


Fig. 4. The discrete real-time piecewise constant model predictive controller stabilizes ϕ and s to zero.

VIII. Conclusions

In this paper we have introduced potential shaping techniques for discrete systems and have shown that these lead to an effective numerical implementation for stabilization in the case of the discrete cart-pendulum model. The method in this paper is related to other discrete methods in control that have a long history; recent papers that use discrete mechanics in the context of optimal control and celestial navigation are [10], [13], and [19]. The full theory of discrete controlled Lagrangians will be developed in a forthcoming paper.

IX. Acknowledgments

The authors would like to thank the reviewers for helpful remarks. The research of AMB was supported by NSF grants DMS-0305837, DMS-0604307, and CMS-0408542. The research of ML was partially supported by NSF grant DMS-0504747 and a University of Michigan Rackham faculty grant. The research of JEM was partially supported by AFOSR Contract FA9550-05-1-0343. The research of DVZ was partially supported by NSF grants DMS-0306017 and DMS-0604108.

References

[1] Auckly, D., L. Kapitanski, & W. White, Control of Nonlinear Underactuated Systems, *Commun. Pure Appl. Math.* **53**, 2000, 354–369.

[2] Bloch, A. M., *Nonholonomic Mechanics and Control*, Interdisciplinary Appl. Math. **24**, Springer-Verlag, 2003.

[3] Bloch, A. M., M. Leok, J. E. Marsden, and D. V. Zenkov [2005], Controlled Lagrangians and Stabilization of the Discrete Cart-Pendulum System, *Proc. CDC* **44**, 6579–6584.

[4] Bloch, A. M., N. Leonard, & J. E. Marsden, Stabilization of Mechanical Systems Using Controlled Lagrangians, *Proc. CDC* **36**, 1997, 2356–2361.

[5] Bloch, A. M., N. Leonard, & J. E. Marsden, Matching and Stabilization by the Method of Controlled Lagrangians, *Proc. CDC* **37**, 1998, 1446–1451.

[6] Bloch, A. M., N. Leonard, & J. E. Marsden, Potential Shaping and the Method of Controlled Lagrangians, *Proc. CDC* **38**, 1999, 1652–1657.

[7] Bloch, A. M., N. E. Leonard, & J. E. Marsden, Controlled Lagrangians and the Stabilization of Mechanical Systems I: The First Matching Theorem, *IEEE Trans. on Systems and Control* **45**, 2000, 2253–2270.

[8] Bloch, A. M., D-E. Chang, N. E. Leonard, & J. E. Marsden, Controlled Lagrangians and the Stabilization of Mechanical Systems II: Potential Shaping, *Trans. IEEE on Autom. Contr.* **46**, 2001, 1556–1571.

[9] Chang, D-E., A. M. Bloch, N. E. Leonard, J. E. Marsden, & C. Woolsey, The Equivalence of Controlled Lagrangian and Controlled Hamiltonian Systems, *Control and the Calculus of Variations (special issue dedicated to J. L. Lions)* **8**, 2002, 393–422.

[10] Guibout, V. & A. M. Bloch, A Discrete Maximum Principle for Solving Optimal Control Problems, *Proc. CDC* **43**, 2004, 1806–1811.

[11] Hamberg, J., General Matching Conditions in the Theory of Controlled Lagrangians, *Proc. CDC* **38**, 1999, 2519–2523.

[12] Hamberg, J., Controlled Lagrangians, Symmetries and Conditions for Strong Matching, *In: Lagrangian and Hamiltonian Methods for Nonlinear Control,* Elsevier, 2000.

[13] Junge, O., J. Marsden, & S. Ober-Blbaum, Discrete Mechanics and Optimal Control, *Proc. of the 16th IFAC World Congress*, 2005.

[14] Kane, C., J. E. Marsden, M. Ortiz, & M. West, Variational Integrators and the Newmark Algorithm for Conservative and Dissipative Mechanical Systems, *Int. J. Numer. Math. Eng.* **49**, 2000, 1295–1325.

[15] Marsden, J. E., *Lectures on Mechanics,* London Mathematical Society Lecture Note Series **174**, Cambridge University Press, 1992.

[16] Marsden, J. E. & T. S. Ratiu, *An Introduction to Mechanics and Symmetry.* Texts in Appl. Math. **17**, Springer-Verlag, 1999.

[17] Marsden, J. E. & M. West, Discrete Mechanics and Variational Integrators, *Acta Numerica* **10**, 2001, 357–514.

[18] Maschke, B., R. Ortega, & A. van der Schaft, Energy-Based Lyapunov Functions for Forced Hamiltonian Systems with Dissipation, *IEEE Trans. Automat. Control* **45**, 2001, 1498–1502.

[19] Sanyal, A., J. Shen, N. H. McClamroch, & A. M. Bloch, Stability and Stabilization of Relative Equilibria of the Dumbbell Satellite in Central Gravity, 2006, *Journal of the American Institute of Aeronautics and Astronautics,* (to appear).

[20] Zenkov, D. V., A. M. Bloch, N. E. Leonard, & J. E. Marsden, Matching and Stabilization of Low-Dimensional Nonholonomic Systems, *Proc. CDC* **39**, 2000, 1289–1295.

[21] Zenkov, D. V., A. M. Bloch, & J. E. Marsden, Flat Nonholonomic Matching, *Proc. ACC*, 2002, 2812–2817.

Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

A Discrete Variational Integrator for Optimal Control Problems on $\mathsf{SO}(3)$

Islam I. Hussein Melvin Leok Amit K. Sanyal Anthony M. Bloch

Abstract— **In this paper we study a discrete variational optimal control problem for the rigid body. The cost to be minimized is the external torque applied to move the rigid body from an initial condition to a pre-specified terminal condition. Instead of discretizing the equations of motion, we use the discrete equations obtained from the discrete Lagrange–d'Alembert principle, a process that better approximates the equations of motion. Within the discrete-time setting, these two approaches are not equivalent in general. The kinematics are discretized using a natural Lie-algebraic formulation that guarantees that the flow remains on the Lie group $\mathsf{SO}(3)$ and its algebra $\mathfrak{so}(3)$. We use Lagrange's method for constrained problems in the calculus of variations to derive the discrete-time necessary conditions. We give a numerical example for a three-dimensional rigid body maneuver.**

I. Introduction

This paper deals with a structure-preserving computational approach to the optimal control problem of minimizing the control effort necessary to perform an attitude transfer from an initial state to a prescribed final state, in the absence of a potential field. The configuration of the rigid body is given by the rotation matrix from the body frame to the spatial frame, which is an element of the group of orientation-preserving isometries in $\mathbb{R}^3$. The state of the rigid body is described by the rotation matrix and its angular velocity.

To motivate the computational approach we adopt in the discrete-time case, we first revisit the variational continuous-time optimal control problem. The continuous-time extremal solutions to this optimal control problem have certain special features, since they arise from variational principles. General numerical integration methods, including the popular Runge-Kutta schemes, typically preserve neither first integrals nor the characteristics of the configuration space. Geometric integrators are the class of numerical integration schemes that preserve such properties, and a good survey can be found in [1]. Techniques particular to Hamiltonian systems are also discussed in [2] and [3].

Our approach to discretizing the optimal control problem is in contrast to traditional techniques such as collocation, wherein the continuous equations of motion are imposed as constraints at a set of collocation points. In our approach, modeled after [4], the discrete equations of motion are derived from a discrete variational principle, and this induces constraints on the configuration at each discrete time step.

This approach yields discrete dynamics that are more faithful to the continuous equations of motion, and consequently yields more accurate numerical solutions to the optimal control problem. This feature is extremely important in computing accurate (sub)optimal trajectories for long-term spacecraft attitude maneuvers. For example, in [5], the authors propose an imaging spacecraft formation design that requires a continuous attitude maneuver over a period of 77 days in a low Earth orbit. Hence, the attitude maneuver has to be very accurate to meet tight imaging constraints over long time ranges. The proposed variational scheme can also be easily extended to other types of Lie groups. For example, in long range inter-planetary orbit transfers (see, for example, [6]), one is interested in computing optimal or suboptimal trajectories on the group of rigid body motions $\mathsf{SE}(3)$ with a high degree of accuracy. Similar requirements also apply to the control of quantum systems. For example, efficient construction of quantum gates is a problem on the unitary Lie group $\mathsf{SU}(N)$. This is an optimal control problem, where one wishes to steer the identity operator to the desired unitary operator (see, for example, [7] and [8]).

Moreover, an important feature of the way we discretize the optimal control problem is that it is $\mathsf{SO}(3)$-equivariant. This is desirable, since it ensures that our numerical results are independent of the choice of coordinates and coordinate frames. This is in contrast to methods based on coordinatizing the rotation group using quaternions, (modified) Rodrigues parameters, and Euler angles, as given in the survey [9]. Even if the optimal cost function is $\mathsf{SO}(3)$-invariant, as in [10], the use of generalized coordinates imposes constraints on the attitude kinematics.

For the purpose of numerical simulation, the corresponding discrete optimal control problem is posed on the discrete state space as a two stage discrete variational problem. In the first step, we derive the discrete dynamics for the rigid body in the context of discrete variational mechanics [11]. This is achieved by considering the discrete Lagrange–d'Alembert variational principle [12] in combination with essential ideas from Lie group methods [13], which yields a Lie group variational integrator [14]. This integrator explicitly preserves the Lie group structure of the configuration space, and is similar to the integrators introduced in [15] for a rigid body in an external field, and in [16] for full body dynamics. These discrete equations are then imposed as constraints to be satisfied by the extremal solutions to the

Islam Hussein is Assistant Professor of Mechanical Engineering at Worcester Polytechnic Institute, ihussein@uiuc.edu.

Melvin Leok is Assistant Professor of Mathematics at Purdue University, West Lafayette, mleok@umich.edu.

Amit Sanyal is Post-Doctoral Research Associate at Arizona State University, Tempe, sanyal@asu.edu.

Anthony Bloch is Professor of Mathematics at the University of Michigan, Ann Arbor, abloch@umich.edu.

1-4244-0171-2/06/$20.00 ©2006 IEEE.

discrete optimal control problem, and we obtain the discrete extremal solutions in terms of the given terminal states.

The paper is organized as follows. As motivation, in Section II, we study the optimal control problem in continuous-time. In Section III, we study the discrete-time case. In particular, in Section III-A we state the optimal control problem and describe our approach. In Section III-B, we derive the discrete-time equations of motion for the rigid body starting with the discrete Lagrange–d'Alembert principle. These equations are used in Section III-C for the optimal control problem. In Section IV, we describe an algorithm for solving the general nonlinear, implicit necessary conditions for $\mathsf{SO}(3)$ and give numerical examples for rest-to-rest and slew-up spacecraft maneuvers.

II. Continuous-Time Results

A. Problem Formulation

In this paper, the natural pairing between $\mathfrak{so}^*(3)$ and $\mathfrak{so}(3)$ is denoted by $\langle \cdot, \cdot \rangle$. Let $\ll \cdot, \cdot \gg$ and $\ll \cdot, \cdot \gg_*$ denote the standard (induced by the Killing form) inner product on $\mathfrak{so}(3)$ and $\mathfrak{so}^*(3)$, respectively. The inner product $\ll \cdot, \cdot \gg_*$ is naturally induced from the standard norm $\ll \boldsymbol{\xi}, \boldsymbol{\omega} \gg = -\frac{1}{2}\mathrm{Tr}(\boldsymbol{\xi}^T\boldsymbol{\omega})$, for all $\boldsymbol{\xi}, \boldsymbol{\omega} \in \mathfrak{so}(3)$, through

$$\begin{aligned} \ll \boldsymbol{\eta}, \boldsymbol{\varphi} \gg_* &= \left\langle \boldsymbol{\eta}, \boldsymbol{\varphi}^\sharp \right\rangle = \langle \boldsymbol{\eta}, \boldsymbol{\omega} \rangle = \left\langle \boldsymbol{\xi}^\flat, \boldsymbol{\omega} \right\rangle \\ &= \ll \boldsymbol{\xi}, \boldsymbol{\omega} \gg, \end{aligned} \tag{1}$$

where $\boldsymbol{\varphi} = \boldsymbol{\omega}^\flat \in \mathfrak{so}^*(3)$ and $\boldsymbol{\eta} = \boldsymbol{\xi}^\flat \in \mathfrak{so}^*(3)$, with $\boldsymbol{\xi}, \boldsymbol{\omega} \in \mathfrak{so}(3)$ and $\flat$ and $\sharp$ are the musical isomorphisms (see §2.5 of [17]) with respect to the standard metric $\ll \cdot, \cdot \gg$. On $\mathfrak{so}(3)$, these isomorphisms correspond to the transpose operation. That is, we have $\boldsymbol{\varphi} = \boldsymbol{\omega}^{\mathrm{T}}$ and $\boldsymbol{\eta} = \boldsymbol{\xi}^{\mathrm{T}}$.

Let $\mathbf{J} : \mathfrak{so}(3) \to \mathfrak{so}^*(3)$ be the positive definite inertia operator. It can be shown that

$$\langle \mathbf{J}(\boldsymbol{\xi}), \boldsymbol{\omega} \rangle = \langle \mathbf{J}(\boldsymbol{\omega}), \boldsymbol{\xi} \rangle . \tag{2}$$

On $\mathfrak{so}(3)$, $\mathbf{J}$ is given by $\mathbf{J}(\boldsymbol{\xi}) = J\boldsymbol{\xi} + \boldsymbol{\xi}J$, where J is a positive definite symmetric matrix (see, for example, [17], [18]). Moreover, we also have $\mathbf{J}(\boldsymbol{\eta}^\sharp)^\sharp = (J\boldsymbol{\eta}^{\mathrm{T}} + \boldsymbol{\eta}^{\mathrm{T}}J)^{\mathrm{T}} = \mathbf{J}(\boldsymbol{\eta})$, which is an abuse of notation since $\boldsymbol{\eta} \in \mathfrak{so}^*(3)$. For the sake of generality and mathematical accuracy, we will use the general definitions, though it is helpful to keep the above identifications for $\mathfrak{so}(3)$ in mind.

In this section we review some continuous-time optimal control results using a simple optimal control example on $\mathsf{SO}(3)$. Here, we minimize the norm squared of the control torque $\boldsymbol{\tau} \in \mathfrak{so}^*(3)$ applied to rotate a rigid body subject to the Lagrange–d'Alembert principle for the rigid body[1] whose configuration is given by $\mathbf{R} \in \mathsf{SO}(3)$ and body angular velocity is given by $\boldsymbol{\Omega} \in \mathfrak{so}(3)$. We require that the system evolve from an initial state $(\mathbf{R}_0, \boldsymbol{\Omega}_0)$ to a final state $(\mathbf{R}_T, \boldsymbol{\Omega}_T)$ at a fixed terminal time T. Hence, we have the following minimum control effort optimal control problem.

[1]This is equivalent to constraining the problem to satisfy the rigid body equations of motion given by equations (7). However, for the sake of generality that will be appreciated in the discrete-time problem, we choose to treat the Lagrange–d'Alembert principle as the constraint as opposed to the rigid body equations of motion. Both are equivalent in the continuous-time case but are generally not equivalent in the discrete-time case.

Problem 2.1: Minimize

$$\mathcal{J} = \frac{1}{2}\int_0^T \ll \boldsymbol{\tau}, \boldsymbol{\tau} \gg_* \mathrm{d}t \tag{3}$$

subject to

1) satisfying *Lagrange–d'Alembert principle*:

$$\boldsymbol{\delta}\int_0^T \frac{1}{2}\langle \mathbf{J}(\boldsymbol{\Omega}), \boldsymbol{\Omega} \rangle \mathrm{d}t + \int_0^T \langle \boldsymbol{\tau}, \mathbf{W} \rangle \mathrm{d}t = 0, \tag{4}$$

subject to $\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\Omega}$, where $\mathbf{W}$ is the variation vector field to be defined below,

2) and the *boundary conditions*

$$\begin{aligned} \mathbf{R}(0) &= \mathbf{R}_0, \ \boldsymbol{\Omega}(0) = \boldsymbol{\Omega}_0, \\ \mathbf{R}(T) &= \mathbf{R}_T, \ \boldsymbol{\Omega}(T) = \boldsymbol{\Omega}_T. \end{aligned} \tag{5}$$

We now show that this is equivalent to the following problem formulation, where the rigid body equations of motion replace the Lagrange–d'Alembert principle.

Problem 2.2: Minimize

$$\mathcal{J} = \frac{1}{2}\int_0^T \ll \boldsymbol{\tau}, \boldsymbol{\tau} \gg_* \mathrm{d}t \tag{6}$$

subject to

1) the *dynamics*

$$\begin{aligned} \dot{\mathbf{R}} &= \mathbf{R}\boldsymbol{\Omega} \\ \dot{\mathbf{M}} &= \mathrm{ad}^*_{\boldsymbol{\Omega}}\mathbf{M} + \boldsymbol{\tau} = [\mathbf{M}, \boldsymbol{\Omega}] + \boldsymbol{\tau}, \end{aligned} \tag{7}$$

where $\mathbf{M} = \mathbf{J}(\boldsymbol{\Omega}) \in \mathfrak{so}^*(3)$ is the momentum,

2) and the *boundary conditions*

$$\begin{aligned} \mathbf{R}(0) &= \mathbf{R}_0, \ \boldsymbol{\Omega}(0) = \boldsymbol{\Omega}_0, \\ \mathbf{R}(T) &= \mathbf{R}_T, \ \boldsymbol{\Omega}(T) = \boldsymbol{\Omega}_T. \end{aligned} \tag{8}$$

In the above, ad^* is the dual of the adjoint representation, ad, of $\mathfrak{so}(3)$ and is given by $\mathrm{ad}^*_{\boldsymbol{\xi}}\boldsymbol{\eta} = -[\boldsymbol{\xi}, \boldsymbol{\eta}] \in \mathfrak{so}^*(3)$, for all $\boldsymbol{\xi} \in \mathfrak{so}(3)$ and $\boldsymbol{\eta} \in \mathfrak{so}^*(3)$. Recall that the bracket is defined by $[\boldsymbol{\xi}, \boldsymbol{\omega}] = \boldsymbol{\xi}\boldsymbol{\omega} - \boldsymbol{\omega}\boldsymbol{\xi}$.

B. The Lagrange–d'Alembert Principle and the Rigid Body Equations of Motion

In this section we derive the forced rigid body equations of motion (equations (7)) from the Lagrange–d'Alembert principle. In dealing with the kinematic constraint, $\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\Omega}$, we may either append it to the Lagrangian using the method of Lagrange multipliers, or we can directly compute the constrained variations (see §13.5 of [19]). Here, we take the direct approach as it yields a more concise derivation.

First, we take variations of the kinematic condition $\boldsymbol{\Omega} = \mathbf{R}^{-1}\mathbf{R}$ to obtain $\boldsymbol{\delta\Omega} = -\mathbf{R}^{-1}(\boldsymbol{\delta}\mathbf{R})\mathbf{R}^{-1}\dot{\mathbf{R}} + \mathbf{R}^{-1}\boldsymbol{\delta}\dot{\mathbf{R}}$. As defined previously, we have $\mathbf{W} = \mathbf{R}^{-1}\boldsymbol{\delta}\mathbf{R}$ and, therefore, $\dot{\mathbf{W}} = -\mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{R}^{-1}\boldsymbol{\delta}\mathbf{R} + \mathbf{R}^{-1}\boldsymbol{\delta}\dot{\mathbf{R}} = -\boldsymbol{\Omega}\mathbf{W} + \mathbf{R}^{-1}\boldsymbol{\delta}\dot{\mathbf{R}}$, since $\boldsymbol{\delta}\dot{\mathbf{R}} = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\delta}\mathbf{R}$ (see [20], p. 52). Hence, we have

$$\boldsymbol{\delta\Omega} = -\mathbf{W}\boldsymbol{\Omega} + \boldsymbol{\Omega}\mathbf{W} + \dot{\mathbf{W}} = \mathrm{ad}_{\boldsymbol{\Omega}}\mathbf{W} + \dot{\mathbf{W}}. \tag{9}$$

Taking variations of the Lagrange–d'Alembert principle we obtain

$$\int_0^T \langle \mathbf{J}(\boldsymbol{\Omega}), \boldsymbol{\delta\Omega} \rangle + \langle \boldsymbol{\tau}, \mathbf{W} \rangle \mathrm{d}t = 0.$$

Using the variation in equation (9) and integrating by parts, we obtain

$$0 = \int_0^T \left\langle -\dot{\mathbf{M}} + \mathrm{ad}^*_{\boldsymbol{\Omega}}\mathbf{M} + \boldsymbol{\tau}, \mathbf{W} \right\rangle \mathrm{d}t + \left[\langle \mathbf{J}(\boldsymbol{\Omega}), \mathbf{W}(t) \rangle\right]_0^T,$$

where we have used the property

$$\langle \boldsymbol{\eta}, \mathrm{ad}_{\boldsymbol{\omega}}\boldsymbol{\xi} \rangle = \langle \mathrm{ad}^*_{\boldsymbol{\omega}}\boldsymbol{\eta}, \boldsymbol{\xi} \rangle , \ \boldsymbol{\eta} \in \mathfrak{so}^*(3), \ \boldsymbol{\omega}, \boldsymbol{\xi} \in \mathfrak{so}(3). \tag{10}$$

This gives the desired result, with $\mathbf{M} = \mathbf{J}(\mathbf{\Omega})$.

C. Continuous-Time Variational Optimal Control Problem

A direct variational approach is used here to derive the necessary conditions for the optimal control Problem (2.2).

A Second Order Direct Approach. "Second order" is used here to reflect the fact that we now study variations of second order dynamical equations as opposed to the kinematic direct approach studied in Section II-B. We now give the resulting necessary conditions using a direct approach as in [19]. We already computed the variations of $\mathbf{R}$ and $\mathbf{\Omega}$. These were as follows: $\boldsymbol{\delta}\mathbf{R} = \mathbf{R}\mathbf{W}$ and $\boldsymbol{\delta}\mathbf{\Omega} = \mathrm{ad}_{\mathbf{\Omega}}\mathbf{W} + \dot{\mathbf{W}}$. We now compute the variation of $\dot{\mathbf{M}}$ with the goal of obtaining the proper variations for $\boldsymbol{\tau}$:

$$\boldsymbol{\delta}\dot{\mathbf{M}} = \mathbf{J}\left(\boldsymbol{\delta}\dot{\mathbf{\Omega}}\right) = \mathbf{J}\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\delta}\mathbf{\Omega} + \mathcal{R}\left(\mathbf{W},\mathbf{\Omega}\right)\mathbf{\Omega}\right),$$

where $\mathcal{R}$ is the curvature tensor on SO(3). The curvature tensor $\mathcal{R}$ arises due to the identity (see [20], page 52)

$$\frac{\partial}{\partial\epsilon}\frac{\partial}{\partial t}\mathbf{Y} - \frac{\partial}{\partial t}\frac{\partial}{\partial\epsilon}\mathbf{Y} = \mathcal{R}(\mathbf{W},\mathbf{Y})\mathbf{\Omega},$$

where $\mathbf{Y} \in \mathsf{TSO}(3)$ is any vector field along the curve $\mathbf{R}(t) \in \mathsf{SO}(3)$. Taking variations of $\dot{\mathbf{M}} = \mathrm{ad}^*_{\mathbf{\Omega}}\mathbf{M} + \boldsymbol{\tau}$ we obtain $\boldsymbol{\delta}\dot{\mathbf{M}} = \mathrm{ad}^*_{\boldsymbol{\delta}\mathbf{\Omega}}\mathbf{M} + \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\delta}\mathbf{M} + \boldsymbol{\delta}\boldsymbol{\tau}$. We now have the desired variation in $\boldsymbol{\tau}$:

$$\begin{aligned}\boldsymbol{\delta}\boldsymbol{\tau} &= \mathbf{J}\left(\mathcal{R}\left(\mathbf{W},\mathbf{\Omega}\right)\mathbf{\Omega}\right) + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{J}\left(\boldsymbol{\delta}\mathbf{\Omega}\right) - \mathrm{ad}^*_{\boldsymbol{\delta}\mathbf{\Omega}}\mathbf{M} \\ &\quad -\mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\delta}\mathbf{M}. \end{aligned} \tag{11}$$

Take variations of the cost functional (6) to obtain:

$$\begin{aligned}\boldsymbol{\delta}\mathcal{J} = \int_0^T \Big\langle &\mathbf{J}(\ddot{\varsigma}) - \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathbf{J}(\dot{\varsigma})\right) + \dot{\boldsymbol{\eta}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{ad}^*_{\varsigma}\mathbf{M}\right) \\ &+ \left[\mathcal{R}\left(\mathbf{J}(\varsigma)^\sharp,\mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\mathrm{ad}^*_{\varsigma}\mathbf{M} - \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\eta}, \mathbf{W}\Big\rangle \mathrm{d}t,\end{aligned}$$

where $\varsigma = \boldsymbol{\tau}^\sharp \in \mathfrak{so}(3)$ and $\boldsymbol{\eta} = \mathbf{J}\left(\mathrm{ad}_{\mathbf{\Omega}}\varsigma\right) \in \mathfrak{so}^*(3)$. Here, we used integration by parts and the boundary conditions (7), equations (9) and (11), and the identities (1), (2) and (10). Hence, we have the following theorem.

Theorem 2.1: The necessary optimality conditions for the problem of minimizing (6) subject to the dynamics (7) and the boundary conditions (8) are given by the single fourth order[2] differential equation

$$\begin{aligned}0 &= \mathbf{J}(\ddot{\varsigma}) - \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathbf{J}(\dot{\varsigma})\right) + \dot{\boldsymbol{\eta}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{ad}^*_{\varsigma}\mathbf{M}\right) \\ &\quad + \left[\mathcal{R}\left(\left(\mathbf{J}(\varsigma)\right)^\sharp,\mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathrm{ad}^*_{\varsigma}\mathbf{M}\right) - \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\eta},\end{aligned}$$

as well as the equations (7) and the boundary conditions (8), where ς and $\boldsymbol{\eta}$ are as defined above.

To obtain above result we used the initial conditions (8), and the fact that the vector fields $\mathbf{\Omega}$ and $\mathbf{W}$ are left-invariant vector fields. The curvature tensor is evaluated at a point $\mathbf{R}(t) \neq \mathbf{I}$. That is, we get $\mathcal{R}_{\mathbf{R}_\epsilon}\left(\frac{\partial\mathbf{R}_\epsilon}{\partial\epsilon}, \frac{\partial\mathbf{R}_\epsilon}{\partial t}\right)\mathbf{\Omega}$. Evaluating this at $\epsilon = 0$ we get: $\mathcal{R}_{\mathbf{R}}\left(\mathbf{R}\mathbf{W},\mathbf{R}\mathbf{\Omega}\right)\mathbf{\Omega}$. Since $\mathbf{R}\mathbf{W}$ and $\mathbf{R}\mathbf{\Omega}$ are left-invariant vector fields at the group element $\mathbf{R}(t)$, by the identification $\mathsf{T}_{\mathbf{R}}\mathsf{SO}(3) \simeq \mathfrak{so}(3)$, we have $\mathcal{R}_{\mathbf{R}}\left(\mathbf{R}\mathbf{W},\mathbf{R}\mathbf{\Omega}\right)\mathbf{\Omega} = \mathcal{R}\left(\mathbf{W},\mathbf{\Omega}\right)\mathbf{\Omega}$, which is the curvature tensor evaluated at the identity element. For a compact semi-simple Lie group G with Lie algebra $\mathfrak{g}$, the curvature tensor, with respect to a bi-invariant metric, is (see [20]):

$$\mathcal{R}\left(\mathbf{X},\mathbf{Y}\right)\mathbf{Z} = \frac{1}{4}\mathrm{ad}_{\mathrm{ad}_{\mathbf{X}}\mathbf{Y}}\mathbf{Z}, \tag{12}$$

for all $\mathbf{X},\mathbf{Y},\mathbf{Z} \in \mathfrak{g}$.

Using a Lagrange multiplier approach, we obtain instead the following theorem.

Theorem 2.2: The necessary optimality conditions for the problem of minimizing (6) subject to the dynamics (7) and the boundary conditions (8) are given by

$$\begin{aligned}\boldsymbol{\tau} &= \mathbf{\Lambda}_2 \\ \dot{\mathbf{\Lambda}}_1 &= \left[\mathcal{R}\left(\mathbf{J}\left(\mathbf{\Lambda}_2\right)^\sharp,\mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\mathbf{\Lambda}_1 \\ \dot{\mathbf{\Lambda}}_2 &= -\mathbf{J}^{-1}\left(\mathbf{\Lambda}_1\right) - \mathrm{ad}_{\mathbf{\Omega}}\mathbf{\Lambda}_2 + \mathbf{J}^{-1}\left(\mathrm{ad}^*_{\mathbf{\Lambda}_2}\mathbf{M}\right)\end{aligned} \tag{13}$$

and the equations (7) and the boundary conditions (8). The Lagrange multipliers $\mathbf{\Lambda}_1 \in \mathfrak{so}^*(3)$, $\mathbf{\Lambda}_2 \in \mathfrak{so}(3)$ correspond to the kinematic and dynamics constraints (7), respectively.

Remark 2.1: Note that the equations of motion that arise from the Lagrange–d'Alembert principle are used to define the dynamic constraints. In effect, we minimize $\mathcal{J}$ subject to satisfying the Lagrange–d'Alembert principle. Analogously, the discrete version of the Lagrange–d'Alembert principle will be used to derive the discrete equations of motion in the discrete optimal control problem to be studied in Section III-C. This view is in line with the approach in [4] in that we do not discretize the equations of motion directly, but, instead, we discretize the Lagrange–d'Alembert principle. These two approaches are not equivalent in general.

Corollary 2.1: The necessary optimality conditions of Theorem 2.1 are equivalent to those of Theorem 2.2.

Proof. In Theorem 2.2, differentiate $\mathbf{\Lambda}_2$ once and then use all three differential equations to replace $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ with expressions involving only $\boldsymbol{\tau}$, $\mathbf{M}$ and $\mathbf{\Omega}$. ■

III. Discrete-Time Results

A. Problem Formulation

In this section we give the discrete version of the problem introduced in Section II-A. So, we consider minimizing the norm squared of the control torque $\boldsymbol{\tau}$ subject to satisfaction of the discrete Lagrange–d'Alembert principle for the rigid body whose configuration and body angular velocity at time step t_k are given by $\mathbf{R}_k \in \mathsf{SO}(3)$ and $\mathbf{\Omega}_k \in \mathfrak{so}(3)$, respectively. The kinematic constraint may be expressed as

$$\mathbf{R}_{k+1} = \mathbf{R}_k \exp\left(h\mathbf{\Omega}_k\right) = \mathbf{R}_k\mathbf{g}_k, \tag{14}$$

where h is the integration time step, $\exp : \mathfrak{so}(3) \to \mathsf{SO}(3)$ is the exponential map and $\mathbf{g}_k = \exp(h\mathbf{\Omega}_k)$. The boundary conditions are given by $(\mathbf{R}^*_0, \mathbf{\Omega}^*_0)$ and $(\mathbf{R}^*_N, \mathbf{\Omega}^*_{N-1})$, where $t_0 = 0$ and $N = T/h$ is such that $t_N = T$.

The reason we constrain $\mathbf{\Omega}$ at $t = h(N-1)$ instead of at $t = hN$ is that a constraint on $\mathbf{\Omega}_k \in \mathfrak{so}(3)$ corresponds, by left translations to a constraint on $\dot{\mathbf{R}}_k \in \mathsf{T}_{\mathbf{R}_k}\mathsf{SO}(3)$. In turn, in the discrete setting and depending on the choice of discretization, this corresponds to a constraint on the neighboring discrete points. With our choice of discretization (equation (14)), this corresponds to constraints on $\mathbf{R}_k$ and $\mathbf{R}_{k+1}$. Hence, to ensure that the effect of the terminal constraint on $\mathbf{\Omega}$ is correctly accounted for, the constraint

[2] Second order in $\boldsymbol{\tau}$ and fourth order in $\mathbf{R}$.

must be imposed on $\mathbf{\Omega}_{N-1}$, which entails some constraints on variations at both $\mathbf{R}_{N-1}$ and $\mathbf{R}_N$. We will return to this point later in the paper.

Equation (14) is just one way of discretizing the kinematics of the rigid body. In the case of planar rigid body dynamics, this leads to the first-order Euler approximation. However, on $\mathsf{SO}(3)$, our approach yields a novel discretization. We make the above choice for discretization as it guarantees, in general, that the angular velocity matrix $\mathbf{\Omega}_k$ remains on the algebra $\mathfrak{so}(3)$ by using the exponential map. This is natural to do in the context of discrete variational numerical solvers (for both initial value and two point boundary value problems). Following the methodology of [4], we have the following optimal control problem.

Problem 3.1: Minimize

$$\mathcal{J} = \sum_{k=0}^{N} \frac{1}{2} \ll \boldsymbol{\tau}_k, \boldsymbol{\tau}_k \gg_* \tag{15}$$

subject to

1) satisfying the *discrete Lagrange–d'Alembert principle*:

$$\boldsymbol{\delta} \sum_{k=0}^{N-1} \frac{1}{2} \left\langle \mathbf{J}\left(\mathbf{\Omega}_k\right), \mathbf{\Omega}_k \right\rangle + \sum_{k=0}^{N} \left\langle \boldsymbol{\tau}_k, \mathbf{W}_k \right\rangle = 0, \tag{16}$$

subject to $\mathbf{R}_0 = \mathbf{R}_0^*$, $\mathbf{R}_N = \mathbf{R}_N^*$ and $\mathbf{R}_{k+1} = \mathbf{R}_k \mathbf{g}_k$, $k = 0, 1, \ldots, N-1$, where $\mathbf{W}_k$ is the variation vector field at time step t_k satisfying $\boldsymbol{\delta}\mathbf{R}_k = \mathbf{R}_k \mathbf{W}_k$,

2) and the *boundary conditions*

$$\begin{aligned} \mathbf{R}_0 &= \mathbf{R}_0^*, \ \mathbf{\Omega}_0 = \mathbf{\Omega}_0^*, \\ \mathbf{R}_N &= \mathbf{R}_N^*, \ \mathbf{\Omega}_{N-1} = \mathbf{\Omega}_{N-1}^*. \end{aligned} \tag{17}$$

The following formulation is equivalent, where the discrete rigid body equations of motion replace the Lagrange–d'Alembert principle constraint.

Problem 3.2: Minimize

$$\mathcal{J} = \sum_{k=0}^{N} \frac{1}{2} \ll \boldsymbol{\tau}_k, \boldsymbol{\tau}_k \gg_* \tag{18}$$

subject to

1) the *discrete dynamics*

$$\begin{aligned} &\mathbf{R}_{k+1} = \mathbf{R}_k \mathbf{g}_k, \ k = 0, \ldots, N-1 \\ &\mathbf{M}_k = \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right), \ k = 1, \ldots, N-1, \\ &\mathbf{M}_k = \mathbf{J}\left(\mathbf{\Omega}_k\right), \ k = 0, \ldots, N-1, \end{aligned} \tag{19}$$

2) and the *boundary conditions*

$$\begin{aligned} \mathbf{R}_0 &= \mathbf{R}_0^*, \ \mathbf{\Omega}_0 = \mathbf{\Omega}_0^*, \\ \mathbf{R}_N &= \mathbf{R}_N^*, \ \mathbf{\Omega}_{N-1} = \mathbf{\Omega}_{N-1}^*. \end{aligned} \tag{20}$$

Regarding terminal velocity conditions, note that in the second of equations (19) if we let $k = N$ we find that $\mathbf{\Omega}_N$ appears in the equation. A constraint on $\mathbf{\Omega}_N$ dictates constraints at the points $\mathbf{R}_N$ and $\mathbf{R}_{N+1}$ through the first equation in (19). Since we only consider time points up to $t = Nh$, we can not allow $k = N$ in the second of equations (19) and hence our terminal velocity constraints are posed in terms of $\mathbf{\Omega}_{N-1}$ instead of $\mathbf{\Omega}_N$.

As mentioned above, $\mathbf{W}_k$ is a variation vector field associated with the perturbed group element $\mathbf{R}_k^\epsilon$. Likewise, we need to define a variation vector field associated with the element $\mathbf{g}_k = \exp(h\mathbf{\Omega}_k)$. First, let the perturbed variable $\mathbf{g}_k^\epsilon$ be defined by

$$\mathbf{g}_k^\epsilon = \mathbf{g}_k \exp(\epsilon h \boldsymbol{\delta}\mathbf{\Omega}_k), \tag{21}$$

where

$$\boldsymbol{\delta}\mathbf{\Omega} = \left.\frac{\partial \mathbf{\Omega}_k^\epsilon}{\partial \epsilon}\right|_{\epsilon=0}.$$

Note that $\mathbf{g}_k^\epsilon\big|_{\epsilon=0} = \mathbf{g}_k$ as desired. Moreover, we have

$$\boldsymbol{\delta}\mathbf{g}_k = \mathbf{g}_k(h\boldsymbol{\delta}\mathbf{\Omega}_k)\exp(\epsilon h \delta\mathbf{\Omega}_k)\big|_{\epsilon=0} = h\mathbf{g}_k\boldsymbol{\delta}\mathbf{\Omega}_k. \tag{22}$$

This will be needed later when taking variations.

B. The Discrete Lagrange–d'Alembert Principle and the Rigid Body Equations of Motion

In this section we derive the discrete forced rigid body equations of motion (equations (19)) starting with the discrete Lagrange–d'Alembert principle. As in the continuous case, we will compute the constrained variation of $\boldsymbol{\delta}\mathbf{\Omega}_k$. We begin by rewriting the kinematic constraint as $\exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) = h\mathbf{\Omega}_k$, which is easier to handle as an expression over the Lie algebra. Take variations to obtain, $-\mathbf{R}_k^{-1}\left(\boldsymbol{\delta}\mathbf{R}_k\right)\mathbf{R}_k^{-1}\mathbf{R}_{k+1} + \mathbf{R}_k^{-1}\boldsymbol{\delta}\mathbf{R}_{k+1} = h\mathbf{g}_k \cdot \boldsymbol{\delta}\mathbf{\Omega}_k$, which is equivalent to $-\mathbf{W}_k\mathbf{g}_k + \mathbf{g}_k\mathbf{W}_{k+1} = h\mathbf{g}_k\boldsymbol{\delta}\mathbf{\Omega}_k$, or

$$\boldsymbol{\delta}\mathbf{\Omega}_k = \frac{1}{h}\left[-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k + \mathbf{W}_{k+1}\right]. \tag{23}$$

Note that this is an expression over the Lie algebra $\mathfrak{so}(3)$.

Taking direct variations of the cost functional we obtain

$$\begin{aligned} 0 = &\left\langle \boldsymbol{\tau}_0 - \frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_0^{-1}}\mathbf{J}\left(\mathbf{\Omega}_0\right), \mathbf{W}_0 \right\rangle \\ &+ \left\langle \boldsymbol{\tau}_N + \frac{1}{h}\mathbf{J}\left(\Omega_{N-1}\right), \mathbf{W}_N \right\rangle \\ &+ \sum_{k=1}^{N-1} \left\langle \boldsymbol{\tau}_k - \frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{J}\left(\mathbf{\Omega}_k\right) + \frac{1}{h}\mathbf{J}\left(\mathbf{\Omega}_{k-1}\right), \mathbf{W}_k \right\rangle. \end{aligned}$$

where we have used equation (23). By the boundary conditions $\mathbf{R}_0 = \mathbf{R}_0^*$ and $\mathbf{R}_N = \mathbf{R}_N^*$, we have $\mathbf{W}_0 = 0$ and $\mathbf{W}_N = 0$. Since $\mathbf{W}_k$, $k = 1, \ldots, N-1$, are arbitrary and independent, we obtain the equivalent of equations (19).

C. Discrete-Time Variational Optimal Control Problem

Analogous to the direct approach in continuous time, here we derive the necessary optimality conditions in a form that does not involve the use of Lagrange multipliers. Using equation (23) and taking variation of the second of equations (19), we obtain

$$\begin{aligned} \boldsymbol{\delta}\boldsymbol{\tau}_k = &\ \mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\left(\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_{k+1} - \mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k\right)\right. \\ &\left. + \frac{1}{h}\left[\mathbf{W}_{k+1} - \mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k, \mathbf{J}\left(\mathbf{\Omega}_k\right)\right]\right) \\ &- \frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_k - \mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\mathbf{W}_{k-1}\right), \end{aligned} \tag{24}$$

for $k = 1, \ldots, N-1$. Taking variations of the cost functional (18) and using equation (24), one obtains after a tedious but straightforward computation an expression for $\boldsymbol{\delta}\mathcal{J}$ in terms of $\boldsymbol{\delta}\boldsymbol{\tau}_k$, which we omit because of space restrictions. When $\boldsymbol{\delta}\mathcal{J}$ is equated to zero, the resulting equation gives (boundary) conditions on $\boldsymbol{\tau}_0, \boldsymbol{\tau}_1, \boldsymbol{\tau}_{N-1}, \boldsymbol{\tau}_N$ as well as discrete evolution equations that are written in

algebraic nonlinear form as:

$$\begin{aligned} 0 &= -\frac{1}{h^2}\Big(\mathbf{J}\left(\boldsymbol{\tau}_k^\sharp\right) - \mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{J}\left(\boldsymbol{\tau}_{k+1}^\sharp\right) \\ &\quad -\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\boldsymbol{\tau}_{k-1}^\sharp\right) + \mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_k^{-1}}\boldsymbol{\tau}_k^\sharp\right)\Big) \\ &\quad -\frac{1}{h}\Big(\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\left[\mathbf{J}\left(\boldsymbol{\Omega}_k\right), \mathrm{Ad}_{\mathbf{g}_k^{-1}}\left(\boldsymbol{\tau}_k^\sharp\right)\right] \\ &\quad -\frac{1}{h}\left[\mathbf{J}\left(\boldsymbol{\Omega}_{k-1}\right), \mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\left(\boldsymbol{\tau}_{k-1}^\sharp\right)\right]\Big), \end{aligned} \tag{25}$$

for $k = 2, \ldots, N-2$.

A Lagrange multiplier approach yields the following equivalent theorem.

Theorem 3.1: The necessary optimality conditions for the discrete Problem 3.2 are

$$\begin{aligned} \mathbf{R}_{k+1} &= \mathbf{R}_k\mathbf{g}_k,\ k = 1, \ldots, N-2 \\ \mathbf{M}_k &= \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right),\ k = 1, \ldots, N-1 \\ 0 &= \boldsymbol{\Lambda}^1_{k-1} - \mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\boldsymbol{\Lambda}^1_k,\ k = 2, \ldots, N-2 \\ 0 &= -\boldsymbol{\Lambda}^1_k + \mathbf{J}\left(\boldsymbol{\Lambda}^2_k\right) - \mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_{k+1}}\boldsymbol{\Lambda}^2_{k+1}\right) \\ &\quad +h\left[\mathbf{M}_k, \boldsymbol{\Lambda}^2_k\right],\ k = 1, \ldots, N-2 \\ \boldsymbol{\tau}_k &= h\left(\mathrm{Ad}_{\mathbf{g}_k}\boldsymbol{\Lambda}^2_k\right)^\flat,\ k = 1, \ldots, N-1 \\ \mathbf{M}_k &= \mathbf{J}\left(\boldsymbol{\Omega}_k\right),\ k = 0, \ldots, N-1, \end{aligned} \tag{26}$$

and the boundary conditions

$$\begin{aligned} \mathbf{R}_0 &= \mathbf{R}^*_0,\ \mathbf{R}_1 = \mathbf{R}^*_0\mathbf{g}^*_0,\ \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}^*_0 \\ \mathbf{R}_N &= \mathbf{R}^*_N,\ \mathbf{R}_{N-1} = \mathbf{R}^*_N\left(\mathbf{g}^*_{N-1}\right)^{-1},\ \boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}^*_{N-1} \\ \boldsymbol{\tau}_0 &= \boldsymbol{\tau}_N = 0, \end{aligned}$$

where $\mathbf{g}^*_0 = \exp(h\boldsymbol{\Omega}^*_0)$ and $\mathbf{g}^*_{N-1} = \exp\left(h\boldsymbol{\Omega}^*_{N-1}\right)$.

IV. Numerical Approach and Results

The first-order optimality equations, equation (25), in combination with the boundary conditions, $\mathbf{R}_0 = \mathbf{R}^*_0$, $\mathbf{R}_N = \mathbf{R}^*_N$, $\boldsymbol{\Omega}_0 = \boldsymbol{\Omega}^*_0$, and $\boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}^*_{N-1}$, leave the torques $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_{N-1}$, and the angular velocities $\boldsymbol{\Omega}_1, \ldots, \boldsymbol{\Omega}_{N-2}$ as unknowns. By substituting the relations $\mathbf{g}_k = \exp(h\boldsymbol{\Omega}_k)$, $\mathbf{M}_k = \mathbf{J}(\boldsymbol{\Omega}_k)$, we can rewrite the necessary conditions (25) as follows,

$$\begin{aligned} 0 = -\frac{1}{h^2}\Big(&\mathbf{J}(\boldsymbol{\tau}_k^\sharp) - \mathrm{Ad}^*_{\exp(-h\boldsymbol{\Omega}_k)}\mathbf{J}(\boldsymbol{\tau}_{k+1}^\sharp) \\ &- \mathbf{J}(\mathrm{Ad}_{\exp(-h\boldsymbol{\Omega}_{k-1})}\boldsymbol{\tau}_{k-1}^\sharp \\ &+ \mathrm{Ad}^*_{\exp(-h\boldsymbol{\Omega}_k)}\mathbf{J}(\mathrm{Ad}_{\exp(-h\boldsymbol{\Omega}_k)}\boldsymbol{\tau}_k^\sharp)\Big) \\ -\frac{1}{h}\Big(&\mathrm{Ad}^*_{\exp(-h\boldsymbol{\Omega}_k)}\left[\mathbf{J}(\boldsymbol{\Omega}_k), \mathrm{Ad}_{\exp(-h\boldsymbol{\Omega}_k)}(\boldsymbol{\tau}_k^\sharp)\right] \\ &- \frac{1}{h}\left[\mathbf{J}(\boldsymbol{\Omega}_{k-1}), \mathrm{Ad}_{\exp(-h\boldsymbol{\Omega}_{k-1})}(\boldsymbol{\tau}_{k-1}^\sharp)\right]\Big), \end{aligned}$$

where $k = 2, \ldots, N-2$, and the discrete evolution equations, given by line 2 of (26), can be written as

$$0 = \mathbf{J}(\boldsymbol{\Omega}_k) - \mathrm{Ad}^*_{\exp(h\boldsymbol{\Omega}_k)}(h\boldsymbol{\tau}_k + \mathbf{J}(\boldsymbol{\Omega}_{k-1})),$$

where $k = 1, \ldots, N-1$. In addition, we use the boundary conditions on $\mathbf{R}_0$ and $\mathbf{R}_N$, together with the update step given by line 1 of (26) to give the last constraint,

$$0 = \log\left(\mathbf{R}_N^{-1}\mathbf{R}_0\exp(h\boldsymbol{\Omega}_0)\ldots\exp(h\boldsymbol{\Omega}_{N-1})\right),$$

where log is the logarithm map on $\mathsf{SO}(3)$.

At this point it should be noted that one important advantage of the manner in which we have discretized the optimal control problem is that it is $\mathsf{SO}(3)$-equivariant. This is to say that if we rotated all the boundary conditions by a fixed rotation matrix, and solved the resulting discrete optimal control problem, the solution we would obtain would simply be the rotation of the solution of the original problem. This can be seen quite clearly from the fact that the discrete problem is expressed in terms of body coordinates, both in terms of body angular velocities and body forces. In addition, the initial and final attitudes $\mathbf{R}_0$ and $\mathbf{R}_N$ only enter in the last equation as a relative rotation.

The $\mathsf{SO}(3)$-equivariance of our numerical method is desirable, since it ensures that our results are independent of the choice of coordinate frames. This is in contrast to methods based on coordinatizing the rotation group using quaternions and Euler angles.

The equations above take values in $\mathfrak{so}(3)$. Consider the Lie algebra isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)$ given by,

$$\mathbf{v} = (v_1, v_2, v_3) \mapsto \hat{\mathbf{v}} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix},$$

which maps 3-vectors to 3×3 skew-symmetric matrices. In particular, we have the following identities,

$$[\hat{\mathbf{u}}, \hat{\mathbf{v}}] = (\mathbf{u}\times\mathbf{v})\hat{}, \quad \mathrm{Ad}_{\mathbf{A}}\hat{\mathbf{v}} = (\mathbf{A}\mathbf{v})\hat{}.$$

Furthermore, we identify $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ by the usual dot product, that is to say if $\boldsymbol{\Pi}$, $\mathbf{v} \in \mathbb{R}^3$, then $\langle\boldsymbol{\Pi}, \hat{\mathbf{v}}\rangle = \boldsymbol{\Pi}\cdot\mathbf{v}$. With this identification, we have that $\mathrm{Ad}^*_{\mathbf{A}^{-1}}\boldsymbol{\Pi} = \mathbf{A}\boldsymbol{\Pi}$. Using the identities above, we write the necessary conditions using matrix-vector products and cross products. Then, each of the equations can be interpreted as 3-vector valued functions, and the system of equations can be considered as a $3(2N-3)$-vector valued function, which is precisely the dimensionality of the unknowns. This reduces the discrete optimal problem to a nonlinear root finding problem.

We used a Newton-Armijo method, a line search algorithm that uses the Newton search direction, and backtracking to ensure sufficient descent of the residual error. The Jacobian is constructed column by column, where the k-th column is computed using the following approximation,

$$\frac{\partial\mathbf{F}}{\partial x_k}(\mathbf{x}) = \frac{1}{\epsilon}\,\mathrm{Im}[\mathbf{F}(\mathbf{x} + i\epsilon\mathbf{e}_k)],$$

where $i = \sqrt{-1}$, $\mathbf{e}_k$ is a basis vector in the x_k direction, and ϵ is of the order of machine epsilon. This method is more accurate than a finite-difference approximation as it does not suffer from round-off errors.

In our numerical simulation, we computed an optimal trajectory for a rest-to-rest maneuver, as illustrated in Figure (1). In this simulation $N = 128$, and essentially identical results were obtained for $N = 64$. It is worth noting that the results are not rotationally symmetric about the midpoint of the simulation interval, which is due to the fact that our choice of update, $\mathbf{R}_{k+1} = \mathbf{R}_k\exp(h\boldsymbol{\Omega}_k)$, does not exhibit time-reversal symmetry. In a forthcoming publication, we will introduce a reversible algorithm to address this issue.

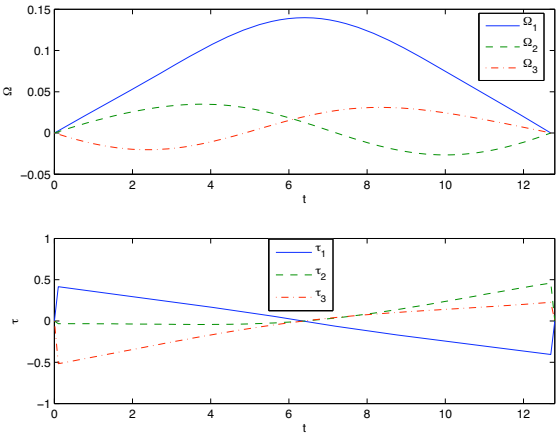


Fig. 1. Discrete optimal rest-to-rest maneuver in $\mathsf{SO}(3)$.

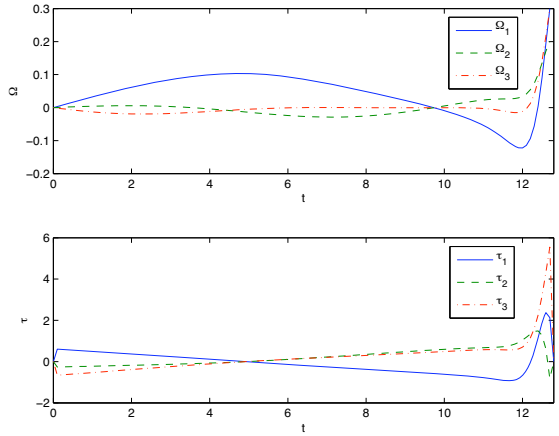


Fig. 2. Discrete optimal slew-up maneuver in $\mathsf{SO}(3)$.

We also present results for an optimal slew-up maneuver in $T = 12.8$ seconds from zero initial angular velocity to a final angular velocity of $\Omega_T = [0.3\ 0.2\ 0.3]^{\mathrm{T}}$, illustrated in Figure (2). The resolution is $N = 128$, and essentially identical results were obtained for $N = 64$.

V. Conclusion

In this paper we studied the continuous- and discrete-time optimal control problem for the rigid body, where the cost to be minimized is the external torque applied to move the rigid body from an initial condition to some pre-specified terminal condition. In the discrete setting, we use the discrete Lagrange–d'Alembert principle to obtain the discrete equations of motion. The kinematics were discretized to guarantee that the flow in phase space remains on the Lie group $\mathsf{SO}(3)$ and its algebra $\mathfrak{so}(3)$. We described how the necessary conditions can be solved for the general three-dimensional case and gave a numerical example for a three-dimensional rigid body maneuver.

Currently, we are investigating the connections with Pontryagin's maximum principle in continuous- and discrete-time. Additionally, we wish to generalize the result to general Lie groups with applications other than rigid body motion on $\mathsf{SO}(3)$. In particular, we are interested in controlling the motion of a rigid body *in space*, which corresponds to motion on the non-compact Lie group $\mathsf{SE}(3)$.

Acknowledgments

The research of Melvin Leok was partially supported by NSF grant DMS-0504747 and a University of Michigan Rackham faculty grant. The research of Anthony Bloch was supported by NSF grants DMS-0305837, CMS-0408542, and DMS-0604307.

References

[1] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*. Berlin: Springer, 2002.

[2] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, ser. Cambridge Monographs on Applied and Computational Mathematics. Cambridge: Cambridge University Press, 2004, vol. 14.

[3] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, ser. Applied Mathematics and Mathematical Computation. London: Chapman and Hall, 1994, vol. 7.

[4] O. Junge, J. E. Marsden, and S. Ober-Blöbaum, "Discrete mechanics and optimal control," *IFAC Congress, Praha*, 2005.

[5] I. I. Hussein, D. J. Scheeres, and D. C. Hyland, "Interferometric observatories in Earth orbit," *Journal of Guidance, Control and Dynamics*, vol. 27, no. 2, pp. 297–301, 2004.

[6] "New horizons: Pluto-kuiper belt mission," Website, March 2006. [Online]. Available: http://www.nasa.gov/mission_pages/newhorizons/main/index.html

[7] J. P. Palao and R. Kosloff, "Quantum computing by an optimal control algorithm for unitary transformations," *Physical Review Letters*, vol. 89, p. 188301, 2002.

[8] N. Khaneja, S. J. Glaser, and R. W. Brockett, "Sub-Riemannian geometry and optimal control of three spin systems," *Physical Review A*, vol. 65, p. 032301, 2002.

[9] S. L. Scrivener and R. C. Thompson, "Survey of time-optimal attitude maneuvers," *Journal of Guidance, Control, and Dynamics*, vol. 17, no. 2, pp. 225–233, 1994.

[10] H. Schaub, J. L. Junkins, and R. D. Robinett, "New attitude penalty functions for spacecraft optimal control problems," *AIAA Guidance, Navigation, and Control Conference*, 1996.

[11] J. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[12] C. Kane, J. E. Marsden, M. Ortiz, and M. West, "Variational integrators and the newmark algorithm for conservative and dissipative mechanical systems," *International Journal of Numerical Methods in Engineering*, vol. 49, no. 10, pp. 1295–1325, 2000.

[13] A. Iserles, H. Munthe-Kaas, S. P. N. rsett, and A. Zanna, "Lie group methods," *Acta Numerica*, vol. 9, pp. 215–265, 2000.

[14] M. Leok, "Generalized galerkin variational integrators," 2004, preprint, arXiv:math.NA/0508360.

[15] T. Lee, M. Leok, and N. H. McClamroch, "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum," *Proceedings of the IEEE Conference on Control Applications*, pp. 962–967, 2005.

[16] ——, "Lie group variational integrators for the full body problem," *Computer Methods in Applied Mechanics and Engineering*, 2005, submitted, arXiv:math.NA/0508365.

[17] A. Bloch, J. Baillieul, P. E. Crouch, and J. E. Marsden, *Nonholonomic Mechanics and Control*. New York, NY: Springer-Verlag, 2003.

[18] I. I. Hussein and A. M. Bloch, "Optimal trajectory tracking on the group of rigid body motions," *2005 IEEE Conference on Decision and Control*, 2005.

[19] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*. New York, NY: Springer-Verlag, 1999.

[20] J. Milnor, *Morse Theory*. Princeton, NJ: Princeton University Press, 1963.

Proceedings of the 2007 American Control Conference
Marriott Marquis Hotel at Times Square
New York City, USA, July 11-13, 2007

Global Attitude Estimation using Single Direction Measurements

Taeyoung Lee*†, Melvin Leok*, N. Harris McClamroch†, and Amit Sanyal

Abstract— **A deterministic attitude estimator for a rigid body under an attitude dependent potential is studied using small error assumptions. This estimator requires only a single direction measurement to a known reference point at each measurement instant. The measurement cannot completely determine the attitude, but an attitude estimation scheme based on this measurement is developed; a feasible set compatible with the measurement is described and it is combined with an attitude dynamics model to obtain an attitude estimate. The attitude is globally represented by a rotation matrix, and the uncertainties are described by ellipsoidal sets. A numerical example for a spacecraft in a circular orbit is presented.**

I. Introduction

The attitude of a rigid body is defined by the orientation of a body-fixed frame with respect to a reference frame, and the attitude is represented by a rotation matrix that is a 3×3 orthogonal matrix with determinant of 1, which transforms a representation of a vector in a body-fixed frame into one represented in the reference frame. Rotation matrices have a group structure denoted by $\mathrm{SO}(3)$. In spacecraft applications, the attitude is usually determined by using a set of direction measurements. The directions to objects such as the sun, stars, and geomagnetic fields, assumed to be known in the reference frame, are measured in the body-fixed frame in order to determine the rotation matrix.

Attitude determination using multiple direction measurements with least squares estimation is known as Wahba's problem [1]. The original solution of Wahba's problem is given in [2], and solutions are expressed in terms of quaternions [3], and in terms of a rotation matrix [4]. Based on these attitude determination schemes, attitude estimation problems are studied in [5], [6] and [7]. The attitude determination/estimation procedures using Wahba's problem formulation require at least two different direction measurements at each measurement instant.

A single direction measurement provides some information about the attitude; the rotation matrix is guaranteed to lie on a one-dimensional subgroup of the three-dimensional special orthogonal group $\mathrm{SO}(3)$, which is diffeomorphic to the one-sphere $\mathbb{S}^1$. The attitude is not completely determined at a single measurement instant. If the process is coupled with an attitude dynamics model, an attitude estimation scheme can be developed using single direction measurements.

Most existing attitude estimation schemes use generalized coordinate representations of the attitude. As is well known, minimal coordinate representations of the rotation group, such as Euler angles, lead to singularities. Non-minimal coordinate representations, like the quaternions, have their own associated problems. Besides the extra constraint of unit norm that one needs to impose on the quaternion, it has an inevitable ambiguity in expressing the attitude.

A stochastic state estimator requires probabilistic models for the state uncertainty and the noise. However, statistical properties of the uncertainty and the noise are often not available. An alternative deterministic approach is to specify bounds on the uncertainty and the measurement noise without assumptions on their distribution. Noise bounds are available in many cases, and deterministic estimation is robust to the noise distribution [8]. An efficient but flexible way to describe the bounds is using ellipsoidal sets, referred to as uncertainty ellipsoids. The deterministic estimation process is based on set theoretic results developed in [9]; optimal deterministic estimation problems using the uncertainty ellipsoids are studied in [10] and [11].

In this paper, a deterministic attitude estimator which requires a single direction measurement at each measurement instant is presented. A feasible set in $\mathrm{SO}(3)$ that is compatible with the measurement is represented by Lie algebra elements and the exponential map. It is compared with the attitude dynamics model to obtain an updated estimate. The estimation scheme presented in this paper has the following distinctive features: the estimator requires only a single direction measurement at each measurement instant, the attitude is represented by a rotation matrix without any local parameterization, the dynamics of the rigid body is considered, and the deterministic estimator is distinguished from a Kalman or extended Kalman filter.

This paper is organized as follows. The attitude dynamics and uncertainty model are given in Section II. The attitude determination scheme and the attitude estimation scheme using single direction measurements are presented in Section III and IV, which is followed by a numerical example in Section V.

II. Attitude dynamics and Uncertainty model

A. Equations of motion

We consider estimation of the attitude dynamics of a rigid body in the presence of an attitude dependent potential,

Taeyoung Lee, Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 tylee@umich.edu

Melvin Leok, Mathematics, Purdue University, West Lafayette, IN 47907 mleok@math.purdue.edu

N. Harris McClamroch, Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 nhm@umich.edu

Amit Sanyal, Mechanical Engineering, University of Hawaii at Manoa, Honolulu, HI 96822 aksanyal@hawaii.edu

*This research has been supported in part by NSF under grant DMS-0504747, and by a grant from the Rackham Graduate School, University of Michigan.

†This research has been supported in part by NSF under grant ECS-0244977 and CMS-0555797.

$U(\cdot) : \mathrm{SO}(3) \mapsto \mathbb{R}$. Systems that can be so modeled include a free rigid body, spacecraft on a circular orbit with gravity gradient effects [12], or a 3D pendulum [13]. The continuous equations of motion are

$$J\dot{\Omega} + \Omega \times J\Omega = M, \tag{1}$$

$$\dot{R} = RS(\Omega), \tag{2}$$

where $J \in \mathbb{R}^{3\times3}$ is the moment of inertia matrix of the rigid body, $\Omega \in \mathbb{R}^3$ is the angular velocity of the body expressed in the body-fixed frame, $R \in \mathrm{SO}(3)$ is the rotation matrix, and $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is an isomorphism between $\mathfrak{so}(3)$ and $\mathbb{R}^3$ defined by the condition $S(x)y = x \times y$ for all $x, y \in \mathbb{R}^3$. The vector $M \in \mathbb{R}^3$ is the moment due to the potential, determined by $S(M) = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}$, or more explicitly,

$$M = r_1 \times v_{r_1} + r_2 \times v_{r_2} + r_3 \times v_{r_3}, \tag{3}$$

where $r_i, v_{r_i} \in \mathbb{R}^{1\times3}$ are the i-th row vectors of R and $\frac{\partial U}{\partial R}$, respectively. Here $\frac{\partial U}{\partial R} \in \mathbb{R}^{3\times3}$ is defined by $[\frac{\partial U}{\partial R}]_{ij} = \frac{\partial U}{\partial [R]_{ij}}$, where $[A]_{ij}$ denotes the i, j-th element of a matrix A.

General numerical integration methods like the popular Runge-Kutta schemes, typically preserve neither first integrals nor the characteristics of the configuration space, $\mathrm{SO}(3)$. In particular, the orthogonal structure of the rotation matrices is not preserved numerically. It is often proposed to parameterize (2) by Euler angles or quaternions instead of integrating (2) directly. However, Euler angles yield only local representations of the attitude and they have singularities. Unit quaternions do not exhibit singularities, but they have the manifold structure of the three-sphere $\mathbb{S}^3$, and double cover $\mathrm{SO}(3)$. Consequently, the unit quaternion representing the attitude is inevitably ambiguous. In addition, general numerical integration methods do not preserve the unit length constraint. Therefore, quaternions exhibit the same numerical drift problem as rotation matrices.

Lie group variational integrators preserve the group structure without the use of local charts, reprojection, or constraints, they are symplectic and momentum preserving, and they exhibit good energy behavior for an exponentially long time period. The following Lie group variational integrator for the attitude dynamics of a rigid body is presented in [13]:

$$hS(J\Omega_k + \frac{h}{2}M_k) = F_k J_d - J_d F_k^T, \tag{4}$$

$$R_{k+1} = R_k F_k, \tag{5}$$

$$J\Omega_{k+1} = F_k^T J\Omega_k + \frac{h}{2}F_k^T M_k + \frac{h}{2}M_{k+1}, \tag{6}$$

where $J_d \in \mathbb{R}^{3\times3}$ is a nonstandard moment of inertia matrix defined by $J_d = \frac{1}{2}\mathrm{tr}[J]\, I_{3\times3} - J$, and $F_k \in \mathrm{SO}(3)$ is the relative attitude between integration steps. The constant $h \in \mathbb{R}$ is the integration step size, and the subscript k denotes the k-th integration step. This integrator yields a map $(R_k, \Omega_k) \mapsto (R_{k+1}, \Omega_{k+1})$ by solving (4) to obtain $F_k \in \mathrm{SO}(3)$ and substituting it into (5) and (6) to obtain R_{k+1} and Ω_{k+1}. We use these discrete equations of motion to propagate the attitude dynamics between measurements during the estimation process.

B. Uncertainty Ellipsoid

We describe uncertainties of the attitude dynamics by using ellipsoidal sets referred to as uncertainty ellipsoids. An uncertainty ellipsoid in $\mathbb{R}^n$ is defined as

$$\mathcal{E}_{\mathbb{R}^n}(\hat{x}, P) = \left\{ x \in \mathbb{R}^n \,\middle|\, (x - \hat{x})^T P^{-1} (x - \hat{x}) \leq 1 \right\}, \tag{7}$$

where $\hat{x} \in \mathbb{R}^n$, and $P \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. We call $\hat{x}$ the center of the uncertainty ellipsoid, and P is the uncertainty matrix that determines the size and the shape of the uncertainty ellipsoid. The size of an uncertainty ellipsoid is measured by $\mathrm{tr}[P]$ which is the sum of the squares of the semi-principal axes of the ellipsoid. We identify $\mathrm{TSO}(3)$ with $\mathrm{SO}(3)\times\mathfrak{so}(3)$ by left-trivialization, and we identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ by the isomorphism $S(\cdot)$. The uncertainty ellipsoid centered at $(\hat{R}, \hat{\Omega}) \in \mathrm{TSO}(3)$ is induced from an uncertainty ellipsoid in $\mathbb{R}^6$;

$$\mathcal{E}(\hat{R}, \hat{\Omega}, P) = \left\{ R \in \mathrm{SO}(3),\, \Omega \in \mathbb{R}^3 \,\middle|\, \begin{bmatrix} \zeta \\ \delta\Omega \end{bmatrix} \in \mathcal{E}_{\mathbb{R}^6}(0_6, P) \right\},$$

where $S(\zeta) = \mathrm{logm}\left(\hat{R}^T R\right) \in \mathfrak{so}(3)$, $\delta\Omega = \Omega - \hat{\Omega} \in \mathbb{R}^3$, and $P \in \mathbb{R}^{6\times6}$ is a symmetric positive definite matrix. An element $(R, \Omega) \in \mathcal{E}(\hat{R}, \hat{\Omega}, P)$ can be written as

$$R = \hat{R} \exp S(\zeta), \quad \Omega = \hat{\Omega} + \delta\Omega,$$

for some $x = [\zeta;\, \delta\Omega] \in \mathbb{R}^6$ satisfying $x^T P^{-1} x \leq 1$.

We assume that the initial conditions are bounded by a prescribed uncertainty ellipsoid

$$(R_0, \Omega_0) \in \mathcal{E}(\hat{R}_0, \hat{\Omega}_0, P_0), \tag{8}$$

where $P_0 \in \mathbb{R}^{6\times6}$ is a symmetric positive definite matrix.

III. Attitude Determination with a Single Direction Measurement

In the attitude determination problem, we measure directions to points in the reference frame. We assume that the directions to these points are known in the reference frame. This either requires that the points are located far away from the spacecraft or the relative location of the spacecraft is known exactly. The directional sensor is fixed in the body-fixed frame, and the measurements are representations of the direction vectors in the body-fixed frame. The representations in the body-fixed frame are transformed into those in the reference frame by multiplication with the rotation matrix that defines the attitude of the rigid body.

A. Exact measurement

Let the direction to a known point in the reference frame be $e \in \mathbb{S}^2$, and let the corresponding vector represented in the body-fixed frame be $b \in \mathbb{S}^2$. We first assume that the direction measurement has no error, so the direction b is exact. Since we only measure a direction to a point, we normalize e and b so that they have unit lengths. The vectors e and b are different representations of the same vector from the spacecraft to the known point, and they are related by the rotation matrix $R \in \mathrm{SO}(3)$ as

$$e = Rb. \tag{9}$$

This equation provides a two-dimensional constraint on the three-dimensional rotation matrix. Consequently, a single direction measurement does not completely determine the attitude. This corresponds to the fact that if we rotate the rigid body about the direction e in the reference frame, then the measured direction b is not changed. The rotation matrix has one-dimensional uncertainty represented by any rotation about the direction e in the reference frame, or equivalently, any rotation about the direction b in the body-fixed frame.

Suppose that $R^\circ \in \mathrm{SO}(3)$ is a particular rotation matrix satisfying (9). This rotation matrix can be represented in several ways. For example, if b and e are not co-linear,

$$R^\circ = \exp\left[\cos^{-1}(b^T e)\ S\left(\frac{b\times e}{\|b\times e\|}\right)\right] \exp S(\theta^\circ b), \quad (10)$$

where the constant $\theta^\circ \in \mathbb{S}^1$ can be arbitrarily chosen. The rotation matrix that represents the attitude of the rigid body can be written in terms of R° as

$$R = R^\circ \exp\left[\theta S(b)\right] \quad (11)$$

for a $\theta \in \mathbb{S}^1$.

In summary, if the single direction to a known point is measured exactly, the rotation matrix lies in the following one-dimensional subgroup of $\mathrm{SO}(3)$:

$$R \in \left\{ R^\circ \exp\left[\theta S(b)\right] \middle| \theta \in \mathbb{S}^1 \right\}. \quad (12)$$

B. Measurement error

We now consider the effects of small measurement errors. Let $\tilde{b} \in \mathbb{S}^2$ be the measured direction of the direction b. Since we only measure directions, we normalize b and $\tilde{b}$ so that they have unit lengths. It is inappropriate to express the measurement error by a vector difference. The measurement error is modeled by rotation of the measured direction

$$b = \exp\left[S(\nu)\right]\tilde{b} \simeq \tilde{b} + S(\nu)\tilde{b}, \quad (13)$$

where $\nu \in \mathbb{R}^3$ is the Euler axis of rotation from $\tilde{b}$ to b, and $\|\nu\|$ is the corresponding rotation angle error in radians. The measurement error is bounded by an uncertainty ellipsoid

$$\nu \in \mathcal{E}_{\mathbb{R}^3}(0_3, S) \quad (14)$$

for a symmetric positive-definite matrix $S \in \mathbb{R}^{3\times 3}$. The magnitude of the measurement error is assumed to be small.

Let $\tilde{R}^\circ \in \mathrm{SO}(3)$ be a rotation matrix obtained by (10) for the measured direction $\tilde{b}$. We express the difference between R° and $\tilde{R}^\circ$ using the exponential map:

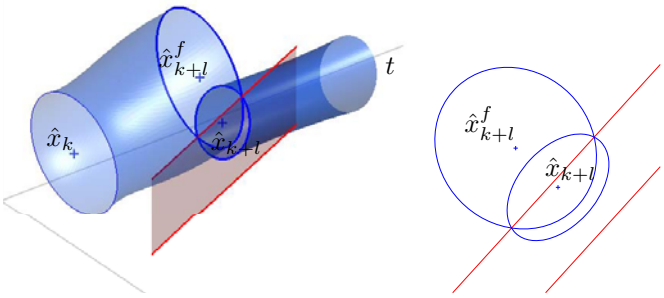
$$R^\circ = \tilde{R}^\circ \exp\left[S(\zeta^\circ)\right] \quad (15)$$

for some $\zeta^\circ \in \mathbb{R}^3$. Since we assume that the measurement error is small, the norm of the vector ζ° is much smaller than π, i.e. $\|\zeta^\circ\| \ll \pi$. Since $e = R^\circ b = \tilde{R}^\circ \tilde{b}$, we obtain

$$\begin{aligned}\tilde{R}^\circ \tilde{b} = R^\circ b &= \tilde{R}^\circ \exp\left[S(\zeta^\circ)\right]\left\{I_{3\times 3} + S(\nu)\right\}\tilde{b},\\ &\simeq \tilde{R}^\circ \left\{I_{3\times 3} + S(\zeta^\circ + \nu)\right\}\tilde{b},\end{aligned}$$

Thus we have $S(\zeta^\circ + \nu)\tilde{b} = 0$, which is equivalent to

$$\zeta^\circ = c\tilde{b} - \nu \quad (16)$$



(a) Propagation of uncertainty ellipsoid (b) Filtering procedure

Fig. 1. Uncertainty ellipsoids

for any constant $c \in \mathbb{R}$. Since $\|\zeta^\circ\| \le |c| + \|\nu\| \ll \pi$, the constant c is much smaller than π, i.e. $|c| \ll \pi$.

Substituting (13), (15), and (16) into (11), we obtain

$$R = \tilde{R}^\circ \exp\left[S(c\tilde{b} - \nu)\right] \exp\left[\theta S\left((I_{3\times 3} + S(\nu))\tilde{b}\right)\right] \quad (17)$$

for constants c and $\theta \in \mathbb{S}^1$.

In summary, if the single direction measurement has a small error represented by (13), then the attitude of the rigid body can be written in terms of the measured direction $\tilde{b}$ and the measurement error ν as (17). This expression includes the uncertainty caused by the measurement error as well as the uncertainty due to the single direction measurement assumption. The constant $\theta^\circ \in \mathbb{S}^1$ to determine R° and $\tilde{R}^\circ$ is specified by the following estimation procedure.

IV. Attitude Estimation with Single Direction Measurements

A. Deterministic estimation

We use deterministic bounded estimation using ellipsoidal sets, referred to as uncertainty ellipsoids, to describe the uncertainty and measurement noise. The estimation process has three steps similar to those in the Kalman filter: prediction, measurement, and filtering steps. We assume that the initial condition lies in a prescribed uncertainty ellipsoid, which is propagated in time using the equations of motion. This defines a prediction step or a flow update. The measurement error bound is described by a measurement uncertainty ellipsoid. Then we can guarantee that the state lies in the intersection of the predicted uncertainty ellipsoid and the measured uncertainty ellipsoid. The intersection of the two ellipsoids is an irregular shape, which is not efficient to compute and store. Instead we find a minimal ellipsoid that contains this intersection. This procedure is repeated whenever new measurements are available.

This deterministic estimation procedure is illustrated in Fig. 1. The left figure shows time evolution of an uncertainty ellipsoid, and the right figure shows a cross section at a fixed measurement instant. At the k-th time step, the state is bounded by an uncertainty ellipsoid centered at $\hat{x}_k$. This initial ellipsoid is propagated through time. Suppose that the state is measured next at the $(k+l)$-th time step. In this single direction measurement estimation, the measurement

ellipsoid degenerates to a strip. In the estimation process, we find a new ellipsoid that contains the intersection, as shown in the right figure. The center of the new ellipsoid, $\hat{x}_{k+l}$ gives a point estimate of the state at time step $(k+l)$, and the magnitude of the new uncertainty ellipsoid measures the estimation accuracy. The deterministic estimates are optimal in the sense that the sizes of the ellipsoids are minimized.

Now we describe the detailed procedures for each step. The superscript f denotes the variables related to the flow update, and the superscript m denotes the variables related to the measurement update. $\tilde{\cdot}$ denotes a measured variable, and $\hat{\cdot}$ denotes an estimated variable.

B. Flow update

Suppose that the attitude and the angular momentum at the k-th step lie in a given uncertainty ellipsoid:

$$(R_k, \Omega_k) \in \mathcal{E}(\hat{R}_k, \hat{\Omega}_k, P_k),$$

and a new measurement is taken at the $(k+l)$-th time step.

The flow update finds the center and the uncertainty matrix that define the uncertainty ellipsoid at the $(k+l)$-th step using the given uncertainty ellipsoid at the k-th step. Since the attitude dynamics of a rigid body is nonlinear, the admissible boundary of the state at the $(k+l)$-th step is not an ellipsoid in general. We assume that the given uncertainty ellipsoid at the k-th step is sufficiently small that attitudes and angular velocities in the uncertainty ellipsoids can be approximated using the linearized equations of motion. Then we can guarantee that the uncertainty set at the $(k+l)$-th step is an ellipsoid, and we can compute its center and its uncertainty matrix at the $(k+l)$-th step separately.

Center: For the given center at step k, $(\hat{R}_k, \hat{\Omega}_k)$, the center of the uncertainty ellipsoid at step $(k+l)$ is $(\hat{R}^f_{k+l}, \hat{\Omega}^f_{k+l})$ given by the discrete equations of motion, (4), (5), and (6):

$$hS(J\hat{\Omega}_k + \frac{h}{2}\hat{M}_k) = \hat{F}_k J_d - J_d \hat{F}_k^T, \tag{18}$$

$$\hat{R}^f_{k+1} = \hat{R}_k \hat{F}_k, \tag{19}$$

$$J\hat{\Omega}^f_{k+1} = \hat{F}_k{}^T \hat{\Omega}_k + \frac{h}{2}\hat{F}_k{}^T \hat{M}_k + \frac{h}{2}\hat{M}_{k+1}. \tag{20}$$

This integrator yields a map $(\hat{R}_k, \hat{\Omega}_k) \mapsto (\hat{R}^f_{k+1}, \hat{\Omega}^f_{k+1})$, and this process is repeatedly applied to find the center at the $(k+l)$-th step, $(\hat{R}^f_{k+l}, \hat{\Omega}^f_{k+l})$.

Uncertainty matrix: We assume that an uncertainty ellipsoid contains small perturbations from the center of the uncertainty ellipsoid. The uncertainty matrix is propagated by using the linearized equation of motion. At the $(k+1)$-th step, the uncertainty ellipsoid is represented by perturbations from the center $(\hat{R}^f_{k+1}, \hat{\Omega}^f_{k+1})$ as

$$R_{k+1} = \hat{R}^f_{k+1} \exp S(\zeta^f_{k+1}),$$
$$\Omega_{k+1} = \hat{\Omega}^f_{k+1} + \delta\Omega^f_{k+1},$$

for some $\zeta^f_{k+1}, \delta\Omega^f_{k+1} \in \mathbb{R}^3$. The uncertainty matrix at the $(k+1)$-th step is obtained by finding a bound on $\zeta^f_{k+1}, \delta\Omega^f_{k+1} \in \mathbb{R}^3$. By assuming that the uncertainty ellipsoid at the k-th step is sufficiently small, $\zeta^f_{k+1}, \delta\Omega^f_{k+1}$ can be obtained from the following linear equations using the results presented in [12]

$$x^f_{k+1} = A^f_k x_k,$$

where $x_k = [\zeta_k; \delta\Omega_k] \in \mathbb{R}^6$, and $A^f_k \in \mathbb{R}^{6\times 6}$ can be suitably defined. Since $(R_k, \Omega_k) \in \mathcal{E}(\hat{R}_k, \hat{\Omega}_k, P_k)$, $x_k \in \mathcal{E}_{\mathbb{R}^6}(0, P_k)$ by the definition of the uncertainty ellipsoid. This implies that $A^f_k x_k$ lies in the following uncertainty ellipsoid

$$A^f_k x_k \in \mathcal{E}_{\mathbb{R}^6}\left(0, A^f_k P_k \left(A^f_k\right)^T\right).$$

The uncertainty matrix at the $(k+1)$-th step is given by

$$P^f_{k+1} = A^f_k P_k \left(A^f_k\right)^T. \tag{21}$$

The above equation is then applied repeatedly to find the uncertainty matrix at the $(k+l)$-th step.

In summary, the uncertainty ellipsoid at the $(k+l)$-th step is computed using (18), (19), (20), and (21) as:

$$(R_{k+l}, \Omega_{k+l}) \in \mathcal{E}(\hat{R}^f_{k+l}, \hat{\Omega}^f_{k+l}, P^f_{k+l}), \tag{22}$$

C. Measurement update

The measurement update finds an uncertainty ellipsoid in the state space using the measurement and the measurement error models described in Section III. Here, we find the measurement uncertainty ellipsoid that contains the feasible set of rotation matrices that are compatible with the single direction measurement as described by (17).

Elements in the measurement uncertainty ellipsoid are expressed as

$$R_{k+1} = \hat{R}^m_{k+l} \exp S(\zeta^m_{k+l}) \tag{23}$$

for the center $\hat{R}^m_{k+l} \in \mathrm{SO}(3)$ and some $\zeta^m_{k+l} \in \mathbb{R}^3$. We omit the subscript $(k+l)$ hereafter for convenience, and it is assumed that the direction is measured at the $(k+l)$-th step.

Center: Comparing (17) and (23), we choose the center of the measurement uncertainty ellipsoid as

$$\begin{aligned}\hat{R}^m &= \tilde{R}^\circ, \\ &= \exp\left[\cos^{-1}(\tilde{b}^T e)\; S\left(\frac{\tilde{b}\times e}{\|\tilde{b}\times e\|}\right)\right] \exp S(\theta^\circ \tilde{b}),\end{aligned} \tag{24}$$

for the constant $\theta^\circ \in S^1$ which is determined by the following filtering procedure.

Uncertainty Matrix: From (17), (23), and (24), we have

$$\exp S(\zeta^m) \exp\Big[-S(c\tilde{b} - \nu)\Big] = \exp\Big[\theta S\left((I_{3\times 3} + S(\nu))\tilde{b}\right)\Big].$$

Since the vectors ζ^m and $\zeta^\circ = c\tilde{b} - \nu$ are assumed to be small, the above equation is approximated as

$$\zeta^m - (c\tilde{b} - \nu) = \theta\tilde{b} + \theta S(\nu)\tilde{b},$$

which can be rewritten as

$$\begin{aligned}\zeta^m - (\theta + c)\tilde{b} &= -\nu - \theta S(\tilde{b})\nu, \\ &= (\theta - 1)\nu - \theta(I_{3\times 3} + S(\tilde{b}))\nu.\end{aligned}$$

Since $\nu \in \mathcal{E}_{\mathbb{R}^3}(0_3, S)$ and $\theta \in \mathbb{S}^1$, the terms on the right-hand side satisfy

$$(\theta - 1)\nu \in \mathcal{E}_{\mathbb{R}^3}(0_3, (1+\pi)^2 S),$$
$$\theta(I_{3\times 3} + S(\tilde{b}))\nu \in \mathcal{E}_{\mathbb{R}^3}(0_3, \pi^2 \mathcal{A}^{m,T} S \mathcal{A}^{m,T}),$$

where $\mathcal{A}^m = I_{3\times 3} + S(\tilde{b}) \in \mathbb{R}^{3\times 3}$. Therefore, the vector $\zeta^m - (\theta + c)\tilde{b}$ lies in an ellipsoid containing the vector sum of the above two ellipsoids. The expressions for the minimal ellipsoid containing the vector sum of two ellipsoids are given in [10]. Using the results, we have

$$\zeta^m - (\theta + c)\tilde{b} \in \mathcal{E}_{\mathbb{R}^3}(0_3, P_0^m), \tag{25}$$

where

$$P_0^m = (1 + q^{-1})Q^1 + (1+q)Q^2, \quad q = \sqrt{\frac{\mathrm{tr}[Q^1]}{\mathrm{tr}[Q^2]}},$$
$$Q^1 = (1+\pi)^2 S, \quad Q^2 = \pi^2 \mathcal{A}^{m,T} S \mathcal{A}^m.$$

From (25), we can guarantee that the vector ζ^m lies in an ellipsoid containing the following union of the sets

$$\mathcal{E}_{\mathbb{R}^3}(-\pi\tilde{b}, P_0^m) \cup \mathcal{E}_{\mathbb{R}^3}(\pi\tilde{b}, P_0^m).$$

This is a consequence of the fact that an ellipsoid is convex and the assumption $|c| \ll \pi$. The ellipsoid that contains the union of two ellipsoids is obtained numerically by the LMI approach presented in [14].

$$\mathcal{E}_{\mathbb{R}^3}(0_3, P^m) \supset \left(\mathcal{E}_{\mathbb{R}^3}(-\pi\tilde{b}, P_0^m) \bigcup \mathcal{E}_{\mathbb{R}^3}(\pi\tilde{b}, P_0^m)\right). \tag{26}$$

In summary, a single direction measurement with small error guarantees that the rotation matrix is expressed as (23), where the center $\hat{R}^m$ is given by (24), and the vector ζ^m lies in the uncertainty ellipsoid given by (26).

D. Filtering procedure

The filtering procedure finds a new uncertainty ellipsoid compatible with both the predicted uncertainty ellipsoid and the measured uncertainty ellipsoid. The intersection of two ellipsoids is generally not an ellipsoid. We find a minimal uncertainty ellipsoid containing the intersection.

The predicted uncertainty ellipsoid is based on $\hat{R}^f$ and the measurement ellipsoid is based on $\hat{R}^m$. In the following development, we assume that the difference between the rotation matrices $\hat{R}^f$ and $\hat{R}^m$ is small. Here, we find a value of $\theta^\circ \in \mathbb{S}^1$ at (24) such that the difference is minimized. Define an index $\mathcal{J} = \mathrm{tr}[I_{3\times 3} - \hat{R}^{f,T}\hat{R}^m]$. A standard variational approach with the use of Rodriguez formula shows that the index is minimized when

$$\theta^\circ = -\tan^{-1}\frac{\mathrm{tr}[\hat{R}^{f,T}\tilde{R}^\Delta S(\tilde{b})]}{\mathrm{tr}[\hat{R}^{f,T}\tilde{R}^\Delta S(\tilde{b})^2]},$$
$$\mathrm{tr}[\hat{R}^{f,T}\tilde{R}^\Delta S(\tilde{b})]\sin\theta^\circ - \mathrm{tr}[\hat{R}^{f,T}\tilde{R}^\Delta S(\tilde{b})^2]\cos\theta^\circ > 0,$$

where $\tilde{R}^\Delta \in \mathrm{SO}(3)$ is the first exponential of (24). The first equation is obtained by the optimality condition $\frac{\partial \mathcal{J}}{\partial \theta^\circ} = 0$, and the second inequality is obtained by $\frac{\partial^2 \mathcal{J}}{\partial (\theta^\circ)^2} > 0$. These conditions define the value of $\theta^0 \in \mathbb{S}^1$ uniquely.

We find a minimal ellipsoid containing the intersection of the predicted uncertainty ellipsoid and the measurement uncertainty ellipsoid. An element in the predicted uncertainty ellipsoid, $(R^f, \Omega^f) \in \mathcal{E}(\hat{R}^f, \hat{\Omega}^f, P^f)$, can be written as

$$R^f = \hat{R}^f e^{S(\zeta^f)}, \tag{27}$$
$$\Omega^f = \hat{\Omega}^f + \delta\Omega^f, \tag{28}$$

for some $(\zeta^f, \delta\Omega^f) \in \mathcal{E}_{\mathbb{R}^6}(0_{6\times 1}, P^f)$. We find an equivalent expression based on the measurement ellipsoid center $\hat{R}^m$. Define $\hat{\zeta}^{mf} \in \mathbb{R}^3$ such that

$$\hat{R}^f = \hat{R}^m e^{S(\hat{\zeta}^{mf})}. \tag{29}$$

Thus, $\hat{\zeta}^{mf}$ represents the difference between the centers of the two ellipsoids. Substituting (29) into (27),

$$R^f = \hat{R}^m e^{S(\hat{\zeta}^{mf})} e^{S(\zeta^f)} \simeq \hat{R}^m e^{S(\hat{\zeta}^{mf} + \zeta^f)},$$

where we assumed that $\hat{\zeta}^{mf}, \zeta^f$ are sufficiently small to obtain the second equality. Thus, the uncertainty ellipsoid obtained by the flow update, $\mathcal{E}(\hat{R}^f, \hat{\Omega}^f, P^f)$ is identified by the center $(\hat{R}^m, \hat{\Omega}^f)$ and the uncertainty ellipsoid in $\mathbb{R}^6$.

$$(\zeta^{mf}, \delta\Omega^f) \in \mathcal{E}_{\mathbb{R}^6}(\hat{x}^{mf}, P^f),$$

where $\hat{x}^{mf} = [\hat{\zeta}^{mf}; 0_3]$, and $S(\zeta^{mf}) = \mathrm{logm}\left(\hat{R}^{m,T}R^f\right) \in \mathfrak{so}(3)$, $\delta\Omega^f = \Omega^f - \hat{\Omega}^f \in \mathbb{R}^3$.

We seek a minimal ellipsoid that contains the intersection of the following uncertainty ellipsoids.

$$\left(\mathcal{E}_{\mathbb{R}^3}(0_3, P^m) \bigcap \mathcal{E}_{\mathbb{R}^6}(\hat{x}^{mf}, P^f)\right) \subset \mathcal{E}_{\mathbb{R}^6}(\hat{x}, P), \tag{30}$$

where $\hat{x} = [\hat{\zeta}^T, \delta\hat{\Omega}^T]^T \in \mathbb{R}^6$. The minimal ellipsoid containing the intersection of two ellipsoids is given in [10]. Using those results, $\hat{x}$ and P are given by

$$\hat{x} = (I_{6\times 6} - LH)\hat{x}^{mf},$$
$$P = \beta(r)[(I - LH)P^f(I - LH)^T + r^{-1}LP^mL^T],$$

where $L \in \mathbb{R}^{6\times 3}$, $H \in \mathbb{R}^{3\times 6}$, and $\beta(r) \in \mathbb{R}$ are given by

$$L = P^f H^T[HP^f H^T + r^{-1}P^m]^{-1},$$
$$H = [I_{3\times 3}, 0_{3\times 3}]^T,$$
$$\beta(r) = 1 + r - (\hat{x}^{mf})^T H^T[HP^f H^T + r^{-1}P^m]^{-1}H\hat{x}^{mf},$$

for a constant r, which is chosen such that $\mathrm{tr}[P]$ is minimized.

In summary, a new uncertainty ellipsoid at the $(k+l)$-th step is given by

$$(R_{k+1}, \Omega_{k+1}) \in \mathcal{E}(\hat{R}_{k+1}, \hat{\Omega}_{k+1}, P_{k+1}), \tag{31}$$

where

$$\hat{R}_{k+1} = \hat{R}^m_{k+1} e^{S(\hat{\zeta})}, \tag{32}$$
$$\hat{\Omega}_{k+1} = \hat{\Omega}^f_{k+1} + \delta\hat{\Omega}, \tag{33}$$
$$P_{k+1} = P. \tag{34}$$

The entire procedure is repeated with each new measurement. The steps outlined above define a dynamic filter.

E. Properties of the estimator

The notable feature of this attitude estimator is that it only requires a single direction measurement. Current attitude estimators based on the solution of Wahba's problem require at least two direction measurements at each instant. A single direction measurement provides only a two-dimensional constraint for the six-dimensional tangent bundle. The information obtained from the attitude dynamics is utilized, together with the measurement, in order to estimate the attitude and the angular velocity of the rigid body. In this paper, it is assumed that the angular velocity is not measured, but the current results can be readily extended to incorporate angular velocity measurements.

This attitude estimator has no singularities since the attitude is represented by a rotation matrix, and the geometric structure of the rotation matrix is preserved since it is updated by the structure-preserving Lie group variational integrator. It is also robust to the distribution of the measurement noise since we only use ellipsoidal bounds on the noise. The measurements need not be periodic, the estimation is repeated whenever new measurements become available.

V. Numerical Example

Numerical simulation results are given for the estimation of the attitude dynamics of an uncontrolled rigid spacecraft in a circular orbit about a large central body, including gravity gradient effects. The details of the on orbit spacecraft model are presented in [12].

The mass, length and time dimensions are normalized by the mass of the spacecraft, the maximum length of the spacecraft, and the orbital angular velocity, respectively. The moment of inertia of the spacecraft is chosen as $J = \mathrm{diag}\,[1,\, 2.8,\, 2]$. The maneuver is a large attitude change completed in a quarter of the orbit. The initial conditions are chosen as

$$R_0 = \begin{bmatrix} 0.707 & -0.707 & 0 \\ 0.707 & 0.707 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Omega_0 = [2.32,\, 0.45,\, -0.59],$$
$$\hat{R}_0 = I_{3\times 3}, \qquad\qquad \hat{\Omega}_0 = [2.12,\, 0.55,\, -0.89].$$

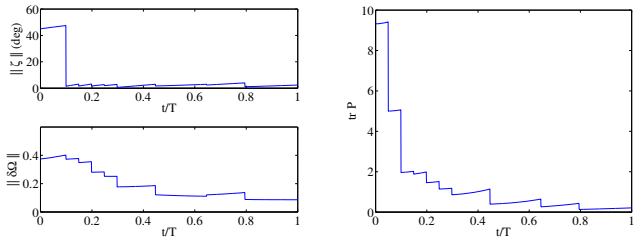
The corresponding initial estimation errors are $\|\zeta_0\| = 45\,\mathrm{deg}$, $\|\delta\Omega_0\| = 21.43\frac{\pi}{180}\,\mathrm{rad/s}$. Note that the actual initial attitude is opposite to the estimated initial attitude. The initial uncertainty matrix is given by

$$P_0 = \mathrm{diag}\,[2.28, 2.28, 2.28, 0.82, 0.82, 0.82]\,,$$

so that $x_0^T P_0^{-1} x_0 = 0.857 \leq 1$.

We assume that the measurements are available twenty times. Seven prescribed inertial directions e are available, and a simple adaptive scheme is developed to choose a best possible inertial direction as the spacecraft rotates. The uncertainty matrix for the measurement noise in (14) is given by $S_k = \left(0.2\frac{\pi}{180}\right)^2 I_{3\times 3}$, and the direction measurement noise is normally distributed.

Fig. 2 shows simulation results, where the left figure shows the attitude estimation error and the angular velocity estimation error, and the right figure shows the size of the uncertainty ellipsoid. The estimation errors and the size of the uncertainty are reduced rapidly after the first few measurements; the estimation error for the angular velocity converges relatively slowly since the angular velocity is not measured directly. The terminal attitude error, and the terminal angular velocity error are less than $2.3\,\mathrm{deg}$, and $0.08\,\mathrm{rad/s}$, respectively.



(a) Estimation error $\|\zeta_k\|$, $\|\delta\Omega_k\|$ (b) Size of uncertainty $\mathrm{tr}[P_k]$

Fig. 2. Estimation errors and uncertainty

References

[1] G. Wahba, "A least squares estimate of satellite attitude, Problem 65-1," *SIAM Review*, vol. 7, no. 5, p. 409, 1965.

[2] J. L. Farrell, J. C. Stuelpnagel, R. H. Wessner, J. R. Velman, and J. E. Brock, "A least squares estimate of satellite attitude, Solution 65-1," *SIAM Review*, vol. 8, no. 3, pp. 384–386, 1966.

[3] M. D. Shuster and S. D. Oh, "Three-axis attitude determination from vector observations," *Journal of Guidance Control and Dynamics*, vol. 4, no. 1, pp. 70–77, 1981.

[4] A. K. Sanyal, "Optimal attitude estimation and filtering without using local coordinates, Part I: Uncontrolled and deterministic attitude dynamics," in *Proceedings of the American Control Conference*, 2006, pp. 5734–5739.

[5] M. D. Shuster, "Kalman filtering of spacecraft attitude and the QUEST model," *Journal of the Astronautical Sciences*, vol. 38, no. 3, pp. 377–393, 1990.

[6] M. L. Psiaki, "Attitude determination filtering via extended quaternion estimation," *AIAA Journal of Guidance, Control and Dynamics*, vol. 23, no. 2, pp. 206–214, 2000.

[7] T. Lee, A. Sanyal, M. Leok, and N. H. McClamroch, "Deterministic global attitude estimation," in *Proceedings of the IEEE Conference on Decision and Control*, 2006.

[8] Y. Theodor, U. Shaked, and C. E. de Souza, "A game theory approach to robust discrete-time H_∞-estimation," *IEEE Transactions on Signal Processing*, vol. 42, no. 6, pp. 1486–1495, 1994.

[9] F. C. Schweppe, "Recursive state estimation: Unknown but bounded errors and system inputs," *IEEE Transactions on Automatic Control*, vol. 13, no. 1, pp. 22–28, 1968.

[10] D. G. Maksarov and J. P. Norton, "State bounding with ellipsoidal set description of the uncertainty," *International Journal of Control*, vol. 65, no. 5, pp. 847–866, 1996.

[11] C. Durieu, E. Walter, and B. Polyak, "Multi-input multi-output ellipsoidal state bounding," *Journal of Optimization Theory and Applications*, vol. 111, no. 2, pp. 273–303, 2001.

[12] T. Lee, M. Leok, and N. H. McClamroch, "Attitude maneuvers of a rigid spacecraft in a circular orbit," in *Proceedings of the American Control Conference*, 2006, pp. 1742–1747.

[13] ——, "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum," in *Proceedings of the IEEE Conference on Control Applications*, 2005, pp. 962–967.

[14] S. Body, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

Proceedings of the 2007 American Control Conference
Marriott Marquis Hotel at Times Square
New York City, USA, July 11-13, 2007

Optimal Attitude Control for a Rigid Body with Symmetry

Taeyoung Lee*†, N. Harris McClamroch†
Department of Aerospace Engineering
University of Michigan, Ann Arbor, MI 48109
{tylee, nhm}@umich.edu

Melvin Leok*
Department of Mathematics
Purdue University, West Lafayette, IN 47907
mleok@math.purdue.edu

Abstract— **Optimal control problems are formulated and efficient computational procedures are proposed for attitude dynamics of a rigid body with symmetry. The rigid body is assumed to act under a gravitational potential and under a structured control moment that respects the symmetry. The symmetry in the attitude dynamics system yields a conserved quantity, and it causes a fundamental singularity in the optimal control problem. The key feature of this paper is its use of computational procedures that are guaranteed to avoid the numerical ill-conditioning that originates from this symmetry. It also preserves the geometry of the attitude dynamics. The theoretical basis for the computational procedures is summarized, and examples of optimal attitude maneuvers for a 3D pendulum are presented.**

I. Introduction

We study a discrete optimal control problem for attitude dynamics of a rigid body with symmetry. The attitude is represented by a rotation matrix, which has a Lie group structure denoted by $\mathrm{SO}(3)$. We assume that the rigid body is acting under an attitude dependent potential, and the potential is invariant under a symmetry action. The external control input is formulated such that it respects the symmetry. This problem provides both a theoretical challenge and a numerical challenge in the sense that the configuration space has a Lie group structure, and the conserved quantity causes ill-conditioning of the numerical optimization.

General purpose numerical integration methods, including the popular Runge-Kutta schemes, typically preserve neither the group structure of the attitude configuration space nor the invariant properties of the dynamics. Geometric structure-preserving integrators are symplectic and momentum preserving, and they exhibit good energy behavior for an exponentially long time period [1]. In particular, Lie group variational integrators have the desirable properties that they preserve the group structure as well as the geometric features, without the need for local parameterization, reprojection, or constraints [2], [3]. The exact geometric properties of the discrete flow not only generate improved qualitative behavior, but also allow for accurate long-time simulation.

Optimal control problems on a Lie group have been studied in [4], [5], [6]. These studies are based on the driftless kinematics of a Lie group. The dynamics are ignored, and elements in the corresponding Lie algebra are considered as control inputs. The discrete optimal control problems of a rigid body are studied in [7], [8], where the dynamics as well as the kinematics equations are explicitly utilized, and an efficient numerical algorithm to solve discrete optimality conditions is presented.

This paper introduces geometrically exact and numerically efficient computational approaches to solve the optimal control problems of the attitude dynamics of a rigid body with symmetry and structured control input. The dynamics are discretized by a Lie group variational integrator, and discrete necessary conditions for optimality are constructed. The utilization of the Lie group variational integrator is justified in this problem, since it preserves the momentum map originating from the symmetry. The rigid body is underactuated since the control input does not act along the symmetry direction. The symmetry of the controlled dynamics causes difficulties in solving the necessary conditions for optimality. A simple numerical approach is presented to overcome this numerical ill-conditioning.

This paper is organized as follows. In Section II, a 3D pendulum is presented as a model of rigid body attitude dynamics, and the symmetry of the 3D pendulum is described. An optimal control problem with symmetry is studied in Section III, and numerical results are given in Section IV.

II. Dynamics of a 3D Pendulum

A 3D pendulum is a rigid body supported by a fixed frictionless pivot acting under the influence of uniform gravitational field [9]. We use a 3D pendulum model to study the optimal control for attitude dynamics of a rigid body, since it has three degrees of rotational freedom, and the gravitational potential has a symmetry: it is invariant under a rotation about the gravity direction.

In this section, the continuous equations of motion are presented. The symmetry of the 3D pendulum are discussed, and the control input structure is described. Discrete equations of motion, referred to as a Lie group variational integrator, are described for a controlled 3D pendulum model.

A. Continuous equations of motion

The configuration space of the 3D pendulum is $\mathrm{SO}(3)$. We identify the tangent bundle $\mathrm{TSO}(3)$ with $\mathrm{SO}(3)\times\mathfrak{so}(3)$ by left-translation, and we identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ by an isomorphism $S(\cdot):\mathbb{R}^3\mapsto\mathfrak{so}(3)$ defined by the condition that $S(x)y=x\times y$ for any $x,y\in\mathbb{R}^3$. We denote the attitude and the angular velocity of the rigid body as $(R,\Omega)\in \mathrm{T}_R\mathrm{SO}(3)$. The rotation matrix $R\in\mathrm{SO}(3)$ transforms a

*This research has been supported in part by NSF under grant DMS-0504747, and by a grant from the Rackham Graduate School, University of Michigan.

†This research has been supported in part by NSF under grant ECS-0244977 and CMS-0555797.

vector represented in the body fixed frame to one represented in the inertial frame.

Let $\rho \in \mathbb{R}^3$ be a vector from the pivot point to the mass center of the rigid body expressed in the body fixed frame, and let $m, g \in \mathbb{R}$ and $J \in \mathbb{R}^{3\times3}$ be the mass of the rigid body, the gravitational acceleration, and the moment of inertia matrix of the rigid body about the pivot point, respectively. The Lagrangian of the 3D pendulum $L : \mathrm{TSO}(3) \mapsto \mathbb{R}$ is given by

$$L(R,\Omega) = \frac{1}{2}\mathrm{tr}\big[S(\Omega)J_d S(\Omega)^T\big] + mge_3^T R\rho,$$

where $J_d \in \mathbb{R}^{3\times3}$ is a nonstandard moment of inertia defined by $J_d = \frac{1}{2}\mathrm{tr}[J]\, I_{3\times3} - J$, and we set the gravitational direction in the inertial frame as $e_3 = [0;0;1]$ in the two-sphere, $\mathbb{S}^2$.

The continuous equations, derived from the Lagrange-d'Alembert principle, are given by

$$\dot{\Pi} + \Omega \times \Pi = mg\rho \times R^T e_3 + M, \tag{1}$$

$$\dot{R} = RS(\Omega), \tag{2}$$

where $\Pi = J\Omega \in \mathbb{R}^3$ is the angular momentum in the body fixed frame, and $M \in \mathbb{R}^3$ is the external control moment.

There are two disjoint equilibria when the direction of gravity in the body fixed frame is collinear with the vector ρ; hanging equilibrium when $R^T e_3 = \rho/\left\|\rho\right\|$, and inverted equilibrium when $R^T e_3 = -\rho/\left\|\rho\right\|$.

B. Symmetry of 3D pendulum

The kinetic energy of the rigid body is left-invariant on $\mathrm{TSO}(3)$, and the gravitational potential energy is invariant under a rotation about the gravity direction, which can be represented by the left action of the subgroup $\big\{\exp S(\theta e_3) \in \mathrm{SO}(3)\big|\theta \in \mathbb{S}^1\big\}$.

As a result, the Lagrangian of the 3D pendulum has a symmetry action by $\mathbb{S}^1$, $\Phi_\theta : \mathbb{S}^1 \times \mathrm{SO}(3) \mapsto \mathrm{SO}(3)$ given by

$$\Phi_\theta(R) = \exp S(\theta e_3)\, R,$$

for $\theta \in \mathbb{S}^1$. It can be shown that $\Phi_\theta^* L(R,\Omega) = L(R,\Omega)$.

Suppose that there is no external control input. Noether's theorem states that a symmetry in the Lagrangian yields conservation of the associated momentum map. For the 3D pendulum, the momentum map of the symmetry action Φ_θ corresponds to the inertial angular momentum of the rigid body about the gravity direction $\pi_3 = e_3^T RJ\Omega \in \mathbb{R}$. It is conserved for the free dynamics of the 3D pendulum.

The structure of the control input respects the symmetry of the uncontrolled free dynamics of the 3D pendulum, namely

$$M = R^T e_3 \times u,$$

for a control parameter $u \in \mathbb{R}^3$. Since the external control moment has no component along the gravity direction, the angular momentum about the gravity direction is also preserved in the controlled dynamics. Such control inputs are physically utilized by actuation mechanisms, such as point mass actuators, that change the center of mass of the 3D pendulum.

Here, we introduce the concept of a geometric phase, and it is used to interpret the numerical optimization result in Section IV. Using the symmetry, the dynamics of the 3D pendulum can be expressed in terms of $\Gamma = R^T e_3$ in the reduced configuration space $\mathrm{SO}(3)/\mathbb{S}^1 \simeq \mathbb{S}^2$. The corresponding flow in the original configuration space is reconstructed by lifting to a level set of the conserved quantity. Suppose that the trajectory in the reduced space is a closed loop, i.e. $\Gamma(0) = \Gamma(T)$ for some $T > 0$, and the value of the angular momentum about the gravity direction is zero. Then, the terminal attitude is related to the initial attitude by a symmetric action. More explicitly, we have

$$R(T) = \Phi_{\theta_{\mathrm{geo}}}(R(0)),$$

where $\theta_{\mathrm{geo}} \in \mathbb{S}^1$ is the geometric phase given by

$$\theta_{\mathrm{geo}} = \int_{\mathcal{B}} \frac{2\left\|J\Gamma\right\|^2 - \mathrm{tr}[J]\,(\Gamma^T J\Gamma)}{(\Gamma^T J\Gamma)^2}\, dA, \tag{3}$$

where $\mathcal{B}$ is a surface in $\mathbb{S}^2$ whose boundary is $\big\{\Gamma(t)\big|t \in [0,T]\big\}$ [10]. Note that the geometric phase only depends on the reduced trajectory of Γ and the characteristics of the rigid body J. It is independent of the velocity $\dot{\Gamma}$.

C. Lie group variational integrator

The attitude of the 3D pendulum is represented by a rotation matrix $R \in \mathrm{SO}(3)$. The conserved quantity, arising from symmetry, is emphasized in this study. However, the most common numerical integration methods, including the widely used Runge-Kutta schemes, neither preserve the Lie group structure nor first integrals. In addition, standard Runge-Kutta methods fail to capture the energy dissipation of a controlled system accurately [11].

For example, if we integrate (2) with a typical Runge-Kutta scheme, the quantity $R^T R$ inevitably drifts from the identity matrix as the simulation time increases. It is often proposed to parameterize (2) using Euler angles or unit quaternions. However, Euler angles are not global expressions of the attitude since they have associated singularities. Unit quaternions do not exhibit singularities, but are constrained to lie on the unit three-sphere $\mathbb{S}^3$, and general numerical integration methods do not preserve the unit length constraint. Therefore, quaternions have the same numerical drift problem. Renormalizing the quaternion vector at each step tends to break other conservation properties. Furthermore, unit quaternions, which are diffeomorphic to $\mathrm{SU}(2)$, double cover $\mathrm{SO}(3)$. So there are inevitable ambiguities in expressing the attitude using quaternions.

In [2], Lie group variational integrators are introduced by explicitly adopting the approach of Lie group methods [12] to the discrete variational principle [11]. They have the desirable property that they are symplectic and momentum preserving, and they exhibit good energy behavior for an exponentially long time period. They also preserve the Lie group structure without the use of local charts, reprojection, or constraints.

Using the results in [2], a Lie group variational integrator on $\mathrm{SO}(3)$ is given for the 3D pendulum by

$$hS(\Pi_k) = F_k J_d - J_d F_k^T, \tag{4}$$

$$R_{k+1} = R_k F_k, \tag{5}$$

$$\Pi_{k+1} = F_k^T \Pi_k + hmg\rho \times R_{k+1}^T e_3 + hR_{k+1}^T e_3 \times u_{k+1}, \tag{6}$$

where the subscript k denotes the kth discrete variable for a fixed integration step size $h \in \mathbb{R}$, and $F_k \in \mathrm{SO}(3)$ is the relative attitude between two adjacent integration steps. For a given (R_k, Π_k) and control inputs, (4) is solved to find F_k. Then (R_{k+1}, Π_{k+1}) is obtained by (5) and (6). This yields a map $(R_k, \Pi_k) \mapsto (R_{k+1}, \Pi_{k+1})$ and this process is repeated. The only implicit part is (4). The actual computation of F_k is done in the Lie algebra $\mathfrak{so}(3)$ of dimension 3, and the rotation matrices are updated by multiplication. So this approach is distinguished from integration of the kinematics equation (2), and there is no excessive computational burden. The properties of these discrete equations of motion are discussed more explicitly in [2], [3]. We use these discrete equations of motion to formulate the following optimal control problem.

III. Optimal control with symmetry

We formulate an optimal attitude control problem for a 3D pendulum with symmetry. Necessary conditions for optimality are developed and computational approaches are presented to solve the corresponding two-point boundary value problem.

A. Problem formulation

A discrete time optimal control problem is formulated as a maneuver of the rigid pendulum body from a given initial attitude $R_0 \in \mathrm{SO}(3)$ and an initial angular momentum $\Pi_0 \in \mathbb{R}^3$ to a desired terminal attitude $R_N^d \in \mathrm{SO}(3)$ and a terminal angular momentum $\Pi_N^d \in \mathbb{R}^3$ during a given maneuver time N. The performance index is the square of the l_2 norm of the control inputs:

$$\text{given: } (R_0, \Pi_0),\ (R_N^d, \Pi_N^d),\ N,$$

$$\min_{u_{k+1}} \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2} \left\| u_{k+1} \right\|^2,$$

$$\text{such that } R_N = R_N^d,\ \Pi_N = \Pi_N^d,$$

$$\text{subject to (4), (5) and (6).}$$

B. Necessary conditions of optimality

Variational models: The necessary conditions of optimality are developed using the standard variational approach. We first derive certain variational formulas. The variation of $R_k \in \mathrm{SO}(3)$ can be expressed in terms of a Lie algebra element $S(\zeta_k) \in \mathfrak{so}(3)$ for $\zeta_k \in \mathbb{R}^3$ and the exponential map as

$$R_k^\epsilon = R_k \exp \epsilon S(\zeta_k).$$

The corresponding infinitesimal variation is given by

$$\delta R_k = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_k \exp \epsilon S(\zeta_k) = R_k S(\zeta_k). \tag{7}$$

This gives an expression for the infinitesimal variation of a Lie group element in terms of its Lie algebra. Then, small perturbations from a given trajectory can be written as

$$\Pi_k^\epsilon = \Pi_k + \epsilon \delta \Pi_k, \tag{8}$$

$$R_k^\epsilon = R_k + \epsilon R_k S(\zeta_k) + \mathcal{O}(\epsilon^2), \tag{9}$$

where $\delta\Pi_k, \zeta_k$ are considered as elements of $\mathbb{R}^3$.

Necessary conditions: Define an augmented performance index as

$$\begin{aligned}\mathcal{J}_a = \sum_{k=0}^{N-1} \frac{h}{2} \left\| u_{k+1} \right\|^2 &+ \lambda_k^{1,T} S^{-1}\big(\mathrm{logm}(F_k - R_k^T R_{k+1})\big) \\ &+ \lambda_k^{2,T} \left\{ -\Pi_{k+1} + F_k^T \Pi_k + hmg\rho \times R_{k+1}^T e_3 \right\} \\ &+ \lambda_k^{2,T} \left\{ R_{k+1}^T e_3 \times u_{k+1} \right\},\end{aligned} \tag{10}$$

where $\lambda_k^1, \lambda_k^2 \in \mathbb{R}^3$ are Lagrange multipliers corresponding to the discrete equations of motion (5) and (6). Here, the matrix logarithm is denoted by logm. The constraint (4) is imposed implicitly by using a constrained variation derived from the infinitesimal version of the constraint (4).

Using the variational models (8) and (9), and the fact that the variations ζ_k, $\delta\Pi_k$ vanish at $k = 0, N$, the infinitesimal variation of the augmented performance index is written as

$$\begin{aligned}\delta\mathcal{J}_a = \sum_{k=1}^{N-1} &h\delta u_k^T \left\{ u_k - R_k^T e_3 \times \lambda_{k-1}^2 \right\} \\ &+ \zeta_k^T \left\{ -\lambda_{k-1}^1 + \mathcal{A}_k^T \lambda_k^1 + \mathcal{C}_k^T \lambda_k^2 - hF_k u_{k+1} e_3^T R_{k+1} \right\} \\ &+ \delta\Pi_k^T \left\{ -\lambda_{k-1}^2 + \mathcal{B}_k^T \lambda_k^1 + \mathcal{D}_k^T \lambda_k^2 - h\mathcal{B}_k^T u_{k+1} e_3^T R_{k+1} \right\},\end{aligned}$$

where

$$\begin{aligned}\mathcal{A}_k &= F_k^T, \\ \mathcal{B}_k &= hF_k^T \left\{ \mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d \right\}^{-1}, \\ \mathcal{C}_k &= hmgS(\rho) S(R_{k+1}^T e_3) F_k^T, \\ \mathcal{D}_k &= F_k^T + S(F_k^T \Pi_k)\mathcal{B}_k + hmgS(\rho) S(R_{k+1}^T e_3)\mathcal{B}_k.\end{aligned}$$

Since $\delta\mathcal{J}_a = 0$ for all variations of $\delta u_k, \zeta_k, \delta\Pi_k$, the expressions in the braces of the above equation are zero. Thus we obtain necessary conditions for optimality as follows.

$$hS(\Pi_k) = F_k J_d - J_d F_k^T, \tag{11}$$

$$R_{k+1} = R_k F_k, \tag{12}$$

$$\Pi_{k+1} = F_k^T \Pi_k + hmg\rho \times R_{k+1}^T e_3 + hR_{k+1}^T e_3 \times u_{k+1}, \tag{13}$$

$$u_{k+1} = R_{k+1}^T e_3 \times \lambda_k^2, \tag{14}$$

$$\begin{bmatrix} \lambda_k^1 \\ \lambda_k^2 \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{k+1}^T & \mathcal{C}_{k+1}^T - hF_{k+1} u_{k+2} e_3^T R_{k+2} \\ \mathcal{B}_{k+1}^T & \mathcal{D}_{k+1}^T - h\mathcal{B}_{k+1}^T u_{k+2} e_3^T R_{k+2} \end{bmatrix} \begin{bmatrix} \lambda_{k+1}^1 \\ \lambda_{k+1}^2 \end{bmatrix}. \tag{15}$$

In the above equations, the implicit parts are (11) and (15). For a given initial condition $(R_0, \Pi_0, \lambda_0^1, \lambda_0^2)$, we can find F_0 by solving (11). Then, R_1 is obtained by (12). Since $u_1 = R_1^T e_3 \times \lambda_0^2$ by (14), Π_1 can be obtained using (13). We solve (11) to obtain F_1 using Π_1. Finally, λ_1^1, λ_1^2 are obtained by solving the implicit equation (15), since $\mathcal{A}_1, \mathcal{B}_1, \mathcal{C}_1, \mathcal{D}_1$ are functions of R_1, Π_1, F_1.

The implicit equation (11) is solved by Newton iteration at the level of the Lie algebra, and the implicit equation (15) is solved by fixed-point iteration. Numerical computations show that two or three iterations are typically sufficient to achieve machine precision.

C. Two-point boundary value problem

The necessary conditions for optimality are given by a 12-dimensional two-point boundary value problem. The problem is to find the optimal discrete flow, multipliers, and control inputs to satisfy the equations of motion (11)–(13), optimality condition (14), multiplier equations (15), and boundary conditions simultaneously.

We substitute the optimality condition (14) into the equations of motion and the multiplier equations, and we apply the shooting method to solve the two-point boundary value problem using sensitivity derivatives. The shooting method is numerically efficient in the sense that the number of iteration parameters is minimized; 6 elements of the initial Lagrange multiplier are iterated. In other approaches, the entire discrete trajectory of the control input and Lagrange multiplier are updated.

The drawback of the shooting method is that the extremal solutions are sensitive to small changes in the unspecified initial multiplier values. The nonlinearity makes it hard to construct an accurate estimate of sensitivity. In addition this problem, the symmetry and the underactuation induce numerical ill-conditioning. Therefore, in order to apply the shooting method, it is important to compute the sensitivities accurately, and the effects of the symmetry should be taken into account.

In this paper, the attitude dynamics of a rigid body is described by the structure-preserving Lie group variational integrator, and the sensitivity is expressed in terms of a Lie algebra element. This approach completely avoids any singularity in the attitude representation, and the discrete flow respects the geometric features. The controlled attitude dynamics is not a Hamiltonian system, but, together with the adjoint multiplier equations, the controlled rigid body dynamics have a Hamiltonian structure which is preserved by the Lie group variational integrator and discrete multiplier equations. The resulting sensitivity derivatives are sufficiently accurate for the shooting method. Furthermore, a simple numerical approach is presented to eliminate the ill-conditioning caused by the symmetry.

Sensitivity derivatives: Taking a variation of the discrete equations of motion and the multiplier equation using the variational models, the linearized equations of motion and the linearized multiplier equations can be written as

$$x_{k+1} = A_k^{11} x_k + A_k^{12} \delta\lambda_k,$$
$$\delta\lambda_k = A_{k+1}^{21} x_{k+1} + \left(A_{k+1}^{11}\right)^T \delta\lambda_{k+1},$$

where $x_k = [\zeta_k; \delta\Pi_k] \in \mathbb{R}^6$, and matrices $A^{ij} \in \mathbb{R}^{6\times 6}$ can be suitably defined. The solution of the linear equations is given by

$$\begin{bmatrix} x_N \\ \delta\lambda_N \end{bmatrix} = \begin{bmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{bmatrix} \begin{bmatrix} x_0 \\ \delta\lambda_0 \end{bmatrix},$$

where $\Psi^{ij} \in \mathbb{R}^{6\times 6}$. For the given two-point boundary value problem, the initial attitude and the initial angular momentum are fixed, and the terminal multiplier is free. Thus, we have the following sensitivity equation for the terminal attitude and the terminal angular momentum with respect to the initial multiplier;

$$x_N = \Psi^{12} \delta\lambda_0. \tag{16}$$

Avoiding numerical ill-conditioning: The symmetry yields a conserved quantity by Noether's theorem, and it causes a fundamental singularity in the sensitivity derivatives for the two-point boundary value problem. At each iteration, we require the inverse of the sensitivity derivative represented by the matrix Ψ^{12} to update the initial multiplier to satisfy the terminal boundary condition. However, this sensitivity matrix has a theoretical rank deficiency of one since the vertical component of the inertial angular momentum is conserved regardless of the initial multiplier variation. Therefore, equation (16) is numerically ill-conditioned.

Here, we presents a simple numerical scheme to avoid the numerical ill-conditioning caused by the symmetry. We decompose the sensitivity derivative into symmetric parts and asymmetric parts. Equation (16) is rewritten as

$$\begin{bmatrix} \zeta_N \\ \delta\Pi_N \end{bmatrix} = \begin{bmatrix} \Psi_1 & \Psi_2 \\ \Psi_3 & \Psi_4 \end{bmatrix} \begin{bmatrix} \delta\lambda_0^1 \\ \delta\lambda_0^2 \end{bmatrix}, \tag{17}$$

where $\Psi_i \in \mathbb{R}^{3\times 3}$ are submatrices of Ψ^{12}. Using the above equation and (7), the infinitesimal variation of the inertial angular momentum is given by

$$\begin{aligned} \delta\pi_N &= \delta(R_N \Pi_N) = \delta R_N \Pi_N + R_N \delta\Pi_N, \\ &= -R_N S(\Pi_N)\zeta_N + R_N \delta\Pi_N, \\ &= -R_N S(\Pi_N)(\Psi_1\lambda_0^1 + \Psi_2\lambda_0^2) + R_N(\Psi_3\lambda_0^1 + \Psi_4\lambda_0^2). \end{aligned}$$

Now, the sensitivity derivative equation (17) can be rewritten in terms of the inertial angular momentum variation as

$$\begin{bmatrix} \zeta_N \\ \delta\pi_N \end{bmatrix} = \begin{bmatrix} \Psi_1 & \Psi_2 \\ R_N(\Psi_3 - S(\Pi_N)\Psi_1) & R_N(\Psi_4 - S(\Pi_N)\Psi_2) \end{bmatrix} \begin{bmatrix} \delta\lambda_0^1 \\ \delta\lambda_0^2 \end{bmatrix}. \tag{18}$$

From the symmetry, the third component of the inertial angular momentum variation is zero; $\delta(\pi_N)_3 = 0$. Thus, the sixth row of the above matrix is zero. (Numerical simulation in the later section shows that the norm of the last row of the transformed sensitivity matrix is at the level of 10^{-15}.) Now, we find an update of the initial multiplier by the pseudo-inverse of the 5×6 matrix;

$$\delta\lambda_0 = \Xi^\dagger x'_N = \Xi^T(\Xi\Xi^T)^{-1} x'_N, \tag{19}$$

where $\Xi \in \mathbb{R}^{5\times 6}$ is composed of the first five rows of the transformed sensitivity derivative in (18), and $x'_N = [\zeta_N; \delta(\pi_N)_1; \delta(\pi_N)_2] \in \mathbb{R}^5$. This approach removes the singularity in the sensitivity derivatives completely, and the resulting optimal control problem is no longer ill-conditioned. Numerical simulations show that the numerical optimization procedure fails without this modification.

Newton iteration: Using the decomposed sensitivity, an initial guess of the unspecified initial conditions is iterated to satisfy the specified terminal boundary conditions in the limit. Any type of Newton iteration can be applied. We use a line search with backtracking algorithm, referred to as Newton-Armijo iteration [13]. The procedure is summarized as follows.

1: Guess an initial multiplier λ_0.
2: Find $\Pi_k, R_k, \lambda_k^1, \lambda_k^2$ using (11)–(15).
3: Compute the terminal B.C. error; Error $= \|x'_N\|$.
4: Set Error^t = Error, $i = 1$.
5: **while** Error $> \epsilon_S$.
6: Find a line search direction; $D = \Xi^\dagger$.
7: Set $c = 1$.
8: **while** Error$^t > (1 - 2\alpha c)$Error
9: Choose a trial multiplier $\lambda_0^t = \lambda_0 + cDx'_N$.
10: Find $\Pi_k, R_k, \lambda_k^1, \lambda_k^2$ using (11)–(15).
11: Compute the error; Error$^t = \left\|x_N'^t\right\|$.
12: Set $c = c/10$, $i = i + 1$.
13: **end while**
14: Set $\lambda_0 = \lambda_0^t$, Error = Error^t. (accept the trial)
15: **end while**

Here, i is the number of iterations, and $\epsilon_S, \alpha \in \mathbb{R}$ are stopping criterion and a scaling factor, respectively. The outer loop finds a search direction by computing the sensitivity derivatives, and the inner loop performs a line search to find the largest step size $c \in \mathbb{R}$ along the search direction. The error in satisfaction of the terminal boundary condition is determined at each inner iteration.

IV. Numerical Examples

Numerical optimization results for the 3D pendulum are given. Two elliptical cylinders, shown in Fig. 1, are used as rigid pendulum models. The properties are chosen as

Body (A): $m = 1$, $J = \mathrm{diag}[0.13, 0.28, 0.17]$, $\rho = 0.3e_3$.
Body (B): $m = 1$, $J = \mathrm{diag}[0.22, 0.23, 0.03]$, $\rho = 0.4e_3$.

Four cases are considered. Each maneuver is from a hanging equilibrium to another hanging equilibrium with a rotation about the vertical axis. The rotation angles are chosen as 90° and 180°. Since the vertical component of the angular momentum is set to zero, the rotation is purely caused by the geometric phase effect given in (3). These problems are challenging in the sense that the desired maneuvers are rotations about the gravity direction, but the control input cannot directly generate any moment about the gravity direction.

(a) Body (A) (b) Body (B)

Fig. 1. Elliptical cylinder

TABLE I
Optimization results

Case	$\mathcal{J}$	$\left\|\mathrm{logm}(R_N^{d,T} R_N)\right\|$	$\left\|\Pi_N^d - \Pi_N\right\|$	ΔT
(i)	5.91	2.30×10^{-14}	1.34×10^{-14}	2.72
(ii)	7.32	4.80×10^{-15}	1.66×10^{-14}	5.25
(iii)	1.73	1.22×10^{-15}	6.55×10^{-14}	4.09
(iv)	3.37	3.06×10^{-14}	3.04×10^{-14}	5.05

ΔT: CPU time in Intel Pentium 1.73GHz processor (min.)

The corresponding boundary conditions are rotations about the gravity direction from a hanging equilibrium to a different hanging equilibrium as follows.

(i) Body (A), 90° rotation about the gravity direction

$$R_0 = I_{3\times 3}, \quad R_N^d = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(ii) Body (A), 180° rotation about the gravity direction

$$R_0 = I_{3\times 3}, \quad R_N^d = \mathrm{diag}[-1, -1, 1]$$

(iii) Body (B), 90° rotation about the gravity direction

$$R_0 = I_{3\times 3}, \quad R_N^d = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(iv) Body (B), 180° rotation about the gravity direction

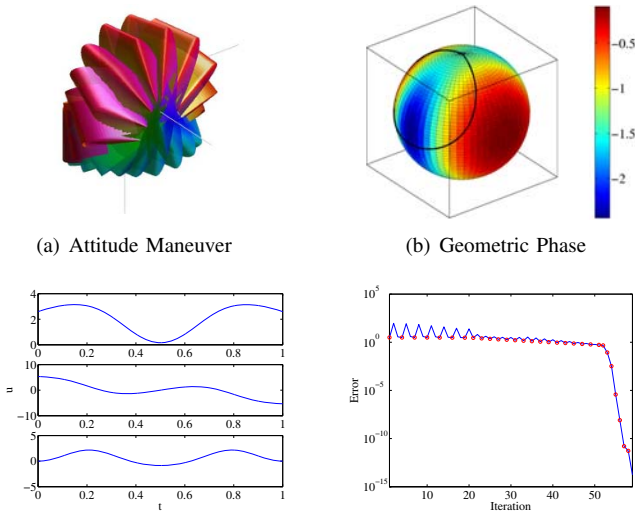
$$R_0 = I_{3\times 3}, \quad R_N^d = \mathrm{diag}[-1, -1, 1]$$

The initial angular momentum and the terminal angular momentum are chosen to be zero, i.e. $\Pi_0 = 0_{3\times 1}$, $\Pi_N^d = 0_{3\times 1}$.

The optimal control results are given in Table I, where the optimized performance index, the error in satisfaction of the terminal boundary condition, and the simulation running time are shown for each case. The terminal error is at the level of machine precision, and the simulation time is less than 6 minutes with an Intel Pentium M 740 1.73GHz processor on MATLAB.

Figures 2 and 3 show snapshots of the attitude maneuvers, reduced trajectory of $\Gamma = R^T e_3$ on a sphere, control input history, and convergence rate for Cases (ii) and (iv). (A simple animation for the attitude maneuver can be seen at `http://www.umich.edu/~tylee`.)

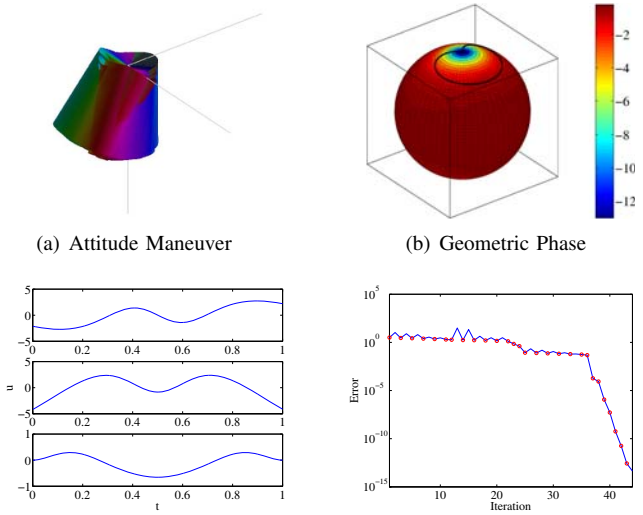
The convergence rate figures show violation of the terminal boundary condition according to the number of iterations in a logarithm scale. Red circles denote outer iterations in Newton-Armijo iteration to compute the sensitivity derivatives. For all cases, the initial guesses of the unspecified initial multiplier are arbitrarily chosen. The error in satisfaction of the terminal boundary condition converges quickly to machine precision after the solution is close to the local minimum at around the 50th iteration. These convergence results



(a) Attitude Maneuver (b) Geometric Phase

(c) Control Input (d) Convergence Rate

Fig. 2. (ii). Body A, 180° rotation about the gravity direction



(a) Attitude Maneuver (b) Geometric Phase

(c) Control Input (d) Convergence Rate

Fig. 3. (iv). Body B, 180° rotation about the gravity direction

are consistent with the quadratic convergence rates expected of Newton methods with accurately computed gradients. The condition number of the decomposed sensitivity derivative given at (19) varies from 10^0 to 10^5. If the sensitivity derivative is not decomposed, then the condition numbers are at the level of 10^{19}, and the numerical iterations fail.

The numerical examples presented in this paper show excellent numerical convergence properties. This is because the proposed computational algorithms on $SO(3)$ are geometrically exact and numerically accurate. In addition, the algorithm incorporates a modification that eliminates the singularity caused by the symmetry.

We interpret the optimization results using the geometric phase formula given by (3). For given initial conditions, the vertical component of the initial angular momentum is zero. Thus, the rotation about the vertical axis is purely caused by the geometric phase. Since the geometric phase is determined by a surface integral on $\mathbb{S}^2$ whose boundary is the reduced trajectory Γ, it is more efficient for the reduced trajectory to enclose the area at which the absolute value of the integrand of (3) is maximized.

In Fig. 2.(b) with Fig. 3.(b), the infinitesimal geometric phase per unit area is shown by color shading. The reduced trajectory, which represents the gravity direction in the body fixed frame, is shown by a solid line. The north pole of the sphere corresponds to the hanging equilibrium manifold, and the reduced trajectory starts and ends at the same north pole for the given boundary conditions.

Comparing Fig. 2.(b) with Fig. 3.(b), it can be seen that Body (A) and Body (B) have different geometric phase characteristics. This is caused by the fact that the geometric phase depends on the moment of inertia of the body. For Body (A), the absolute value of the infinitesimal geometric phase is maximized at a point on the equator, and for Body (B), it is maximized at the north pole. We see that the optimized reduced trajectories try to enclose those points.

As a result, the optimized attitude maneuver of Body (A) is distinguished from that of Body (B). The attitude maneuver of Body (A) is relatively more aggressive than that of Body (B) since the reduced trajectory passes near the equator corresponding to a horizontal position. Body (B) does not have to move far away from the hanging equilibrium since the infinitesimal geometric phase is maximized at that point. The resulting attitude maneuver is relatively benign.

REFERENCES

[1] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*. Springer, 2002.

[2] T. Lee, M. Leok, and N. H. McClamroch, "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum," in *Proceedings of the IEEE Conference on Control Applications*, 2005, pp. 962–967.

[3] ——, "Lie group variational integrators for the full body problem in orbital mechanics," *Celestial Mechanics and Dynamical Astronomy*, 2007, accepted, DOI : 10.1007/s10569-007-9073-x.

[4] K. Spindler, "Optimal control on Lie groups with applications to attitude control," *Mathematics of Control, Signals, and Systems*, vol. 11, pp. 197–219, 1998.

[5] S. Sastry, "Optimal control on Lie groups," in *Proceedings of the Third International Congress on Industrial and Applied Mathematics (ICIAM)*, 1995.

[6] V. Jurdjevic, *Geometric Control Theory*. Cambridge University, 1997.

[7] T. Lee, M. Leok, and N. H. McClamroch, "Optimal attitude control of a rigid body using geometrically exact computations on SO(3)," *Journal of Optimization Theory and Applications*, 2006, submitted. [Online]. Available: http://arxiv.org/math.OC/0601424

[8] ——, "Optimal control of a rigid body using geometrically exact computations on SE(3)," in *Proceedings of the IEEE Conference on Decision and Control*, 2006, pp. 2170–2175.

[9] J. Shen, A. K. Sanyal, N. A. Chaturvedi, D. Bernstein, and N. H. McClamroch, "Dynamics and control of a 3D pendulum," in *Proceedings of 43rd IEEE Conference on Decision and Control*, Dec. 2004, pp. 323–328.

[10] J. E. Marsden, R. Montgomery, and T. S. Ratiu, *Reduction, Symmetry and Phases in Mechanics*. American Mathematical Society, 1990.

[11] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[12] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, "Lie-group methods," *Acta Numerica*, vol. 9, pp. 215–365, 2000.

[13] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.

Proceedings of the
46th IEEE Conference on Decision and Control
New Orleans, LA, USA, Dec. 12-14, 2007

Propagation of Uncertainty in Rigid Body Attitude Flows

Taeyoung Lee*†, Nalin A. Chaturvedi†, Amit Sanyal, Melvin Leok*, and N. Harris McClamroch†

Abstract— **Motivated by attitude control and attitude estimation problems for a rigid body, computational methods are proposed to propagate uncertainties in the angular velocity and the attitude. Uncertainties in the angular velocity and attitude are described in terms of ellipsoidal sets that are propagated through this nonlinear attitude flow. Computational methods are proposed, one method based on a local linearization of the attitude flow and two methods based on propagation of a small (unscented) sample selected from the initial uncertainty ellipsoid. Each of these computational methods is constructed using a Lie group variational integrator, viewed as a discretization of the attitude dynamics. Computational results are obtained that indicate (1) the strongly nonlinear attitude flow characteristics and (2) the limitations of each of these methods, and indeed any method, in providing effective global bounds on the nonlinear attitude flow.**

I. Introduction

As an integrable system, the attitude dynamics of the free rigid body are reasonably well understood. However, if there is an attitude dependent potential that influences the rigid body, then the dynamics can be surprisingly complex. In this paper, such attitude dynamics are studied for a rigid body with an inertially fixed pivot acting under the influence of uniform and constant gravity; this model is subsequently referred to as the 3D pendulum [1].

The objective of this paper is to study the nonlinear attitude flow of the 3D pendulum dynamics by characterizing the propagation of uncertainty for the attitude and angular velocity in a deterministic sense; bounds on the uncertainty are specified and propagated without an assumption on their distribution. This line of research was motivated by our prior work on deterministic attitude estimation [2], [3], which is composed of deterministic uncertainty propagation, measurement update, and filtering procedures.

Uncertainty propagation can also be studied within a probabilistic framework using the Liouville partial differential equation, a special case of the Fokker-Planck equation, but this leads to significant solution difficulties [4]. However, statistical properties of uncertainties are often unavailable or are difficult to obtain in practice. It has been shown that deterministic estimation is more robust to noise distribution [5].

Taeyoung Lee, Nalin A. Chaturvedi, and N. Harris McClamroch, Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 {tylee,nalin,nhm}@umich.edu

Amit Sanyal, Mechanical Engineering, University of Hawaii at Manoa, Honolulu, HI 96822 aksanyal@hawaii.edu

Melvin Leok, Mathematics, Purdue University, West Lafayette, IN 47907 mleok@math.purdue.edu

*This research has been supported in part by NSF under grants DMS-0504747 and DMS-0726263.

†This research has been supported in part by NSF under grant CMS-0555797.

In this paper, three computational methods for deterministic attitude uncertainty propagation are numerically compared; one method based on a local linearization of the attitude flow and two methods based on propagation of a small (unscented) sample selected from the initial uncertainty ellipsoid. Our approach in this paper is to make use of geometrically-exact computational tools that respect the Lie group structure of the configuration space for the rigid body attitude dynamics. This approach is based on the recently introduced Lie group variational integrator for the 3D pendulum and on a framework for propagating suitably defined uncertainty ellipsoids [6].

It turns out that these computational tools are useful in some cases, but they have important limitations that arise from the complex attitude dynamics of the 3D pendulum. The results in this paper demonstrate the complex global dynamics that can occur for an uncontrolled 3D pendulum. In this way, the challenges of attitude control and estimation are made clear, at least in the case where global results are desired.

This paper is organized as follows. The 3D pendulum model is described in Section II. Three computation approach for deterministic attitude uncertainty propagation are presented in Section III, and their computational properties are shown in Section IV. The global dynamic characteristics of the 3D pendulum are discussed in Section V.

II. 3D Pendulum

A rigid 3D pendulum is a rigid body supported by a fixed, frictionless pivot, acted on by uniform gravitational forces [1]. The supporting pivot allows the pendulum three rotational degrees of freedom.

Two reference frames are introduced. An inertial reference frame has its origin at the pivot; the first two axes lie in the horizontal plane and the third axis is vertical in the direction of gravity. A reference frame fixed to the pendulum body is also introduced. The origin of this body-fixed frame is also located at the pivot.

The configuration space is the special orthogonal group $SO(3)$,

$$SO(3) = \left\{ R \in \mathbb{R}^{3\times 3} \,|\, R^T R = I_{3\times 3},\, \det R = 1 \right\},$$

where the rotation matrix $R \in SO(3)$ represents the linear transformation from the body-fixed frame to the inertial frame.

The dynamics of the 3D pendulum are given by the Euler rigid body equation that includes the moment due to gravity:

$$J\dot{\Omega} = J\Omega \times \Omega + mg\rho \times R^T e_3, \tag{1}$$

where the angular velocity in the body-fixed frame is denoted by $\Omega \in \mathbb{R}^3$, the moment of inertia matrix is denoted by $J \in \mathbb{R}^{3\times3}$, and the vector $\rho \in \mathbb{R}^3$ represents the location of the center of mass in the body-fixed frame. The constants m and g denote the mass of the pendulum and the gravitational acceleration, respectively. The kinematic equation is

$$\dot{R} = RS(\Omega). \tag{2}$$

For a given vector $a \in \mathbb{R}^3$, the 3×3 skew-symmetric matrix $S(a)$ is defined so that $S(a)b = a \times b$ for all $b \in \mathbb{R}^3$.

There are two disjoint equilibria when the direction of gravity in the body-fixed frame is collinear with the vector ρ. We define

$$H = \left\{(R,\Omega) \in TSO(3) \,\middle|\, R^T e_3 = \rho/\left\|\rho\right\|, \Omega = 0\right\},$$
$$I = \left\{(R,\Omega) \in TSO(3) \,\middle|\, R^T e_3 = -\rho/\left\|\rho\right\|, \Omega = 0\right\}$$

as hanging equilibria and inverted equilibria, respectively. Here, the tangent bundle of $SO(3)$ is denoted by $TSO(3)$, which can be identified with $SO(3) \times \mathbb{R}^3$.

III. Uncertainty propagation

An uncertainty ellipsoid on $TSO(3)$ is defined by

$$\mathcal{E}(\hat{R},\hat{\Omega},P) = \left\{(R,\Omega) \in TSO(3) \,\middle|\, x^T P^{-1} x \le 1\right\},$$

where $x = [\zeta; \delta\Omega] \in \mathbb{R}^6$ and $\zeta = \log(\hat{R}^T R), \delta\Omega = \Omega - \hat{\Omega} \in \mathbb{R}^3$. The center of the ellipsoid is given by the rotation matrix $\hat{R}$ and the angular velocity $\hat{\Omega}$; the matrix $P \in \mathbb{R}^{6\times6}$ is the uncertainty matrix that characterizes the size and the shape of the ellipsoid. In particular, if the initial attitude and angular velocity are known to lie within an initial uncertainty ellipsoid, we seek computational methods to propagate the uncertainty so as to obtain an uncertainty ellipsoid, for the current time, within which the current attitude and angular velocity of the 3D pendulum are expected to lie.

We study three methods to propagate the uncertainty set over the time interval $[0, T]$. All three methods make use of the Lie group variational integrator that is described by the following discrete update equations [6],

$$hS(J\Omega_k + \frac{h}{2}M_k) = F_k J_d - J_d F_k^T, \tag{3}$$
$$R_{k+1} = R_k F_k, \tag{4}$$
$$J\Omega_{k+1} = F_k^T J\Omega_k + \frac{h}{2}F_k^T M_k + \frac{h}{2}M_{k+1}, \tag{5}$$

to propagate the angular velocity and the attitude. The subscript k denotes a variable corresponding to the kth discrete timestep for a fixed integration step size $h \in \mathbb{R}$, and $F_k \in SO(3)$ is the relative attitude between two adjacent integration steps. The nonstandard moment of inertia matrix is given by $J_d = \frac{1}{2}\mathrm{tr}\left[J\right] I_{3\times3} - J \in \mathbb{R}^{3\times3}$, and the moment due to gravity is denoted by $M_k = mg\rho \times R_k^T e_3 \in \mathbb{R}^3$. For a given (R_k, Ω_k), (3) is solved to find $F_k \in SO(3)$. Then (R_{k+1}, Ω_{k+1}) is obtained by (4) and (5). This yields a map $(R_k, \Omega_k) \mapsto (R_{k+1}, \Omega_{k+1})$ and this process is repeated. The use of the Lie group variational integrator is desirable since it preserves the orthogonal structure of $SO(3)$ without need of local parameterization or constraints.

A. Uncertainty propagation using linearization on $[0, T]$

In this method, the uncertainty ellipsoid is propagated by updating both the center of the ellipsoid and the the uncertainty matrix P. The center of the uncertainty ellipsoid, denoted by $\hat{R}_k, \hat{\Omega}_k$, is propagated according to the Lie group variational integrator, initialized by the center of the initial uncertainty ellipsoid. We then assume that the uncertainty ellipsoid is sufficiently small that the flow of points within the uncertainty ellipsoid is well-approximated by the linearized flow of the Lie group variational integrator about this center solution, which we denote by

$$x_{k+1} = A_k x_k,$$

where $x_k = [\zeta_k; \delta\Omega_k] \in \mathbb{R}^6$ and the matrix $A_k \in \mathbb{R}^{6\times6}$ depends on $\hat{R}_k, \hat{\Omega}_k$. The expression for A_k can be found in [7]. Using this linearization, the uncertainty matrix is propagated according to

$$P_{k+1} = A_k P_k A_k^T \tag{6}$$

with initial condition given by the initial uncertainty matrix. In this way, the uncertainty ellipsoid is propagated and determined at time T.

B. Uncertainty propagation using unscented method on $[0, T]$

In this method, we compute the propagated uncertainty ellipsoid by using the Lie group variational integrator to propagate a small sample of points selected from the initial uncertainty ellipsoid. Following the conventional wisdom, we choose the 12 points corresponding to the intersection of the boundary of the initial uncertainty ellipsoid and its principal axes. This choice is informed by the fact that if the initial uncertainty ellipsoid is propagated by a linear flow, it will remain an ellipsoid, and it will coincide with the minimal volume ellipsoid containing the the 12 intersection points propagated by the same linear flow.

More explicitly, suppose that the initial uncertainty ellipsoid is given by

$$\mathcal{E}(R_0, \Omega_0, P_0).$$

Let $\lambda^i \in \mathbb{R}$ and $\phi^i = [\phi_R^i; \phi_\Omega^i] \in \mathbb{R}^6$ be the i-th eigenvalue and eigenvector, respectively, of the uncertainty matrix $P_0 \in \mathbb{R}^{6\times6}$ for $i \in \{1, 2\ldots, 6\}$. The intersection of the corresponding principal axis and the ellipsoid boundary is $\sqrt{\lambda_i}\phi_i$. Then, the 12 intersection points are given by,

$$\left\{\left(R_0 \exp(\pm\sqrt{\lambda^i}\hat{\phi}_R^i), \Omega_0 \pm \sqrt{\lambda^i}\phi_\Omega^i\right)\right\} \quad i \in \{1, 2, \ldots, 6\}.$$

We propagate these 12 initial conditions using the Lie group variational integrator to determine the 12 values of the attitude and angular velocity at time T. We then construct a minimal volume ellipsoid that contains all 12 values of the attitude and angular velocity at time T [8]. From this minimal volume ellipsoid, the center of the ellipsoid, and the uncertainty matrix at time T can be computed.

C. Uncertainty propagation using unscented method with re-sampling on $[0, T]$

The unscented method propagates the same sampled points obtained from the *initial* uncertainty ellipsoid throughout the entire time period $[0, T]$. In this modification, we partition the time period $[0, T]$ into subintervals and we use the Lie group variational integrator to propagate the 12 sample points throughout each subinterval. At the end of each subinterval, we construct a minimal volume ellipsoid that contains all 12 values of the attitude and angular velocity at the end of that subinterval. We then select a new set of sample points located on the principal axes of this new uncertainty ellipsoid. On the first subinterval, the 12 initial values of the attitude and angular velocity are obtained from the initial uncertainty ellipsoid. In contrast to the previous method which initializes the sample points at the start of the subsequent subinterval by taking the propagated points from the end of the prior subinterval, the current method obtains sample points by resampling from the propagated uncertainty ellipsoid.

If the flow is linear, the re-sampling technique will yield the same estimates as the usual unscented method. If, however, the flow is nonlinear, the re-sampling technique will tend to yield a more conservative estimate, since it captures the deformation of the ellipsoid in the nonlinear flow more explicitly. Indeed, the difference between the two estimates gives an indication of how nonlinear the flow is.

IV. Numerical examples

We apply these methods to the attitude dynamics of the 3D pendulum. The pendulum body is chosen as an elliptic cylinder, and its properties are given by

$$J = \mathrm{diag}[0.13, 0.28, 0.17]\,\mathrm{kg}\cdot\mathrm{m}^2,\ m = 1\,\mathrm{kg},\ \rho = 0.3e_3\,\mathrm{m}.$$

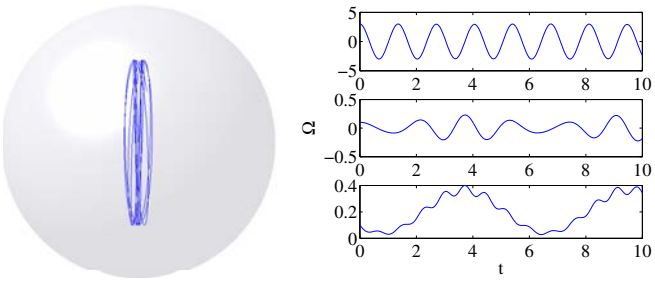
We consider two initial conditions; the first initial condition results in a near-oscillatory attitude flow while the second initial condition results in a highly irregular attitude flow.

Oscillatory attitude flow: The initial uncertainty ellipsoid is characterized by its initial center attitude and angular velocity and initial uncertainty matrix, which are given by

$$R_0 = I_{3\times 3},\quad \Omega_0 = [3.0, 0.1, 0.1]\,\mathrm{rad/s},$$
$$P_0 = \mathrm{diag}\left[(5\frac{\pi}{180})^2[1,1,1],\ 0.01^2[1,1,1]\right].$$

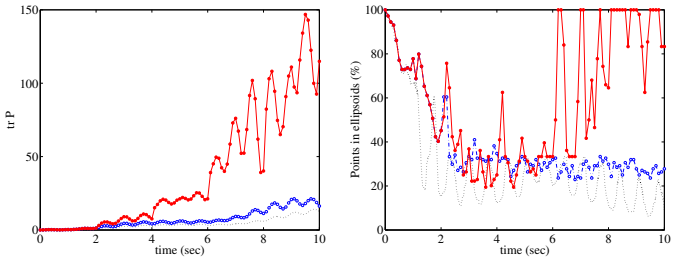
It is convenient for visualization to plot the flow of the reduced attitude $R^T e_3$ on the two-sphere $\mathbb{S}^2$. This is because the 3D pendulum has an $\mathbb{S}^1$ symmetry given by rotations about the vertical axis, which allows the configuration space to be reduced to the quotient space $SO(3)/\mathbb{S}^1 \simeq \mathbb{S}^2$ [9]. The reduced attitude denotes the direction of gravity in the body-fixed frame. Plots of the reduced attitude and the angular velocity responses, corresponding to the initial conditions, are shown in Fig. 1 for a time period of 10 seconds.

We apply the three methods to propagate the initial uncertainty through this oscillatory attitude flow for 10 seconds. The integration step size is $h = 0.005$, corresponding to 2000 time steps over the length of the simulation. For the unscented method, we find the minimal volume covering



(a) Reduced attitude Γ (b) Angular velocity Ω

Fig. 1. Reduced attitude and angular velocity for oscillatory attitude flow (The center of the sphere is the hanging equilibrium where $R^T e_3 \quad e_3$.)



(a) Magnitude of uncertainty matrix (b) Percentage of sample points contained in the estimated uncertainty (%)

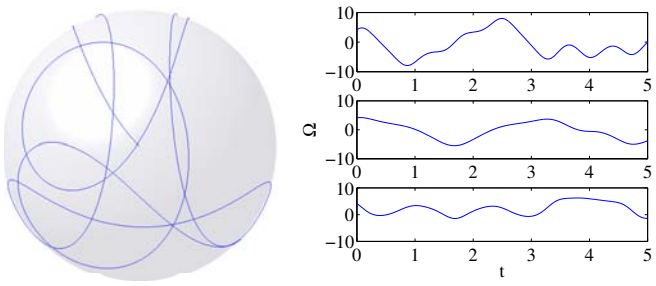
Fig. 2. Uncertainty propagation through oscillatory attitude flow (Linearization: dotted, Unscented: blue, Unscented with re-sampling: red)

ellipsoid every 0.1 seconds. For the unscented method with re-sampling, we find the minimal volume covering ellipsoid every 0.1 seconds, and we choose new set of sample points every 2.0 seconds. To provide a baseline for comparing the performance of each method, we choose sample points in the interior of the initial uncertainty ellipsoid. In particular, we choose 144 points on the level set defined by $x^T P_0^{-1} x = 0.8$, and we numerically integrate each of them.

The magnitudes of the uncertainty matrix and the percentage of sample points (out of 144) contained in the computed uncertainty ellipsoids are shown in Fig. 2. The propagation of uncertainties in the reduced attitude are shown in Fig. 6.

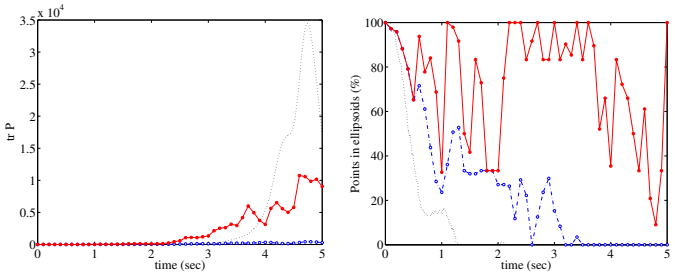
For a short time period, the properties of all methods are similar. The uncertainty ellipsoid computed by the unscented method is larger than for the linearization method, reflecting the nonlinear nature of the dynamics. The re-sampling in the unscented method with re-sampling has the effect of enlarging the uncertainty ellipsoids. Thus, the magnitude of the uncertainty ellipsoid increases rapidly, and the uncertainty ellipsoid from the re-sampling method contains a greater proportion of the sample points after 6 seconds.

The mean precentages of sample points contained in the computed uncertainty ellipsoids are 28.9%, 39.8%, and 59.8%, respectively. This suggests that even for this oscillatory attitude flow, the nonlinear effects are so strong that it is difficult to accurately propagate the uncertainty. The computation times are 3.24, 45.18, and 45.84 seconds, respectively on an Intel Pentium M 1.73GHz processor.



(a) Reduced attitude Γ (b) Angular velocity Ω

Fig. 3. Reduced attitude and angular velocity for irregular attitude flow (The center of the sphere is the hanging equilibrium where $R^T e_3 \quad e_3$.)



(a) Magnitude of uncertainty matrix (b) Percentage of sample points contained in the estimated uncertainty (%)

Fig. 4. Uncertainty propagation through irregular attitude flow (Linearization: dotted, Unscented: blue, Unscented with re-sampling: red)

Irregular attitude flow: The initial uncertainty ellipsoid is given by

$$R_0 = I_{3\times 3}, \quad \Omega_0 = [4.14, 4.14, 4.14]\,\mathrm{rad/s},$$
$$P_0 = \mathrm{diag}\left[(5\frac{\pi}{180})^2[1,1,1],\ 0.01^2[1,1,1]\right].$$

The reduced attitude on $\mathbb{S}^2$ and the angular velocity responses for 10 seconds corresponding to the center initial conditions are shown in Fig. 3.

We apply the three methods to propagate the initial uncertainty through this irregular attitude flow for 5 seconds. The integration step size is $h = 0.002$, corresponding to 2500 time steps over the length of the simulation. For the unscented method, we find the minimal volume covering ellipsoid every 0.1 seconds. For the unscented method with re-sampling, we find the minimal volume covering ellipsoid every 0.1 seconds, and we choose new set of sample points every 0.5 seconds. To compare the properties of each method, we choose 144 initial sample points on the level set of $x^T P_0^{-1} x = 0.8$, and we numerically integrate them.

The magnitudes of the uncertainty matrix and the percentage of sample points (out of 144) contained in the computed uncertainty ellipsoid are shown in Fig. 4. The propagation of uncertainties in the reduced attitude are shown in Fig. 7.

For a short time period, the properties of all methods are similar. The uncertainty ellipsoid computed by the linearization method grows rapidly, but it encloses few points after 1.5 second. The unscented method encloses more points with smaller uncertainty ellipsoids than for the linearization method, but it encloses few points after 3.5 second. The re-sampling in the unscented method with re-sampling has the effect of enlarging the uncertainty ellipsoids. Thus, the size of the uncertainty ellipsoids increase rapidly, and the uncertainty ellipsoid contains more sample points than the other two methods.

The mean numbers of sample points contained in the computed uncertainty ellipsoid are 10.7%, 26.3%, and 73.55%, respectively. The computation times are 4.97, 43.50, and 45.56 seconds, respectively on an Intel Pentium M 1.73GHz processor.

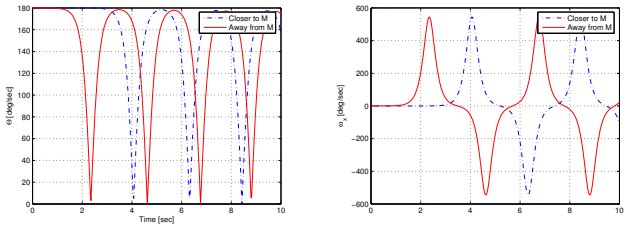
V. Global Features of the Attitude Flow

The numerical results presented in the previous section demonstrate the difficulty in obtaining accurate global bounds on attitude solutions that are initialized in an uncertainty ellipsoid. It is claimed that the source of this difficulty is the nonlinear attitude flow of the 3D pendulum, especially the fact that the flow can exhibit chaos and extreme sensitivity to initial conditions. A conceptual description of certain global features of the attitude flow is now provided.

As described in [9], the global dynamics of the 3D pendulum are complicated. There is a 1D hanging equilibrium submanifold of the 3D configuration manifold, consisting of hanging equilibria that differ by a rotation about the vertical. There is also a 1D inverted equilibrium submanifold consisting of inverted equilibria. Each hanging equilibrium is stable in the sense of Lyapunov. Each inverted equilibrium is unstable, with a 2D stable manifold, a 2D unstable manifold, and a 2D center manifold. Let M denote the union of all the 2D stable manifolds corresponding to inverted equlibria. This 3D set M plays an important role in understanding the global dynamics of the attitude flow.

Every trajectory in M converges to the inverted equilibrium manifold. Although the set M has measure zero, its existence influences the dynamics of the 3D pendulum attitude flow near M. Since M is constructed as the union of the stable manifolds of unstable equilibria, trajectories near M remain near M for an extended period of time. In particular, the closer a trajectory is to M the longer it remains near M. In fact, there are trajectories that remain close to M for arbitrarily long time periods. This property is due to the saddle character of each inverted equilibrium.

This can be illustrated by a numerical simulation of the 3D pendulum. Fig. 5 show the attitude deviation from the hanging equilibrium, measured by the quantity $\left\|S^{-1}(\log R)\right\|$, and the angular velocity response for two close initial conditions near the inverted equilibrium; the initial angular velocities are the same, and the initial attitudes only differ by 0.0066 degrees. One initial condition is chosen to be slightly nearer the set M. Note that both the attitude and the angular velocity responses have different characteristics; the solution that starts closer to M, denoted by dashed lines, remains for more than 2 seconds near the inverted equilibrium, and the angular velocity responses have almost opposite phase. These differences are further increased as the initial condition is chosen to be closer to the set M.



(a) Attitude deviation from the hanging equilibrium (b) First component of angular velocity (Ω_1)

Fig. 5. Effect of a stable manifold of an unstable equilibrium for the 3D pendulum

It should be mentioned that it is difficult to determine exactly the set M. One can make use of linear attitude equations near an inverted equilibrium to approximate the tangent space to the stable manifold of that equilibrium. However, this provides only local information about M; the non-local properties of the set M are not understood. In practice, to accurately compute the global structure of a stable manifold, one relies on either (i) extremely high-order Taylor approximations of the nonlinear stable manifold for a neighborhood of the equilibrium, which is used to obtain sample points on the stable manifold that are then propagated backwards in time in order to compute the global structure of the stable manifold [10], or (ii) set-oriented techniques based on representing the nonlinear flow map for short times as a Markov chain [11].

This argument demonstrates that the set M has a strong influence on the 3D pendulum dynamics near M, with high shearing and thus high sensitivity of the attitude flow near M. This is one of the mechanisms leading to the complex nonlinear dynamics of the 3D pendulum and makes it impossible to efficiently compute accurate global bounds on attitude solutions that are initialized in an uncertainty ellipsoid.

VI. Conclusions

The Lie group variational integrator is known to provide accurate long-term solutions of the rigid body equations in the presence of an external potential for a given initial attitude and angular velocity; these computed solutions exactly conserve the theoretical conservation properties, namely the symplectic structure, and the angular momentum component about the vertical axis in the case for the 3D pendulum. In addition, it exhibits very good energy behavior, with only a very small bounded energy oscillation, for exponentially long times. Furthermore, the Lie group variational integrator is also known to exactly conserve orthogonality of the computed attitude as a rotation matrix.

It is particularly important to use symplectic methods to propagate individual trajectories in determining uncertainty propagation in Hamiltonian systems since the Gromov non-squeezing theorem [12] from symplectic geometry places fundamental limits on how the uncertainty of a Hamiltonian mechanical system evolves [13]. Consequently, inaccuracies in the approximate ellipsoidal bounds computed according to the three computational methods introduced do not arise from computational difficulties with the Lie group variational integrator. Rather, the inaccuracies in the approximate ellipsoidal bounds arise from the fact that the attitude flow dynamics are highly nonlinear, with regions wherein the dynamics cannot be adequately approximated by linear dynamics.

With these qualifications, it is clear that the unscented method with resampling is the most accurate of the three proposed methods in propagating the uncertainty. This method can provide a basis for analysis of control and estimation problems for attitude systems such as the 3D pendulum.

For example, attitude estimation operates open loop between measurement times; the analysis in this paper demonstrates the importance of the choice of inter-measurement time in obtaining accurate propagation of the attitude flow between measurement times.

The bottom line demonstrated by the development in this paper is that guaranteed global bounds for attitude dynamics defined by the 3D pendulum, or indeed for any attitude dynamics with a nontrivial potential, are not achievable. That is, there is no universal approach to global and uniform approximation of the attitude flow dynamics.

References

[1] J. Shen, A. Sanyal, N. Chaturvedi, D. Bernstein, and N. McClamroch, "Dynamics and control of a 3D pendulum," *Proceedings of the IEEE Conference on Decision and Control*, pp. 323–328, 2004.

[2] T. Lee, A. Sanyal, M. Leok, and N. H. McClamroch, "Deterministic global attitude estimation," in *Proceedings of the IEEE Conference on Decision and Control*, Dec 2006, pp. 3174–3179.

[3] A. Sanyal, T. Lee, M. Leok, and N. H. McClamroch, "Global optimal attitude estimation using uncertainty ellipsoids," *Systems and Control Letters*, 2007, accepted. [Online]. Available: http://arxiv.org/abs/math.OC/0606083

[4] M. Kumar, S. Chakravorty, and J. L. Junkins, "A homotopic approach to domain determination and solution refinement for the Fokker-Planck equation," in *Proceedings of the American Control Conference*, Jul 2007, pp. 1045–1050.

[5] Y. Theodor, U. Shaked, and C. E. de Souza, "A game theory approach to robust discrete-time H_∞-estimation," *IEEE Transactions on Signal Processing*, vol. 42, no. 6, pp. 1486–1495, 1994.

[6] T. Lee, M. Leok, and N. H. McClamroch, "A Lie group variational integrator for the attitude dynamics of a rigid body with application to the 3D pendulum," in *Proceedings of the IEEE Conference on Control Application*, Aug 2005, pp. 962–967.

[7] ——, "Attitude maneuvers of a rigid spacecraft in a circular orbit," in *Proceedings of the American Control Conference*, Jun 2006, pp. 1742–1747.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[9] N. Chaturvedi, T. Lee, M. Leok, and N. H. McClamroch, "Nonlinear dynamics of the 3D pendulum," *SIAM Journal on Applied Dynamical Systems*, 2007, submitted. [Online]. Available: http://arxiv.org/abs/0707.1196

[10] G. Gómez, W. S. Koon, M. W. Lo, J. E. Marsden, J. Masdemont, and S. D. Ross, "Connecting orbits and invariant manifolds in the spatial restricted three-body problem," *Nonlinearity*, vol. 17, no. 5, pp. 1571–1606, 2004.

[11] M. Dellnitz, G. Froyland, and O. Junge, "The algorithms behind GAIO-set oriented numerical methods for dynamical systems," in *Ergodic theory, analysis, and efficient simulation of dynamical systems*. Berlin: Springer, 2001, pp. 145–174, 805–807.

[12] M. Gromov, "Pseudo holomorphic curves in symplectic manifolds," *Invent. Math.*, vol. 82, no. 2, pp. 307–347, 1985.

[13] F. Hsiao and D. J.Scheeres, "Fundamental constraints on uncertainty evolution in Hamiltonian systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 4, pp. 686–691, 2007.

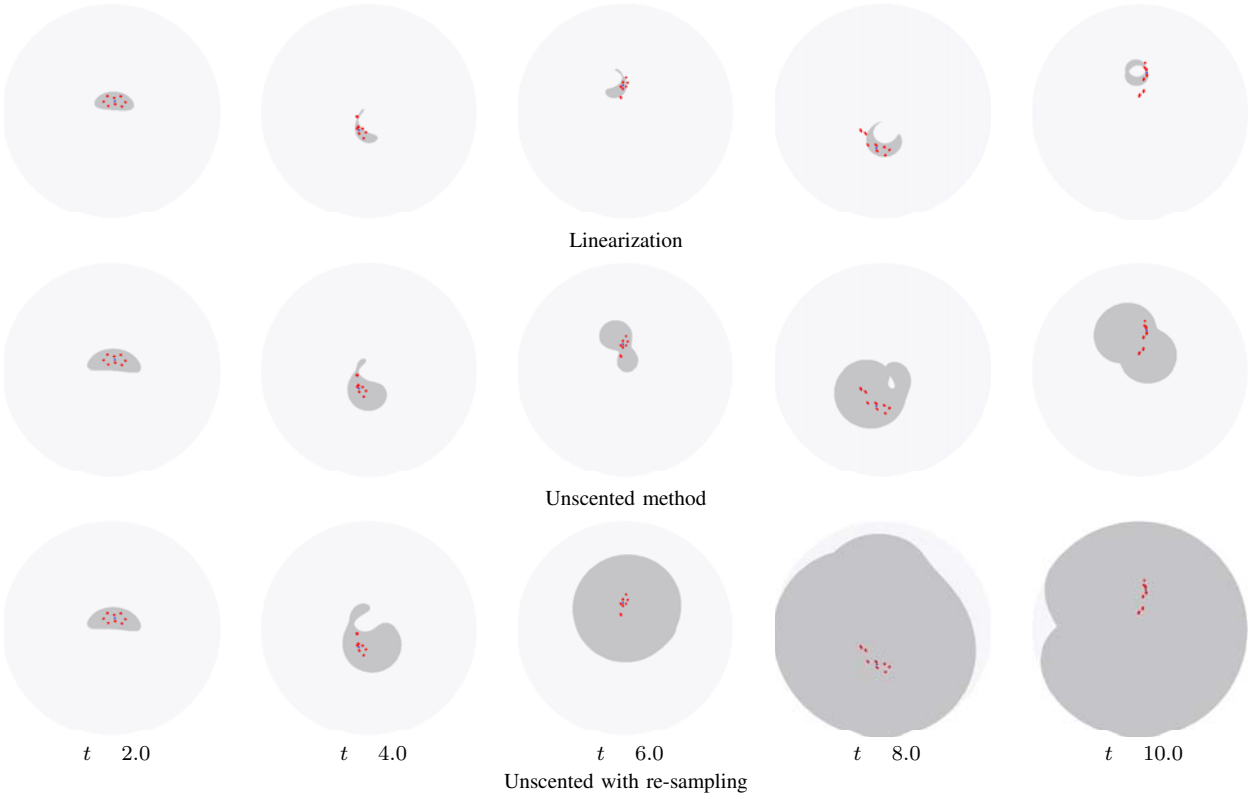


Fig. 6. Uncertainty projected onto the reduced attitude on $\mathbb{S}^2$ for oscillatory attitude flow (The center of the sphere is the hanging equilibrium where $R^T e_3 \quad e_3$.)

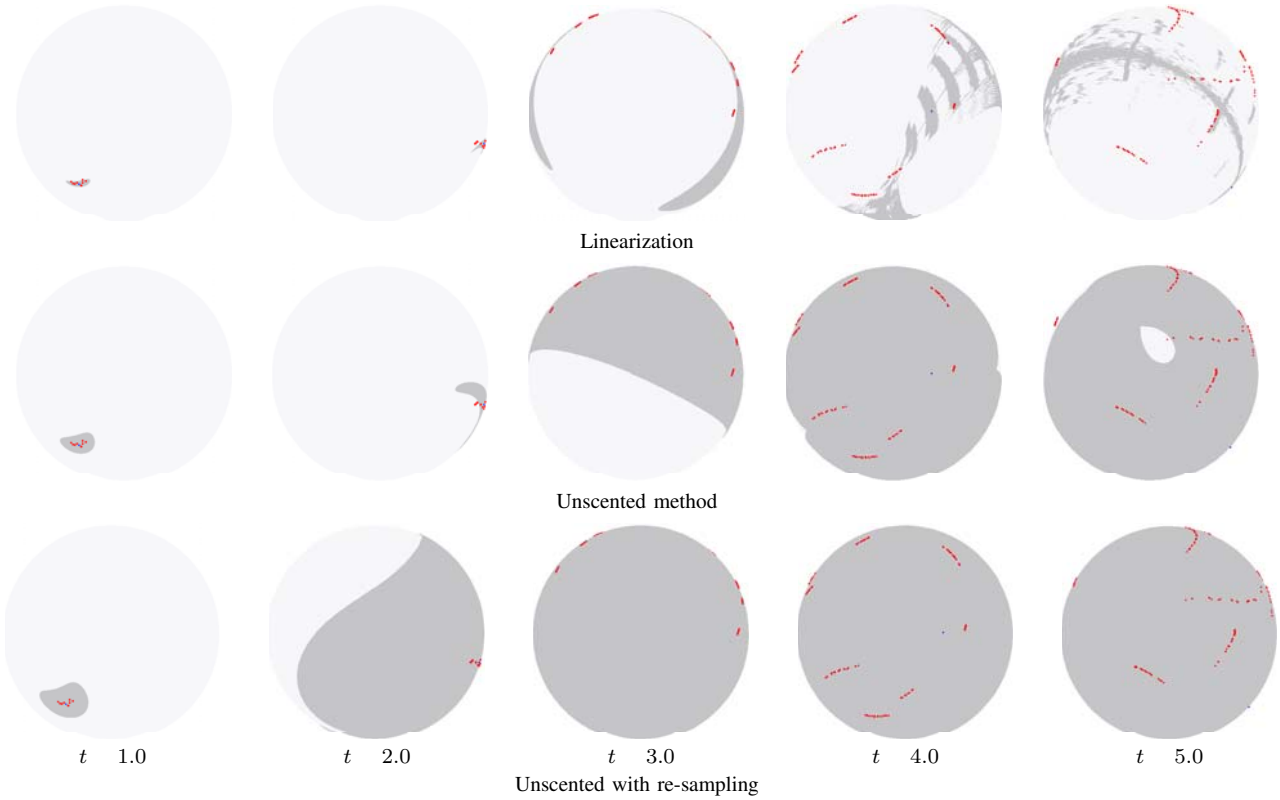


Fig. 7. Uncertainty projected onto the reduced attitude on $\mathbb{S}^2$ for irregular attitude flow (The center of the sphere is the hanging equilibrium where $R^T e_3 \quad e_3$.)

Time Optimal Attitude Control for a Rigid Body

Taeyoung Lee*†, Melvin Leok*, and N. Harris McClamroch†

***Abstract*— A time optimal control problem is studied for the attitude dynamics of a rigid body. The objective is to minimize the maneuver time to rotate the rigid body to a desired attitude and angular velocity while subject to constraints on the control input. Necessary conditions for optimality are developed directly on the special orthogonal group using rotation matrices. They completely avoid singularities associated with local parameterizations such as Euler angles, and they are expressed as compact vector equations. In addition, a discrete-time control method based on a geometric numerical integrator, referred to as a Lie group variational integrator, is proposed to compute the optimal control input that respects the underlying geometric properties of the rigid body. The proposed method is demonstrated by a large-angle time optimal maneuver for an elliptic cylinder rigid body.**

I. Introduction

The time optimal control of spacecraft has received consistent interest as rapid attitude maneuvers are critical to various space missions such as military observation and satellite communication. The objective is to reorient the attitude of the spacecraft in a minimal maneuver time with constrained control moments. To accomplish many space missions, large-angle attitude maneuvering capabilities are required.

Time optimal attitude maneuvers have been extensively studied in the literature [1]. The time optimal solution is found for a single degree of freedom system, where the attitude maneuver is constrained to an eigen-axis rotation, in [2]. It is known that the eigen-axis rotation is not generally time optimal [3], [4]. The attitude dynamics is often simplified in an optimality analysis, e.g., by assuming an inertially symmetric rigid body model [3], [4], [5], linearization [6] and constant magnitude angular velocity [5].

The attitude is defined as the orientation of a body-fixed frame with respect to a reference frame, and it is represented by a rotation matrix that lies in the special orthogonal group, $\mathrm{SO}(3)$. However, most existing optimal control schemes for the dynamics of a rigid body uses coordinate representations such as Euler angles or quaternions. The minimal attitude representations, such as Euler angles and Rodrigues parameters, have singularities, so they are not desirable for large-angle maneuvers. The non-minimal attitude representations, such as quaternions, have associated problems; besides the unit norm constraint, the quaternion representation double covers $\mathrm{SO}(3)$. So, it has an inevitable ambiguity in describing attitude. There are several possible representations for the same boundary condition, and they may result in different computational properties in finding the optimal control input.

The objective of this paper is to solve the time optimal attitude control problem directly on $\mathrm{SO}(3)$ using rotation matrices without need of any attitude parameterization. Using a specific property of the special orthogonal group, namely that the Lie algebra $\mathfrak{so}(3)$ which consists of skew-symmetric matrices is isomorphic to $\mathbb{R}^3$, necessary conditions for optimality are developed and represented as vector equations on $\mathbb{R}^3$. This avoids singularities completely, and the resulting necessary optimality conditions are more compact than expressions obtained in terms of quaternions.

The remainder of this paper is focused on developing a computational approach to solve this optimal control problem. The dynamics of a rigid body have certain geometric features; in addition to the configuration space being a Lie group, the dynamics are characterized by symplectic, momentum and energy preserving properties. Most commonly used numerical integration methods, including the widely used (non-symplectic) explicit Runge–Kutta schemes, do not preserve these geometric properties.

Lie group variational integrators are geometric numerical integrators that preserve these geometric features of the rigid body dynamics [7]. Based on this structure-preserving numerical integrator, computational approaches have been proposed to solve various optimal control problems for the dynamics of rigid bodies [8], [9]. In this paper, the time optimal attitude control problem is discretized at the level of the initial problem formulation, and discrete necessary conditions for optimality are developed. This provides a geometrically exact but computationally efficient approach.

This paper is organized as follows. The time optimal attitude control problem is formulated, and continuous-time necessary conditions for optimality are developed in Section II. In a parallel fashion, a discrete-time optimal control method is presented in Section III, followed by numerical examples in Section IV.

II. Time Optimal Attitude Control

A. Equations of Motion

We consider the attitude dynamics of a rigid body. The configuration space is the special orthogonal group

$$\mathrm{SO}(3) = \left\{ R \in \mathbb{R}^{3\times 3} \,\middle|\, R^T R = I_{3\times 3}, \quad \det R = 1 \right\},$$

where the rotation matrix $R \in \mathrm{SO}(3)$ represents the linear transformation from the body-fixed frame to the inertial frame.

Taeyoung Lee, and N. Harris McClamroch, Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 {tylee,nhm}@umich.edu

Melvin Leok, Mathematics, Purdue University, West Lafayette, IN 47907 mleok@math.purdue.edu

*This research has been supported in part by NSF under grants DMS-0504747 and DMS-0726263.

†This research has been supported in part by NSF under grant CMS-0555797.

The continuous equations of motion for the attitude dynamics of a rigid body are given by

$$J\dot{\Omega} + \Omega \times J\Omega = u, \tag{1}$$

$$\dot{R} = R\hat{\Omega}, \tag{2}$$

where the matrix $J \in \mathbb{R}^{3\times3}$ is the inertia matrix, the vector $\Omega \in \mathbb{R}^3$ is the angular velocity expressed in the body-fixed frame, and the external control moment is denoted by $u \in \mathbb{R}^3$. The *hat map* $\hat{\cdot} : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is an isomorphism from $\mathbb{R}^3$ to skew-symmetric matrices $\mathfrak{so}(3)$, and is defined by the condition $\hat{x}y = x \times y$ for all $x, y \in \mathbb{R}^3$. The inverse map is denoted by the *vee map* $(\cdot)^\vee : \mathfrak{so}(3) \mapsto \mathbb{R}^3$.

B. Time Optimal Attitude Control Problem

The objective of the time optimal attitude control problem is to transfer a given initial attitude and angular velocity $(R_\circ, \Omega_\circ)$ of the rigid body to desired values (R_f, Ω_f) within a minimal maneuver time t_f with constrained control moment $\|u\| \le \overline{u}$ for a given control limit $\overline{u} \in \mathbb{R}$.

$$\begin{gathered}\text{For given: } (R_\circ, \Omega_\circ), (R_f, \Omega_f), \bar{u}\\ \min_u \left\{ \mathcal{J} = \int_0^{t_f} 1\, dt \right\},\\ \text{such that } R(t_f) = R_f,\ \Omega(t_f) = \Omega_f,\\ \text{subject to } \|u(t)\| \le \bar{u}\ \ \forall t \in [0, t_f] \text{ and } (1), (2).\end{gathered}$$

C. Necessary Conditions for Optimality

We solve this optimal control problem using variational principles applied on $\mathrm{SO}(3)$. Expressions for variations of a rotation matrix, and transversality conditions are presented, and necessary conditions for optimality are developed.

Expressions for variations: We represent a variation of a rotation matrix using the exponential map, $\exp : \mathfrak{so}(3) \mapsto \mathrm{SO}(3)$

$$R^\epsilon = R \exp \epsilon\hat{\eta}, \tag{3}$$

where $\epsilon \in (-c, c)$ for $c > 0$, and $\hat{\eta} \in \mathfrak{so}(3)$ for $\eta \in \mathbb{R}^3$. Since the exponential map is a local diffeomorphism, this expression is well-defined for some constant c for given $\hat{\eta}$. The infinitesimal variation of the rotation matrix is given by

$$\delta R = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} R \exp \epsilon\hat{\eta} = R\hat{\eta}. \tag{4}$$

The infinitesimal variation of $R^T\dot{R}$ is obtained from (2) and (4) as

$$\begin{aligned}\delta(R^T\dot{R}) &= \delta R^T\dot{R} + R^T\delta\dot{R} = -\hat{\eta}R^T\dot{R} + R^T(\dot{R}\hat{\eta} + R\hat{\dot{\eta}})\\ &= \hat{\dot{\eta}} + \hat{\Omega}\hat{\eta} - \hat{\eta}\hat{\Omega} = (\dot{\eta} + \Omega \times \eta)\widehat{\ }.\end{aligned} \tag{5}$$

The variational expressions given by (4) and (5) are the key ingredients to developing necessary conditions for optimality for an arbitrary optimal attitude maneuver.

Transversality conditions: The differentials in the terminal attitude and the terminal angular velocity are composed of the variation for a fixed time and a term due to the terminal time variation. Since the terminal boundary conditions are fixed, the transversality conditions are

$$\delta R(t_f) + \dot{R}(t_f)dt_f = R(t_f)\hat{\eta}(t_f) + \dot{R}(t_f)dt_f = 0, \tag{6}$$

$$\delta\Omega(t_f) + \dot{\Omega}(t_f)dt_f = 0. \tag{7}$$

Necessary conditions for optimality: Define the augmented performance index as

$$\begin{aligned}\mathcal{J}_a = \int_0^{t_f} &1 + \lambda^\Omega \cdot (u - \Omega \times J\Omega - J\dot{\Omega})\\ &+ \lambda^R \cdot (\hat{\Omega} - R^T\dot{R})^\vee\, dt,\end{aligned}$$

where $\lambda^\Omega, \lambda^R \in \mathbb{R}^3$ are Lagrange multipliers.

Using (5), the infinitesimal variation of the augmented performance index is given by

$$\begin{aligned}\delta\mathcal{J}_a = \int_0^{t_f} &\lambda^\Omega \cdot (\delta u - \delta\Omega \times J\Omega - \Omega \times J\delta\Omega - J\delta\dot{\Omega})\\ &+ \lambda^R \cdot (\delta\Omega - \dot{\eta} - \Omega \times \eta)\, dt\\ + \left\{1 + \lambda^\Omega \cdot (u - \Omega \times J\Omega - J\dot{\Omega}) + \lambda^R \cdot (\hat{\Omega} - R^T\dot{R})^\vee\right\}&\Big|_{t_f} dt_f.\end{aligned}$$

Using integration by parts, we obtain

$$\begin{aligned}\delta\mathcal{J}_a = \int_0^{t_f} &\lambda^\Omega \cdot (\delta u - \delta\Omega \times J\Omega - \Omega \times J\delta\Omega) + \dot{\lambda}^\Omega \cdot J\delta\Omega\\ &+ \lambda^R \cdot (\delta\Omega - \Omega \times \eta) + \dot{\lambda}^R \cdot \eta\, dt\\ &- \{\lambda^\Omega \cdot J\delta\Omega + \lambda^R \cdot \eta\}\Big|_0^{t_f}\\ + \left\{1 + \lambda^\Omega \cdot (u - \Omega \times J\Omega - J\dot{\Omega}) + \lambda^R \cdot (\hat{\Omega} - R^T\dot{R})^\vee\right\}&\Big|_{t_f} dt_f.\end{aligned}$$

Since the initial attitude and the initial angular velocity are fixed, we have $\eta(0) = 0$, $\delta\Omega(0) = 0$. Substituting and rearranging, the infinitesimal variation of the augmented performance index is given by

$$\begin{aligned}\delta\mathcal{J}_a = \int_0^{t_f} &\delta\Omega \cdot \{-J\Omega \times \lambda^\Omega - J(\lambda^\Omega \times \Omega) + J\dot{\lambda}^\Omega + \lambda^R\}\\ &+ \eta \cdot \left\{\Omega \times \lambda^R + \dot{\lambda}^R\right\} + \delta u \cdot \lambda^\Omega\, dt\\ &+ \left\{1 + \lambda^\Omega \cdot (u - \Omega \times J\Omega) + \lambda^R \cdot \Omega\right\}\Big|_{t_f} dt_f.\end{aligned}$$

We choose multiplier equations and boundary conditions such that the expressions in the braces in the above equations are identically zero. Then, we have

$$\delta\mathcal{J}_a = \int_0^{t_f} \delta u \cdot \lambda^\Omega\, dt.$$

The optimal control input u must satisfy

$$\lambda^\Omega \cdot \delta u \ge 0, \tag{8}$$

for all admissible δu in $t \in [0, t_f]$. If $\lambda^\Omega = 0$ on a finite time period, the control input is not determined by (8). Such solutions are referred to as singular arcs. Later, it is shown that there is no singular arc in this optimal control problem.

In summary, the necessary conditions for optimality are given by

- Multiplier equations

$$J\dot{\lambda}^{\Omega} + J(\Omega \times \lambda^{\Omega}) - J\Omega \times \lambda^{\Omega} + \lambda^{R} = 0, \tag{9}$$

$$\dot{\lambda}^{R} + \Omega \times \lambda^{R} = 0, \tag{10}$$

- Optimality condition

$$u = -\bar{u}\left(\lambda^{\Omega} / \left\|\lambda^{\Omega}\right\|\right), \tag{11}$$

- Boundary and transversality conditions

$$(R(0), \Omega(0)) = (R_{\circ}, \Omega_{\circ}), \tag{12}$$

$$(R(t_f), \Omega(t_f)) = (R_f, \Omega_f), \tag{13}$$

$$\left\{1 + \lambda^{\Omega} \cdot (u - \Omega \times J\Omega) + \lambda^{R} \cdot \Omega\right\}\Big|_{t_f} = 0, \tag{14}$$

Assuming that the rigid body is inertially symmetric, $J = I_{3\times 3}$, the multiplier equation (9) is reduced to $\dot{\lambda}^{\Omega} + \lambda^{R} = 0$.

Remarks: Use of the rotation matrix is often avoided, since it is thought that representing a 3-dimensional attitude using 9 real elements with 6 constraints is inefficient. This redundancy is eliminated by using the exponential map; in practice, the analytical development is carried out in $\mathbb{R}^3$ and the resulting necessary conditions are written as compact vector equations on $\mathbb{R}^3$. These equations are more compact than the necessary conditions expressed in terms of quaternions. The proposed necessary conditions have the benefit that the rotation matrix has no singularity or ambiguity.

D. Singular arcs

In this subsection, we show that singular arcs do not exist along a solution of this time optimal control problem. Suppose that there exists a singular interval, i.e. $\lambda^{\Omega}(t) = 0$ for a finite time period in $[0, t_f]$. Then, the minimum principle given by (8) does not lead to a well-defined condition for the optimal control input. Instead, the control input is determined by the requirement that the time derivative of λ^{Ω} is zero.

Let an integer q be the order of the singular arc [10]. In other words, the $2q$-th time derivative of λ^{Ω} is the lowest order derivative in which the control input u appears explicitly with a coefficient that is not identically zero on the singular interval. Here, due to the special linear structure of this multiplier equation, the singular arc must have infinite order. If the condition $\lambda^{\Omega} = \dot{\lambda}^{\Omega} = 0$ is satisfied at a single point along the trajectory, $\lambda^{R} = \dot{\lambda}^{R} = 0$, and these are satisfied identically throughout the trajectory. In this case, it is clear that the boundary condition (14) cannot be satisfied. Thus, there is no singular arc.

III. Discrete-time Time Optimal Attitude Control

In this section, we present a computational approach, referred to as discrete optimal control of discrete Lagrangian systems [11], to solve the time optimal attitude control problem numerically. In this approach, the dynamics of the rigid body is discretized using the discrete Hamilton's principle, in order to obtain a Lie group variational integrator [7]. The corresponding discrete equations of motion are imposed as dynamic constraints to be satisfied by using Lagrange multipliers, and necessary conditions for optimality, expressed as discrete-time multiplier equations, are obtained.

This method yields substantial computational advantages in finding an optimal control solution. The discrete-time dynamics are faithful to the continuous equations of motion, and consequently more accurate solutions to the optimal control problem are obtained. It has been shown that the discrete-time dynamics is more reliable even for controlled system as it computes the energy dissipation rate more accurately [12]. In particular, the discrete flow of the Lie group variational integrator remains on $\mathrm{SO}(3)$.

Optimal solutions, computed using an indirect approach, are usually sensitive to small variations of the multipliers. This causes difficulties, such as numerical ill-conditioning, when solving the necessary conditions for optimality expressed as a two-point boundary value problem. Sensitivity derivatives, computed using the discrete-time necessary conditions, are not corrupted by numerical dissipation caused by conventional numerical integration schemes. Thus, the proposed computational approach is numerically robust, and the necessary conditions can be solved efficiently.

A. Lie Group Variational Integrator

Since the dynamics of a rigid body has the structure of a Lagrangian or Hamiltonian system, they are symplectic, momentum and energy preserving. These geometric features determine the qualitative behavior of the rigid body dynamics.

In contrast, the most common numerical integration methods, including the widely used (non-symplectic) explicit Runge–Kutta schemes, preserve neither the Lie group structure nor these geometric properties. Additionally, if we integrate (2) using a typical Runge–Kutta scheme, the quantity $R^T R$ inevitably drifts from the identity matrix as the simulation time increases.

In [7], Lie group variational integrators are constructed by explicitly adapting Lie group methods [13] to the discrete variational principle [12]. They have the desirable property that they are symplectic and momentum preserving, and they exhibit good energy behavior for an exponentially long time period. They also preserve the Lie group structure. These geometrically exact numerical integration methods yield highly efficient and accurate computational algorithms for rigid body dynamics, and avoid singularities and ambiguities.

Using the results presented in [7], a Lie group variational integrator on $\mathrm{SO}(3)$ for equations (1), (2) is given by

$$h\widehat{J\Omega_k} = F_k J_d - J_d F_k^T, \tag{15}$$

$$R_{k+1} = R_k F_k, \tag{16}$$

$$J\Omega_{k+1} = F_k^T J\Omega_k + h u_{k+1}, \tag{17}$$

where the subscript k denotes the k-th step for a fixed integration step size $h \in \mathbb{R}$. The matrix $J_d \in \mathbb{R}^{3\times 3}$ is a nonstandard inertia matrix defined by $J_d = \frac{1}{2}\mathrm{tr}[J] I_{3\times 3} - J \in \mathbb{R}^{3\times 3}$. The matrix $F_k \in \mathrm{SO}(3)$ denotes the relative attitude between adjacent integration steps.

For given (R_k, x_k) and control input, (15) is solved to find F_k. Then (R_{k+1}, Ω_{k+1}) are obtained by (16) and (17). This yields a map $(R_k, \Omega_k) \mapsto (R_{k+1}, \Omega_{k+1})$, and this process is repeated. The only implicit part is (15). In essence, the attitude update F_k is computed in $\mathbb{R}^3$; we represent F_k in terms of the Lie algebra $\mathfrak{so}(3) \simeq \mathbb{R}^3$ by using the Cayley transformation, and we rewrite (15) as an equivalent vector equation, to which Newton iteration is applied. Since F_k represents a small attitude update near the identity matrix, it can be easily computed by using the Cayley transformation. This approach avoids the 6 dimensional nonlinear constraints on the rotation matrix, as one numerically solves an equivalent vector equation in $\mathbb{R}^3$.

One of the distinct features of the Lie group variational integrator is that it preserves both the symplectic property and the Lie group structure of the rigid body dynamics. As such, it exhibits substantially improved computational accuracy and efficiency compared with other geometric integrators that preserve only one of these properties such as non-symplectic Lie group methods [14]. The symplectic property is important even in the case of controlled dynamics, since the dissipation rate of the total energy is typically computed inaccurately by non-symplectic integrators [12].

B. Discrete-time Time Optimal Attitude Control Problem

The objective is to transfer the rigid body in a prescribed way within a minimal discrete maneuver time N.

$$\text{For given: } (R_\circ, \Omega_\circ), (R_f, \Omega_f), \bar{u}$$
$$\min_{u_{k+1}} \left\{ \mathcal{J} = \sum_{k=0}^{N-1} 1 \right\},$$
$$\text{such that } R_N = R_f,\ \Omega_N = \Omega_f,$$
$$\text{subject to } \|u_{k+1}\| \le \bar{u}\ \ \forall k \in [0, N-1] \text{ and } (15)-(17).$$

C. Discrete-Time Necessary Conditions for Optimality

Expressions for variations: Similar to (4), the variation of rotation matrices R_k and F_k are expressed as

$$\delta R_k = R_k \hat{\eta}_k, \quad \delta F_k = F_k \hat{\xi}_k \tag{18}$$

for $\eta_k, \xi_k \in \mathbb{R}^3$. Using this and (16), the variation of $R_k^T R_{k+1}$ is given by

$$\begin{aligned}\delta(R_k^T R_{k+1}) &= \delta R_k^T R_{k+1} + R_k^T \delta R_{k+1} \\ &= -\hat{\eta} F_k + F_k \hat{\eta}_{k+1} \\ &= F_k(-F_k^T \eta_k + \eta_{k+1})\widehat{\ }, \end{aligned} \tag{19}$$

where the property $\widehat{F^T x} = F^T \hat{x} F$ for any $x \in \mathbb{R}^3$ and $F \in \mathrm{SO}(3)$ is used in the last step.

Now we develop an expression for a constrained variation corresponding to (15). Taking a variation of (15), we obtain

$$h\widehat{J\delta\Omega_k} = F_k \hat{\xi}_k J_d + J_d \hat{\xi}_k F_k^T.$$

Using the property, $\hat{x}A + A^T \hat{x} = (\{\mathrm{tr}[A]\, I_{3\times 3} - A\}\, x)^\wedge$ for all $x \in \mathbb{R}^3$, $A \in \mathbb{R}^{3\times 3}$, the above equation can be written as

$$\begin{aligned} h J \delta\hat{\Omega}_k &= \widehat{F_k \xi_k} F_k J_d + J_d F_k^T \widehat{F_k \xi_k} \\ &= (\{\mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d\}\, F_k \xi_k)^\wedge. \end{aligned}$$

Thus, the vector ξ_k is expressed in terms of $\delta\Omega_k$ as

$$\xi_k = \mathcal{B}_k J \delta\Omega_k, \tag{20}$$

where $\mathcal{B}_k = h F_k^T \{\mathrm{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d\}^{-1} \in \mathbb{R}^{3\times 3}$. This shows the relationship between $\delta\Omega_k$ and δF_k.

Transversality conditions: Similar to (7), we choose the transversality conditions for the angular velocity as

$$\delta\Omega_N + (\Omega_N - \Omega_{N-1})\delta N = 0. \tag{21}$$

The variation of the terminal attitude due to the terminal time change is expressed as

$$\begin{aligned} R_N &\left\{ \frac{1}{2} R_{N-1}^T (R_N - R_{N-1}) + \frac{1}{2} R_N^T (R_N - R_{N-1}) \right\} \delta N \\ &= \frac{1}{2} R_N \left\{ F_{N-1} - F_{N-1}^T \right\} \delta N. \end{aligned}$$

This expression is chosen such that it respects the skew-symmetry of a Lie algebra $\mathfrak{so}(3)$ element. Using this, the transversality conditions for the attitude are given by

$$R_N \hat{\eta}_N + \frac{1}{2} R_N \left\{ F_{N-1} - F_{N-1}^T \right\} \delta N = 0. \tag{22}$$

Necessary conditions for optimality: Define the augmented performance index as

$$\begin{aligned} \mathcal{J}_a = \sum_{k=0}^{N-1} & 1 + \lambda_k^\Omega \cdot \left\{ -J\Omega_{k+1} + F_k^T J \Omega_k + h u_{k+1} \right\} \\ & + \lambda_k^R \cdot \frac{1}{2} \left((F_k - F_k^T)^\vee - (R_k^T R_{k+1} - R_{k+1}^T R_k)^\vee \right). \end{aligned}$$

Here we assume that the time step size h is small so that the relative attitude rotation between adjacent integration steps is less than $\frac{\pi}{2}$, i.e. $\|(\log F_k)^\vee\| < \frac{\pi}{2}$. Then, F_k is equal to $R_k^T R_{k+1}$ if and only if their skew parts are identical, which can be easily shown using Rodrigues' formula. Equation (15) is imposed implicitly using a constrained variation.

Using (19), the infinitesimal variation of the augmented performance index is given by

$$\begin{aligned} \delta\mathcal{J}_a = & \sum_{k=0}^{N-1} \lambda_k^\Omega \cdot \left\{ h\delta u_{k+1} - J\delta\Omega_{k+1} + \delta F_k^T J \Omega_k + F_k^T J \delta\Omega_k \right\} \\ & + \lambda_k^R \cdot \frac{1}{2} \Big\{ F_k (\xi_k + F_k^T \eta_k - \eta_{k+1})^\wedge \\ & \qquad + (\xi_k + F_k^T \eta_k - \eta_{k+1})^\wedge F_k^T \Big\}^\vee \\ & + \{1 + \lambda_{N-1}^\Omega \cdot \left\{ -J\Omega_N + F_{N-1}^T J \Omega_{N-1} + h u_N \right\}\} \delta N \\ & + \lambda_{N-1}^R \cdot \frac{1}{2} (F_{N-1} - F_{N-1}^T)^\vee \delta N \\ & - \lambda_{N-1}^R \cdot \frac{1}{2} (R_{N-1}^T R_N - R_N^T R_{N-1})^\vee \delta N. \end{aligned}$$

Several algebraic manipulations are required here; (i) using the property $\hat{x}A + A^T\hat{x} = (\{\mathrm{tr}[A]\, I_{3\times 3} - A\}\, x)^\wedge$ for all $x \in \mathbb{R}^3$ and $A \in \mathbb{R}^{3\times 3}$, the expression in the second braces is written as a vector form, (ii) equation (20) is substituted to express ξ_k in terms of $\delta\Omega_k$, and (iii) using the fact that $\eta_0 = 0$, $\delta\Omega_0 = 0$, the summation indices for the variables at

the $k+1$-th step are rewritten, which is the discrete analog of integration by parts. Then, we obtain

$$\begin{aligned}
\delta\mathcal{J}_a = &\sum_{k=0}^{N-1} \lambda_k^\Omega \cdot h\delta u_{k+1} \\
&+ \sum_{k=1}^{N-1} \delta\Omega_k \cdot \Big\{ -J\lambda_{k-1}^\Omega + J(F_k - \mathcal{B}_k^T \widehat{F_k^T J\Omega_k})\lambda_k^\Omega \\
&\qquad + \frac{1}{2} J\mathcal{B}_k^T (\mathrm{tr}[F_k]\, I - F_k)\lambda_k^R \Big\} \\
&+ \sum_{k=1}^{N-1} \eta_k \cdot \Big\{ \frac{1}{2}(\mathrm{tr}[F_{k-1}]\, I - F_{k-1})\lambda_{k-1}^R \\
&\qquad - \frac{1}{2} F_k (\mathrm{tr}[F_k]\, I - F_k)\lambda_k^R \Big\} \\
&- \lambda_{N-1}^\Omega \cdot J\delta\Omega_N - \lambda_{N-1}^R \cdot \frac{1}{2}(\mathrm{tr}[F_{N-1}]\, I - F_{N-1}^T)\eta_N \\
&+ \{1 + \lambda_{N-1}^\Omega \cdot \{-J\Omega_N + F_{N-1}^T J\Omega_{N-1} + hu_N\}\}\delta N \\
&+ \lambda_{N-1}^R \cdot \frac{1}{2}(F_{N-1} - F_{N-1}^T)^\vee \delta N \\
&- \lambda_{N-1}^R \cdot \frac{1}{2}(R_{N-1}^T R_N - R_N^T R_{N-1})^\vee \delta N. \qquad (23)
\end{aligned}$$

Substituting the transversality conditions (21) and (22), all of the expressions in the last four lines of the above equation are reduced to

$$\begin{aligned}
\Big\{ &1 + \lambda_{N-1}^\Omega \cdot \{-J\Omega_{N-1} + F_{N-1}^T J\Omega_{N-1} + hu_N\} \\
&+ \lambda_{N-1}^R \cdot \frac{1}{4}\left((F_{N-1})^2 - (F_{N-1}^T)^2\right)^\vee \Big\}\delta N. \qquad (24)
\end{aligned}$$

We choose discrete multiplier equations such that the expressions in the first two braces in (23) are identically zero, and we choose transversality condition such that the expression given by (24) is equal to zero. Then, we have

$$\delta\mathcal{J}_a = \sum_{k=0}^{N-1} \lambda_k^\Omega \cdot h\delta u_{k+1}.$$

The optimal control input u_{k+1} must satisfy

$$\lambda_k^\Omega \cdot \delta u_{k+1} \geq 0,$$

for all admissible δu_{k+1} and $k \in \{0, \cdots, N-1\}$. There is no singular arc in the optimal control problem as presented in Section II-D. In summary, the discrete-time necessary conditions for optimality are given by

- Multiplier equations

$$\begin{aligned}
&-J\lambda_{k-1}^\Omega + J(F_k - \mathcal{B}_k^T \widehat{F_k^T J\Omega_k})\lambda_k^\Omega \\
&\qquad + \frac{1}{2} J\mathcal{B}_k^T (\mathrm{tr}[F_k]\, I - F_k)\lambda_k^R = 0, \qquad (25)
\end{aligned}$$

$$(\mathrm{tr}[F_{k-1}]\, I - F_{k-1})\lambda_{k-1}^R - F_k(\mathrm{tr}[F_k]\, I - F_k)\lambda_k^R = 0. \qquad (26)$$

- Optimality condition

$$u_{k+1} = -\bar{u}\left(\lambda_k^\Omega / \left\|\lambda_k^\Omega\right\|\right) \qquad (27)$$

- Boundary and transversality conditions

$$(R_0, \Omega_0) = (R_\circ, \Omega_\circ), \qquad (28)$$

$$(R_N, \Omega_N) = (R_f, \Omega_f), \qquad (29)$$

$$\begin{aligned}
&1 + \lambda_{N-1}^\Omega \cdot \{-J\Omega_{N-1} + F_{N-1}^T J\Omega_{N-1} + hu_N\} \\
&\quad + \lambda_{N-1}^R \cdot \frac{1}{4}\left((F_{N-1})^2 - (F_{N-1}^T)^2\right)^\vee = 0. \qquad (30)
\end{aligned}$$

In the above equations, the only implicit part is (15). For a given initial condition $\{(R_0, \Omega_0), (\lambda_0^R, \lambda_0^\Omega)\}$, we solve (15) to obtain F_0, and we find the control input u_1 by (27). Then, (R_1, Ω_1) are obtained by (16) and (17). Using Ω_1, we solve (15) to obtain F_1. Finally, $(\lambda_1^R, \lambda_1^\Omega)$ are obtained by (26) and (25). This yields a map $\{(R_0, \Omega_0), (\lambda_0^R, \lambda_0^\Omega)\} \mapsto \{(R_1, \Omega_1), (\lambda_1^R, \lambda_1^\Omega)\}$, and this process is repeated.

The discrete-time necessary conditions for optimality are given as a two-point boundary value problem. This is to find the optimal discrete flow, multipliers, control input, and terminal maneuver time to satisfy the equations of motion (15)–(17), multiplier equations (25), (26), optimality condition (27), and boundary and transversality conditions (28)–(30) simultaneously.

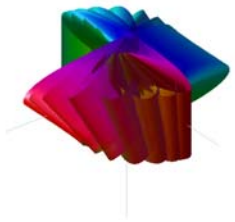
We use a neighboring extremal method [15]. A nominal solution satisfying all of the necessary conditions except the boundary conditions is chosen. The unspecified initial multiplier is updated so as to satisfy the specified terminal boundary conditions in the limit. This is also referred to as a shooting method. The main advantage of the neighboring extremal method is that the number of iteration variables is small. In other approaches, the initial guess of control input history or multiplier variables are iterated, so the number of optimization parameters are proportional to the number of discrete time steps.

A potential difficulty is that the extremal solutions are sensitive to small changes in the unspecified initial multiplier values. The nonlinearities also make it hard to construct an accurate estimate of sensitivity, and it may result in numerical ill-conditioning. By adopting a geometric numerical integrator approach, sensitivity derivatives along the discrete-time necessary conditions are not corrupted by numerical dissipation. Thus, the necessary conditions are solved in a computationally efficient manner.

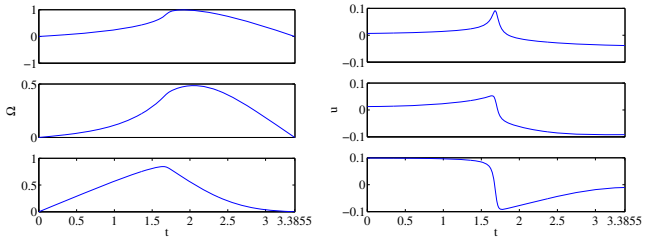
IV. Numerical Example

We choose an elliptic cylinder for a rigid body model with semi-major axis $0.8\,\mathrm{m}$, semi-minor axis $0.2\,\mathrm{m}$, height $0.6\,\mathrm{m}$, mass $1,\mathrm{kg}$. The inertia matrix is $J = \mathrm{diag}[0.04,\ 0.19,\ 0.17]\,\mathrm{kgm}^2$, and the maximum control inputs is chosen as $\bar{u} = 0.1\,\mathrm{Nm}$. The desired attitude maneuver is a rest-to-rest large angle rotation given by $(R_\circ, \Omega_\circ) = (I_{3\times3}, 0)$, $(R_f, \Omega_f) = (\exp\theta v, 0)$, where $v = \frac{1}{\sqrt{3}}[1,\ 1,\ 1] \in \mathbb{R}^3$, and θ is varied as 120° and 180°.

In the numerical computation, we fix the number of steps as $N = 1000$ in this particular numerical example, and we vary the timestep h. In essence, we find the seven parameters, initial multipliers $(\lambda_0^R, \lambda_0^\Omega)$ and the time step h, satisfying the seven-dimensional terminal boundary conditions (28)–(30) under the discrete-time equations of motion, the multiplier equation, and the optimality condition.

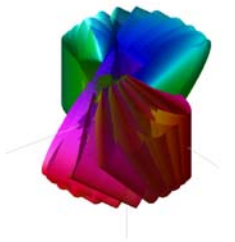


(a) Attitude maneuver

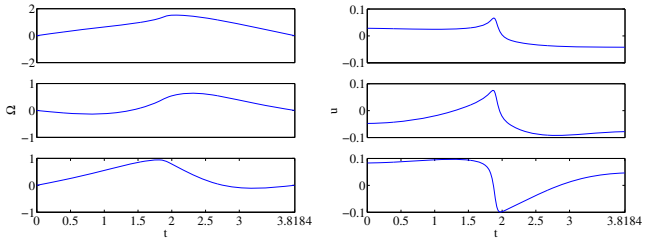


(b) Angular velocity Ω (c) Control input u

Fig. 1. Time optimal attitude maneuver, $\theta = 120°$



(a) Attitude maneuver



(b) Angular velocity Ω (c) Control input u

Fig. 2. Time optimal attitude maneuver, $\theta = 180°$

We solve this two-point boundary value problem, interpreted as a nonlinear equation by the shooting method, using a general nonlinear equation solver, namely the Matlab `fsolve` function. The multipliers are initialized randomly, and the time step is initialized as $h = 0.002$ seconds. The optimal solutions are found in 94 and 211 seconds, respectively, on Intel Pentinum M 1.73 GHz processor; the boundary condition errors are less than 10^{-15}.

The optimized attitude maneuver, angular velocity, and control input history are presented in Figures 1 and 2. (Simple animations which show these maneuvers of the rigid body are available at `http://www.umich.edu/~tylee`.) The optimized maneuver times are 3.39 and 3.82 seconds, respectively.

V. Conclusions

A time optimal attitude control problem to rotate a rigid body within a minimal time with constrained control input is studied. Necessary conditions for optimality are developed on $\mathrm{SO}(3)$ using rotation matrices without need of attitude parameterizations such as Euler angles and quaternions. This provides a globally applicable and compact form of necessary conditions for optimality. For overall computational accuracy and efficiency, a discrete optimal control method is proposed using a Lie group variational integrator.

In this paper, the two-norm of the control moment is constrained, and consequently, there is no singular arc in the optimal solution. The proposed necessary conditions for optimality can be directly applied, without modification, to the case where the absolute value of each component of the control moment is bounded. In this case, the expressions for optimal singular control can be developed, for example, by following the approach given in [4], using the compact multiplier equations presented here.

References

[1] S. L. Scrivener and R. C. Thompson, "Survey of time-optimal attitude maneuvers," *Journal of Guidance, Control, and Dynamics*, vol. 17, no. 2, pp. 225–233, 1994.

[2] J. R. Etter, "A solutions of the time optimal Euler rotation problem," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, 1989, pp. 1441–1449.

[3] K. D. Bilimoria and B. Wie, "Time-optimal three-axis reorientation of a rigid spacecraft," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 3, pp. 446–452, 1993.

[4] H. Seywald and R. R. Kumar, "Singular control in minimum time spacecraft reorientation," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 4, pp. 686–694, 1993.

[5] M. Modgalya and S. P. Bhat, "Time-optimal attitude reorientation at constant angular velocity magnitude with bounded angular acceleration," in *Proceedings of IEEE Conference on Decision and Control*, 2006, pp. 223–228.

[6] R. M. Byers and S. R. Vadali, "Quasi-closed form solution to the time-optimal rigid spacecraft reorientation problem," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 3, pp. 453–461, 1993.

[7] T. Lee, M. Leok, and N. H. McClamroch, "Lie group variational integrators for the full body problem," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, pp. 2907–2924, 2007.

[8] ——, "Optimal control of a rigid body using geometrically exact computations on SE(3)," in *Proceedings of the IEEE Conference on Decision and Control*, 2006, pp. 2170–2175.

[9] ——, "A combinatorial optimal control problem for spacecraft formation reconfiguration," in *Proceedings of the IEEE Conference on Decision and Control*, 2007, pp. 5370–5375.

[10] D. J. Bell and D. H. Jacobson, *Singular Optimal Control Problems*. Academic Press, 1975.

[11] T. Lee, M. Leok, and N. H. McClamroch, "Discrete control systems," in *The Encyclopedia of Complexity and System Science*. Springer, 2007, to appear. [Online]. Available: http://arxiv.org/abs/0705.3868

[12] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[13] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, "Lie-group methods," *Acta Numerica*, vol. 9, pp. 215–365, 2000.

[14] T. Lee, M. Leok, and N. H. McClamroch, "Lie group variational integrators for the full body problem in orbital mechanics," *Celestial Mechanics and Dynamical Astronomy*, vol. 98, pp. 121–144, 2007.

[15] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control*. Hemisphere Publishing Corporation, 1975.

Global Symplectic Uncertainty Propagation on $\mathrm{SO}(3)$

Taeyoung Lee*†, Melvin Leok*, and N. Harris McClamroch†

Abstract— **This paper introduces a global uncertainty propagation scheme for rigid body dynamics, through a combination of numerical parametric uncertainty techniques, noncommutative harmonic analysis, and geometric numerical integration. This method is distinguished from prior approaches, as it allows one to consider probability densities that are global, and are not supported on only a single coordinate chart on the manifold. The use of Lie group variational integrators, that are symplectic and stay on the Lie group, as the underlying numerical propagator ensures that the advected probability densities respect the geometric properties of uncertainty propagation in Hamiltonian systems, which arise as consequence of the Gromov nonsqueezing theorem from symplectic geometry. We also describe how the global uncertainty propagation scheme can be applied to the problem of global attitude estimation.**

I. Introduction

A nonlinear uncertainty propagation scheme is developed for the dynamics of a rigid body, viewed as evolving on the special orthogonal group $\mathrm{SO}(3)$. Rigid body dynamics is a Hamiltonian flow on a Lie group, and most current attitude uncertainty propagation schemes [1] do not properly take these characteristics into account. Typically, the attitude is represented by unit quaternions, which is problematic for global uncertainty propagation due to the ambiguity introduced by the double cover of $\mathrm{SO}(3)$ by the three-sphere $\mathbb{S}^3$ of unit quaternions. Furthermore, the dynamics are often simplified to kinematic equations, thereby ignoring uncertainties in the angular velocity. As such, most existing techniques are only valid over time periods when the uncertainties are small.

The method introduced in this paper is focused on developing a global uncertainty propagation method for a rigid body by explicitly considering the characteristics of the Hamiltonian flow and the configuration manifold, without implicitly assuming that the uncertainty is localized, nor that the uncertainty distribution is fully supported in a single coordinate chart on the manifold.

Although it is not widely known in the engineering community, the theories of probability and stochastic processes on manifolds have been studied by theoretical statisticians [2], [3], [4]. Earlier works on attitude estimation on $\mathrm{SO}(3)$ include [5], where a probability density function is expressed using noncommutative harmonic analysis. This idea of using fourier analysis on manifolds has been applied in [6], [7], [8], [9], and recently, results in [5] are extended to include the effects of process noise and sensor parameters in [10].

The Liouville equation describes the evolution of a probability density function in the absence of external diffusion, and can be viewed as the deterministic analogue of the Fokker-Planck equation. When the flow that is advecting the probability density is Hamiltonian, the Liouville equation reduces to an ordinary differential equation [11]. Thus, a probability density function can be propagated using the flow map of the Hamiltonian system. In [12], an attitude estimation scheme is developed by linearizing the attitude dynamics along the mean obtained from the probability density function propagated by this property. However, nonlinearities of the flow imply that numerical methods for propagating uncertainties using linearization rapidly degrade in performance, unless frequent physical measurements are available [13]. It is therefore desirable to construct efficient numerical methods for solving the Liouville equation that describes the evolution of a probability density advected by a prescribed Hamiltonian flow.

In this paper, a global attitude uncertainty propagation scheme is developed in the absence of process noise. This is achieved through a synthesis of numerical parametric uncertainty analysis techniques [14], [15], noncommutative harmonic analysis [16], and geometric numerical integration [17]. We backpropagate sample points along the Hamiltonian flow of the rigid body dynamics, and we use the advected probabilities to reconstruct the probability density using noncommutative harmonic analysis on $\mathrm{SO}(3)$. This is in contrast to Monte Carlo methods, where the sample trajectories are used to compute statistical properties of the advected density.

The Gromov nonsqueezing theorem [18] from symplectic geometry places fundamental limits on how the uncertainty of a Hamiltonian system evolves [19]. It is therefore essential that symplectic methods be used to propagate individual trajectories when analyzing uncertainty propagation in Hamiltonian systems. We use a geometric numerical integrator, referred to as a Lie group variational integrator, that preserves the symplectic property of the Hamiltonian dynamics and the group structure of the configuration space [20], [21]. The purpose of this paper is to develop a computational method to propagate uncertainties under a Hamiltonian flow on a Lie group. Our development is for a specific Hamiltonian system, namely the 3D pendulum, that evolves on the Lie group $\mathrm{SO}(3)$. We also comment on how this development is applicable to attitude estimation. The methodology proposed

Taeyoung Lee, N. Harris McClamroch, Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 {tylee,nhm}@umich.edu

Melvin Leok, Mathematics, Purdue University, West Lafayette, IN 47907 mleok@math.purdue.edu

*This research has been supported in part by NSF under grants DMS-0504747, DMS-0726263, and DMS-0714223.

†This research has been supported in part by NSF under grant CMS-0555797.

in this paper should be distinguished from the numerical methods that compute a specific realization of a stochastic Hamiltonian system [22], [23], [24].

II. Attitude Dynamics of a Rigid Body

A. 3D Pendulum

The 3D pendulum is nontrivial example of a Hamiltonian system that evolves on a Lie group; this example is used in the subsequent development. A rigid 3D pendulum is a rigid body supported by a fixed, frictionless pivot, acted on by uniform gravitational force [25]. The supporting pivot allows the pendulum three rotational degrees of freedom.

Two reference frames are introduced. An inertial reference frame has its origin at the pivot; the first two axes lie in the horizontal plane and the third axis is vertical in the direction of gravity. A reference frame fixed to the pendulum body is also introduced. The origin of this body-fixed frame is also located at the pivot. The configuration manifold is the special orthogonal group $\mathrm{SO}(3)$,

$$\mathrm{SO}(3) = \left\{ R \in \mathbb{R}^{3\times 3} \,|\, R^T R = I_{3\times 3},\, \det R = 1 \right\},$$

where the rotation matrix $R \in \mathrm{SO}(3)$ represents the linear transformation from the body-fixed frame to the inertial frame.

The dynamics of the 3D pendulum are given by the Euler rigid body equation that includes the moment due to gravity:

$$J\dot{\Omega} = J\Omega \times \Omega + mg\rho \times R^T e_3, \tag{1}$$

where the angular velocity in the body-fixed frame is denoted by $\Omega \in \mathbb{R}^3$, the inertia matrix is denoted by $J \in \mathbb{R}^{3\times 3}$, and the vector $\rho \in \mathbb{R}^3$ represents the location of the center of mass in the body-fixed frame. The constants m and g denote the mass of the pendulum and the gravitational acceleration, respectively. The kinematic equation is

$$\dot{R} = RS(\Omega). \tag{2}$$

For a given vector $x \in \mathbb{R}^3$, the 3×3 skew-symmetric matrix $S(x)$ is defined so that $S(x)y = x \times y$ for all $y \in \mathbb{R}^3$.

There are two disjoint equilibria when the direction of gravity in the body fixed frame is collinear with the vector ρ; the hanging equilibrium when $R^T e_3 = \rho/\left\|\rho\right\|$, and the inverted equilibrium when $R^T e_3 = -\rho/\left\|\rho\right\|$. The 3D pendulum exhibits surprisingly rich and complicated attitude dynamics [13], and is therefore particularly appropriate for demonstrating the properties of our global attitude uncertainty propagation scheme.

B. Lie Group Variational Integrator

Lie group variational integrators preserve the group structure without the use of local charts, reprojection, or constraints, they are symplectic and momentum preserving, and they exhibit good energy behavior for an exponentially long time period. The following Lie group variational integrator for the attitude dynamics of a rigid body was presented in [20], [21]:

$$hS(J\Omega_k + \frac{h}{2}M_k) = F_k J_d - J_d F_k^T, \tag{3}$$

$$R_{k+1} = R_k F_k, \tag{4}$$

$$J\Omega_{k+1} = F_k^T J\Omega_k + \frac{h}{2}F_k^T M_k + \frac{h}{2}M_{k+1}, \tag{5}$$

where $J_d \in \mathbb{R}^{3\times 3}$ is a nonstandard moment of inertia matrix given by $J_d = \frac{1}{2}\mathrm{tr}[J]\, I_{3\times 3} - J$, and $F_k \in \mathrm{SO}(3)$ is the relative attitude between integration steps. The moment due to the potential is denoted by $M_k = mg\rho \times R_k^T e_3$. The constant $h \in \mathbb{R}$ is the integration step size, and the subscript k denotes the k-th integration step. This integrator yields a map $(R_k, \Omega_k) \mapsto (R_{k+1}, \Omega_{k+1})$ by solving (3) to obtain $F_k \in \mathrm{SO}(3)$ and substituting it into (4) and (5) to obtain R_{k+1} and Ω_{k+1}. We use these discrete-time equations of motion to propagate the attitude dynamics.

III. Global Symplectic Uncertainty Propagation on $\mathrm{SO}(3)$

Suppose that a probability density function for the attitude and the angular velocity of a rigid body at a time t_k is given as $p_k(R, \Omega) : \mathrm{SO}(3) \times \mathbb{R}^3 \to \mathbb{R}$. In this section, we develop a computational method to propagate this density along the Hamiltonian attitude dynamics assuming that there is no process noise. We represent the propagated probability density function using noncommutative harmonic analysis. A method for visualizing the attitude uncertainty is also discussed.

A. Symplectic Uncertainty Propagation for a Hamiltonian System

In [11], it is shown that the probability density function is preserved along a Hamiltonian flow on Euclidean space. In this subsection, we generalize this result to a general Hamiltonian system evolving on a manifold.

Consider a Hamiltonian system on a $2n$-dimensional symplectic manifold (Q, ω), where Q is a $2n$-dimensional manifold and $\omega : TQ \times TQ \to \mathbb{R}$ is a nondegenerate symplectic two-form on Q [26]. The Liouville volume form $\mu : (TQ)^{2n} \to \mathbb{R}$ is defined as the n-fold wedge product of the symplectic two-form with itself, $\mu = \frac{(-1)^{n(n-1)/2}}{n!}\omega \wedge \cdots \wedge \omega \quad (n \text{ times})$. In local coordinates, this corresponds to the usual notion of the volume element in Euclidean spaces. Let $\mathcal{L}_X \mu$ be the Lie derivative of the volume form μ along a vector field $X : Q \to TQ$. The divergence of a vector field X on Q is defined as $\mathcal{L}_X \mu = \mathrm{div}_\mu(X)\, \mu$. Thus, the divergence $\mathrm{div}_\mu(X)$ represents the rate of change of a unit volume along the vector field X.

In [27], it is shown that the Fokker-Planck equation for a dynamic system on a manifold in the absence of diffusion terms can be written as,

$$\frac{\partial p}{\partial t} + \mathrm{div}_\mu\,(pX) = 0. \tag{6}$$

Note that this has the same structure as the Euler equation for the density of compressible fluids. The existence and uniqueness of the solution of the equations of motion provides a property analogous to mass conservation in fluid dynamics. Since $\operatorname{div}_\mu(pX) = p\operatorname{div}_\mu(X) + \mathcal{L}_X p$, the time derivative of the probability density function is given by

$$\frac{d}{dt}p = \frac{\partial p}{\partial t} + \mathcal{L}_X p = -p\operatorname{div}_\mu(X). \tag{7}$$

If the vector field X is the Hamiltonian vector field on (Q,ω), the divergence vanishes, $\operatorname{div}_\mu(X) = 0$ according to the Liouville theorem [26]. Therefore, the Fokker-Planck equation for a deterministic Hamiltonian system is represented by the ordinary differential equation

$$\frac{d}{dt}p = 0. \tag{8}$$

This states that the probability density function is preserved along a Hamiltonian flow without stochastic diffusion effects. More precisely, (8) implies that the propagated probability density function at t_{k+1} can be explicitly expressed as a composition of the backward flow map and the given probability density function at t_k

$$p_{k+1}(R,\Omega) = p_k(\mathcal{F}^{-1}(R,\Omega)), \tag{9}$$

where $\mathcal{F}^k : \mathrm{SO}(3)\times\mathbb{R}^3 \to \mathrm{SO}(3)\times\mathbb{R}^3$ represents the k step discrete flow of the attitude dynamics. We can apply this equation recursively to propagate the probability density function over any time interval.

B. Noncommutative Harmonic Analysis on $\mathrm{SO}(3)$

Equation (9) provides a method to compute the probability density function at any time based on the probability density function at some prior time. As the flow is nonlinear, it is inefficient to characterize the density using its moments, since the moment expansion may decay slowly. Here, we propose a computational scheme that represents the probability density function for the attitude dynamics using noncommutative harmonic analysis on $\mathrm{SO}(3)$.

Noncommutative harmonic analysis is a generalization of Fourier analysis on Euclidean spaces to manifolds [27]. This is particularly useful since it provides a mathematical tool to approximate a probability density function based on samples of that function. More precisely, the propagated probability density function for the attitude dynamics of a rigid body can be expressed as

$$p(R,\Omega) = \sum_{l=0}^{\infty} \frac{2l+1}{(2\pi)^3} \int_{\mathbb{R}^3} \exp(j\theta\cdot\Omega)\operatorname{tr}\left[P^l(\theta)U^l(R)\right]\, d\theta, \tag{10}$$

where the vector $\theta \in \mathbb{R}^3$ and non-negative integer $l \in \mathbb{N}\bigcup\{0\}$ are Fourier parameters, and the set of complex matrices $\{P^l(\theta)\in\mathbb{C}^{(2l+1)\times(2l+1)}\}_{l=0}^{\infty}$ is the Fourier spectrum of the density $p(R,\Omega)$. We denote the l-th irreducible unitary representation matrix of $\mathrm{SO}(3)$ by $U^l(R)\in\mathbb{C}^{(2l+1)\times(2l+1)}$. A representation of a group is a homomorphism from the group to the set of invertible matrices, and by the Peter-Weyl theorem [28], the irreducible unitary representations form a complete orthonormal basis for the set of square-integrable functions on the group. The irreducible unitary representations can be expressed in various ways. For example, it can be expressed in terms of the *wigner-d* functions using 3-1-3 Euler angles α,β,γ as

$$U^l_{m,n}(R(\alpha,\beta,\gamma)) = i^{m-n}e^{-i(m\alpha+n\gamma)}d^l_{mn}(\cos\beta) \tag{11}$$

for $-l \le m,n \le l$ [29]. A few *wigner-d* functions are given by

$$d^0(\cos\beta) = 1,$$
$$d^1(\cos\beta) = \begin{bmatrix} \frac{1+\cos\beta}{2} & -\frac{\sin\beta}{\sqrt{2}} & \frac{1-\cos\beta}{2} \\ \frac{\sin\beta}{\sqrt{2}} & \cos\beta & -\frac{\sin\beta}{\sqrt{2}} \\ \frac{1-\cos\beta}{2} & \frac{\sin\beta}{\sqrt{2}} & \frac{1+\cos\beta}{2} \end{bmatrix}.$$

Higher order *wigner-d* functions can be obtained using a recursive formula [27].

From the orthnormal property of the irreducible unitary representation, the Fourier spectrum is computed by

$$P^l(\theta) = \int_{\mathrm{SO}(3)}\int_{\mathbb{R}^3} p(R,\Omega)\exp(-j\theta\cdot\Omega)U^l(R^{-1})d\Omega\, dR. \tag{12}$$

The volume element for the rotation matrix dR represents the Haar measure of $\mathrm{SO}(3)$; it can be written in terms of 3-1-3 Euler angles α,β,γ as $dR(\alpha,\beta,\gamma) = \frac{1}{8\pi^2}\sin\beta\, d\alpha d\beta d\gamma$.

Substituting (9), we can compute the Fourier spectrum of the propagated probability density function. The propagated distribution can be reconstructed by (10). This is a global particle-based method to construct the propagated probability density function on $\mathrm{SO}(3)\times\mathbb{R}^3$.

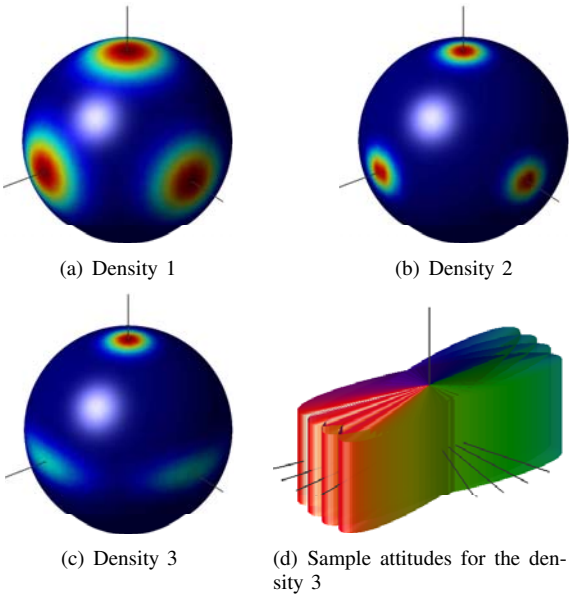
C. Visualization of the Attitude Uncertainty

Let $p_R : \mathrm{SO}(3) \to \mathbb{R}$ be a probability density function on $\mathrm{SO}(3)$. For example, it can be obtained by integrating (10) over $\mathbb{R}^3$, i.e. $p_R(R) = \int_{\mathbb{R}^3} p(R,\Omega)d\Omega$. We propose a method for visualizing probability densities on $\mathrm{SO}(3)$ using three copies of two-spheres. The rotation matrix represents a linear transformation from a body fixed frame to an inertial frame. Therefore, the i-th column of the rotation matrix Re_i represent the direction of the i-th body fixed axis in the inertial frame. Since the vector Re_i lies on the two-sphere $\mathbb{S}^2$, we can visualize uncertainties of Re_i on a sphere either by color shading or by contour lines. Three copies of these spheres, one for each of the body fixed axes, can be used to visualize uncertainties on $\mathrm{SO}(3)$.

We find the marginal probability density of p_R for each column of the rotation matrix. For given $r\in\mathbb{S}^2$, define a subset of $\mathrm{SO}(3)$ as

$$H_i(r) = \left\{R\in\mathrm{SO}(3)\,\middle|\, Re_i = r\right\}. \tag{13}$$

This is a subgroup of $\mathrm{SO}(3)$ diffeomorphic to $\mathbb{S}^1$, and it represents the set of rotation matrices whose i-th column is equal to the given direction r. This subgroup can be parameterized by $\theta\in\mathbb{S}^1$. More explicitly, we can find an element $R_i^\circ(r)$ of $H_i(r)$ (for example, we have $R_i^\circ(r) =$



(a) Density 1 (b) Density 2

(c) Density 3 (d) Sample attitudes for the density 3

Fig. 1. Attitude uncertainty visualization example

$\exp(\frac{\text{acos}(r\cdot e_i)}{\|e_i\times r\|}S(e_i\times r))$ if $e_i\times r\neq 0$). Then, the subgroup $H_i(r)$ is parameterized as

$$H_i(r)=\left\{R_i^\circ(r)\exp S(\theta e_i)\,\middle|\,\theta\in\mathbb{S}\right\}. \tag{14}$$

The corresponding quotient space is the two-sphere, $\text{SO}(3)/H_i(r)\simeq\mathbb{S}^2$. Using the properties of integration on a quotient space of a Lie group, we have

$$\begin{aligned}1&=\int_{\text{SO}(3)}p_R(R)\,dR\\&=\int_{r\in S^2}\left(\frac{1}{2\pi}\int_{\theta\in\mathbb{S}^1}p_R(R_i^\circ(r)\exp S(\theta e_i)\,d\theta\right)dr.\end{aligned} \tag{15}$$

Therefore, the marginal probability density for the i-th column of the rotation matrix, $p_R^i:\mathbb{S}^2\to\mathbb{R}$ is given by

$$p_R^i(r)=\frac{1}{2\pi}\int_{\theta\in\mathbb{S}^1}p_R(R_i^\circ(r)\exp(\theta\hat{e}_i))\,d\theta. \tag{16}$$

We plot these marginal probability density functions $p_R^i(r)$, that represent the probability density of the direction of the body fixed axes, on three two-spheres. If the magnitude of uncertainties is small, we can plot uncertainties for each body fixed axis on a single sphere, from which we can intuitively understand the attitude uncertainty of the rigid body. Fig. 1 shows examples for the attitude probability density visualization. It is easily to see that the second density in Fig. 1(b) has smaller variation than the first density in Fig. 1(b). Fig. 1(d) shows some sample attitudes that we are likely to obtain from the third density given in Fig. 1(c), which reflect the fact that the direction of the z-axis of the body in the inertial frame is well localized, but there is greater uncertainty in the direction of the x- and y-axes.

IV. Numerical Computations

In this section, we propagate an initial probability density along the nontrivial dynamics of the 3D pendulum, and we visualize the attitude uncertainty. The properties of the pendulum are given by

$$\begin{gathered}J=\text{diag}[0.13,0.28,0.17]\,\text{kgm}^2,\quad m=1\,\text{kg},\\ \rho=0.3e_3\,\text{m},\quad g=9.81\,\text{m/s}^2.\end{gathered}$$

The von Mises distribution (also known as the circular normal distribution) is a continuous probability density on the circle, which can be thought of as the circular analogue of the normal density [30].

$$p(\theta)=\frac{1}{2\pi I_0(\kappa)}\exp\left(\kappa\cos(\theta-\overline{\theta})\right), \tag{17}$$

where I_0 is the zeroth order modified Bessel function of the first kind, given by $I_0(\kappa)=\sum_{i=0}^{\infty}\frac{(1/4\kappa^2)^i}{(\kappa!)^2}$ with parameters $\kappa,\overline{\theta}$ that determine the shape of the density; as κ increases, the density approaches a normal density with mean $\overline{\theta}$ and variance $\frac{1}{\kappa}$.

For two given rotation matrices $R,\overline{R}\in\text{SO}(3)$, the quantity $\frac{1}{2}(\text{tr}\left[\overline{R}^TR\right]-1)$ represents the cosine of the rotation angle between the two attitudes. Using this, we define a probability density function on $\text{SO}(3)$ from the von Mises distribution. The probability distribution at the initial time is chosen as

$$\begin{aligned}p_0(R,\Omega)=\frac{1}{c}&\exp\left\{\frac{1}{2}\kappa\left(\text{tr}\left[\overline{R}_0^TR\right]-1\right)\right\}\\&\times\exp\left\{-\frac{1}{2}(\Omega-\overline{\Omega}_0)^T\Sigma^{-1}(\Omega-\overline{\Omega}_0)\right\},\end{aligned} \tag{18}$$

where $\overline{R}_0=I_{3\times3}$, $\overline{\Omega}_0=[4.14,4.14,4.14]\,\text{rad/s}$, $\Sigma=0.1414^2I_{3\times3}$, and $\kappa=8$. The constant c is a scaling factor chosen such that $\int p(R,\Omega)\,dRd\Omega=1$. The corresponding mean $(\overline{R}_0,\overline{\Omega}_0)$ yields an irregular, perhaps chaotic, attitude response [13].

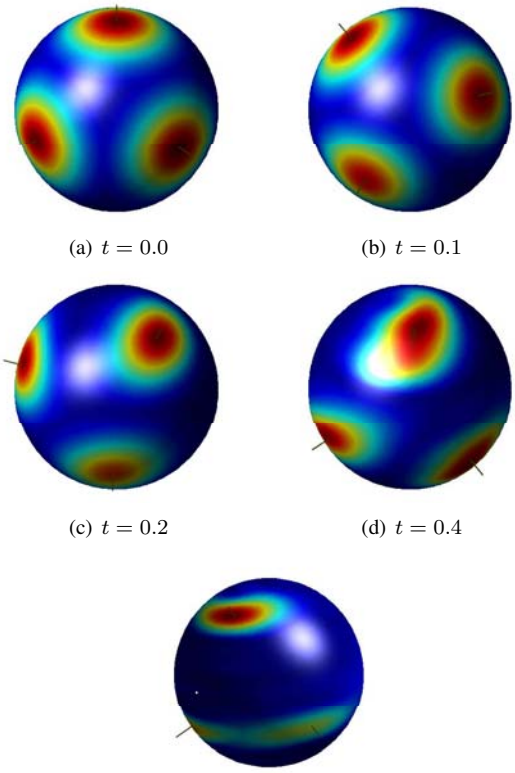
We propagate this initial distribution using (9) and compute the Fourier spectrum using (12). The volume integration is approximated by Simpson's rule, and the flow map is computed using the Lie group variational integrator. The application of the Lie group variational integrator is particularly useful since it is symplectic, group structure preserving, and time reversible.

We have developed a parallel computing code using the MPI (Message Passing Interface) library, where the domain of integration is divided uniformly and distributed to each processor. This is desirable in terms of minimizing communication between processors and balancing the computational load among the processors. This algorithm has been implemented on 32 AMD Opeteron processors. Fig. 2 illustrates the propagated attitude uncertainty.

V. Comments on Global Attitude Estimation

The presented uncertainty propagation method can be applied to develop an attitude estimation scheme using Bayes rule. We first define a measurement model. We assume that the attitude and the angular velocity are measured by sensors to obtain

$$z_k=H(R_k,\Omega_k)+v_k, \tag{19}$$



(a) $t = 0.0$ (b) $t = 0.1$

(c) $t = 0.2$ (d) $t = 0.4$

(e) $t = 1.0$

Fig. 2. Propagated attitude uncertainty and max mean attitude

where $H : \mathrm{SO}(3) \times \mathbb{R}^3 \to \mathbb{R}^m$ is a measurement function, $z_k \in \mathbb{R}^m$ is the measured value, and $v_k \in \mathbb{R}^m$ is measurement noise. For example, if we measure a direction to a known object $a \in \mathbb{S}^2$ and the angular velocity, the measurement function can be written as $H(R, \Omega) = [R^T a; \Omega]$. We assume the probability density of the measurement noise v_k is given by $p_{z_k|k}$, and it is an independent process.

The set of all measurements from the initial time to t_k is denoted by the capital letter $Z_k = [z_0, z_1, \ldots, z_k]$. Suppose that we have a probability density function at the k-th time conditioned by Z_k, i.e. the expression for $p_{k|Z_k}$ is known, and a new measurement is obtained z_{k+1} at t_{k+1}. Estimation can be described as finding an expression for $p_{k+1|Z_{k+1}}$ given $p_{k|Z_k}$ and $Z_{k+1} = [Z_k, z_{k+1}]$. Using Bayes rule [31], we have

$$p_{k+1|Z_{k+1}}(R, \Omega|[Z, z]) = \frac{p_{z_{k+1}|k+1,Z_k}(z|R, \Omega, Z)\ p_{k+1|Z_k}(R, \Omega|Z)}{p_{z_{k+1}|Z_k}(z|Z)}. \tag{20}$$

Since the measurement processes are independent, we have $p_{z_{k+1}|k+1,Z_k} = p_{z_{k+1}|k+1}$, and the propagated density is given by (9), i.e. $p_{k+1|Z_k}(R, \Omega|Z) = p_{k|Z_k}(\mathcal{F}^{-1}(R, \Omega)|Z)$. The denominator is a normalizing constant that can be computed to satisfy $p_{z_{k+1}|Z_k}(z|Z) = \int_{\mathrm{SO}(3)\times\mathbb{R}^3} p_{z_{k+1}|k+1}(R, \Omega) p_{k+1|Z_k}(R, \Omega|Z) dR d\Omega$. In summary, the propagated probability density conditioned by the new measurement is given by

$$p_{k+1|Z_{k+1}}(R, \Omega|[Z, z]) = \frac{1}{c} p_{z_{k+1}|k+1}(z|R, \Omega)\ p_{k|Z_k}(\mathcal{F}^{-1}(R, \Omega)|Z) \tag{21}$$

for a constant c. This also can be represented using the harmonic analysis as in (10).

This expression is the key to solving the attitude estimation problem. This expression for $p_{x_{k+1}|Z_{k+1}}$ contains all the statistical information that can be derived from the attitude dynamics and the available measurements.

VI. Conclusions

This paper addresses the problem of propagating the attitude uncertainty of a rigid body which evolves on the rotation group. The use of noncommutative harmonic analysis techniques to represent the uncertainty distribution in a global fashion overcomes a fundamental limitation of existing techniques, which implicitly assume that the uncertainty is localized or small. By exploiting the fact that the Liouville equation for a Hamiltonian system reduces to an ordinary differential equation, we are able to adopt a particle based approach for computing the advected probability density, thereby avoiding the computational expense of solving the equation as a numerical partial differential equation. By adopting a Lie group variational integrator as the underlying numerical scheme, we ensure that the resulting uncertainty propagation method inherits the geometric properties of the time evolution of probability densities in Hamiltonian systems, that arise from the symplectic geometry of the phase space.

A natural application of the proposed scheme is to the problem of global attitude estimation, particularly when the dynamics of the rigid body are extremely nonlinear, and the attitude measurements are relatively infrequent. Estimation problems with such characteristics are problematic for traditional techniques, such as the Kalman filter, which require frequent measurements, and relatively benign dynamics, in order to justify the localization and linearization assumptions built into the method.

References

[1] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 1, pp. 12–28, 2007.

[2] P. Diaconis, *Group Representations in Probability and Statistics*. Institute of Mathematical Statistics, 1988.

[3] K. D. Elworthy, *Stochastic Differential Equations on Manifolds*. Cambridge University Press, 1982.

[4] M. Emery, *Stochastic Calculus in Manifolds*. Springer, 1989.

[5] J. T. H. Lo and L. R. Eshleman, "Exponential Fourier densities on SO(3) and optimal estimation and detection for rotational processes," *SIAM Journal on Applied Mathematics*, vol. 36, no. 1, pp. 73–82, 1979.

[6] H. Hendriks, "Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions," *The Annals of Statistics*, vol. 18, no. 2, pp. 832–849, 1990.

[7] P. T. Kim, "Deconvolution density estimation on SO(n)," *The Annals of Statistics*, vol. 26, no. 3, pp. 1083–1102, 1998.

[8] W. Park, J. S. Kim, Y. Zhou, N. J. Cowan, A. M. Okamura, and G. S. Chirikjian, "Diffusion-based motion planning for a nonholonomic flexible needle model," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, 2005.

[9] Y. Wang, Y. Zhou, D. K. Maslen, and G. S. Chirikjian, "Solving phase-noise Fokker-Planck equations using the motion-group Fourier transform," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 868–877, 2006.

[10] F. L. Markley, "Attitude filtering on SO(3)," in *Proceedings of the AAS Malcolm D. Shuster Astronautics Symposium*, 2005, AAS 05-460.

[11] D. J. Scheeres, F. Y. Hsiao, R. S. Park, B. F. Villac, and J. M. Maruskin, "Fundamental limits on spacecraft orbit uncertainty and distribution propagation," in *Proceedings of the AAS Malcolm D. Shuster Astronautics Symposium*, 2005, AAS 05-471.

[12] J. M. Valpiani and P. L. Palmer, "Nonliner symplectic attitude estimation for small satellitese," in *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 2006, AIAA 2006-6159.

[13] T. Lee, N. Chaturvedi, A. Sanyal, M. Leok, and N. H. McClamroch, "Propagation of uncertainty in rigid body attitude flows," in *Proceedings of the IEEE Conference on Decision and Control*, 2007, pp. 2689–2694.

[14] D. Xiu, "Efficient collocational approach for parametric uncertainty analysis," *Commun. Comput. Phys.*, vol. 2, no. 2, pp. 293–309, 2007.

[15] D. Xiu and J. Hesthaven, "High-order collocation methods for differential equations with random inputs," *SIAM J. Sci. Comput.*, vol. 27, no. 3, pp. 1118–1139 (electronic), 2005.

[16] G. Chirikjian and A. Kyatkin, *Engineering applications of noncommutative harmonic analysis*. Boca Raton, FL: CRC Press, 2001.

[17] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*, 2nd ed., ser. Springer Series in Computational Mathematics. Springer-Verlag, 2006, vol. 31.

[18] M. Gromov, "Pseudo holomorphic curves in symplectic manifolds," *Invent. Math.*, vol. 82, no. 2, pp. 307–347, 1985.

[19] F. Hsiao and D. J.Scheeres, "Fundamental constraints on uncertainty evolution in Hamiltonian systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 4, pp. 686–691, 2007.

[20] T. Lee, M. Leok, and N. H. McClamroch, "Lie group variational integrators for the full body problem," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, pp. 2907–2924, May 2007.

[21] ——, "Lie group variational integrators for the full body problem in orbital mechanics," *Celestial Mechanics and Dynamical Astronomy*, vol. 98, no. 2, pp. 121–144, June 2007.

[22] J.-A. Lazaro-Cami and J.-P. Ortega, "Stochastic hamiltonian dynamical systems," 2007. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:math/0702787

[23] N. Bou-Rabee and H. Owhadi, "Stochastic variational integrators," 2007. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:0708.2187

[24] L. Wang, "Variational integrators and generating functions for stochastic hamiltonian systems," Ph.D. dissertation, Universität Karlsruhe, 2007.

[25] J. Shen, A. Sanyal, N. Chaturvedi, D. Bernstein, and N. McClamroch, "Dynamics and control of a 3D pendulum," *Proceedings of the IEEE Conference on Decision and Control*, pp. 323–328, 2004.

[26] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*. Springer, 1999.

[27] G. S. Chirikjian and A. B. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis*. CRC Press, 2000.

[28] F. Peter and H. Weyl, "Die vollständigkeit der primitiven darstellungen einer geschlossenen kontinuierlichen gruppe," *Math. Ann.*, vol. 97, pp. 735–755, 1927.

[29] L. C. Biedenharn and J. D. Louck, *Angular Momentum in Quantum Physics*. Addison-Wesley, 1981.

[30] von Mises, "Über die Ganzzahligkeit der Atomgewicht und verwandte Fragen," *Physikal, Z*, vol. 19, pp. 490–500, 1918.

[31] A. H. Jazwinski, *Stochastic Process and Filtering Theory*. Academic Press, 1970.

GEOMETRIC STRUCTURE-PRESERVING OPTIMAL CONTROL OF THE RIGID BODY

ANTHONY M. BLOCH, ISLAM I. HUSSEIN, MELVIN LEOK, AND AMIT K. SANYAL

Abstract. In this paper we study a discrete variational optimal control problem for the rigid body. The cost to be minimized is the external torque applied to move the rigid body from an initial condition to a pre-specified terminal condition. Instead of discretizing the equations of motion, we use the discrete equations obtained from the discrete Lagrange–d'Alembert principle, a process that better approximates the equations of motion. Within the discrete-time setting, these two approaches are not equivalent in general. The kinematics are discretized using a natural Lie-algebraic formulation that guarantees that the flow remains on the Lie group $\mathsf{SO}(3)$ and its algebra $\mathfrak{so}(3)$. We use Lagrange's method for constrained problems in the calculus of variations to derive the discrete-time necessary conditions. We give a numerical example for a three-dimensional rigid body maneuver.

1. Introduction

This paper deals with a structure-preserving computational approach to the optimal control problem of minimizing the control effort necessary to perform an attitude transfer from an initial state to a prescribed final state, in the absence of a potential field. The configuration of the rigid body is given by the rotation matrix from the body frame to the spatial frame, which is an element of the group of orientation-preserving isometries in $\mathbb{R}^3$. The state of the rigid body is described by the rotation matrix and its angular velocity.

To motivate the computational approach we adopt in the discrete-time case, we first revisit the variational continuous-time optimal control problem. The continuous-time extremal solutions to this optimal control problem have certain special features, since they arise from variational principles. General numerical integration methods, including the popular Runge-Kutta schemes, typically preserve neither first integrals nor the characteristics of the configuration space. Geometric integrators are the class of numerical integration schemes that preserve such properties, and a good survey can be found in [5]. Techniques particular to Hamiltonian systems are also discussed in [16] and [24].

Our approach to discretizing the optimal control problem is in contrast to traditional techniques such as collocation, wherein the continuous equations of motion are imposed as constraints at a set of collocation points. In our approach, modeled after [11], the discrete equations of motion are derived from a discrete variational principle, and this induces constraints on the configuration at each discrete time step.

This approach yields discrete dynamics that are more faithful to the continuous equations of motion, and consequently yields more accurate solutions to the optimal

Date: Received: date / Revised: date.

control problem that is being approximated. This feature is extremely important in computing accurate (sub)optimal trajectories for long-term spacecraft attitude maneuvers. For example, in [9], the authors propose an imaging spacecraft formation design that requires a continuous attitude maneuver over a period of 77 days in a low Earth orbit. Hence, attitude maneuver has to be very accurate to meet tight imaging constraints over long time ranges.

While the discrete optimal control method presented here is illustrated using the Lie group $\mathsf{SO}(3)$ of rotation matrices, and its corresponding Lie algebra $\mathfrak{so}(3)$ of skew-symmetric matrices, we have derived the method with sufficient generality to address the problem of optimal control on arbitrary Lie groups with the drift vector field given by geodesic flow on the group, and it therefore widely applicable. For example, in inter-planetary orbit transfers (see, for example, [1]), one is interested in computing optimal or suboptimal trajectories on the group of rigid body motions $\mathsf{SE}(3)$ with a high degree of accuracy. Similar requirements also apply to the control of quantum systems. For example, efficient construction of quantum gates is a problem on the unitary Lie group $\mathsf{SU}(N)$. This is an optimal control problem, where one wishes to steer the identity operator to the desired unitary operator (see, for example, [13] and [23]).

Moreover, an important feature of the way we discretize the optimal control problem is that it is $\mathsf{SO}(3)$-equivariant. The $\mathsf{SO}(3)$-equivariance of our numerical method is desirable, since it ensures that our results do not depend on the choice of coordinates and coordinate frames. This is in contrast to methods based on coordinatizing the rotation group using quaternions, (modified) Rodrigues parameters, and Euler angles, as given in the survey [26]. Even if the optimal cost function is $\mathsf{SO}(3)$-invariant, as in [25], the use of generalized coordinates imposes constraints on the attitude kinematics.

For the purpose of numerical simulation, the corresponding discrete optimal control problem is posed on the discrete state space as a two stage discrete variational problem. In the first step, we derive the discrete dynamics for the rigid body in the context of discrete variational mechanics [20]. This is achieved by considering the discrete Lagrange–d'Alembert variational principle [12] in combination with essential ideas from Lie group methods [10], which yields a Lie group variational integrator [17]. This integrator explicitly preserves the Lie group structure of the configuration space, and is similar to the integrators introduced in [14] for a rigid body in an external field, and in [15] for full body dynamics. These discrete equations are then imposed as constraints to be satisfied by the extremal solutions to the discrete optimal control problem, and we obtain the discrete extremal solutions in terms of the given terminal states.

The paper is organized as follows. As motivation, in Section 2, we study the minimum control effort optimal control problem in continuous-time. In Section 3, we study the corresponding discrete-time optimal control problem. In Section 3.1 we state the optimal control problem and describe our approach. In Section 3.2, we derive the discrete-time equations of motion for the rigid body starting with the discrete Lagrange–d'Alembert principle. These equations are used in Section 3.3 to obtain the solution to the discrete optimal control problem. In Section 5, we describe an algorithm for solving the general nonlinear, implicit necessary conditions for $\mathsf{SO}(3)$ and give numerical examples for rest-to-rest and slew-up spacecraft maneuvers.

2. Continuous-Time Results

2.1. **Problem Formulation.** In this paper, the natural pairing between $\mathfrak{so}^*(3)$ and $\mathfrak{so}(3)$ is denoted by $\langle\cdot,\cdot\rangle$. Let $\ll\cdot,\cdot\gg$ and $\ll\cdot,\cdot\gg_*$ denote the standard (induced by the Killing form) inner product on $\mathfrak{so}(3)$ and $\mathfrak{so}^*(3)$, respectively. The inner product $\ll\cdot,\cdot\gg_*$ is naturally induced from the standard norm $\ll\boldsymbol{\xi},\boldsymbol{\omega}\gg = -\frac{1}{2}\mathrm{Tr}(\boldsymbol{\xi}^T\boldsymbol{\omega})$, for all $\boldsymbol{\xi},\boldsymbol{\omega}\in\mathfrak{so}(3)$, through

$$\begin{aligned} \ll\boldsymbol{\eta},\boldsymbol{\varphi}\gg_* &= \left\langle\boldsymbol{\eta},\boldsymbol{\varphi}^\sharp\right\rangle = \langle\boldsymbol{\eta},\boldsymbol{\omega}\rangle = \left\langle\boldsymbol{\xi}^\flat,\boldsymbol{\omega}\right\rangle \\ &= \ll\boldsymbol{\xi},\boldsymbol{\omega}\gg, \end{aligned} \tag{1}$$

where $\boldsymbol{\varphi}=\boldsymbol{\omega}^\flat\in\mathfrak{so}^*(3)$ and $\boldsymbol{\eta}=\boldsymbol{\xi}^\flat\in\mathfrak{so}^*(3)$, with $\boldsymbol{\xi},\boldsymbol{\omega}\in\mathfrak{so}(3)$ and $\flat$ and $\sharp$ are the musical isomorphisms with respect to the standard metric $\ll\cdot,\cdot\gg$. On $\mathfrak{so}(3)$, these isomorphisms correspond to the transpose operation. That is, we have $\boldsymbol{\varphi}=\boldsymbol{\omega}^{\mathrm{T}}$ and $\boldsymbol{\eta}=\boldsymbol{\xi}^{\mathrm{T}}$.

Let $\mathbf{J}:\mathfrak{so}(3)\to\mathfrak{so}^*(3)$ be the positive definite inertia operator. It can be shown that

$$\langle\mathbf{J}(\boldsymbol{\xi}),\boldsymbol{\omega}\rangle = \langle\mathbf{J}(\boldsymbol{\omega}),\boldsymbol{\xi}\rangle\,. \tag{2}$$

On $\mathfrak{so}(3)$, $\mathbf{J}$ is given by $\mathbf{J}(\boldsymbol{\xi})=J\boldsymbol{\xi}+\boldsymbol{\xi}J$, where J is a positive definite symmetric matrix (see, for example, [3, 8]). Moreover, we also have $\mathbf{J}(\boldsymbol{\eta}^\sharp)^\sharp = (J\boldsymbol{\eta}^{\mathrm{T}}+\boldsymbol{\eta}^{\mathrm{T}}J)^{\mathrm{T}} = \mathbf{J}(\boldsymbol{\eta})$, which is an abuse of notation since $\boldsymbol{\eta}\in\mathfrak{so}^*(3)$. For the sake of generality and mathematical precision we will use the general definitions, though it helps to keep the above identifications for $\mathfrak{so}(3)$ in mind.

In this section we review some continuous-time optimal control results using a simple optimal control example on $\mathsf{SO}(3)$. The problem we consider is that of minimizing the norm squared of the control torque $\boldsymbol{\tau}\in\mathfrak{so}^*(3)$ applied to rotate a rigid body subject to the Lagrange–d'Alembert principle for the rigid body[1] whose configuration is given by $\mathbf{R}\in\mathsf{SO}(3)$ and body angular velocity is given by $\boldsymbol{\Omega}\in\mathfrak{so}(3)$. We require that the system evolve from an initial state $(\mathbf{R}_0,\boldsymbol{\Omega}_0)$ to a final state $(\mathbf{R}_T,\boldsymbol{\Omega}_T)$ at a fixed terminal time T.

Before proceeding with a statement of the optimal control problem, we first define variations of the rigid body configuration $\mathbf{R}$ and its velocity $\boldsymbol{\Omega}$. Let $\mathbf{W}(t)\in\mathfrak{so}(3)$ be the variation vector field associated with a curve $\mathbf{R}(t)$ on $\mathsf{SO}(3)$ [2, 28]. The vector field $\mathbf{W}(t)$ satisfies

$$\boldsymbol{\delta}\mathbf{R}(t)=\mathbf{R}\mathbf{W}\in\mathsf{T}_{\mathbf{R}(t)}\mathsf{SO}(3),\ \mathbf{W}(0)=\mathbf{0},\ \mathbf{W}(T)=\mathbf{0},$$

where $\boldsymbol{\delta}\mathbf{R}$ is defined by $\boldsymbol{\delta}\mathbf{R}(t)=\partial\mathbf{R}_\epsilon(t)/\partial\epsilon\big|_{\epsilon=0}$, with $\mathbf{R}_\epsilon(t):=\mathbf{R}(t,\epsilon)$ is the variation of the curve $\mathbf{R}(t)$ that satisfies $\mathbf{R}(t,0)=\mathbf{R}(t)$. The variation in the velocity vector field is denoted $\boldsymbol{\delta\Omega}$. For a deeper understanding of variations of general vector fields, see for example the treatment in [7].

We now state the minimum control effort optimal control problem.

[1]This is equivalent to constraining the problem to satisfy the rigid body equations of motion given by equations (7). However, for the sake of generality that will be appreciated in the discrete-time problem, we choose to treat the Lagrange–d'Alembert principle as the constraint as opposed to the rigid body equations of motion. Both are equivalent in the continuous-time case but are generally not equivalent in the discrete-time case.

Problem 2.1. *Minimize*

$$\mathcal{J} = \frac{1}{2}\int_0^T \ll \boldsymbol{\tau}, \boldsymbol{\tau} \gg_* \mathrm{d}t \tag{3}$$

subject to

(1) *satisfying* Lagrange–d'Alembert principle*:*

$$\boldsymbol{\delta}\int_0^T \frac{1}{2}\left\langle \mathbf{J}\left(\boldsymbol{\Omega}\right), \boldsymbol{\Omega}\right\rangle \mathrm{d}t + \int_0^T \left\langle \boldsymbol{\tau}, \mathbf{W}\right\rangle \mathrm{d}t = 0, \tag{4}$$

for a variation vector field $\mathbf{W}(t)$*, and subject to* $\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\Omega}$*,*

(2) *and the* boundary conditions

$$\begin{aligned} \mathbf{R}(0) &= \mathbf{R}_0,\ \boldsymbol{\Omega}(0) = \boldsymbol{\Omega}_0, \\ \mathbf{R}(T) &= \mathbf{R}_T,\ \boldsymbol{\Omega}(T) = \boldsymbol{\Omega}_T. \end{aligned} \tag{5}$$

We now show that the constraint of satisfying the Lagrange–d'Alembert principle leads to the following problem formulation, where the rigid body equations of motion replace the Lagrange–d'Alembert principle.

Problem 2.2. *Minimize*

$$\mathcal{J} = \frac{1}{2}\int_0^T \ll \boldsymbol{\tau}, \boldsymbol{\tau} \gg_* \mathrm{d}t \tag{6}$$

subject to

(1) *the* dynamics

$$\begin{aligned} \dot{\mathbf{R}} &= \mathbf{R}\boldsymbol{\Omega} \\ \dot{\mathbf{M}} &= \mathrm{ad}^*_{\boldsymbol{\Omega}}\mathbf{M} + \boldsymbol{\tau} = [\mathbf{M}, \boldsymbol{\Omega}] + \boldsymbol{\tau}, \end{aligned} \tag{7}$$

where $\mathbf{M} = \mathbf{J}(\boldsymbol{\Omega}) \in \mathfrak{so}^*(3)$ *is the momentum,*

(2) *and the* boundary conditions

$$\begin{aligned} \mathbf{R}(0) &= \mathbf{R}_0,\ \boldsymbol{\Omega}(0) = \boldsymbol{\Omega}_0, \\ \mathbf{R}(T) &= \mathbf{R}_T,\ \boldsymbol{\Omega}(T) = \boldsymbol{\Omega}_T. \end{aligned} \tag{8}$$

In the above, ad^* is the dual of the adjoint representation, ad, of $\mathfrak{so}(3)$ and is given by $\mathrm{ad}^*_{\boldsymbol{\xi}}\boldsymbol{\eta} = -[\boldsymbol{\xi}, \boldsymbol{\eta}] \in \mathfrak{so}^*(3)$, for all $\boldsymbol{\xi} \in \mathfrak{so}(3)$ and $\boldsymbol{\eta} \in \mathfrak{so}^*(3)$. Recall that the bracket is defined by $[\boldsymbol{\xi}, \boldsymbol{\omega}] = \boldsymbol{\xi}\boldsymbol{\omega} - \boldsymbol{\omega}\boldsymbol{\xi}$.

2.2. **The Lagrange–d'Alembert Principle and the Rigid Body Equations of Motion.** In this section we derive the forced rigid body equations of motion (equations (7)) from the Lagrange–d'Alembert principle. We begin by appending the constraint $\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\Omega}$ to the Lagrangian

$$\begin{aligned} 0 = \boldsymbol{\delta}\int_0^T \left(\frac{1}{2}\left\langle \mathbf{J}\left(\boldsymbol{\Omega}\right), \boldsymbol{\Omega}\right\rangle + \left\langle \boldsymbol{\Lambda}, \mathbf{R}^{-1}\dot{\mathbf{R}} - \boldsymbol{\Omega}\right\rangle\right)\mathrm{d}t \\ + \int_0^T \left\langle \boldsymbol{\tau}, \mathbf{W}\right\rangle \mathrm{d}t, \end{aligned}$$

where $\boldsymbol{\Lambda} \in \mathfrak{so}^*(3)$ is a Lagrange multiplier. Taking variations, we obtain

$$\begin{aligned} 0 = \int_0^T \left(\left\langle -\dot{\boldsymbol{\Lambda}} - [\boldsymbol{\Omega}, \boldsymbol{\Lambda}] + \boldsymbol{\tau}, \mathbf{W}\right\rangle + \left\langle -\boldsymbol{\Lambda} + \mathbf{J}\left(\boldsymbol{\Omega}\right), \boldsymbol{\delta}\boldsymbol{\Omega}\right\rangle\right)\mathrm{d}t \\ + \left[\left\langle \boldsymbol{\Lambda}, \mathbf{W}(t)\right\rangle\right]_0^T. \end{aligned}$$

The term outside the integral vanishes by virtue of the boundary conditions on $\mathbf{W}(t)$. Since $\boldsymbol{\delta\Omega}$ and $\mathbf{W}$ are arbitrary and independent, we must have $\left\langle -\boldsymbol{\Lambda} + \mathbf{J}(\boldsymbol{\Omega}), \boldsymbol{\delta\Omega}\right\rangle = 0$ and hence $\boldsymbol{\Lambda} = \mathbf{J}(\boldsymbol{\Omega})$ which is the body angular momentum. For the remainder of this section, we set $\mathbf{M} = \boldsymbol{\Lambda} = \mathbf{J}(\boldsymbol{\Omega})$. The first term in the integrand gives us the forced second order dynamics of the rigid body which is

$$\dot{\mathbf{M}} = [\mathbf{M}, \boldsymbol{\Omega}] + \boldsymbol{\tau}.$$

This completes the proof that the Problem (2.1) is equivalent to Problem (2.2).

A Direct Approach. We now give a direct derivation that does not involve the use of Lagrange multipliers. This approach can be found in Section 13.5 in [21]. First, we take variations of the kinematic condition $\boldsymbol{\Omega} = \mathbf{R}^{-1}\dot{\mathbf{R}}$ to obtain $\boldsymbol{\delta\Omega} = -\mathbf{R}^{-1}(\boldsymbol{\delta R})\mathbf{R}^{-1}\dot{\mathbf{R}} + \mathbf{R}^{-1}\left(\boldsymbol{\delta}\dot{\mathbf{R}}\right)$. As defined previously, we have $\mathbf{W} = \mathbf{R}^{-1}\boldsymbol{\delta R}$ and, therefore, $\dot{\mathbf{W}} = -\mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{R}^{-1}\boldsymbol{\delta R} + \mathbf{R}^{-1}\boldsymbol{\delta}\dot{\mathbf{R}} = -\boldsymbol{\Omega}\mathbf{W} + \mathbf{R}^{-1}\boldsymbol{\delta}\dot{\mathbf{R}}$, since $\boldsymbol{\delta}\dot{\mathbf{R}} = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\delta R}$ (see for example [22], page 52). Hence, we have

$$\boldsymbol{\delta\Omega} = -\mathbf{W}\boldsymbol{\Omega} + \boldsymbol{\Omega}\mathbf{W} + \dot{\mathbf{W}} = \mathrm{ad}_{\Omega}\mathbf{W} + \dot{\mathbf{W}}. \tag{9}$$

Taking variations of the Lagrange–d'Alembert principle we obtain

$$\int_0^T \langle \mathbf{J}(\boldsymbol{\Omega}), \boldsymbol{\delta\Omega}\rangle + \langle \boldsymbol{\tau}, \mathbf{W}\rangle \,\mathrm{d}t = 0.$$

Using the variation in equation (9) and integrating by parts, we obtain

$$0 = \int_0^T \left\langle -\dot{\mathbf{M}} + \mathrm{ad}^*_{\Omega}\mathbf{M} + \boldsymbol{\tau}, \mathbf{W}\right\rangle \mathrm{d}t + \left[\langle \mathbf{J}(\boldsymbol{\Omega}), \mathbf{W}(t)\rangle\right]_0^T,$$

where we have used the identity

$$\langle \boldsymbol{\eta}, \mathrm{ad}_{\boldsymbol{\omega}}\boldsymbol{\xi}\rangle = \langle \mathrm{ad}^*_{\boldsymbol{\omega}}\boldsymbol{\eta}, \boldsymbol{\xi}\rangle, \ \boldsymbol{\eta} \in \mathfrak{so}^*(3), \ \boldsymbol{\omega}, \boldsymbol{\xi} \in \mathfrak{so}(3). \tag{10}$$

This gives the desired result, with $\mathbf{M} = \mathbf{J}(\boldsymbol{\Omega})$.

In Section 2.3, we demonstrate how the necessary conditions for Problem (2.2) are derived using a variational approach.

2.3. **Continuous-Time Variational Optimal Control Problem.** A direct variational approach is used here to obtain the differential equation that satisfies the optimal control Problem (2.2).

A Second Order Direct Approach. "Second order" is used here to reflect the fact that we now study variations of second order dynamical equations as opposed to the kinematic direct approach studied in Section 2.2. We now give the resulting necessary conditions using a direct approach as in [21]. We already computed the variations of $\mathbf{R}$ and $\boldsymbol{\Omega}$. These were as follows: $\boldsymbol{\delta R} = \mathbf{RW}$ and $\boldsymbol{\delta\Omega} = \mathrm{ad}_{\boldsymbol{\Omega}}\mathbf{W} + \dot{\mathbf{W}}$. We now compute the variation of $\dot{\mathbf{M}}$ with the goal of obtaining the proper variations for $\boldsymbol{\tau}$:

$$\boldsymbol{\delta}\dot{\mathbf{M}} \quad = \quad \mathbf{J}\left(\boldsymbol{\delta}\dot{\boldsymbol{\Omega}}\right) = \mathbf{J}\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\delta\Omega} + \mathcal{R}(\mathbf{W}, \boldsymbol{\Omega})\boldsymbol{\Omega}\right),$$

where $\mathcal{R}$ is the curvature tensor on SO(3). The curvature tensor $\mathcal{R}$ arises due to the identity (see [22], page 52)

$$\frac{\partial}{\partial\epsilon}\frac{\partial}{\partial t}\mathbf{Y} - \frac{\partial}{\partial t}\frac{\partial}{\partial\epsilon}\mathbf{Y} = \mathcal{R}(\mathbf{W}, \mathbf{Y})\boldsymbol{\Omega},$$

where $\mathbf{Y} \in \mathsf{TSO}(3)$ is any vector field along the curve $\mathbf{R}(t) \in \mathsf{SO}(3)$. Taking variations of $\dot{\mathbf{M}} = \mathrm{ad}^*_{\mathbf{\Omega}}\mathbf{M} + \boldsymbol{\tau}$, we obtain $\boldsymbol{\delta}\dot{\mathbf{M}} = \mathrm{ad}^*_{\boldsymbol{\delta\Omega}}\mathbf{M} + \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\delta}\mathbf{M} + \boldsymbol{\delta\tau}$. We now have the desired variation in $\boldsymbol{\tau}$:

$$\boldsymbol{\delta\tau} = \mathbf{J}\left(\mathcal{R}\left(\mathbf{W},\mathbf{\Omega}\right)\mathbf{\Omega}\right) + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{J}\left(\boldsymbol{\delta\Omega}\right) - \mathrm{ad}^*_{\boldsymbol{\delta\Omega}}\mathbf{M} - \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\delta}\mathbf{M}. \tag{11}$$

Taking variations of the cost functional (6) we obtain:

$$\begin{aligned}\boldsymbol{\delta}\mathcal{J} = \int_0^T \Big(&\big\langle \mathbf{J}(\ddot{\varsigma}) - \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathbf{J}(\dot{\varsigma})\right) + \dot{\boldsymbol{\eta}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{ad}^*_{\boldsymbol{\varsigma}}\mathbf{M}\right) \\ &+ \left[\mathcal{R}\left(\mathbf{J}(\boldsymbol{\varsigma})^\sharp,\mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\mathrm{ad}^*_{\boldsymbol{\varsigma}}\mathbf{M} - \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\eta}, \mathbf{W}\big\rangle\Big)\mathrm{d}t,\end{aligned}$$

where $\boldsymbol{\varsigma} = \boldsymbol{\tau}^\sharp \in \mathfrak{so}(3)$ and $\boldsymbol{\eta} = \mathbf{J}\left(\mathrm{ad}_{\mathbf{\Omega}}\boldsymbol{\varsigma}\right) \in \mathfrak{so}^*(3)$. In obtaining the above expression, we have used integration by parts and the boundary conditions (8), equations (9) and (11), and the identities (1), (2) and (10). Hence, we have the following theorem.

Theorem 2.1. *The necessary optimality conditions for the problem of minimizing (6) subject to the dynamics (7) and the boundary conditions (8) are given by the single fourth order*[2] *differential equation*

$$\begin{aligned}0 = &\ \mathbf{J}(\ddot{\varsigma}) - \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathbf{J}(\dot{\varsigma})\right) + \dot{\boldsymbol{\eta}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{ad}^*_{\boldsymbol{\varsigma}}\mathbf{M}\right) \\ &+ \left[\mathcal{R}\left(\left(\mathbf{J}(\boldsymbol{\varsigma})\right)^\sharp,\mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathrm{ad}^*_{\boldsymbol{\varsigma}}\mathbf{M}\right) - \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\eta},\end{aligned}$$

as well as the equations (7) and the boundary conditions (8), where $\boldsymbol{\varsigma}$ and $\boldsymbol{\eta}$ are as defined above.

Note that for a compact semi-simple Lie group G with Lie algebra $\mathfrak{g}$, the curvature tensor, with respect to a bi-invariant metric, is given by (see [22]):

$$\mathcal{R}\left(\mathbf{X},\mathbf{Y}\right)\mathbf{Z} = \frac{1}{4}\mathrm{ad}_{\mathrm{ad}_{\mathbf{X}}\mathbf{Y}}\mathbf{Z}, \tag{12}$$

for all $\mathbf{X},\mathbf{Y},\mathbf{Z} \in \mathfrak{g}$.

Using a Lagrange multiplier approach as in Section 2.2, we may show that the result of Theorem 2.1 is equivalent to the following theorem.

Theorem 2.2. *The necessary optimality conditions for the problem of minimizing (6) subject to the dynamics (7) and the boundary conditions (8) are given by*

$$\begin{aligned}\boldsymbol{\tau} &= \mathbf{\Lambda}_2 \\ \dot{\mathbf{\Lambda}}_1 &= \left[\mathcal{R}\left(\mathbf{J}\left(\mathbf{\Lambda}_2\right)^\sharp,\mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\mathbf{\Lambda}_1 \\ \dot{\mathbf{\Lambda}}_2 &= -\mathbf{J}^{-1}\left(\mathbf{\Lambda}_1\right) - \mathrm{ad}_{\mathbf{\Omega}}\mathbf{\Lambda}_2 + \mathbf{J}^{-1}\left(\mathrm{ad}^*_{\mathbf{\Lambda}_2}\mathbf{M}\right)\end{aligned} \tag{13}$$

as well as the equations (7) and the boundary conditions (8), where the Lagrange multipliers $\mathbf{\Lambda}_1 \in \mathfrak{so}^*(3)$, $\mathbf{\Lambda}_2 \in \mathfrak{so}(3)$ correspond to the kinematic and dynamics constraints (7), respectively.

[2]Second order in $\boldsymbol{\tau}$ and fourth order in $\mathbf{R}$.

Remark 2.1. *Note that the equations of motion that arise from the Lagrange–d'Alembert principle are used to define the dynamic constraints. So, in effect, we are minimizing $\mathcal{J}$ subject to satisfying the Lagrange–d'Alembert principle for the rigid body. Analogously, the discrete version of the Lagrange–d'Alembert principle will be used to derive the discrete equations of motion in the discrete optimal control problem to be studied in Section 3.3. This view is in line with the approach in* [11] *in that we do not discretize the equations of motion directly, but, instead, we discretize the Lagrange–d'Alembert principle. These two approaches are not equivalent in general.*

Corollary 2.1. *The necessary optimality conditions of Theorem 2.1 are equivalent to the necessary conditions of Theorem 2.2.*

Proof. In Theorem 2.2, differentiate $\mathbf{\Lambda}_2$ once and then use all three differential equations to replace $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ with expressions involving only $\boldsymbol{\tau}$, $\mathbf{M}$ and $\mathbf{\Omega}$. □

3. Discrete-Time Results

3.1. **Problem Formulation.** In this section we give the discrete version of the problem introduced in Section 2.1. So, we consider minimizing the norm squared of the control torque $\boldsymbol{\tau}$ subject to satisfaction of the discrete Lagrange–d'Alembert principle for the rigid body whose configuration and body angular velocity at time step t_k are given by $\mathbf{R}_k \in \mathsf{SO}(3)$ and $\mathbf{\Omega}_k \in \mathfrak{so}(3)$, respectively. The kinematic constraint may be expressed as

$$\mathbf{R}_{k+1} = \mathbf{R}_k \exp\left(h\mathbf{\Omega}_k\right) = \mathbf{R}_k \mathbf{g}_k, \tag{14}$$

where h is the integration time step, $\exp : \mathfrak{so}(3) \to \mathsf{SO}(3)$ is the exponential map and $\mathbf{g}_k = \exp(h\mathbf{\Omega}_k)$. The boundary conditions are given by $(\mathbf{R}_0^*, \mathbf{\Omega}_0^*)$ and $(\mathbf{R}_N^*, \mathbf{\Omega}_{N-1}^*)$, where $t_0 = 0$ and $N = T/h$ is such that $t_N = T$.

More generally, one considers the ansatz $\mathbf{R}_{k+1} = \mathbf{R}_k \exp\left(\mathbf{\Omega}(h)\right)$, where $\mathbf{\Omega}(\cdot)$ is an interpolatory curve in $\mathfrak{so}(3)$ parameterized by the angular velocity at internal nodal points. This allows one to construct Lie group variational integrators of arbitrarily high order [17]. To simplify the subsequent treatment, we adopt (14) as the kinematic constraint, which yields a first-order accurate Lie symplectic Euler method, which will nevertheless have effective order two as it is symplectically conjugate to the second-order accurate Lie Störmer–Verlet method (see, §4).

The reason we constrain $\mathbf{\Omega}$ at $t = h(N-1)$ instead of at $t = hN$ will become clear when we derive the discrete equations of motion in Section 3.2. A simple explanation for this is that a constraint on $\mathbf{\Omega}_k \in \mathfrak{so}(3)$ corresponds, by left translations to a constraint on $\dot{\mathbf{R}}_k \in \mathsf{T}_{\mathbf{R}_k}\mathsf{SO}(3)$. In turn, in the discrete setting and depending on the choice of discretization, this corresponds to a constraint on the neighboring discrete points $\ldots, \mathbf{R}_{k-2}, \mathbf{R}_{k-1}, \mathbf{R}_{k+1}, \mathbf{R}_{k+2}, \ldots$. With our choice of discretization (equation (14)), this corresponds to constraints on $\mathbf{R}_k$ and $\mathbf{R}_{k+1}$. Hence, to ensure that the effect of the terminal constraint on $\mathbf{\Omega}$ is correctly accounted for, the constraint must be imposed on $\mathbf{\Omega}_{N-1}$, which entails some constraints on variations at both $\mathbf{R}_{N-1}$ and $\mathbf{R}_N$. We will return to this point later in the paper.

The discrete kinematic constraint ensures that the sequence $\mathbf{R}_k$ stays on the rotation group, since the exponential of the angular velocity matrix $\mathbf{\Omega}_k$, which is in the algebra $\mathfrak{so}(3)$, is a rotation matrix, and the rotation group is closed under matrix multiplication. This is natural to do in the context of discrete variational numerical solvers (for both initial value and two point boundary value problems).

Following the methodology of [11], we have the following optimal control problem.

Problem 3.1. *Minimize*

$$\mathcal{J} = \sum_{k=0}^{N} \frac{1}{2} \ll \boldsymbol{\tau}_k, \boldsymbol{\tau}_k \gg_* \tag{15}$$

subject to

(1) *satisfying the* discrete Lagrange–d'Alembert principle*:*

$$\boldsymbol{\delta} \sum_{k=0}^{N-1} \frac{1}{2} \left\langle \mathbf{J}\left(\boldsymbol{\Omega}_k\right), \boldsymbol{\Omega}_k \right\rangle + \sum_{k=0}^{N} \left\langle \boldsymbol{\tau}_k, \mathbf{W}_k \right\rangle = 0, \tag{16}$$

subject to $\mathbf{R}_0 = \mathbf{R}_0^*$, $\mathbf{R}_N = \mathbf{R}_N^*$ *and* $\mathbf{R}_{k+1} = \mathbf{R}_k \mathbf{g}_k$, $k = 0, 1, \ldots, N-1$, *where* $\mathbf{W}_k$ *is the variation vector field at time step* t_k *satisfying* $\boldsymbol{\delta}\mathbf{R}_k = \mathbf{R}_k \mathbf{W}_k$,

(2) *and the* boundary conditions

$$\begin{aligned} \mathbf{R}_0 &= \mathbf{R}_0^*, \ \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_0^*, \\ \mathbf{R}_N &= \mathbf{R}_N^*, \ \boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}_{N-1}^*. \end{aligned} \tag{17}$$

In Problem 3.1, the discrete Lagrange–d'Alembert principle is used to derive the equations of motion for the rigid body with initial and terminal configuration constraints. Hence, we get a two point boundary value problem. The full configuration and velocity boundary conditions come into the picture when we study the optimal control problem. We will show that the constraint of satisfying the Lagrange–d'Alembert principle in Problem 3.1 leads to the following problem formulation, where the discrete rigid body equations of motion replace the Lagrange–d'Alembert principle constraint. Only when addressing the following optimal control problem will we need to include the velocity boundary conditions in the derivation.

Problem 3.2. *Minimize*

$$\mathcal{J} = \sum_{k=0}^{N} \frac{1}{2} \ll \boldsymbol{\tau}_k, \boldsymbol{\tau}_k \gg_* \tag{18}$$

subject to

(1) *the* discrete dynamics

$$\begin{aligned} &\mathbf{R}_{k+1} = \mathbf{R}_k \mathbf{g}_k, \ k = 0, \ldots, N-1 \\ &\mathbf{M}_k = \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right), \ k = 1, \ldots, N-1, \\ &\mathbf{M}_k = \mathbf{J}\left(\boldsymbol{\Omega}_k\right), \ k = 0, \ldots, N-1, \end{aligned} \tag{19}$$

(2) *and the* boundary conditions

$$\begin{aligned} \mathbf{R}_0 &= \mathbf{R}_0^*, \ \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_0^*, \\ \mathbf{R}_N &= \mathbf{R}_N^*, \ \boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}_{N-1}^*. \end{aligned} \tag{20}$$

Regarding terminal velocity conditions, note that in the second of equations (19) if we let $k = N$ we find that $\boldsymbol{\Omega}_N$ appears in the equation. A constraint on $\boldsymbol{\Omega}_N$ dictates constraints at the points $\mathbf{R}_N$ and $\mathbf{R}_{N+1}$ through the first equation in (19). Since we only consider time points up to $t = Nh$, we can not allow $k = N$ in the second of equations (19) and hence our terminal velocity constraints are posed in terms of $\boldsymbol{\Omega}_{N-1}$ instead of $\boldsymbol{\Omega}_N$.

As mentioned above, $\mathbf{W}_k$ is a variation vector field associated with the perturbed group element $\mathbf{R}_k^\epsilon$. Likewise, we need to define a variation vector field associated with the element $\mathbf{g}_k = \exp(h\boldsymbol{\Omega}_k)$. First, let the perturbed variable $\mathbf{g}_k^\epsilon$ be defined by

$$\mathbf{g}_k^\epsilon = \mathbf{g}_k \exp(\epsilon h \boldsymbol{\delta\Omega}_k), \tag{21}$$

where

$$\boldsymbol{\delta\Omega}_k = \left.\frac{\partial \boldsymbol{\Omega}_k^\epsilon}{\partial \epsilon}\right|_{\epsilon=0}.$$

Note that $\mathbf{g}_k^\epsilon\big|_{\epsilon=0} = \mathbf{g}_k$ as desired. Moreover, we have

$$\boldsymbol{\delta}\mathbf{g}_k = \mathbf{g}_k(h\boldsymbol{\delta\Omega}_k)\exp(\epsilon h\delta\boldsymbol{\Omega}_k)\big|_{\epsilon=0} = h\mathbf{g}_k\boldsymbol{\delta\Omega}_k. \tag{22}$$

This will be needed later when taking variations.

3.2. **The Discrete Lagrange–d'Alembert Principle and the Rigid Body Equations of Motion.** In this section we derive the discrete forced rigid body equations of motion (equations (19)) starting with the discrete Lagrange–d'Alembert principle. We begin by rewriting the kinematic constraint as $\exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) = h\boldsymbol{\Omega}_k$, which is an expression on the Lie algebra $\mathfrak{so}(3)$, and appending it to the Lagrangian

$$\begin{aligned} 0 \;=\; & \sum_{k=0}^{N} \langle \boldsymbol{\tau}_k, \mathbf{W}_k \rangle + \boldsymbol{\delta} \sum_{k=0}^{N-1} \left(\frac{1}{2} \left\langle \mathbf{J}\left(\boldsymbol{\Omega}_k\right), \boldsymbol{\Omega}_k \right\rangle \right. \\ & \left. + \left\langle \mathbf{M}_k, \frac{1}{h} \exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) - \boldsymbol{\Omega}_k \right\rangle \right), \end{aligned}$$

where $\mathbf{M}_k \in \mathfrak{so}^*(3)$, $k = 0, 1, \ldots, N-1$, are Lagrange multipliers. These multipliers enforce the kinematic discrete equations. We could have added additional terms to enforce the configuration boundary conditions, allowing for $\mathbf{W}_0$ and $\mathbf{W}_N$ to be arbitrary and non-zero. Instead, we elect to enforce the constraints by requiring $\mathbf{W}_0 = 0$ and $\boldsymbol{\delta\Omega}_0 = 0$. These two approaches are of course equivalent.

Taking variations of $\frac{1}{h}\exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) - \boldsymbol{\Omega}_k = 0$ is equivalent to taking variations of the original expression $\mathbf{R}_k^{-1}\mathbf{R}_{k+1} = \exp\left(h\boldsymbol{\Omega}_k\right)$, and is easier to compute since it as an expression over the Lie algebra. Once the variations are computed, one can easily obtain the Lie algebra-equivalent of the variations as follows. First take variations of the kinematics (14) to get $-\mathbf{R}_k^{-1}\left(\boldsymbol{\delta}\mathbf{R}_k\right)\mathbf{R}_k^{-1}\mathbf{R}_{k+1} + \mathbf{R}_k^{-1}\boldsymbol{\delta}\mathbf{R}_{k+1} = h\mathbf{g}_k \cdot \boldsymbol{\delta\Omega}_k$, which is equivalent to $-\mathbf{W}_k\mathbf{g}_k + \mathbf{g}_k\mathbf{W}_{k+1} = h\mathbf{g}_k\boldsymbol{\delta\Omega}_k$, or

$$\boldsymbol{\delta\Omega}_k = \frac{1}{h}\left[-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k + \mathbf{W}_{k+1}\right]. \tag{23}$$

Note that this is an expression over the Lie algebra $\mathfrak{so}(3)$.

After simple algebraic and re-indexing operations, the Lagrange–d'Alembert principle gives

$$\begin{aligned} 0 = &\left\langle \boldsymbol{\tau}_0 - \frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_0^{-1}}\mathbf{M}_0, \mathbf{W}_0 \right\rangle + \left\langle \boldsymbol{\tau}_N + \frac{1}{h}\mathbf{M}_{N-1}, \mathbf{W}_N \right\rangle \\ &+ \sum_{k=0}^{N-1} \left\langle \mathbf{J}\left(\boldsymbol{\Omega}_k\right) - \mathbf{M}_k, \boldsymbol{\delta\Omega}_k \right\rangle \\ &+ \sum_{k=1}^{N-1} \left\langle \boldsymbol{\tau}_k - \frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{M}_k + \frac{1}{h}\mathbf{M}_{k-1}, \mathbf{W}_k \right\rangle. \end{aligned}$$

By the boundary conditions $\mathbf{R}_0 = \mathbf{R}_0^*$ and $\mathbf{R}_N = \mathbf{R}_N^*$, we have $\mathbf{W}_0 = 0$ and $\mathbf{W}_N = 0$. Since $\boldsymbol{\delta\Omega}_k$, $k = 0, \ldots, N-1$, and $\mathbf{W}_k$, $k = 1, \ldots, N-1$, are arbitrary and independent, then the Lagrange–d'Alembert principle requires that the equations (19) hold true. The variables $\mathbf{M}_k$, $k = 0, \ldots, N-1$, are of course nothing but the discrete angular momentum of the rigid body. The equations (19) can be viewed in two ways. The first is to consider the two point boundary value problem where we retain the terminal condition on $\mathbf{R}_N$. In this case a (constrained) variety of a combination of control torques $\boldsymbol{\tau}_k$, $k = 0, \ldots, N$, and initial velocity conditions $\boldsymbol{\Omega}_0$ can be chosen to drive the rigid body from the initial condition $\mathbf{R}_0$ to the terminal condition $\mathbf{R}_N$. The second view is to treat it as an initial value problem by ignoring any terminal configuration constraints. In this case $\mathbf{W}_N \neq 0$ and any combination of control torques $\boldsymbol{\tau}_k$, $k = 0, \ldots, N$, and initial velocity conditions $\boldsymbol{\Omega}_0$ can be chosen freely.

A Direct Approach. Taking direct variations of the cost functional we obtain

$$\begin{aligned} 0 = &\left\langle \boldsymbol{\tau}_0 - \frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_0^{-1}}\mathbf{J}\left(\boldsymbol{\Omega}_0\right), \mathbf{W}_0 \right\rangle \\ &+ \left\langle \boldsymbol{\tau}_N + \frac{1}{h}\mathbf{J}\left(\Omega_{N-1}\right), \mathbf{W}_N \right\rangle \\ &+ \sum_{k=1}^{N-1} \left\langle \boldsymbol{\tau}_k - \frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{J}\left(\boldsymbol{\Omega}_k\right) + \frac{1}{h}\mathbf{J}\left(\boldsymbol{\Omega}_{k-1}\right), \mathbf{W}_k \right\rangle. \end{aligned}$$

where we have used equation (23). This gives the same equations of motion as those in equation (19).

Simulation Results. To test our results, we re-write the discrete equations (19) for the subgroup $\mathsf{SO}(2)$. For $\mathsf{SO}(2)$ we have

$$\mathbf{R}_k = \begin{bmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{bmatrix}, \ \boldsymbol{\Omega}_k = \begin{bmatrix} 0 & -\omega_k \\ \omega_k & 0 \end{bmatrix} \tag{24}$$

and

$$\exp\left(\boldsymbol{\Omega}_k\right) = \begin{bmatrix} \cos\omega_k & -\sin\omega_k \\ \sin\omega_k & \cos\omega_k \end{bmatrix}. \tag{25}$$

The inertia operation is simply given by

$$\mathbf{J}\left(\boldsymbol{\Omega}_k\right) = \begin{bmatrix} 0 & -I\omega_k \\ I\omega_k & 0 \end{bmatrix}, \tag{26}$$

where I is the mass moment of inertia of the body about the out-of-plane axis. One can check that $\mathrm{Ad}_{\exp(\boldsymbol{\omega})}\boldsymbol{\xi} = \boldsymbol{\xi}$ and that $\mathrm{Ad}^*_{\exp(\boldsymbol{\omega})}\boldsymbol{\eta} = \boldsymbol{\eta}$, for all $\boldsymbol{\xi},\ \boldsymbol{\omega} \in \mathfrak{so}(2)$ and $\boldsymbol{\eta} \in \mathfrak{so}^*(2)$.

Then the equations (19) (treated as an initial value problem) are given for $\mathsf{SO}(2)$ by

$$\begin{aligned} \theta_{k+1} &= \theta_k + h\omega_k,\ k = 0, \ldots, N-1 \\ \omega_k &= \frac{h}{I}\tau_k + \omega_{k-1},\ k = 1, \ldots, N-1 \end{aligned} \tag{27}$$

in addition to the initial conditions $\theta_0 = \theta_0^*,\ \omega_0 = \omega_0^*$.

To verify the accuracy of our numerical computation, we give the corresponding continuous-time equations of motion for the planar rigid body on $\mathsf{SO}(2)$ using equations (7). The Lie bracket on $\mathsf{SO}(2)$ is identically equal to zero. Hence, one can check that the equations (7) are given by $\dot{\theta} = \omega$, $\dot{\omega} = \frac{\tau}{I}$, where θ, ω and τ are the continuous time angular position, velocity and torque, respectively. We integrate the equations using the torque $\tau(t) = \sin\left(\frac{\pi t}{2}\right),\ t \in [0, T]$. We use the following parameters for our simulations: $T = 10$, $I = 1$, $\theta(0) = 3$, $\omega(0) = 4$ and we try three different time steps corresponding to $N = 1000, 1500, 2000$. The error between the continuous- and discrete-time values of θ and ω are given in Figure (1). Note that the accuracy of the simulation improves with increasing N.

Remark 3.1. *Note that the discrete-time equations (27) correspond to the Euler approximation for the equations of motion. This is a check that our method returns something familiar for a simple example as the planar rigid body. However, we emphasize that on $\mathsf{SO}(3)$ the discretization will not necessarily be equivalent to any of the classical discretization schemes. The discretization will generally result in a set of nonlinear implicit algebraic equations.*

3.3. **Discrete-Time Variational Optimal Control Problem.** We now address Problem 3.2 by first forming the appended cost functional:

$$\begin{aligned} \mathcal{J} &= \sum_{k=0}^{N} \frac{1}{2} \ll \boldsymbol{\tau}_k, \boldsymbol{\tau}_k \gg_* \\ &\quad + \sum_{k=0}^{N-1} \left\langle \boldsymbol{\Lambda}_k^1, -h\boldsymbol{\Omega}_k + \exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) \right\rangle \\ &\quad + \sum_{k=1}^{N-1} \left\langle \mathbf{M}_k - \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right), \boldsymbol{\Lambda}_k^2 \right\rangle. \end{aligned}$$

Writing $\mathbf{M}_k = \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right)$ as $\mathbf{M}_k = \mathbf{g}_k^{-1}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right)\mathbf{g}_k$ and taking variations of this expression, we obtain

$$\begin{aligned} \boldsymbol{\delta}\mathbf{M}_k &= -h\boldsymbol{\delta\Omega}_k\mathbf{g}_k^{-1}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right)\mathbf{g}_k \\ &\quad + \mathbf{g}_k^{-1}\left(h\boldsymbol{\delta\tau}_k + \boldsymbol{\delta}\mathbf{M}_{k-1}\right)\mathbf{g}_k \\ &\quad + h\mathbf{g}_k^{-1}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right)\mathbf{g}_k\boldsymbol{\delta\Omega}_k \\ &= -h\boldsymbol{\delta\Omega}_k\mathbf{M}_k + \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\delta\tau}_k + \mathbf{J}\left(\boldsymbol{\delta\Omega}_{k-1}\right)\right) \\ &\quad + h\mathbf{M}_k\boldsymbol{\delta\Omega}_k \\ &= \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\delta\tau}_k + \mathbf{J}\left(\boldsymbol{\delta\Omega}_{k-1}\right)\right) + h\left[\mathbf{M}_k, \boldsymbol{\delta\Omega}_k\right]. \end{aligned}$$

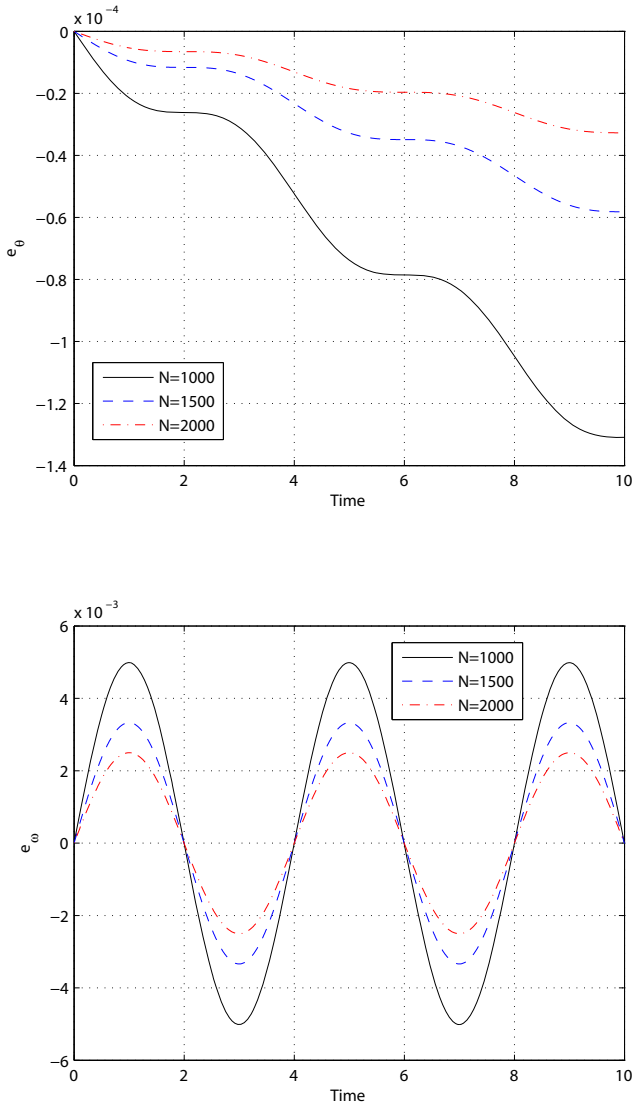


FIGURE 1. Error dynamics on $\mathsf{SO}(2)$.

In obtaining this expression we used the facts that $\boldsymbol{\delta}\mathbf{g}_k = h\mathbf{g}_k\boldsymbol{\delta}\boldsymbol{\Omega}_k$ and $\boldsymbol{\delta}\left(\mathbf{g}_k^{-1}\right) = -h\boldsymbol{\delta}\boldsymbol{\Omega}_k\mathbf{g}_k^{-1}$. The latter equality is obtained as follows. Taking variations of $\left(\mathbf{g}_k^{\epsilon}\right)^{-1}\left(\mathbf{g}_k^{\epsilon}\right) = \mathbf{I}$, we obtain

$$0 \quad = \quad \boldsymbol{\delta}\left(\mathbf{g}_k^{-1}\right)\mathbf{g}_k + h\mathbf{g}_k^{-1}\mathbf{g}_k\boldsymbol{\delta}\boldsymbol{\Omega}_k$$

which implies that

$$\boldsymbol{\delta}\left(\mathbf{g}_k^{-1}\right) = -h\boldsymbol{\delta}\boldsymbol{\Omega}_k\mathbf{g}_k^{-1},$$

which is the desired result.

We now also consider the velocity boundary conditions $\boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_0^*$ and $\boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}_{N-1}^*$. Note that variations in $\boldsymbol{\delta\Omega}_k$ directly induce variations in $\mathbf{W}_{k+1}$. In particular, if $k = 0$ and we have the initial constraints $\mathbf{R}_0 = \mathbf{R}_0^*$ and $\boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_0^*$, then $\mathbf{W}_0 = 0$ and $\boldsymbol{\delta\Omega}_0 = 0$. Using these two equations in equation (23), we find that

$$\mathbf{W}_1 = 0. \tag{28}$$

At $k = N, N-1$, the constraints $\mathbf{R}_N = \mathbf{R}_N^*$ and $\boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}_{N-1}^*$ imply that $\mathbf{W}_N = 0$ and $\boldsymbol{\delta\Omega}_{N-1} = 0$. Using these two equations in equation (23), we find that

$$\mathbf{W}_{N-1} = 0. \tag{29}$$

The observations stated in equations (28) and (29) are equivalent to having

$$\mathbf{R}_1 = \mathbf{R}_0^* \exp\left(h\boldsymbol{\Omega}_0^*\right), \ \mathbf{R}_{N-1} = \mathbf{R}_N^* \exp\left(-h\boldsymbol{\Omega}_{N-1}^*\right). \tag{30}$$

Taking variations of the cost functional, we obtain

$$\begin{aligned}
\boldsymbol{\delta}\mathcal{J} = & \sum_{k=0}^{N} \left\langle \boldsymbol{\delta\tau}_k, \boldsymbol{\tau}_k^\sharp \right\rangle \\
& + \sum_{k=0}^{N-1} \left\langle \boldsymbol{\Lambda}_k^1, -\boldsymbol{\delta\Omega}_k + \frac{1}{h}\left[-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k + \mathbf{W}_{k+1}\right] \right\rangle \\
& + \sum_{k=1}^{N-1} \left\langle \mathbf{J}\left(\boldsymbol{\delta\Omega}_k\right) - \mathrm{Ad}_{\mathbf{g}_k}^*\left(h\boldsymbol{\delta\tau}_k + \mathbf{J}\left(\boldsymbol{\delta\Omega}_{k-1}\right)\right) \right. \\
& \qquad \left. -h\left[\mathbf{M}_k, \boldsymbol{\delta\Omega}_k\right], \boldsymbol{\Lambda}_k^2 \right\rangle,
\end{aligned}$$

where we have replaced $\boldsymbol{\delta}\mathbf{M}_k$ with $\mathbf{J}\left(\boldsymbol{\delta\Omega}_k\right)$. Collecting terms, setting $\boldsymbol{\delta}\mathcal{J}$ to zero, and using the conditions $\mathbf{W}_0 = \mathbf{W}_1 = \mathbf{W}_{N-1} = \mathbf{W}_N = \boldsymbol{\delta\Omega}_0 = \boldsymbol{\delta\Omega}_{N-1} = 0$ and the fact that $\mathbf{J}(\cdot)$ is self-adjoint, we obtain the following theorem.

Theorem 3.1. *The necessary optimality conditions for the discrete Problem 3.2 are*

$$\begin{aligned}
\mathbf{R}_{k+1} &= \mathbf{R}_k\mathbf{g}_k, \ k = 1, \ldots, N-2 \\
\mathbf{M}_k &= \mathrm{Ad}_{\mathbf{g}_k}^*\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right), \ k = 1, \ldots, N-1 \\
0 &= \boldsymbol{\Lambda}_{k-1}^1 - \mathrm{Ad}_{\mathbf{g}_k^{-1}}^*\boldsymbol{\Lambda}_k^1, \ k = 2, \ldots, N-2 \\
0 &= -\boldsymbol{\Lambda}_k^1 + \mathbf{J}\left(\boldsymbol{\Lambda}_k^2\right) - \mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_{k+1}}\boldsymbol{\Lambda}_{k+1}^2\right) \\
&\quad + h\left[\mathbf{M}_k, \boldsymbol{\Lambda}_k^2\right], \ k = 1, \ldots, N-2 \\
\boldsymbol{\tau}_k &= h\left(\mathrm{Ad}_{\mathbf{g}_k}\boldsymbol{\Lambda}_k^2\right)^\flat, \ k = 1, \ldots, N-1 \\
\mathbf{M}_k &= \mathbf{J}\left(\boldsymbol{\Omega}_k\right), \ k = 0, \ldots, N-1,
\end{aligned} \tag{31}$$

and the boundary conditions

$$\begin{aligned}
\mathbf{R}_0 &= \mathbf{R}_0^*, \ \mathbf{R}_1 = \mathbf{R}_0^*\mathbf{g}_0^*, \ \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_0^* \\
\mathbf{R}_N &= \mathbf{R}_N^*, \ \mathbf{R}_{N-1} = \mathbf{R}_N^*\left(\mathbf{g}_{N-1}^*\right)^{-1}, \ \boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}_{N-1}^* \\
\boldsymbol{\tau}_0 &= \boldsymbol{\tau}_N = 0,
\end{aligned}$$

where $\mathbf{g}_0^* = \exp(h\boldsymbol{\Omega}_0^*)$ *and* $\mathbf{g}_{N-1}^* = \exp\left(h\boldsymbol{\Omega}_{N-1}^*\right)$.

A Second Order Direct Approach. Analogous to the direct approach in continuous time, here we derive the necessary optimality conditions in a form that does not involve the use of Lagrange multipliers. Using equation (23) and taking variation of the second of equations (19), we obtain

$$\begin{aligned}\boldsymbol{\delta\tau}_k &= \mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\left(\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_{k+1}-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k\right)\right.\\ &\quad\left.+\frac{1}{h}\left[\mathbf{W}_{k+1}-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k,\mathbf{J}\left(\boldsymbol{\Omega}_k\right)\right]\right)\\ &\quad -\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_k-\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\mathbf{W}_{k-1}\right),\end{aligned} \tag{32}$$

for $k=1,\ldots,N-1$. Taking variations of the cost functional (18) and substituting from equation (32) one obtains after a tedious but straight forward computation an expression for $\boldsymbol{\delta}\mathcal{J}$ in terms of $\boldsymbol{\delta\tau}_k$:

$$\begin{aligned}\boldsymbol{\delta}\mathcal{J} &= \sum_{k=1}^{N-1}\left[\left\langle \mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\left(\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_{k+1}-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k\right)\right.\right.\right.\\ &\quad\left.\left.+\frac{1}{h}\left[\mathbf{W}_{k+1}-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k,\mathbf{J}\left(\boldsymbol{\Omega}_k\right)\right]\right)-\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_k-\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\mathbf{W}_{k-1}\right),\boldsymbol{\tau}_k^\sharp\right\rangle\right]\\ &\quad+\left\langle\boldsymbol{\delta\tau}_0,\boldsymbol{\tau}_0^\sharp\right\rangle+\left\langle\boldsymbol{\delta\tau}_N,\boldsymbol{\tau}_N^\sharp\right\rangle\end{aligned}$$

.

When $\boldsymbol{\delta}\mathcal{J}$ is equated to zero (and after some algebraic rearrangement), one can obtain the boundary conditions on $\boldsymbol{\tau}_0,\boldsymbol{\tau}_1,\boldsymbol{\tau}_{N-1},\boldsymbol{\tau}_N$ from the resulting equations below:

$$\begin{aligned}\boldsymbol{\tau}_0 &= 0\\ 0 &= -\frac{1}{h^2}\left(\mathbf{J}\left(\boldsymbol{\tau}_1^\sharp\right)+\mathrm{Ad}^*_{\mathbf{g}_1^{-1}}\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_1^{-1}}\boldsymbol{\tau}_1^\sharp\right)\right)\\ &\quad-\frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_1^{-1}}\left[\mathbf{J}\left(\boldsymbol{\Omega}_1\right),\mathrm{Ad}_{\mathbf{g}_1^{-1}}\left(\boldsymbol{\tau}_1^\sharp\right)\right]\\ 0 &= -\frac{1}{h^2}\left(\mathbf{J}\left(\boldsymbol{\tau}_{N-1}^\sharp\right)+\mathrm{Ad}^*_{\mathbf{g}_{N-1}^{-1}}\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_{N-1}^{-1}}\boldsymbol{\tau}_{N-1}^\sharp\right)\right)\\ &\quad-\frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_{N-1}^{-1}}\left[\mathbf{J}\left(\boldsymbol{\Omega}_{N-1}\right),\mathrm{Ad}_{\mathbf{g}_{N-1}^{-1}}\left(\boldsymbol{\tau}_{N-1}^\sharp\right)\right]\\ \boldsymbol{\tau}_N &= 0\end{aligned}$$

as well as discrete evolution equations that are written in algebraic nonlinear form as:

$$\begin{aligned}0 &= -\frac{1}{h^2}\left(\mathbf{J}\left(\boldsymbol{\tau}_k^\sharp\right)-\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{J}\left(\boldsymbol{\tau}_{k+1}^\sharp\right)\right.\\ &\quad\left.-\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\boldsymbol{\tau}_{k-1}^\sharp\right)+\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_k^{-1}}\boldsymbol{\tau}_k^\sharp\right)\right)\\ &\quad-\frac{1}{h}\left(\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\left[\mathbf{J}\left(\boldsymbol{\Omega}_k\right),\mathrm{Ad}_{\mathbf{g}_k^{-1}}\left(\boldsymbol{\tau}_k^\sharp\right)\right]\right.\\ &\quad\left.-\frac{1}{h}\left[\mathbf{J}\left(\boldsymbol{\Omega}_{k-1}\right),\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\left(\boldsymbol{\tau}_{k-1}^\sharp\right)\right]\right),\end{aligned} \tag{33}$$

for $k = 2, \ldots, N-2$.

The following section shows that while our discrete approximation (14) is formally first-order accurate, it is symplectically equivalent to the second-order accurate Störmer–Verlet method, and hence has effective order two.

4. Lie Symplectic Euler and Symplectic Equivalence

Notice that the discrete Lagrangian adopted in our paper is obtained by approximating the velocity as a constant over the timestep h, and by approximating the integral in time by $\int_{t_1}^{t_2} f(t)dt \approx (t_2 - t_1)f(t_1)$. In the Lie group setting, the constant angular velocity approximation corresponds to the condition,

$$\mathbf{R}_{k+1} = \mathbf{R}_k \exp(h\boldsymbol{\Omega}_k)$$

or equivalently,

$$\boldsymbol{\Omega}_k = \frac{1}{h}\exp^{-1}(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}).$$

When we let $G = \mathbb{R}^n$, and we adopt the notation $(\mathbf{q}, \mathbf{v}) \in T\mathbb{R}^n$, we obtain,

$$\mathbf{v}_k = \frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h},$$

which is a usual finite-difference approximation for the velocity. Consider then a Lagrangian of the form,

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

Approximating the action integral from 0 to h using a constant velocity approximation and a quadrature formula, yields,

$$\int_0^h L(\mathbf{q}(t), \mathbf{v}(t))dt \approx \int_0^h L\Big(\mathbf{q}(t), \frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big)dt \approx hL\Big(\mathbf{q}_k, \frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big).$$

We then choose as our discrete Lagrangian,

$$L_d(\mathbf{q}_k, \mathbf{q}_{k+1}) = hL\Big(\mathbf{q}_k, \frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big) = h\left[\frac{1}{2}\Big(\frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big)^T M\Big(\frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big) - V(\mathbf{q}_k)\right].$$

The discrete Euler–Lagrange equations,

$$D_2L_d(\mathbf{q}_{k-1}, \mathbf{q}_k) + D_1L_d(\mathbf{q}_k, \mathbf{q}_{k+1}) = 0,$$

yields,

$$M\Big(\frac{\mathbf{q}_k - \mathbf{q}_{k-1}}{h}\Big) - M\Big(\frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big) - h\frac{\partial V}{\partial \mathbf{q}}(\mathbf{q}_k) = 0,$$

which induces an implicit update map $(\mathbf{q}_{k-1}, \mathbf{q}_k) \mapsto (\mathbf{q}_k, \mathbf{q}_{k+1})$. To obtain the corresponding Hamiltonian update map, we push-forward this algorithm to T^*Q by using the discrete fiber derivative $\mathbb{F}L_d : Q \times Q \to T^*Q$, which takes $(\mathbf{q}_k, \mathbf{q}_{k+1}) \mapsto (\mathbf{q}_{k+1}, D_2L_d(\mathbf{q}_k, \mathbf{q}_{k+1}))$. In particular, we have that,

$$\mathbf{p}_{k+1} = D_2L_d(\mathbf{q}_k, \mathbf{q}_{k+1}) = M\Big(\frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big),$$

which implies

$$\mathbf{q}_{k+1} = \mathbf{q}_k + hM^{-1}\mathbf{p}_{k+1}. \tag{34}$$

This allows us to rewrite the discrete Euler–Lagrange equations as,

$$\mathbf{p}_k - \mathbf{p}_{k+1} - h\frac{\partial V}{\partial q}(\mathbf{q}_k) = 0,$$

or equivalently,

$$\mathbf{p}_{k+1} = \mathbf{p}_k - h\frac{\partial V}{\partial q}(\mathbf{q}_k). \tag{35}$$

Now, (34) and (35) are precisely the symplectic Euler method applied to the corresponding Hamiltonian vector field, as we shall see.

The corresponding Hamiltonian is given by,

$$H(\mathbf{q},\mathbf{p}) = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + V(\mathbf{q}).$$

Hamilton's equations yield,

$$\begin{pmatrix}\dot{\mathbf{q}}\\ \dot{\mathbf{p}}\end{pmatrix} = \begin{pmatrix}\frac{\partial H}{\partial \mathbf{p}}\\ -\frac{\partial H}{\partial \mathbf{q}}\end{pmatrix} = \begin{pmatrix}M^{-1}\mathbf{p}\\ -\frac{\partial V}{\partial \mathbf{q}}\end{pmatrix}.$$

The symplectic Euler method has the form,

$$\begin{aligned}\mathbf{q}_{k+1} &= \mathbf{q}_k + h\dot{\mathbf{q}}(\mathbf{q}_k,\mathbf{p}_{k+1}),\\ \mathbf{p}_{k+1} &= \mathbf{p}_k + h\dot{\mathbf{p}}(\mathbf{q}_k,\mathbf{p}_{k+1}),\end{aligned}$$

which yields,

$$\begin{aligned}\mathbf{q}_{k+1} &= \mathbf{q}_k + hM^{-1}\mathbf{p}_{k+1},\\ \mathbf{p}_{k+1} &= \mathbf{p}_k + h\left(-\frac{\partial V}{\partial q}(\mathbf{q}_k)\right),\end{aligned}$$

which is precisely what we obtained in (34) and (35). This demonstrates that our method is the generalization of the symplectic Euler method to Lie groups, which has important numerical consequences. While symplectic Euler is formally first-order accurate, it is symplectically equivalent [27, 18] to the second-order accurate Störmer–Verlet method [6]. This means that one can obtain the Störmer–Verlet method F_{SV} by conjugating the symplectic Euler method F_{E} with a symplectic transformation T,

$$F_{\mathrm{SV}} = TF_{\mathrm{E}}T^{-1}.$$

In particular, numerical trajectories of symplectic Euler will shadow numerical trajectories obtained using Störmer–Verlet. Consider the implications of this symplectic equivalence for our discrete optimal control problem. Let the boundary conditions be specified by $\mathbf{q}_0, \mathbf{q}_N$, and assume that we use Störmer–Verlet to propagate the solution, then the boundary condition is expressed as, $\mathbf{q}_N = F_{\mathrm{SV}}^N\mathbf{q}_0 = (TF_{\mathrm{E}}T^{-1})^N\mathbf{q}_0 = TF_{\mathrm{E}}^NT^{-1}\mathbf{q}_0$, which is equivalent to $\tilde{q}_N = T^{-1}\mathbf{q}_N = F_{\mathrm{E}}^NT^{-1}\mathbf{q}_0 = F_{\mathrm{E}}^N\tilde{q}_0$. This implies that if we preprocess the boundary conditions $\mathbf{q}_0, \mathbf{q}_N$, to obtain $\tilde{q}_0 = T^{-1}\mathbf{q}_0, \tilde{q}_N = T^{-1}\mathbf{q}_N$, we could use symplectic Euler at the internal stages to propagate the states and costates, and then postprocess them to obtain the trajectory one would have obtained by using Störmer–Verlet.

In practice, the shadowing result imparts the symplectic Euler method with the same desirable qualitative properties as Störmer–Verlet, and it is not necessary to postprocess the numerical solutions in order to achieve accurate results. Since on an appropriate choice of charts, our Lie symplectic Euler method reduces to symplectic Euler in coordinates, it follows that there is a corresponding second-order Lie Störmer–Verlet method that our method is symplectically equivalent to, and in particular, our method has effective order two.

5. Numerical Approach and Results

The multiplier-free version of the first-order optimality equations, equation (33), in combination with the boundary conditions,

$$\mathbf{R}_0 = \mathbf{R}_0^*, \quad \mathbf{R}_N = \mathbf{R}_N^*, \quad \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_0^*, \quad \text{and } \boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}_{N-1}^*,$$

leave the torques $\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_{N-1}$, and the angular velocities $\boldsymbol{\Omega}_1, \dots, \boldsymbol{\Omega}_{N-2}$ as unknowns. By substituting the relations $\mathbf{g}_k = \exp(h\boldsymbol{\Omega}_k)$, $\mathbf{M}_k = \mathbf{J}(\boldsymbol{\Omega}_k)$, we can rewrite the necessary conditions (33) as follows,

$$\begin{aligned}
0 = &-\frac{1}{h^2}\Big(\mathbf{J}(\boldsymbol{\tau}_k^\sharp) - \mathrm{Ad}^*_{\exp(-h\boldsymbol{\Omega}_k)}\mathbf{J}(\boldsymbol{\tau}_{k+1}^\sharp) \\
&\qquad - \mathbf{J}(\mathrm{Ad}_{\exp(-h\boldsymbol{\Omega}_{k-1})}\boldsymbol{\tau}_{k-1}^\sharp \\
&\qquad + \mathrm{Ad}^*_{\exp(-h\boldsymbol{\Omega}_k)}\mathbf{J}(\mathrm{Ad}_{\exp(-h\boldsymbol{\Omega}_k)}\boldsymbol{\tau}_k^\sharp)\Big) \\
&-\frac{1}{h}\Big(\mathrm{Ad}^*_{\exp(-h\boldsymbol{\Omega}_k)}\Big[\mathbf{J}(\boldsymbol{\Omega}_k), \mathrm{Ad}_{\exp(-h\boldsymbol{\Omega}_k)}(\boldsymbol{\tau}_k^\sharp)\Big] \\
&\qquad - \frac{1}{h}\Big[\mathbf{J}(\boldsymbol{\Omega}_{k-1}), \mathrm{Ad}_{\exp(-h\boldsymbol{\Omega}_{k-1})}(\boldsymbol{\tau}_{k-1}^\sharp)\Big]\Big),
\end{aligned}$$

where $k = 2, \dots, N-2$, and the discrete evolution equations, given by line 2 of (31), can be written as

$$0 = \mathbf{J}(\boldsymbol{\Omega}_k) - \mathrm{Ad}^*_{\exp(h\boldsymbol{\Omega}_k)}(h\boldsymbol{\tau}_k + \mathbf{J}(\boldsymbol{\Omega}_{k-1})),$$

where $k = 1, \dots, N-1$. In addition, we use the boundary conditions on $\mathbf{R}_0$ and $\mathbf{R}_N$, together with the update step given by line 1 of (31) to give the last constraint,

$$0 = \log\Big(\mathbf{R}_N^{-1}\mathbf{R}_0 \exp(h\boldsymbol{\Omega}_0)\dots\exp(h\boldsymbol{\Omega}_{N-1})\Big),$$

where log is the logarithm map on $\mathsf{SO}(3)$.

Note that while we use the direct variational approach to obtain the discrete extremal solutions, an alternate way to obtain the discrete extremal solutions would be to use Pontryagin's maximum principle. In particular, Bonnans and Laurent-Varin [4] show that these two approaches are equivalent in the context of symplectic partitioned Runge-Kutta schemes.

At this point it should be noted that one important advantage of the manner in which we have discretized the optimal control problem is that it is $\mathsf{SO}(3)$-equivariant. This is to say that if we rotated all the boundary conditions by a fixed rotation matrix, and solved the resulting discrete optimal control problem, the solution we would obtain would simply be the rotation of the solution of the original problem. This can be seen quite clearly from the fact that the discrete problem is expressed in terms of body coordinates, both in terms of body angular velocities and body forces. In addition, the initial and final attitudes $\mathbf{R}_0$ and $\mathbf{R}_N$ only enter in the last equation as a relative rotation.

The $\mathsf{SO}(3)$-equivariance of our numerical method is desirable, since it ensures that our results do not depend on the choice of coordinate frames. This is in contrast to methods based on coordinatizing the rotation group using quaternions and Euler angles.

Each of the equations above take values in $\mathfrak{so}(3)$. Consider the Lie algebra isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)$ given by the hat map,

$$\mathbf{v} = (v_1, v_2, v_3) \mapsto \hat{\mathbf{v}} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix},$$

which maps 3-vectors to 3×3 skew-symmetric matrices. In particular, we have the following identities,

$$[\hat{\mathbf{u}}, \hat{\mathbf{v}}] = (\mathbf{u} \times \mathbf{v})\hat{}, \quad \mathrm{Ad}_{\mathbf{A}} \hat{\mathbf{v}} = (\mathbf{A}\mathbf{v})\hat{}.$$

Furthermore, we identify $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ by the usual dot product, that is to say if $\mathbf{\Pi}$, $\mathbf{v} \in \mathbb{R}^3$, then $\langle \mathbf{\Pi}, \hat{\mathbf{v}} \rangle = \mathbf{\Pi} \cdot \mathbf{v}$. With this identification, we have that $\mathrm{Ad}^*_{\mathbf{A}^{-1}} \mathbf{\Pi} = \mathbf{A}\mathbf{\Pi}$. Using the identities above, we write the necessary conditions using matrix-vector products and cross products. Then, each of the equations can be interpreted as 3-vector valued functions, and the system of equations can be considered as a $3(2N-3)$-vector valued function, which is precisely the dimensionality of the unknowns. This reduces the discrete optimal problem to a nonlinear root finding problem.

The nonlinear system of equations was solved in MATLAB using the `fsolve` routine, where the Jacobian is constructed column by column, and the k-th column is computed using the following approximation [19],

$$\frac{\partial \mathbf{F}}{\partial x_k}(\mathbf{x}) = \frac{1}{\epsilon} \mathrm{Im}[\mathbf{F}(\mathbf{x} + i\epsilon \mathbf{e}_k)],$$

where $i = \sqrt{-1}$, $\mathbf{e}_k$ is a basis vector in the x_k direction, and ϵ is of the order of machine epsilon. This method is preferable to a finite-difference approximation, since it does not suffer from round-off errors, which would otherwise limit how small ϵ can be.

In our numerical simulation, we computed an optimal trajectory for a rest-to-rest maneuver, as illustrated in Figure (2). Here, the maneuver time is 12.8sec, $N = 128$, and the moment of inertia is given by

$$\mathbf{J} = \begin{bmatrix} 13.25 & -7.80 & -11.40 \\ -7.80 & 16.25 & 4.71 \\ -11.40 & 4.71 & 18.37 \end{bmatrix}.$$

The prescribed maneuver corresponds to a rotation by $\frac{\pi}{3}$ about the x-axis. Since the moment of inertia tensor is not a multiple of the identity, and the x-axis does not correspond to the axis of minimal inertia, the optimal trajectory does not just involve a pure rotation about the x-axis. It is worth noting that the results are not rotationally symmetric about the midpoint of the simulation interval, which is due to the fact that our choice of update, $\mathbf{R}_{k+1} = \mathbf{R}_k \exp(h\mathbf{\Omega}_k)$, does not exhibit time-reversal symmetry. In a forthcoming publication, we will introduce a reversible algorithm to address this issue. In particular, this will involve explicitly computing the stationarity conditions for the discrete optimal control problem constrained by the time-symmetric Lie Störmer–Verlet method.

We also present results for an optimal slew-up maneuver, illustrated in Figure (3). This uses the same moment of inertia tensor as in the previous simulation, and the desired maneuver involves a rotation of $\frac{\pi}{6}$ about the x-axis from rest to a final

angular velocity of $\Omega_{N-1} = \begin{bmatrix}0.3 & 0.2 & 0.3\end{bmatrix}^{\mathrm{T}}$, over a maneuver time of 12.8sec, and $N = 128$.

6. Conclusion

In this paper we studied the continuous- and discrete-time optimal control problem for the rigid body, where the cost to be minimized is the external torque applied to move the rigid body from an initial condition to some pre-specified terminal condition. In the discrete setting, we use the discrete Lagrange–d'Alembert principle to obtain the discrete equations of motion. The kinematics were discretized to guarantee that the flow in phase space remains on the Lie group $\mathsf{SO}(3)$ and its algebra $\mathfrak{so}(3)$. We described how the necessary conditions can be solved for the general three-dimensional case and gave a numerical example for a three-dimensional rigid body maneuver.

The synthesis of variational mechanics with discrete-time optimal control is particularly advantageous from the point of view of computational efficiency, since the symplectic Euler method is symplectically conjugate to the Störmer–Verlet method, and hence has effective order two. Consequently, for our discrete-time optimal control method, the cost functional converges at a rate which is characteristic of a second-order method, while being based on a first-order method that is computationally cheaper.

Currently, we are investigating the use of the Pontryagin's maximum principle with Lie group methods in continuous- and discrete-time to obtain the necessary conditions. Additionally, we wish to generalize the result to general Lie groups that have applications other than the rigid body motion on $\mathsf{SO}(3)$. In particular, we are interested in controlling the motion of a rigid body *in space*, which corresponds to motion on the non-compact Lie group $\mathsf{SE}(3)$.

Acknowledgments

The research of Islam Hussein was supported by a WPI Faculty Development Grant. The research of Melvin Leok was partially supported by NSF grants DMS-0504747 and DMS-0726263 and a University of Michigan Rackham faculty grant. The research of Anthony Bloch was supported by NSF grants DMS-030583, and CMS-0408542.

References

[1] New horizons: Pluto-kuiper belt mission. Website (2006). URL http://www.nasa.gov/mission_pages/newhorizons/main/index.html

[2] Agrachev, A.A., Sachkov, Y.: Control Theory from the Geometric Viewpoint. Springer-Verlag, New York, NY (2004)

[3] Bloch, A.: Nonholonomic Mechanics and Control. Springer-Verlag, New York, NY (2003)

[4] Bonnans, J., Laurent-Varin, J.: Computation of order conditions for symplectic partitioned runge-kutta schemes with application to optimal control. Num. Math. **103**, 1–10 (2006)

[5] Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration. Springer, Berlin (2002)

[6] Hairer, E., Lubich, C., Wanner, G.: Geometric numerical integration illustrated by the Störmer-Verlet method. Acta Numer. **12**, 399–450 (2003)

[7] Hussein, I.I., Bloch, A.M.: Optimal control on Riemannian manifolds with potential fields. 43rd IEEE Conference on Decision and Control (2004). 1982–1987

[8] Hussein, I.I., Bloch, A.M.: Optimal trajectory tracking on the group of rigid body motions. 2005 IEEE Conference on Decision and Control (2005)

[9] Hussein, I.I., Scheeres, D.J., Hyland, D.C.: Interferometric observatories in Earth orbit. Journal of Guidance, Control and Dynamics **27**(2), 297–301 (2004)

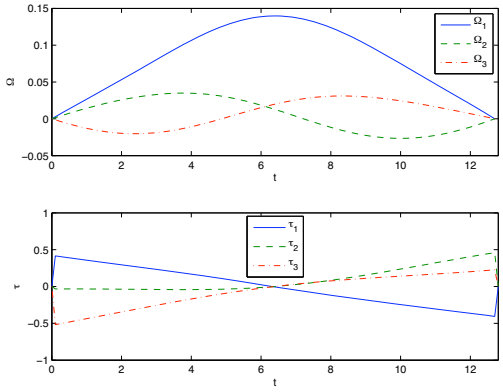
[10] Iserles, A., Munthe-Kaas, H., Nørsett, S.P., Zanna, A.: Lie group methods. Acta Numerica **9**, 215–265 (2000)
[11] Junge, O., Marsden, J.E., Ober-Blöbaum, S.: Discrete mechanics and optimal control. IFAC Congress, Praha (2005)
[12] Kane, C., Marsden, J.E., Ortiz, M., West, M.: Variational integrators and the newmark algorithm for conservative and dissipative mechanical systems. International Journal of Numerical Methods in Engineering **49**(10), 1295–1325 (2000)
[13] Khaneja, N., Glaser, S.J., Brockett, R.W.: Sub-Riemannian geometry and optimal control of three spin systems. Physical Review A **65**, 032,301 (2002)
[14] Lee, T., Leok, M., McClamroch, N.H.: A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum. Proceedings of the IEEE Conference on Control Applications pp. 962–967 (2005)
[15] Lee, T., Leok, M., McClamroch, N.H.: Lie group variational integrators for the full body problem. Computer Methods in Applied Mechanics and Engineering (2005). Submitted, arXiv:math.NA/0508365
[16] Leimkuhler, B., Reich, S.: Simulating Hamiltonian Dynamics, *Cambridge Monographs on Applied and Computational Mathematics*, vol. 14. Cambridge University Press, Cambridge (2004)
[17] Leok, M.: Generalized galerkin variational integrators (2004). Preprint, arXiv:math.NA/0508360
[18] Littell, T., Skeel, R., Zhang, M.: Error analysis of symplectic multiple time stepping. SIAM J. Numer. Anal. **34**(5), 1792–1807 (1997)
[19] Lyness, J., Moler, C.: Numerical differentiation of analytic functions. SIAM J. Numer. Anal. **4**, 202–210 (1967)
[20] Marsden, J., West, M.: Discrete mechanics and variational integrators. Acta Numerica **10**, 357–514 (2001)
[21] Marsden, J.E., Ratiu, T.S.: Introduction to Mechanics and Symmetry. Springer-Verlag, New York, NY (1999)
[22] Milnor, J.: Morse Theory. Princeton University Press, Princeton, NJ (1963)
[23] Palao, J.P., Kosloff, R.: Quantum computing by an optimal control algorithm for unitary transformations. Physical Review Letters **89**, 188,301 (2002)
[24] Sanz-Serna, J.M., Calvo, M.P.: Numerical Hamiltonian Problems, *Applied Mathematics and Mathematical Computation*, vol. 7. Chapman and Hall, London (1994)
[25] Schaub, H., Junkins, J.L., Robinett, R.D.: New attitude penalty functions for spacecraft optimal control problems. AIAA Guidance, Navigation, and Control Conference (1996)
[26] Scrivener, S.L., Thompson, R.C.: Survey of time-optimal attitude maneuvers. Journal of Guidance, Control, and Dynamics **17**(2), 225–233 (1994)
[27] Suzuki, M.: Improved Trotter-like formula. Phys. Lett. A **180**(3), 232–234 (1993)
[28] Warner, F.: Foundations of Differentiable Manifolds and Lie Groups. Scott, Foresman, Glenview, IL (1973)

Alexander Ziwet Collegiate Professor of Mathematics and Department Chair, Mathematics, University of Michigan
E-mail address: abloch@umich.edu

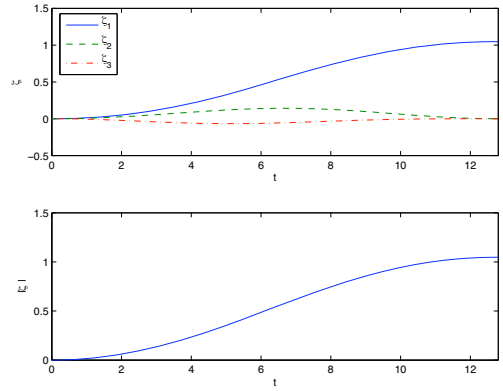
Assistant Professor, Mechanical Engineering, Worcester Polytechnic Institute
E-mail address: ihussein@wpi.edu

Assistant Professor, Mathematics, Purdue University
E-mail address: mleok@math.purdue.edu

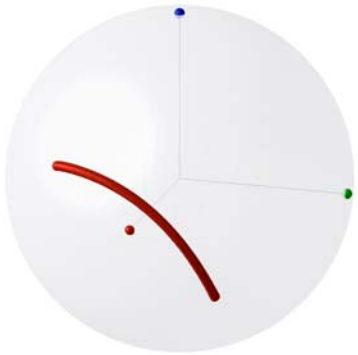
Assistant Professor, Mechanical Engineering, University of Hawaii
E-mail address: aksanyal@hawaii.edu



(a) Angular velocity and control torques

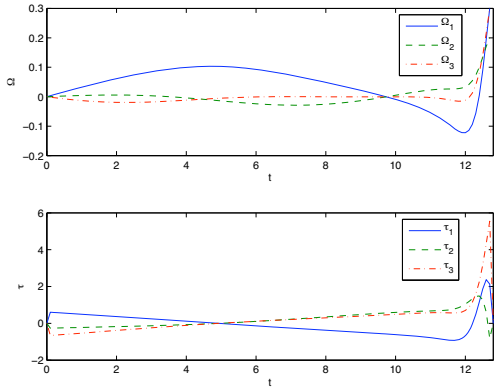


(b) Principal axis and angle

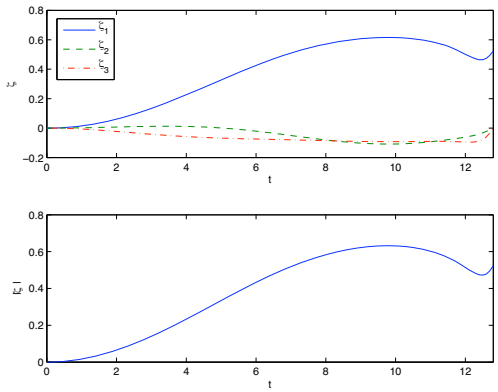


(c) Instantaneous rotation axis

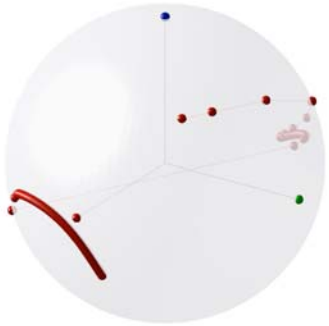
FIGURE 2. Discrete optimal rest-to-rest maneuver in $\mathsf{SO}(3)$.



(a) Angular velocity and control torques



(b) Principal axis and angle



(c) Instantaneous rotation axis

FIGURE 3. Discrete optimal slew-up maneuver in $\mathsf{SO}(3)$.

COMPUTATIONAL GEOMETRIC OPTIMAL CONTROL OF RIGID BODIES

TAEYOUNG LEE, MELVIN LEOK, AND N. HARRIS MCCLAMROCH

Dedicated to Roger W. Brockett on his 70th birthday

Abstract. This paper formulates optimal control problems for rigid bodies in a geometric manner and it presents computational procedures based on this geometric formulation for numerically solving these optimal control problems. The dynamics of each rigid body is viewed as evolving on a configuration manifold that is a Lie group. Discrete-time dynamics of each rigid body are developed that evolve on the configuration manifold according to a discrete version of Hamilton's principle so that the computations preserve geometric features of the dynamics and guarantee evolution on the configuration manifold; these discrete-time dynamics are referred to as Lie group variational integrators. Rigid body optimal control problems are formulated as discrete-time optimization problems for discrete Lagrangian/Hamiltonian dynamics, to which standard numerical optimization algorithms can be applied. This general approach is illustrated by presenting results for several different optimal control problems for a single rigid body and for multiple interacting rigid bodies. The computational advantages of the approach, that arise from correctly modeling the geometry, are discussed.

1. Introduction

This paper utilizes methods from geometric mechanics and optimal control to develop new computational procedures for geometric optimal control of rigid bodies. The emphasis is on formulating a discrete-time optimal control problem that inherits important conservation properties of rigid body dynamics; this is achieved by combining variational integrators [36] and Lie group methods [16] to evolve the mechanical configuration. This approach leads to Lie group variational integrators that define the discrete-time rigid body dynamics which the optimal control computations are based upon [27, 28].

Most of the prior work related to optimal control of a rigid body is based on local coordinates on SO(3) or quaternions [2, 11, 41, 42]. Minimal representations of the attitude of a rigid body, such as Euler angles, exhibit coordinate singularities, and require manipulating complicated trigonometric

Taeyoung Lee, Department of Aerospace Engineering, University of Michigan. tylee@umich.edu
Melvin Leok, Assistant Professor, Department of Mathematics, Purdue University. mleok@math.purdue.edu
N. Harris McClamroch, Professor, Department of Aerospace Engineering, University of Michigan. nhm@umich.edu .

expressions. Nonminimal representations such as quaternions have no coordinate singularities, but they also introduce certain complications. In particular, the group of unit quaternions $\mathsf{SU}(2) \simeq \mathsf{S}^3$ double covers SO(3), so there is an ambiguity in representing an attitude of a rigid body. Furthermore, the Hamiltonian structure of rigid body attitude dynamics is unnecessarily complicated when it is expressed in terms of quaternions [32].

By considering rigid body translation and rotation as evolution on a Lie group, optimal control problems defined on Lie groups were introduced by Roger Brockett [8, 9] and by John Baillieul [1]. They emphasized the use of Lie group structures to characterize controllability and existence of optimal controls; they also obtained analytical results for the solution of certain types of optimal control problems. An optimal control problem for a generalized rigid body on $\mathrm{SO}(n)$ was considered in [3], and a general theory of optimal control problems on a Lie group was developed in [18, 19, 20] together with reachability and controllability conditions. Although these papers viewed rigid body translation and rotation as motion on a Lie group, their results are limited to optimal control problems that can be formulated solely in terms of kinematics. In particular, they do not include dynamics in their analysis, and assume that the controls enter directly at the level of the Lie algebra.

The approach of computational geometric optimal control is focused on developing numerical algorithms, for optimal control problems, that preserve the geometric properties of the dynamics and the optimal control problem [22]. The essential idea is to apply geometric optimal control theory to discrete-time mechanical systems obtained using geometric numerical integrators. A discrete-time version of the generalized rigid body equations and their formulation as an optimal control problem are presented in [4, 5], and discrete-time optimal control problems for the dynamics of a rigid body are considered in [6, 30, 25]. A direct optimal control approach is applied to discrete-time mechanical systems in [17], and it is referred to as *Discrete Mechanics and Optimal Control*.

This paper presents the approach of computational geometric optimal control for the dynamics of rigid bodies on a Lie group. We take the same geometric perspective as in the work of Roger Brockett [8, 9], viewing evolution on a Lie group as fundamental. However, the emphasis in the present paper is on geometric formulations of both the kinematics and dynamics in the optimal control formulation and the role of geometric methods in optimal control computations.

The development in the paper makes clear that there are important advantages in formulating the optimal control problem as a discrete-time optimal control problem using Lie group variational integrators and then applying standard computational methods to solve the resulting discrete-time optimization problem. This is in contrast with approaches that construct continuous-time necessary conditions and then make use of computational methods to solve these necessary conditions. The

paper demonstrates that for for the optimal control of rigid bodies, the proposed approach exhibits important advantages.

The main contributions of this paper can be summarized as follows: (i) the analytical and computational results presented in this paper are coordinate free; they avoid the singularities, ambiguity, and complications associated with local coordinates, and they provide a global insight into rigid body dynamics, (ii) a geometric optimal control problem is formulated for nontrivial rigid body dynamics that evolve on a Lie group, and (iii) a computational geometric optimal control approach is developed based on a geometric numerical integrator.

Section 2 provides a summary of Lie group variational integrators for rigid bodies that evolve on a Lie group. The resulting discrete-time rigid body dynamics are used as a basis for formulating a discrete-time optimal control problem. In Section 3 and 4, four different examples of rigid body optimal control problems are studied in some detail. First, optimal orbit and attitude maneuvers for a rigid dumbbell spacecraft in orbit about a large central body are studied. Then, optimal attitude maneuvers for a 3D pendulum acting under uniform gravity are studied; the control input conserves the component of the vertical component of the angular momentum thereby requiring a careful computational treatment that avoids numerical ill-conditioning. The third example is a 3D pendulum attached to a cart that can move in a horizontal plane; optimal reconfiguration maneuvers are studied for this cart and pendulum system. The fourth example involves optimal attitude maneuvers of two rigid bodies connected by a universal joint; the control input conserves angular momentum and the resulting controlled system exhibits a symmetry that has to be taken into account in the numerical approach in order to avoid numerical ill-conditioning.

2. Mathematical formulation for optimal control of rigid bodies

The dynamics of rigid bodies exhibit important geometric features. The configuration of a rigid body can be described by the position vector of its center of mass in the Euclidean space $\mathbb{R}^3$ and by the attitude of the rigid body represented by a rotation matrix in the special orthogonal group $\mathrm{SO}(3) = \{R \in \mathbb{R}^{3\times 3} \,|\, R^T R = I, \det R = 1\}$. Thus, the general motion of a rigid body is described by the special Euclidean group $\mathrm{SE}(3) = \mathrm{SO}(3) \circledS \mathbb{R}^3$. The configuration manifold for the class of multiple rigid bodies can be represented as a product involving $\mathbb{R}^3, \mathrm{SO}(3)$, and $\mathrm{SE}(3)$. Therefore, the configuration manifold of rigid bodies is a Lie group. Furthermore, the dynamics of rigid bodies, viewed as Lagrangian or Hamiltonian systems, are characterized by symplectic, momentum and energy preserving properties. These geometric features determine the qualitative behavior of the rigid body dynamics.

In this paper, we study optimal control problems for rigid bodies while carefully considering the geometric features of the dynamics in both the analysis and numerical computations. In particular, discrete-time dynamics of rigid bodies are developed that evolve on the configuration manifold according to a discrete version of Hamilton's principle. The resulting geometric numerical integrator, referred to as a Lie group variational integrator, preserves geometric features of the dynamics and guarantees evolution on the configuration manifold. Based on the discrete-time rigid bodies dynamics, a discrete-time optimal control problem for rigid bodies is formulated. Standard numerical optimization algorithms can then be applied to solve this discrete-time optimal control problem.

Thus, our approach to discrete-time optimal control is characterized by discretizing the continuous-time optimal control problem at the problem formulation stage using Lie group variational integrators. This is in contrast to traditional techniques wherein discretization only arises at the last stage when numerically solving the continuous-time optimality conditions. Since the geometric properties of the dynamics of rigid bodies are preserved by using a Lie group variational integrator, this optimal control approach yields geometrically-exact optimal control inputs and accurate trajectories that are efficiently computed [4, 17, 30, 25].

In this section, we first describe the fundamental procedure to develop a Lie group variational integrator and its computational properties. Then, a discrete-time optimal control problem is formulated using the Lie group variational integrator, and computational approaches are presented to solve it numerically.

2.1. **Lie group variational integrator.** Geometric numerical integrators are numerical integration algorithms that preserve features of the continuous-time dynamics such as invariants, symplecticity, and the configuration manifold [14]. The geometrically exact properties of the discrete-time flow generate improved qualitative behavior. In this paper, we view a Lie group variational integrator as an intrinsically discrete-time dynamical system.

Numerical integration methods that preserve the simplecticity of a Hamiltonian system have been studied extensively [39, 32]. One traditional approach is to carefully choose the coefficients of a Runge-Kutta method to satisfy a simplecticity criterion and order conditions in order to obtain a symplectic Runge-Kutta method. However, it can be difficult to construct such integrators, and it is not guaranteed that other invariants of the system, such as momentum maps, are preserved. Alternatively, variational integrators are constructed by discretizing Hamilton's principle, rather than discretizing the continuous Euler-Lagrange equation [38, 36]. The resulting integrators have the desirable property that they are symplectic and momentum preserving, and they exhibit good energy behavior for exponentially long times. Lie group methods are numerical integrators that

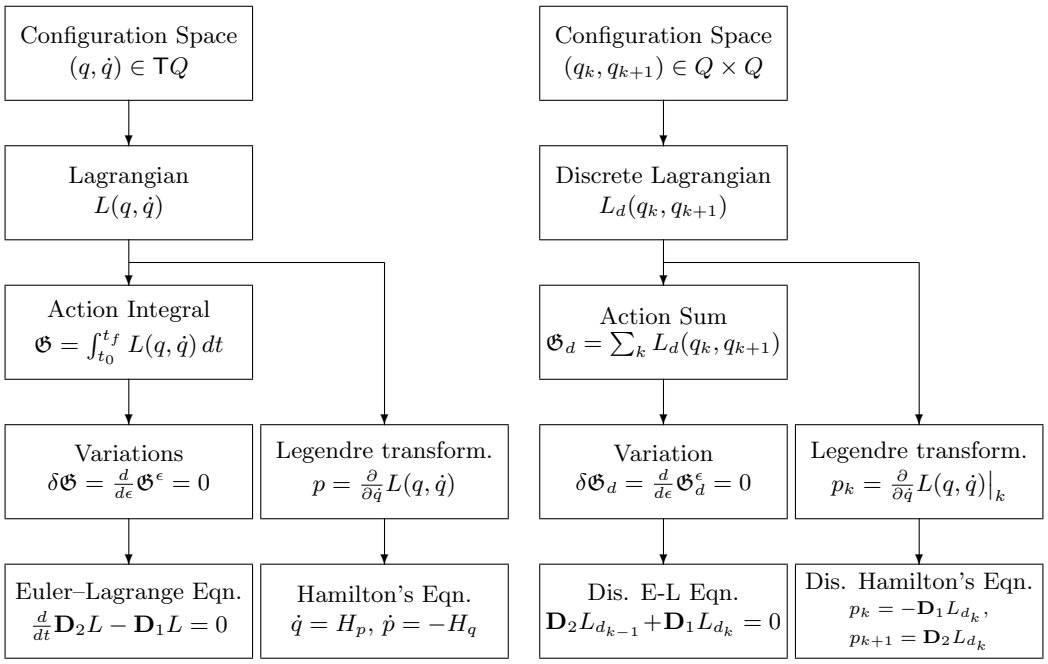


FIGURE 1. Procedures to derive continuous-time and discrete-time equations of motion

preserve the Lie group structure of the configuration manifold [16]. Recently, these two approaches have been unified to obtain Lie group variational integrators that preserve the geometric properties of the dynamics as well as the Lie group structure of the configuration manifold without the use of local charts, reprojections, or constraints [34, 33, 27, 28].

We now summarize the derivation of a Lie group variational integrator. In Lagrangian mechanics, the equations of motion are derived by finding the path that extremizes the action integral, which is the integral of the Lagrangian over time. The Legendre transformation provides an alternative description that leads to Hamilton's equations. Discrete-time Lagrangian and Hamiltonian mechanics, referred to as variational integrators, have been developed by reformulating Lagrangian and Hamiltonian mechanics in a discrete-time setting [36].

Discrete-time mechanics has a parallel structure with the mechanics described in continuous-time, as summarized in Figure 1. The phase variables of the continuous-time Lagrangian are replaced by two copies of the discrete-time configuration variables and a discrete-time Lagrangian that approximates a segment of the action integral is chosen. An action sum is defined using the discrete-time Lagrangian such that it approximates the action integral. This is the only approximation made in the development of discrete-time mechanics. Discrete-time Euler-Lagrange equations are obtained

by setting the variation of the action sum to zero. The discrete-time Legendre transformation yields the equivalent of Hamilton's equations. Lie group variational integrators are developed to preserve the structure of the Lie group configurations as well as the geometric properties of the continuous-time dynamics. The basic idea for all Lie group methods is to express the update map for group elements in the configuration manifold in terms of the group operation, so that the group structure is preserved automatically without need of parameterizations, constraints, or reprojections.

More explicitly, consider a mechanical system whose configuration manifold is a Lie group G and is described by a Lagrangian $L : \mathsf{T}G \to \mathbb{R}$. The discrete update for the configuration is chosen as

$$g_{k+1} = g_k f_k, \tag{1}$$

where $g_k, g_{k+1} \in G$ are configuration variables, and the subscript k denotes the value of a variable at the time $t = kh$ for a fixed timestep $h \in \mathbb{R}$. The discrete-time update map is represented by a right group action of $f_k \in G$ on g_k. Since the group element is updated by a group action, the group structure is preserved.

The expression for the flow map in discrete-time is obtained from the discrete variational principle on a Lie group, as presented in Figure 1. A discrete Lagrangian $L_d : G \times G \to \mathbb{R}$ approximates the integral of the Lagrangian over a time step along the solution of the Euler-Lagrange equation

$$L_d(g_k, f_k) \approx \int_{kh}^{(k+1)h} L(g(t), \dot{g}(t))\, dt, \tag{2}$$

where a curve $g(t) : [kh, (k+1)h] \to G$ satisfies the Euler–Lagrange equation in the time interval $[k, (k+1)h]$ with boundary conditions $g(kh) = g_k$ and $g((k+1)h) = g_k f_k = g_{k+1}$. Analogous to the action integral, the action sum is defined as

$$\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(g_k, f_k). \tag{3}$$

The discrete Lagrange-d'Alembert principle, which is a modification of Hamilton's principle to include the effect of control inputs, states that the sum of the variation of the action sum and the virtual work done by the control inputs is zero. But, the infinitesimal variation of a Lie group element must be carefully expressed to respect the structure of the Lie group. For example, it can be expressed in terms of the exponential map $\exp : \mathfrak{g} \to G$ as

$$\delta g = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} g \exp \epsilon\eta = g\eta, \tag{4}$$

for a Lie algebra element $\eta \in \mathfrak{g}$. From the discrete Lagrange-d'Alembert principle, we obtain

$$\delta \mathfrak{G}_d + \sum_{k=0}^{N-1} \left[u_k^+ \cdot \eta_{k+1} + u_k^- \cdot \eta_k \right] = 0 \tag{5}$$

for any δg_k, and for given discrete Lagrangian forces $u_{d_k}^+, u_{d_k}^- \in \mathfrak{g}^*$. This yields the generalized discrete Euler–Poincaré equation

$$\begin{aligned} \mathsf{T}_e^* \mathsf{L}_{f_k} \cdot \mathbf{D}_2 L_d(g_k, f_k) - \mathrm{Ad}_{f_k}^* \cdot (\mathsf{T}_e^* \mathsf{L}_{f_{k+1}} \cdot \mathbf{D}_2 L_d(g_{k+1}, f_{k+1})) \\ + \mathsf{T}_e^* \mathsf{L}_{g_{k+1}} \cdot \mathbf{D}_1 L_d(g_{k+1}, f_{k+1}) + u_{d_{k-1}}^+ + u_{d_k}^- = 0. \end{aligned} \tag{6}$$

Here $\mathsf{L}_f : G \to G$ denotes the left translation map given by $\mathsf{L}_f g = fg$ for $f, g \in G$, $\mathsf{T}_g \mathsf{L}_f : \mathsf{T}_g G \to \mathsf{T}_{fg} G$ is the tangent map for the left translation and $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is the adjoint map. A dual map is denoted by a superscript * (see [22] for detailed definitions and developments).

This approach has been applied to the rotation group SO(3) and to the special Euclidean group SE(3) for dynamics of rigid bodies in [27, 24, 28] and the generalization to abstract Lie groups are summarized here, thereby generating a unified geometric integrator for the class of multiple generalized rigid bodies whose configuration manifold can be expressed as a Lie group, which includes products involving $\mathbb{R}^3$, SO(3), and SE(3) as special cases.

2.2. **Discrete-time optimal control.** Optimal control problems involve finding a control input such that a certain optimality objective is achieved under prescribed constraints. Here, the control inputs are parameterized by their values at each discrete time step, and the discrete-time equations of motion, including the control inputs, are obtained from (6). Any standard numerical algorithm for constrained optimization can be applied to this discrete-time system.

An indirect approach to solving a discrete-time optimal control problem is based on solving discrete-time necessary conditions for optimality. The resulting two-point boundary value problem can be solved by using standard numerical root finding techniques; one such approach is the shooting method that iterates on initial values of the multipliers. Alternatively, a direct approach formulates the discrete-time optimal control problem as a nonlinear programming problem, which is solved using standard numerical optimization algorithms such as a sequential quadratic programming algorithm; one such approach is the DMOC (Discrete Mechanics and Optimal Control) approach [17].

Explicit time-discretization prior to numerical optimization has significant computational advantages. As discussed in the previous section, the discrete-time dynamics are faithful representations

of the continuous-time dynamics, and consequently more accurate solutions to the optimal control problems are typically obtained. The external control inputs may break the Lagrangian and Hamiltonian system structure; for example, the total energy may not be conserved for a controlled mechanical system. But, the computational superiority of the discrete mechanics formulation still holds for controlled systems. In particular, it has been demonstrated in [36] that the discrete-time dynamics derived from the discrete Lagrange-d'Alembert principle accurately computes the energy dissipation rate of controlled systems. For example, this feature is extremely important in accurately computing optimal trajectories for spacecraft orbit and attitude maneuvers for which the control authority is low and the maneuver time is large.

The proposed discrete-time optimal control formulation provides a framework for accurate computations. In most indirect optimal control approaches, the optimal solutions are sensitive to small variations in the initial values of the multipliers. This may cause difficulties, such as numerical ill-conditioning, in solving the necessary conditions for optimality expressed as a two-point boundary value problem. Numerically computed sensitivity derivatives, using Lie group variational integrators, do not exhibit numerical dissipation, which typically arises in conventional numerical integration schemes. Thus, the proposed approach leads to numerical robustness and efficient numerical computations. This indirect computational approach exhibits the quadratic convergence rate that is typical of Newton methods when it is applied to an optimal attitude control problem [29]; the error in satisfaction of the optimality condition converges to machine precision superlinearly. For the direct optimal control approach, the optimal control inputs can be parametrized using fewer degrees of freedom, thereby reducing the computational overhead.

Several optimal control problems involving rigid bodies have been previously studied by the authors. Minimum-fuel and time-optimal control of spacecraft large-angle attitude maneuvers are studied in [23, 15, 30, 31]. The optimal orbit transfer of a dumbbell spacecraft, wherein the rotational attitude dynamics are non-trivially coupled to the translational dynamics, is studied in [25]. An underactuated optimal control problem for the attitude maneuver of a 3D pendulum is studied in [29]. An optimal formation reconfiguration of multiple rigid body spacecraft is studied in [26]. An optimal control problem for a dynamic system evolving on an abstract Lie group is developed in [22], thereby generating a unified approach for optimal control problems of multiple rigid bodies.

In this paper, we summarize results for two optimal control problems for a single rigid body in Section 3 and results for two optimal control problems for multiple rigid bodies in Section 4. Each of these optimal control problems treats complex dynamics of a single or multiple rigid bodies, demonstrating the value of the proposed geometric optimal control approach.

3. Optimal control problems for a single rigid body

3.1. **Optimal maneuver of a dumbbell spacecraft on** SE(3)**.** We develop an optimal 3D translational and rotational maneuver of a rigid dumbbell spacecraft in orbit about a large central body. The dumbbell spacecraft is composed of two spheres connected by a massless rod. An interesting feature of the dumbbell spacecraft is that there is coupling between its translational dynamics and its rotational dynamics due to the presence of both gravity forces and gravity moments that act on the dumbbell spacecraft.

The configuration manifold is the special Euclidean group $\mathrm{SE}(3) = \mathrm{SO}(3) \circledS \mathbb{R}^3$. For $(R, x) \in \mathrm{SE}(3)$, the linear transformation from the body-fixed frame to the inertial frame is denoted by the rotation matrix $R \in \mathrm{SO}(3)$, and the position of the mass center in the inertial frame is denoted by a vector $x \in \mathbb{R}^3$. The vectors $\Omega, v \in \mathbb{R}^3$ are the angular velocity in the body-fixed frame, and the translational velocity in the inertial frame, respectively. Let $m \in \mathbb{R}$ and $J \in \mathbb{R}^{3\times 3}$ be the mass and the moment of inertia matrix of a rigid body. We assume that external control force $u^f \in \mathbb{R}^3$ and control moment $u^m \in \mathbb{R}^3$ act on the dumbbell spacecraft. Control inputs are parameterized by their values at each time step.

Define a $f_k = (F_k, Y_k) \in \mathrm{SE}(3)$ such that $g_{k+1} = (R_{k+1}, x_{k+1})$ is equal to $g_k f_k$, i.e. $(R_{k+1}, x_{k+1}) = (R_k, x_k) \circ (F_k, Y_k) = (R_k F_k, x_k + R_k Y_k)$. The rotation matrix F_k represent the relative update of the attitude between integration steps. The gravitational potential is denoted by $U : \mathrm{SE}(3) \to \mathbb{R}$. We choose the following discrete Lagrangian

$$L_d(R_k, x_k, F_k, Y_k) = \frac{1}{2h} m Y_k^T Y_k + \frac{1}{h}\mathrm{tr}[(I - F_k)J_d] - hU(R_k F_k, x_k + R_k Y_k), \tag{7}$$

where $J_d \in \mathbb{R}^{3\times 3}$ is a non-standard moment of inertia matrix defined as $J_d = \frac{1}{2}\mathrm{tr}[J] I_{3\times 3} - J$. Substituting this discrete Lagrangian into (6), we obtain the following discrete equations of motion (see [22] for detailed development).

$$h\widehat{J\Omega_k} = F_k J_d - J_d F_k^T, \tag{8}$$

$$R_{k+1} = R_k F_k, \tag{9}$$

$$x_{k+1} = x_k + h v_k, \tag{10}$$

$$J\Omega_{k+1} = F_k^T J \Omega_k + h(M_{k+1} + u^m_{k+1}), \tag{11}$$

$$m v_{k+1} = m v_k - h\frac{\partial U_{k+1}}{\partial x_{k+1}} + h u^f_{k+1}, \tag{12}$$

where the hat map $\hat{\cdot}$ is an isomorphism from $\mathbb{R}^3$ to 3×3 skew-symmetric matrices $\mathfrak{so}(3)$, defined such that $\hat{x}y = x \times y$ for any $x, y \in \mathbb{R}^3$. The moment $M \in \mathbb{R}^3$ due to the potential is given by,

$$\hat{M} = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}, \tag{13}$$

where the matrix $\frac{\partial U}{\partial R} \in \mathbb{R}^{3\times 3}$ is defined by $[\frac{\partial U}{\partial R}]_{ij} = \frac{\partial U}{\partial [R]_{ij}}$ for $i, j \in \{1, 2, 3\}$, and the i, j-th element of a matrix is denoted by $[\cdot]_{ij}$.

For a given $(R_k, x_k, \Omega_k, v_k)$, we solve the implicit equation (8) to find $F_k \in \mathrm{SO}(3)$. Then, the configuration at the next step (R_{k+1}, x_{k+1}) is obtained from (9) and (10). Using the computed moment M_{k+1} and force $-\frac{\partial U_{k+1}}{\partial x_{k+1}}$, velocities Ω_{k+1}, v_{k+1} are obtained from (11) and (12). This defines a discrete flow map, $(R_k, x_k, \Omega_k, v_k) \mapsto (R_{k+1}, x_{k+1}, \Omega_{k+1}, v_{k+1})$, and this process can be repeated.

Since this Lie group variational integrator is obtained by discretizing Hamilton's principle, it is symplectic and preserves the momentum map associated with the symmetry of the Lagrangian. In the absence of external forces and moments, the total energy oscillates around its initial value with small bounds on a comparatively short timescale, but there is no tendency for the mean of the oscillation in the total energy to drift (increase or decrease) from the initial value for exponentially long times.

The discrete flow map also preserves the group structure. By using the given computational approach, the matrix F_k, representing the change in relative attitude change over a time step, is guaranteed to be a rotation matrix. The rotation matrix R_{k+1} is obtained by the group operation in (9), so that it evolves on SO(3). Therefore, the orthogonal structure of the rotation matrices is preserved, and the attitude of each rigid body is determined accurately and globally.

This geometrically exact numerical integration method yields a highly efficient computational algorithm. The self-adjoint discrete Lagrangian used to derive this Lie group variational integrator guarantees that this integrator has second-order accuracy, while requiring only one function evaluation per integration step. Higher-order methods can be easily constructed using a composition method [14].

An implicit equation (8) must be solved at each time step to determine the attitude update. However the computational effort to solve each implicit equation is negligible; the relative attitude update is expressed at the Lie algebra level isomorphic to $\mathbb{R}^3$, and the corresponding Newton iteration converges to machine precision within two or three iterations. This method could be considered *almost explicit* when the computational cost is compared with explicit integrators with the same order of accuracy [27].

Optimal control problem. The objective is to transfer the spacecraft from a given initial condition $(R_0, x_0, \Omega_0, v_0)$ to a desired terminal condition $(R^f, x^f, \Omega^f, v^f)$ during a fixed maneuver time Nh, while minimizing the square of the l_2 norm of the control inputs.

$$\min_{u_{k+1}} \left\{ \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2} (u_{k+1}^f)^T W_f u_{k+1}^f + \frac{h}{2} (u_{k+1}^m)^T W_m u_{k+1}^m \right\}, \tag{14}$$

where $W_f, W_m \in \mathbb{R}^{3\times 3}$ are symmetric positive-definite matrices.

Necessary conditions for optimality. An indirect optimization method is used to determine the optimal solution, based on necessary conditions for optimality derived using variational arguments; the optimal control is characterized as a solution of a two-point boundary value problem. The augmented cost function to be minimized is

$$\begin{aligned} \mathcal{J}_a = \sum_{k=0}^{N-1} & \frac{h}{2} (u_{k+1}^f)^T W^f u_{k+1}^f + \frac{h}{2} (u_{k+1}^m)^T W^m u_{k+1}^m \\ & + \lambda_k^{1,T} \left\{ -x_{k+1} + x_k + h v_k \right\} + \lambda_k^{2,T} \left\{ -m v_{k+1} + m v_k - h \frac{\partial U_{k+1}}{\partial x_{k+1}} + h u_{k+1}^f \right\} \\ & + \lambda_k^{3,T} \left(\operatorname{logm}(F_k - R_k^T R_{k+1}) \right)^\vee + \lambda_k^{4,T} \left\{ -J\Omega_{k+1} + F_k^T J \Omega_k + h \left(M_{k+1} + u_{k+1}^m \right) \right\}, \end{aligned} \tag{15}$$

where $\lambda_k^1, \lambda_k^2, \lambda_k^3, \lambda_k^4 \in \mathbb{R}^3$ are Lagrange multipliers. The matrix logarithm is denoted by $\operatorname{logm} : \mathrm{SO}(3) \to \mathfrak{so}(3)$ and the vee map $\vee : \mathfrak{so}(3) \to \mathbb{R}^3$ is the inverse of the hat map. The logarithmic form of (9) is used, and the constraint (8) is implicitly imposed using constrained variations. Using similar expressions for the variations given in (4), the infinitesimal variation of the cost can be written as

$$\delta \mathcal{J}_a = \sum_{k=1}^{N-1} h \delta u_k^{f,T} \left\{ W_f u_k^f + \lambda_{k-1}^2 \right\} + h \delta u_k^{m,T} \left\{ W_m u_k^m + \lambda_{k-1}^4 \right\} + z_k^T \left\{ -\lambda_{k-1} + A_k^T \lambda_k \right\}, \tag{16}$$

where $\lambda_k = [\lambda_k^1; \lambda_k^2; \lambda_k^3; \lambda_k^4] \in \mathbb{R}^{12}$ is the vector of Lagrange multipliers, and $z_k \in \mathbb{R}^{12}$ represents the infinitesimal variation of $(R_k, x_k, \Omega_k, v_k)$, given by $z_k = [\operatorname{logm}(R_k^T \delta R_k)^\vee; \delta x_k, \delta \Omega_k, \delta v_k]$. The matrix $A_k \in \mathbb{R}^{12\times 12}$ is expressed in terms of $(R_k, x_k, \Omega_k, v_k)$ [25]. Thus, necessary conditions for optimality are given by

$$u_{k+1}^f = -W_f^{-1} \lambda_k^2, \tag{17}$$

$$u_{k+1}^m = -W_m^{-1} \lambda_k^4, \tag{18}$$

$$\lambda_k = A_{k+1}^T \lambda_{k+1} \tag{19}$$

together with the discrete equations of motion and the boundary conditions.

Computational approach. Necessary conditions for optimality are expressed in terms of a two-point boundary problem. This problem is to find the optimal discrete flow, multipliers, and control inputs that simultaneously satisfies the equations of motion, optimality conditions, multiplier equations, and boundary conditions. We use a neighboring extremal method [10], and choose a nominal solution satisfying all of the necessary conditions except the boundary conditions. The unspecified initial multiplier is updated by successive linearization so as to satisfy the specified terminal boundary conditions in the limit. This is also referred to as a shooting method. The main advantage of the neighboring extremal method is that the number of iteration variables is small.

The difficulty is that the extremal solutions are sensitive to small changes in the unspecified initial multiplier values. The nonlinearities also make it hard to construct an accurate estimate of sensitivity, thereby resulting in numerical ill-conditioning. Therefore, it is important to compute the sensitivities accurately in the neighboring extremal method. Here, the optimality conditions (17) and (18) are substituted into the equations of motion and the multiplier equations, which are linearized to obtain

$$\begin{bmatrix} z_N \\ \delta\lambda_N \end{bmatrix} = \begin{bmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{bmatrix} \begin{bmatrix} z_0 \\ \delta\lambda_0 \end{bmatrix},$$

where $\Psi^{ij} \in \mathbb{R}^{6\times 6}$ for $i, j \in \{1, 2\}$ represents a computable linear operator. For the given two-point boundary value problem, $z_0 = 0$ since the initial condition is fixed. The terminal multipliers are free. Thus, we obtain

$$z_N = \Psi^{12}\delta\lambda_0.$$

The linear operator Ψ^{12} represents the sensitivity of the specified terminal boundary conditions with respect to the unspecified initial multiplier. Using this sensitivity, a guess of the unspecified initial multipliers is iterated to satisfy the specified terminal conditions in the limit. Any type of Newton iteration can be applied. We use a line search with backtracking algorithm, referred to as the Newton-Armijo iteration [21].

Numerical example. We study a maneuver of a rigid spacecraft under a central gravity field. We assume that the mass of the spacecraft is negligible compared to the mass of a central body, and we consider a fixed frame attached to the central body as an inertial frame. The resulting model is a Restricted Full Two Body Problem (RF2BP) [40].

The spacecraft is modeled as a dumbbell, which consists of two equally massive spheres and a massless rod. The gravitational potential is given by

$$U(R,x) = -\frac{GMm}{2}\sum_{q=1}^{2}\frac{1}{\|x+R\rho^q\|}, \tag{20}$$

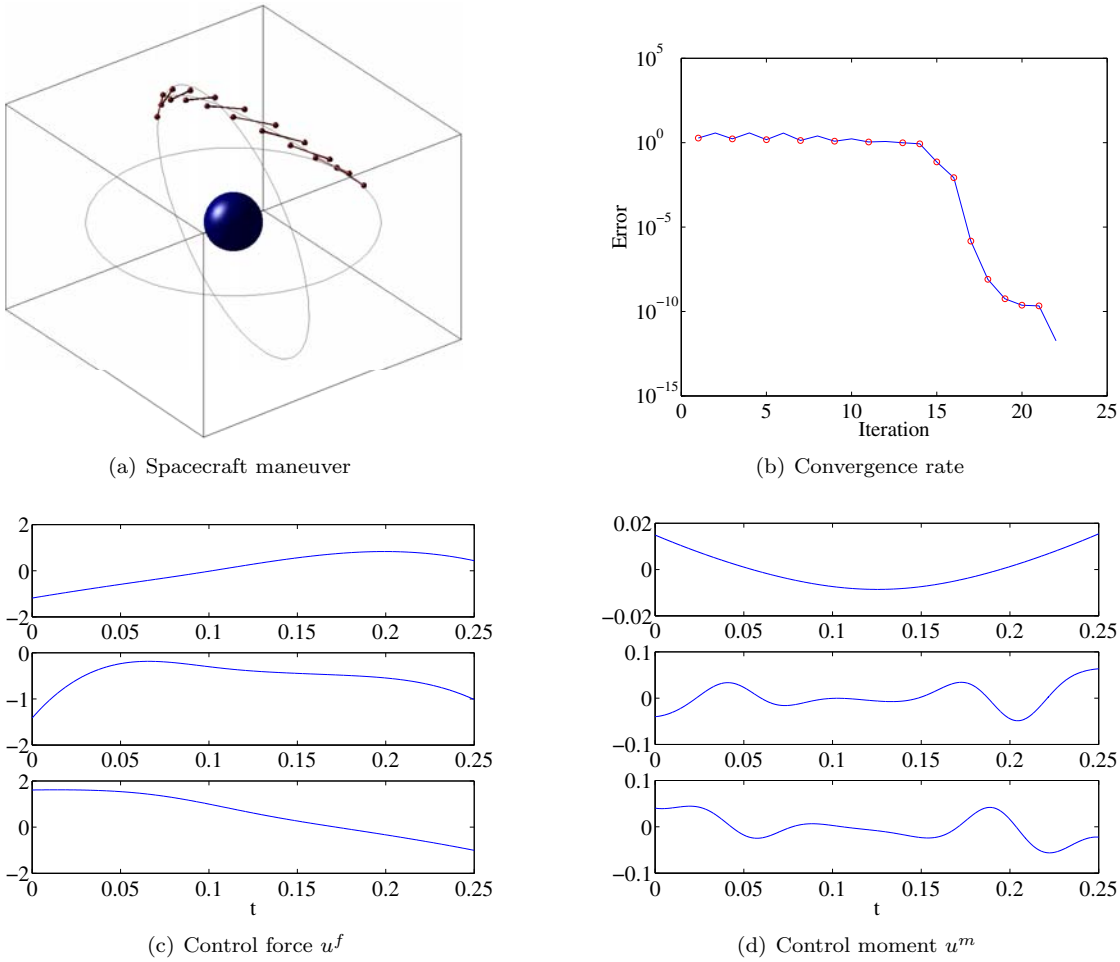
where $G \in \mathbb{R}$ is the gravitational constant, $M, m \in \mathbb{R}$ are the mass of the central body, and the mass of the dumbbell, respectively. The vector $\rho^q \in \mathbb{R}^3$ is the position of the qth sphere from the mass center of the dumbbell expressed in the body fixed frame ($q \in \{1,2\}$). The mass, length, and time dimensions are normalized by the mass of the dumbbell, the radius of a reference circular orbit, and its orbital period.

Initially, the spacecraft is on a circular orbit. The desired maneuver is to increase the orbital inclination by 60°. We explicitly consider the coupling effect between the orbital motion and the rotational attitude maneuver of the spacecraft. The maneuver time is chosen to be a quarter of the orbital period of the initial circular orbit. The boundary conditions are as follows,

$$\begin{aligned}
&x_0 = [1,0,0], && x^f = [-0.3536, 0.3536, 0.8660],\\
&R_0 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, && R^f = \begin{bmatrix} -0.7071 & 0.3535 & 0.6123 \\ -0.7071 & -0.3535 & -0.6123 \\ 0 & -0.8660 & 0.5 \end{bmatrix},\\
&\dot{x}_0 = [0, 0.9835, 0], && \dot{x}^f = [-0.6954, -0.6954, 0],\\
&\Omega_0 = [0,0,0.9835], && \Omega^f = [0,0,0.9835].
\end{aligned}$$

Figure 2 illustrates the optimal spacecraft maneuver, convergence rate, and optimal control inputs. The optimal cost and the violation of the terminal boundary conditions are 13.03, and 9.32×10^{-15} respectively. Figure 2(b) shows the violation of the terminal boundary conditions versus the number of iterations on a semi-logarithmic scale. Red circles denote outer iterations of the Newton-Armijo iteration where the sensitivity derivatives are computed, and inner iterations correspond to backtracking in the line search routine. The initial guess of the unspecified initial multipliers is arbitrarily chosen. The error in satisfaction of the terminal boundary condition converges quickly to machine precision after the 20th iteration. These convergence results are consistent with the quadratic convergence rates expected of Newton methods with accurately computed gradients.

The shooting method may be prone to numerical ill-conditioning, as a small change in the initial multiplier can cause highly nonlinear behavior of the terminal conditions. However, as shown in Figure 2(b), the computational geometric optimal control approach exhibits excellent numerical



(a) Spacecraft maneuver

(b) Convergence rate

(c) Control force u^f

(d) Control moment u^m

FIGURE 2. Optimal orbit transfer of a dumbbell spacecraft

convergence properties. This is because the proposed computational algorithms are geometrically exact and numerically accurate. There is no numerical dissipation introduced by the numerical algorithm, and therefore, the sensitivity derivatives are more accurately computed.

3.2. **Optimal attitude reorientation of an underactuated 3D pendulum on** SO(3) [29]**.** A 3D pendulum is a rigid body supported by a fixed frictionless pivot acting under the influence of a uniform gravitational field [43]. The rigid body has three rotational degrees of freedom, and the configuration manifold is SO(3). The linear transformation from the body fixed frame and the inertial frame is denoted by $R \in \mathrm{SO}(3)$, and the angular velocity represented in the body fixed frame is denoted by $\Omega \in \mathbb{R}^3$. Let $e_3 \in \mathbb{R}^3$ be the gravity direction in the inertial frame, and $J \in \mathbb{R}^{3\times 3}$ be

the moment of inertia matrix of the rigid body with respect to the pivot point. The vector from the pivot point to the mass center, represented in the body fixed frame is given by $\rho \in \mathbb{R}^3$.

The Lagrangian of the 3D pendulum is invariant under a rotation about the gravity direction, and therefore the 3D pendulum has a S^1 symmetry action. Consequently, the angular momentum about the gravity direction, represented by $e_3^T R J \Omega$, is preserved.

We study an optimal attitude control of the 3D pendulum with symmetry. An external control moment is chosen such that it does not have any component about the gravity direction. The structure of the control moment is chosen as $R^T e_3 \times u$ for a control parameter $u \in \mathbb{R}^3$. Thus, the angular momentum about the gravity direction is conserved along the controlled dynamics of the 3D pendulum. Such control inputs are physically realized by actuation mechanisms, such as point mass actuators, that change the center of mass of the 3D pendulum.

The discrete Lagrangian of the 3D pendulum is chosen to be

$$L_d(R_k, F_k) = \frac{1}{h}\mathrm{tr}[(I - F_k)J_d] + hmge_3^T R\rho.$$

The resulting Lie group variational integrator, including an external control input, is given by

$$h\widehat{J\Omega_k} = F_k J_d - J_d F_k^T, \tag{21}$$

$$R_{k+1} = R_k F_k, \tag{22}$$

$$J\Omega_{k+1} = F_k^T J\Omega_k + hM_{k+1} + hR^T e_3 \times u_{k+1}. \tag{23}$$

Optimal control problem. The objective of the optimal control problem is to transfer the 3D pendulum from a given initial condition (R_0, Ω_0) to a desired terminal condition (R^f, Ω^f) during a fixed maneuver time Nh, while minimizing the square of the l_2 norm of the control inputs.

$$\min_{u_{k+1}} \left\{ \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2}(u_{k+1})^T W u_{k+1} \right\}, \tag{24}$$

where $W \in \mathbb{R}^{3\times 3}$ is a symmetric positive-definite matrix. In particular, we choose attitude maneuvers that can be described by rest-to-rest rotations about the unactuated gravity direction. The resulting optimal attitude maneuver exhibits the geometric phase effect [35], which in the zero group momentum case directly relates the group motion to the curvature enclosed by the trajectory in shape space.

Necessary conditions for optimality. We solve this optimal control problem by using an indirect optimization method, where necessary conditions for optimality are derived using variational arguments, and a solution of the corresponding two-point boundary value problem provides the optimal

control. The augmented cost function to be minimized is

$$\mathcal{J}_a = \sum_{k=0}^{N-1} \frac{h}{2} u_{k+1}^T W u_{k+1} + \lambda_k^{1,T} \left(\operatorname{logm}(F_k - R_k^T R_{k+1})\right)^\vee + \lambda_k^{2,T} \left\{-J\Omega_{k+1} + F_k^T J\Omega_k + hM_{k+1} + hR_{k+1}^T e_3 \times u_{k+1}\right\}, \tag{25}$$

where $\lambda_k^1, \lambda_k^2 \in \mathbb{R}^3$ are Lagrange multipliers. The infinitesimal variation can be written as

$$\delta\mathcal{J}_a = \sum_{k=1}^{N-1} h\delta u_k^T \left\{W u_k - R_k^T e_3 \times \lambda_{k-1}^2\right\} + z_k^T \left\{-\lambda_{k-1} + A_k^T \lambda_k\right\}, \tag{26}$$

where $\lambda_k = [\lambda_k^1; \lambda_k^2] \in \mathbb{R}^6$, and $z_k \in \mathbb{R}^6$ represents the infinitesimal variation of (R_k, Ω_k), given by $z_k = [\operatorname{logm}(R_k^T \delta R_k)^\vee; \delta\Omega_k]$. The matrix $A_k \in \mathbb{R}^{6\times 6}$ can be expressed in terms of $(R_k, \Omega_k), \lambda_k$. Thus, necessary conditions for optimality are given by

$$u_{k+1} = W^{-1}(R_{k+1}^T e_3 \times \lambda_k^2), \tag{27}$$

$$\lambda_k = A_{k+1}^T \lambda_{k+1} \tag{28}$$

together with the discrete equations of motion and the boundary conditions.

Computational approach. We apply the neighboring extremal method described in Section 3.1; the optimality condition is substituted into the equations of motion and the multiplier equation, and sensitivity derivatives of the optimal solution with respect to the initial multiplier are obtained, and the initial multiplier is iterated to satisfy the terminal boundary condition.

However, the underactuated control input, that respects the symmetry of the 3D pendulum, causes a fundamental singularity in the sensitivity derivatives, since the controlled system inherits the S^1 symmetry, and the cost functional is invariant under the lifted action of S^1. Consequently, the sensitivity derivatives vanish in the group direction. At each iteration, we need to compute inverse of a matrix of sensitivity derivatives to update the initial multiplier. However, the sensitivity matrix has a theoretical rank deficiency of one since the vertical component of the inertial angular momentum is conserved regardless of the initial multiplier variation. Therefore, this matrix inversion is numerically ill-conditioned.

We present a simple numerical scheme to avoid the numerical ill-conditioning caused by this symmetry. At each step, we decompose the matrix of sensitivity derivatives into a symmetric part and an anti-symmetric part. The symmetric part describes the sensitivity of the conserved angular momentum component due to the symmetry, and therefore it is zero and does not depend on the initial multiplier values. An update for the initial multipliers is determined using the matrix inverse

of the anti-symmetric part; this matrix inverse is not ill-conditioned. This approach removes the singularity in the sensitivity derivatives completely, and the resulting optimal control problem is no longer ill-conditioned.

Numerical example. Properties of the 3D pendulum are chosen as,

$$m = 1\,\text{kg}, \quad J = \text{diag}[0.13, 0.28, 0.17]\,\text{kgm}^2, \quad \text{and} \quad \rho = [0, 0, 0.3]\,\text{m}.$$

The desired maneuver is a 180° rotation about the vertical axis from a hanging equilibrium to another hanging equilibrium. The corresponding boundary conditions are given by

$$\begin{aligned} R_0 &= I, & R^f &= \text{diag}[-1, -1, 1], \\ \Omega_0 &= [0, 0, 0], & \Omega^f &= [0, 0, 0]. \end{aligned}$$

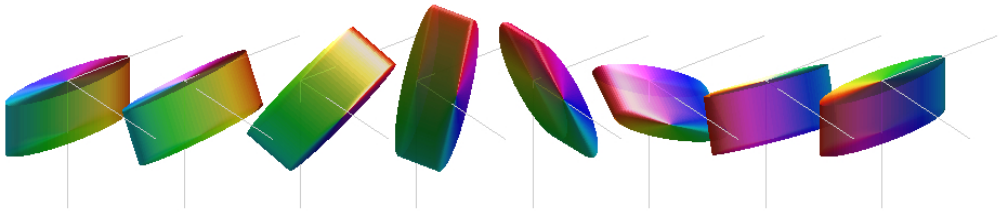
The maneuver time is 1 second, and the time step is $h = 0.001$. Since the vertical component of the angular momentum is zero, the rotation is a consequence of the geometric phase effect [35]. This problem is challenging in the sense that the desired maneuvers are rotations about the gravity direction, but the control input does not directly generate any moment about the gravity direction.

Figure 3 illustrates the optimal pendulum maneuver, convergence rate, and optimal control inputs. The optimal cost and the violation of the terminal boundary conditions are 7.32, and 4.80×10^{-15} respectively. As shown in Figure 3(c), the error in satisfaction of the terminal boundary condition converges to machine precision after the 50th iteration. The condition number of the decomposed sensitivity derivative varies from 10^0 to 10^5. If the sensitivity derivative is not decomposed, then the condition numbers are at the level of 10^{19}, and the numerical iterations fail to converge. This numerical example demonstrates the excellent numerical convergence properties of the computational geometric optimal control approach that is achieved by incorporating a modification that eliminates the numerical ill-conditioning introduced by the symmetry.

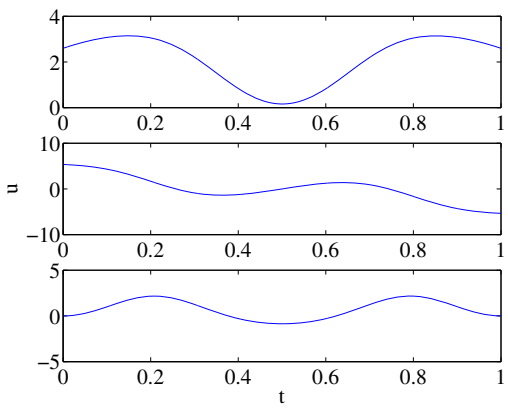
4. Optimal control problems for multiple rigid bodies

4.1. **Optimal maneuver of a 3D pendulum on a 2D cart on** $\text{SO}(3) \times \mathbb{R}^2$**.** Consider a 3D pendulum whose pivot is attached to a cart that can translate on a horizontal plane. This is a generalization of the popular planar pendulum on a cart model (see, for example, [7]), where the pendulum has three rotational degrees of freedom, and the cart moves on a two dimensional horizontal plane.

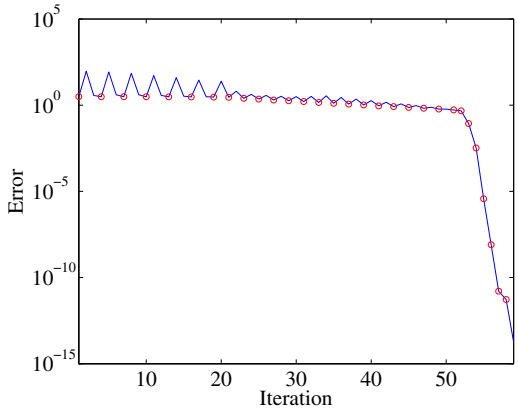
We define two frames; an inertial frame and a body fixed frame for the 3D pendulum whose origin is located at the moving pivot point. Define



(a) 3D pendulum maneuver



(b) Control moment u



(c) Convergence rate

FIGURE 3. Optimal control of a 3D pendulum with symmetry

$x \in \mathbb{R}$	Displacement of the cart along the e_1 direction in the reference frame
$y \in \mathbb{R}$	Displacement of the cart along the e_2 direction in the reference frame
$R \in \mathrm{SO}(3)$	Rotation matrix from the body fixed frame to the reference frame
$\Omega \in \mathbb{R}^3$	Angular velocity of the pendulum represented in the body fixed frame
$d \in \mathbb{R}^3$	Vector from the pivot to the mass center of the pendulum represented in the body fixed frame
$m \in \mathbb{R}$	Mass of the pendulum
$M \in \mathbb{R}$	Mass of the cart

The configuration manifold is $\mathrm{SO}(3) \times \mathbb{R}^2$. We assume that external control forces $u_x, u_y \in \mathbb{R}$ are applied to the cart.

The Lagrangian of the 3D pendulum on a cart is invariant under a rotation about the gravity direction. Therefore, it has a symmetry of S^1 action, and the total angular momentum about the gravity direction is preserved. The external control forces acting on the cart break this symmetry,

and the controlled system is not symmetric. In particular, the total angular momentum is not preserved in the controlled dynamics. Therefore, this optimal control problem should be distinguished from the optimal control of a 3D pendulum with symmetry, discussed in Section 3.2, where a symmetry-preserving control input is chosen.

The discrete Lagrangian for the 3D pendulum on a 2D cart is

$$\begin{aligned} L_d(R_k, x_k, y_k, R_{k+1}, x_{k+1}, y_{k+1}) &= \frac{1}{2h}(M+m)((x_{k+1}-x_k)^2 + (y_{k+1}-y_k)^2) \\ &+ \frac{1}{h}\mathrm{tr}[(I-F_k)J_d] + \frac{m}{h}(x_{k+1}-x_k)e_1^T(R_{k+1}-R_k)d + \frac{m}{h}(y_{k+1}-y_k)e_2^T(R_{k+1}-R_k)d \\ &+ \frac{h}{2}mge_3^TR_kd + \frac{h}{2}mge_3^TR_{k+1}d. \end{aligned} \tag{29}$$

From (6), the Lie group variational integrator for the 3D pendulum on a cart is given by the discrete-time equations

$$p_{x_k} = \frac{1}{h}(M+m)(x_{k+1}-x_k) + \frac{m}{h}e_1(R_{k+1}-R_k)d, \tag{30}$$

$$p_{y_k} = \frac{1}{h}(M+m)(y_{k+1}-y_k) + \frac{m}{h}e_2(R_{k+1}-R_k)d, \tag{31}$$

$$\hat{p}_{\Omega_k} = \frac{1}{h}(F_kJ_d - J_dF_k^T) + \left\{\frac{m}{h}(x_{k+1}-x_k)\hat{d}R_k^Te_1 + \frac{m}{h}(y_{k+1}-y_k)\hat{d}R_k^Te_2 - \frac{h}{2}mg\hat{d}R_k^Te_3\right\}^\wedge, \tag{32}$$

$$R_{k+1} = R_kF_k, \tag{33}$$

$$p_{x_{k+1}} = p_{x_k} + hu_{x_{k+1}}, \tag{34}$$

$$p_{y_{k+1}} = p_{y_k} + hu_{y_{k+1}}, \tag{35}$$

$$\hat{p}_{\Omega_{k+1}} = \frac{1}{h}(J_dF_k - F_k^TJ_d) + \left\{\frac{m}{h}(x_{k+1}-x_k)\hat{d}R_{k+1}^Te_1 + \frac{m}{h}(y_{k+1}-y_k)\hat{d}R_{k+1}^Te_2 + \frac{h}{2}mg\hat{d}R_{k+1}^Te_3\right\}^\wedge. \tag{36}$$

The momenta variables $p_\Omega \in \mathbb{R}^3$, $p_x, p_y \in \mathbb{R}$ are given by

$$\begin{bmatrix} p_\Omega \\ p_x \\ p_y \end{bmatrix} = \begin{bmatrix} J & m\hat{d}R^Te_1 & m\hat{d}R^Te_2 \\ -me_1^TR\hat{d} & M+m & 0 \\ -me_2^TR\hat{d} & 0 & M+m \end{bmatrix} \begin{bmatrix} \Omega \\ \dot{x} \\ \dot{y} \end{bmatrix}. \tag{37}$$

The detailed derivation of this Lie group variational integrator is available in [22]. For given $(R_k, x_k, y_k, \Omega_k, \dot{x}_k, \dot{y}_k)$, we compute $(p_{\Omega_k}, p_{x_k}, p_{y_k})$ by (37). We use a fixed-point iteration to

compute R_{k+1}. For an initial guess for R_{k+1}, the corresponding x_{k+1}, y_{k+1} are obtained by using (30),(31). Then, we can find F_k by solving (32). The updated value for R_{k+1} is given by (33). This is repeated until R_{k+1} converges. Then, x_{k+1}, y_{k+1} are obtained from (30),(31), and $(p_{\Omega_{k+1}}, p_{x_{k+1}}, p_{y_{k+1}})$ are obtained by (34),(35), and (36). The velocities $(\Omega_{k+1}, \dot{x}_{k+1}, \dot{y}_{k+1})$ are obtained from (37). This yields a flow map,

$$(R_k, x_k, y_k, \Omega_k, \dot{x}_k, \dot{y}_k) \mapsto (R_{k+1}, x_{k+1}, y_{k+1}, \Omega_{k+1}, \dot{x}_{k+1}, \dot{y}_{k+1}).$$

Optimal control problem. The objective of the optimal control problem is to transfer the 3D pendulum on a cart from a given initial condition $(R_0, x_0, y_0, \Omega_0, \dot{x}_0, \dot{y}_0)$ to a desired terminal condition $(R^f, x^f, y^f, \Omega^f, \dot{x}^f, \dot{y}^f)$ during a fixed maneuver time Nh, while minimizing the square of the l_2 norm of the control inputs.

$$\min_{u_{k+1}} \left\{ \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2} u_{k+1}^T W u_{k+1} \right\}, \tag{38}$$

where $u_k = [u_{x_k}; u_{y_k}] \in \mathbb{R}^2$, and $W \in \mathbb{R}^{2\times 2}$ is a symmetric positive-definite matrix. The 3D pendulum on a cart is underactuated, since only the planar motion of the cart in its horizontal plane is actuated.

Computational approach. We apply a direct optimal control approach. The control inputs are parameterized by several points that are uniformly distributed over the maneuver time, and control inputs between these points are approximated using a cubic spline interpolation. For given control input parameters, the value of the cost is given by (38), and the terminal conditions are obtained by the discrete-time equations of motion given by (30)-(36). The control input parameters are optimized using constrained nonlinear parameter optimization to satisfy the terminal boundary conditions while minimizing the cost.

This approach is computationally efficient when compared to the usual collocation methods, where the continuous-time equations of motion are imposed as constraints at a set of collocation points. Using the proposed discrete-time optimal control approach, optimal control inputs can be obtained by using a large step size, thereby resulting in efficient total computations. Since the computed optimal trajectories do not have numerical dissipation caused by conventional numerical integration schemes, they are numerically more robust. Furthermore, the corresponding gradient information is accurately computed, which improves the convergence properties of the numerical optimization procedure.

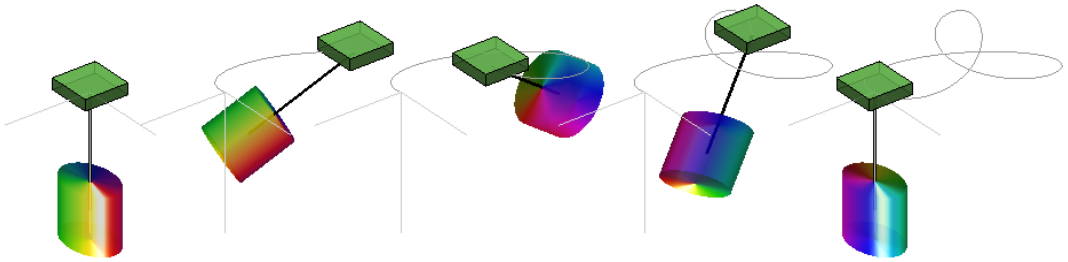
Numerical example. Properties of the 3D pendulum and the cart are chosen as,

$$M = m = 1\,\text{kg}, \quad J = \text{diag}[1.03, 1.04, 0.03]\,\text{kgm}^2, \quad \text{and} \quad d = [0, 0, 1]\,\text{m}.$$

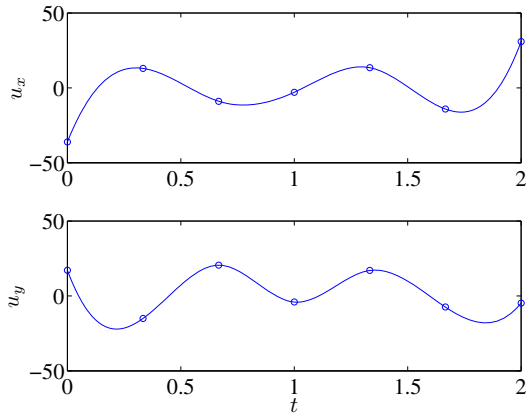
The desired maneuver is a rest-to-rest 180° rotation of the pendulum about the vertical axis, while the cart returns to the initial location at the terminal time. The corresponding boundary conditions are given by

$$R_0 = I, \quad \Omega_0 = [0, 0, 0], \quad x_0 = y_0 = 0, \quad \dot{x}_0 = \dot{y}_0 = 0,$$
$$R^f = \text{diag}[-1, -1, 1], \quad \Omega^f = [0, 0, 0], \quad x^f = y^f = 0, \quad \dot{x}^f = \dot{y}^f = 0.$$

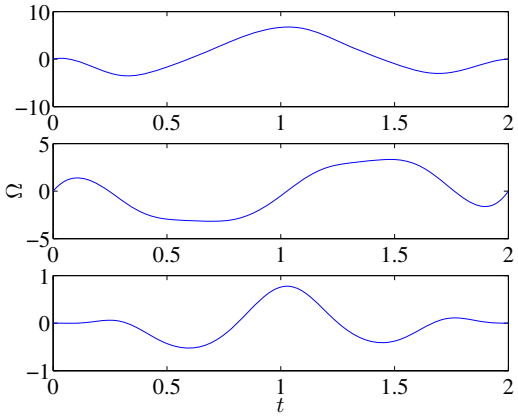
The maneuver time is 2 seconds, and the time step is $h = 0.01$. Since only the planar motion of the cart is actuated, the rotation of the 3D pendulum is caused by the nonlinear coupling between the cart and the pendulum.



(a) Optimal maneuver of a 3D pendulum on a cart



(b) Control force $u = (u_x, u_y)$



(c) Angular velocity Ω

FIGURE 4. Optimal control of a 3D pendulum on a cart

Each component of the control inputs is parameterized by 7 points. The resulting 14 control input parameters are optimized using sequential quadratic programming. Figure 4 illustrates the optimal maneuver of the pendulum and the cart, angular velocity, and optimal control inputs. The blue circles denote the optimized control input parameters. The optimal cost and the violation of the terminal boundary conditions are 297.43, and 1.83×10^{-8}, respectively. The optimal motion of the cart on the horizontal plane consists of a triangular-shaped loop, and the optimal maneuver of the 3D pendulum consists of large angle rotations. This also demonstrates the advantages of the computational geometric optimal control approach: it is difficult to study this kind of aggressive maneuvers of a multibody system using local coordinates, due to the coordinate singularities and the complexity of the equations in local coordinates. The presented computational geometric optimal control approach accurately characterizes the nonlinear coupling between the cart and the pendulum dynamics to obtain a nontrivial optimal maneuver of the 3D pendulum on a cart.

4.2. **Optimal attitude reorientation of two connected rigid bodies on** $\mathrm{SO}(3) \times \mathrm{SO}(3)$**.** Consider two rigid bodies connected with a ball joint that has three rotational degrees of freedom. This represents a freely rotating system of coupled rigid bodies. The relative equilibria structure of this rigid body dynamics has been studied in [44]. We introduce three frames; an inertial frame and two body-fixed frames. Define

$x \in \mathbb{R}^3$	Position of the ball joint in a reference frame
$R_i \in \mathrm{SO}(3)$	Rotation matrix from the i-th body-fixed frame to a reference frame
$d_i \in \mathbb{R}^3$	Vector from the joint to the mass center of the i-th body in the i-th body-fixed frame
$m_i \in \mathbb{R}$	Mass of the i-th body

for $i \in \{1, 2\}$. The configuration manifold is $\mathrm{SO}(3) \times \mathrm{SO}(3) \times \mathbb{R}^3$. In the absence of the potential field, the connected rigid body model has two symmetries; a symmetry of the translational action of $\mathbb{R}^3$, and a symmetry of the rotational action of $\mathrm{SO}(3)$.[1] Due to these symmetries, the total linear momentum and the total angular momentum are preserved, and the configuration manifold can be reduced to a quotient space.

In this optimal control problem, we reduce the configuration manifold to $\mathrm{SO}(3) \times \mathrm{SO}(3)$ using the symmetry of the translational action of $\mathbb{R}^3$. The corresponding value of the total linear momentum is set to zero. The resulting connected rigid bodies model with a fixed mass center is closely related to the falling cat problem [13]. An appropriate cyclic change in the shape of the body yields a

[1]These can be considered as a single symmetry of the translational and rotational action of SE(3), but they are considered separately in this optimal control problem. By the general theory of reduction by stages [12], the two approaches are equivalent.

rotation in the orientation of the cat in accordance with the geometric phase effect [37]. In contrast to other models of the falling cat, which typically introduce two one-dimensional rotational joints, with a shape space given by $S^1 \times S^1$, we consider instead a single ball joint with a shape space given by $SO(3)$.

Similar to the falling cat problem, we assume that an internal control moment $u \in \mathbb{R}^3$ is applied at the joint, so that it controls the relative attitude between two rigid bodies. More precisely, the control input u represents the control moment applied to the first rigid body, represented in the reference frame. The equal and opposite control moment is applied to the second rigid body. Therefore the control moment changes the shape of the system. The total angular momentum is conserved for the controlled dynamics as the control input is an internal moment of the connected rigid bodies system. This optimal control problem is similar to the optimal control problem of the 3D pendulum discussed in Section 3.2, as the control input respects the symmetry, and the corresponding momentum is preserved in the controlled dynamics.

The discrete Lagrangian for the two connected rigid bodies is

$$\begin{aligned} L_d(R_{1_k}, F_{1_k}, R_{2_k}, F_{2_k}, x_k, x_{k+1}) =& \frac{m_1+m_2}{2h}(x_{k+1}-x_k)\cdot(x_{k+1}-x_k) + \frac{1}{h}\mathrm{tr}[(I_{3\times 3}-F_{1_k})J_{d_1}] \\ &+ \frac{1}{h}\mathrm{tr}[(I_{3\times 3}-F_{2_k})J_{d_2}] + \frac{1}{h}\mathrm{tr}\big[m_1 R_{1_k}(F_{1_k}-I_{3\times 3})d_1(x_{k+1}-x_k)^T\big] \\ &+ \frac{1}{h}\mathrm{tr}\big[m_2 R_{2_k}(F_{2_k}-I_{3\times 3})d_2(x_{k+1}-x_k)^T\big]. \end{aligned} \tag{39}$$

From (6), we obtain the Lie group variational integrator, viewed as discrete-time equations of motion on $\mathrm{SO}(3)\times\mathrm{SO}(3)\times\mathbb{R}^3$. Since we are only interested in rotational maneuvers, we derive the following reduced equations of motion on $\mathrm{SO}(3)\times\mathrm{SO}(3)$ using the fact that the linear momentum is conserved.

$$\begin{aligned} \hat{p}_{1_k} =& \frac{1}{h}\left\{F_{1_k}(J_{d_1}-\alpha m_1 d_1 d_1^T) - (J_{d_1}-\alpha m_1 d_1 d_1^T)F_{1_k}^T\right\} \\ &- \beta\frac{m_1}{h}(R_{1_k}^T R_{2_k} F_{2_k} d_2 d_1^T - d_1 d_2^T F_{2_k}^T R_{2_k}^T R_{1_k}) + \beta\frac{m_1}{h}(R_{1_k}^T R_{2_k} d_2 d_1^T - d_1 d_2^T R_{2_k}^T R_{1_k}), \end{aligned} \tag{40}$$

$$\begin{aligned} \hat{p}_{2_k} =& \frac{1}{h}\left\{F_{2_k}(J_{d_2}-\beta m_2 d_2 d_2^T) - (J_{d_2}-\beta m_2 d_2 d_2^T)F_{2_k}^T\right\} \\ &- \alpha\frac{m_2}{h}(R_{2_k}^T R_{1_k} F_{1_k} d_1 d_2^T - d_2 d_1^T F_{1_k}^T R_{1_k}^T R_{2_k}) + \alpha\frac{m_2}{h}(R_{2_k}^T R_{1_k} d_1 d_2^T - d_2 d_1^T R_{1_k}^T R_{2_k}), \end{aligned} \tag{41}$$

$$R_{1_{k+1}} = R_{1_k} F_{1_k}, \tag{42}$$

$$R_{2_{k+1}} = R_{2_k} F_{2_k}, \tag{43}$$

$$p_{1_{k+1}} = F_{1_k}^T(p_{1_k} - (B_{1_k} - B_{1_k}^T)^\vee) + hR_{1_{k+1}}^T u_{k+1}, \tag{44}$$

$$p_{2_{k+1}} = F_{2_k}^T (p_{2_k} - (B_{2_k} - B_{2_k}^T)^\vee) - hR_{2_{k+1}}^T u_{k+1}, \tag{45}$$

where $\alpha = \frac{m_1}{m_1+m_2}$, $\beta = \frac{m_2}{m_1+m_2} \in \mathbb{R}$, and the matrix $B_{i_k} \in \mathbb{R}^{3\times 3}$ for $i \in \{1, 2\}$ is defined as

$$B_{i_k} = \frac{m_i}{h}(F_{i_k} - I)d_i \left\{-\alpha R_{1_k}(F_{1_k} - I)d_1 - \beta R_{2_k}(F_{2_k} - I)d_2\right\}^T R_{i_k}. \tag{46}$$

The momenta variables $p_1, p_2 \in \mathbb{R}^3$ are given by

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} J_1 + \alpha m_1 \hat{d}_1^T & \beta m_1 \hat{d}_1 R_1^T R_2 \hat{d}_2 \\ \alpha m_2 \hat{d}_2 R_2^T R_1 \hat{d}_1 & J_2 + \beta m_2 \hat{d}_2^2 \end{bmatrix} \begin{bmatrix} \Omega_1 \\ \Omega_2 \end{bmatrix}. \tag{47}$$

For given $(R_{1_k}, R_{2_k}, \Omega_{1_k}, \Omega_{2_k})$, we find p_{1_k}, p_{2_k} by (47). We solve the implicit equations (40), (41) to obtain F_{1_k}, F_{2_k}. Then, $R_{1_{k+1}}, R_{2_{k+1}}$ are obtained from (42),(43), and $p_{1_{k+1}}, p_{2_{k+1}}$ are obtained by (44),(45). Finally, $\Omega_{1_{k+1}}, \Omega_{2_{k+1}}$ are computed from (47). This yields a discrete flow map $(R_{1_k}, R_{2_k}, \Omega_{1_k}, \Omega_{2_k}) \mapsto (R_{1_{k+1}}, R_{2_{k+1}}, \Omega_{1_{k+1}}, \Omega_{2_{k+1}})$.

Optimal control problem. The objective of the optimal control problem is to transfer the connected rigid bodies from a given initial condition $(R_{1_0}, R_{2_0}, \Omega_{1_0}, \Omega_{2_0})$ to a desired terminal condition $(R_1^f, R_2^f, \Omega_1^f, \Omega_2^f)$ during a fixed maneuver time Nh, while minimizing the square of the l_2 norm of the control inputs.

$$\min_{u_{k+1}} \left\{ \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2} u_{k+1}^T W u_{k+1} \right\}, \tag{48}$$

where $W \in \mathbb{R}^{3\times 3}$ is a symmetric positive-definite matrix. In particular, we choose an attitude maneuver that is described by a rest-to-rest rotation of the entire system while the relative attitude configuration at the terminal time is the same as at the initial time.

Computational approach. We apply a direct optimal control approach. For a given control input, the value of the cost is given by (48), and the terminal conditions are obtained by the discrete-time equations of motion given by (40)-(46). We use constrained nonlinear parameter optimization to minimize the cost function subject to the terminal boundary condition obtained by the discrete-time equations of motion.

Since the total angular momentum is conserved regardless of the control input, the terminal constraints introduces a singularity due to the rotational symmetry. This ill-conditioning can be avoided by disregarding the terminal angular velocity constraint for the second body. For the given boundary conditions, the terminal angular velocity condition is automatically satisfied if the remaining terminal constraints are satisfied, due to the angular momentum conservation property. By

formulating the optimization process this way, we eliminate the source of numerical ill-conditioning. This is similar to the modified computational approach discussed in Section 3.2.

Numerical example. Properties of the rigid bodies are chosen as

$$m_1 = 1.5\text{kg}, \quad J_1 = \begin{bmatrix} 0.18 & 0.32 & 0.32 \\ 0.32 & 1.88 & -0.06 \\ 0.32 & -0.06 & 1.86 \end{bmatrix} \text{kg} \cdot \text{m}^2, \quad d_1 = [-1.08, 0.20, 0.20]\text{m},$$

$$m_2 = 1\text{kg}, \quad J_2 = \begin{bmatrix} 0.11 & -0.18 & -0.18 \\ -0.18 & 0.89 & -0.04 \\ -0.18 & -0.04 & 0.88 \end{bmatrix} \text{kg} \cdot \text{m}^2, \quad d_2 = [0.9, 0.2, 0.2]\text{m}.$$

The desired maneuver is a rest-to-rest 180° rotation about the x axis.

$$R_{1_0} = I, \quad \Omega_{1_0} = 0, \quad R_{2_0} = I, \quad \Omega_{2_0} = 0,$$
$$R_1^f = \text{diag}[1, -1, -1], \quad \Omega_1^f = [0, 0, 0], \quad R_2^f = \text{diag}[1, -1, -1], \quad \Omega_2^f = [0, 0, 0].$$

The maneuver time is 4 seconds, and the step size is $h = 0.01$.

We parameterize each component of the control input at 7 discrete points, and the control inputs are reconstructed by cubic spline interpolation. The resulting 21 control input parameters are optimized by a sequential quadratic programming method to satisfy the terminal boundary conditions while minimizing the cost function.

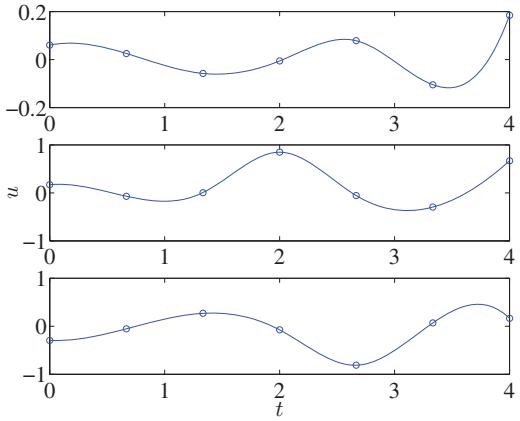
Figure 5 shows the optimal maneuver of the rigid bodies, angular velocity, and optimal control inputs. The blue circles denote the optimized control input parameters. The optimal cost and the violation of the terminal boundary conditions are 0.574, and 2.48×10^{-8}, respectively.

The optimal maneuver consists of large angle rotations of the two rigid bodies. Throughout this complicated maneuver, the total angular momentum is zero, and the rotation about the e_1 axis depends on the geometric phase effect. This also demonstrates the advantages of the computational geometric optimal control approach. The Lie group variational integrator computes the weak geometric phase effect accurately, so that the iterations converge to a nontrivial optimal maneuver of the coupled rigid bodies.

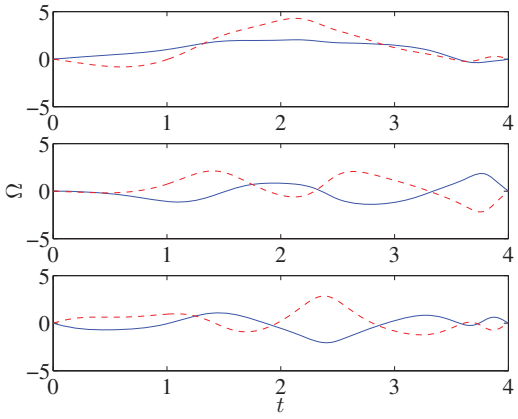
5. Conclusions

In this paper, a computational geometric approach for an optimal control problem of rigid body dynamics has been developed. The essential idea is formulating a discrete-time optimal control problem using a structure-preserving geometric numerical integrator, referred to as a Lie group

(a) Optimal maneuver



(b) Control input u



(c) Angular velocity (Ω_1:solid, Ω_2:dashed)

FIGURE 5. Optimal control of two connected rigid bodies

variational integrator, and applying standard optimal control approaches, such as an indirect optimal control and a direct optimal control, to discrete-time equations of motion. This method is in contrast to the usual optimal control approach, where the discretization appears only in the last stage when numerically computing the optimal control inputs.

The computational geometric optimal control approach has substantial advantages in terms of preserving the geometric properties of optimality conditions. The discrete flow of Lie group variational integrators has desirable geometric properties, such as symplecticity and momentum preservation, and it is more reliable and robust over longer time periods. The computational geometric optimal control approach inherits the desirable properties of the Lie group variational integrator. In the necessary conditions for optimality, the multiplier equations are dual to the linearized equations of motion. Since the linearized flow of a Lagrangian/Hamiltonian system is

also symplectic, the multiplier equations inherit certain geometric properties. The discrete-time necessary conditions preserve the geometric properties of the optimality conditions, as they are derived from a discrete-time analogue of Hamilton's variational principle that yields a symplectic discrete-time flow.

The computational geometric optimal control approach allows us to find the optimal control input more efficiently. In the indirect optimal control, the shooting method may be prone to numerical ill-conditioning, as a small change in the initial multiplier can cause highly nonlinear behavior of the terminal condition. However, as shown in Figure 2(b) and Figure 3(c), the computational geometric optimal control approach exhibits excellent numerical convergence properties. This is because the proposed computational algorithms are geometrically exact and numerically accurate. There is no numerical dissipation introduced by the numerical algorithm, and therefore, we are more accurately characterizing the sensitivities along the solution.

Another advantage of the computational geometric optimal control of rigid bodies is that the method is directly developed on a Lie group. There is no ambiguity or singularity in representing the configuration of rigid bodies globally, and the resulting equations of motion are more compact than those written in terms of local coordinates. As illustrated by Figure 4(a) and Figure 5(a), the presented computational geometric optimal control approach utilizes the effects of the nonlinear coupling and the weak geometric phase of a multibody system to obtain nontrivial aggressive maneuvers of the rigid bodies. These results are independent of a specific choice of local coordinates, and they completely avoid any singularity, ambiguity, and complexity associated with local coordinates. Furthermore, the numerical results are group-equivariant, and are independent of the choice of inertial frame, which is in contrast to methods based on local coordinate representations. By formulating the problem in a global and intrinsic fashion, the algorithms presented are able to explore the space of control strategies which extend beyond a single coordinate chart, thereby providing a deeper insight into the global controlled dynamics of systems of rigid bodies.

References

[1] J. Baillieul, "Geometric methods for nonlinear optimal control problems," *Journal of Optimization Theory and Applications*, vol. 25, pp. 519–548, 1978.

[2] K. Bilimoria and B. Wie, "Time-optimal three-axis reorientation of a rigid spacecraft," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 3, pp. 446–452, 1993.

[3] A. Bloch and P. Crouch, "Optimal control and geodesic flows," *System and Control Letters*, vol. 28, pp. 65–72, 1996.

[4] A. Bloch, P. Crouch, J. Marsden, and T. Ratiu, "Discrete rigid body dynamics and optimal control," in *IEEE Conference on Decision and Control*, 1998, pp. 2249–2254.

[5] ——, "The symmetric representation of the rigid body equations and their discretization," *Nonlinearity*, vol. 15, no. 4, pp. 1309–1341, 2002.

[6] A. Bloch, I. Hussein, M. Leok, and A. Sanyal, "Geometric structure-preserving optimal control of the rigid body," *Journal of Dynamical and Control Systems*, 2007, submitted.

[7] A. Bloch, M. Leok, J. Marsden, and D. Zenkov, "Controlled Lagrangians and stabilization of the discrete cart-pendulum system," in *Proceedings of the IEEE Conference on Decision and Control*, 2005, pp. 6579–6584.

[8] R. Brockett, "System theory on group manifolds and coset spaces," *SIAM Journal on Control*, vol. 10, pp. 265–284, 1972.

[9] ——, "Lie theory and control systems defined on spheres," *SIAM Journal on Applied Mathematics*, vol. 25, pp. 213–225, 1973.

[10] A. Bryson and Y. Ho, *Applied Optimal Control*. Hemisphere Publishing Corporation, 1975.

[11] R. Byers and S. Vadali, "Quasi-closed form solution to the time-optimal rigid spacecraft reorientation problem," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 3, pp. 453–461, 1993.

[12] H. Cendra, J. E. Marsden, and T. S. Ratiu, "Lagrangian reduction by stages," *Mem. Amer. Math. Soc.*, vol. 152, no. 722, 2001.

[13] M. Enos, Ed., *Dynamics and Control of Mechanical Systems: the falling cat and related problems*, ser. Fields Institute Communications. American Mathematical Society, 1993.

[14] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*, 2nd ed., ser. Springer Series in Computational Mathematics. Springer-Verlag, 2006, vol. 31.

[15] I. Hussein, M. Leok, A. Sanyal, and A. Bloch, "A discrete variational integrator for optimal control problems on SO(3)," in *IEEE Conference on Decision and Control*, 2006, pp. 6636–6641.

[16] A. Iserles, H. Munthe-Kaas, S. Nørsett, and A. Zanna, "Lie-group methods," in *Acta Numerica*. Cambridge University Press, 2000, vol. 9, pp. 215–365.

[17] O. Junge, J. Marsden, and S. Ober-Blöbaum, "Discrete mechanics and optimal control," in *IFAC Congress*, Praha, 2005.

[18] V. Jurdjevic, "Optimal control, geometry, and mechanics," in *Mathematical Control Theory*, J. Baillieul and J. C. Willems, Eds. Springer, 1998, ch. 7, pp. 227–267.

[19] ——, "Optimal control problems on Lie groups: Crossroads between geometry and mechanics," in *Geometry of Feedback and Optimal Control*, ser. Pure and Applied Mathematics: A series of monographs and textbooks, B. Jackubczyk and W. Respondek, Eds. CRC, 1998, ch. 7, pp. 257–304.

[20] ——, *Geometric Control Theory*. Cambridge University, 1997.

[21] C. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.

[22] T. Lee, "Computational geometric mechanics and control of rigid bodies," Ph.D. dissertation, University of Michigan, 2008.

[23] T. Lee, M. Leok, and N. H. McClamroch, "Attitude maneuvers of a rigid spacecraft in a circular orbit," in *Proceedings of the American Control Conference*, 2005, pp. 1742–1747.

[24] ——, "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum," in *Proceedings of the IEEE Conference on Control Applications*, 2005, pp. 962–967.

[25] ——, "Optimal control of a rigid body using geometrically exact computations on SE(3)," in *Proceedings of the IEEE Conference on Decision and Control*, 2006, pp. 2170–2175.

[26] ——, "A combinatorial optimal control problem for spacecraft formation reconfiguration," in *Proceedings of the IEEE Conference on Decision and Control*, 2007, pp. 5370–5375.

[27] ——, "Lie group variational integrators for the full body problem," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, pp. 2907–2924, May 2007.

[28] ——, "Lie group variational integrators for the full body problem in orbital mechanics," *Celestial Mechanics and Dynamical Astronomy*, vol. 98, no. 2, pp. 121–144, June 2007.

[29] ——, "Optimal attitude control for a rigid body with symmetry," in *Proceedings of the American Control Conference*, 2007, pp. 1073–1078.

[30] ——, "Optimal attitude control of a rigid body using geometrically exact computations on SO(3)," *Journal of Dynamical and Control Systems*, 2007, accepted. [Online]. Available: http://arxiv.org/abs/math.OC/0601424

[31] ——, "Time optimal attitude control for a rigid body," in *Proceedings of the American Control Conference*, 2008, submitted, available at http://arxiv.org/abs/0709.2514.

[32] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, ser. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2004, vol. 14.

[33] M. Leok, "Foundations of Computational Geometric Mechanics," Ph.D. dissertation, California Instittute of Technology, 2004.

[34] J. Marsden, S. Pekarsky, and S. Shkoller, "Discrete Euler–Poincaré and Lie–Poisson equations," *Nonlinearity*, vol. 12, no. 6, pp. 1647–1662, 1999.

[35] J. Marsden and T. Ratiu, *Introduction to Mechanics and Symmetry*, 2nd ed., ser. Texts in Applied Mathematics. Springer-Verlag, 1999, vol. 17.

[36] J. Marsden and M. West, "Discrete mechanics and variational integrators," in *Acta Numerica*. Cambridge University Press, 2001, vol. 10, pp. 317–514.

[37] R. Montgomery, "Guage theory of the falling cat," in *Dynamics and control of mechanical systems: the falling cat and related problems*, ser. Fields Institute Communications, M. Enos, Ed. American Mathematical Society, 1991, pp. 193–218.

[38] J. Moser and A. Veselov, "Discrete versions of some classical integrable systems and factorization of matrix polynomials," *Communications in Mathematical Physics*, vol. 139, pp. 217–243, 1991.

[39] J. Sanz-Serna, "Symplectic integrators for Hamiltonian problems: an overview," in *Acta Numerica*. Cambridge University Press, 1992, vol. 1, pp. 243–286.

[40] D. Scheeres, "Stability in the full two body problem," *Celestial Mechanics and Dynamical Astronomy*, vol. 83, pp. 155–169, 2002.

[41] S. Scrivener and R. Thompson, "Survey of time-optimal attitude maneuvers," *Journal of Guidance, Control, and Dynamics*, vol. 17, no. 2, pp. 225–233, 1994.

[42] H. Seywald and R. Kumar, "Singular control in minimum time spacecraft reorientation," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 4, pp. 686–694, 1993.

[43] J. Shen, A. Sanyal, N. Chaturvedi, D. Bernstein, and N. H. McClamroch, "Dynamics and control of a 3D pendulum," in *Proceedings of 43rd IEEE Conference on Decision and Control*, Dec. 2004, pp. 323–328.

[44] L. Wang, "Geometry, dynamics and control of coupled systems," Ph.D. dissertation, University of Maryland, 1990.

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2008; **00**:1–27 *Prepared using* nmeauth.cls *[Version: 2002/09/18 v2.02]*

Lagrangian Mechanics and Variational Integrators on Two-Spheres

Taeyoung Lee[1,*], Melvin Leok[2] and N. Harris McClamroch[1]

[1] *Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109*
[2] *Department of Mathematics, Purdue University, West Lafayette, IN 47907*

SUMMARY

Euler-Lagrange equations and variational integrators are developed for Lagrangian mechanical systems evolving on a product of two-spheres. The geometric structure of a product of two-spheres is carefully considered in order to obtain global equations of motion. Both continuous equations of motion and variational integrators completely avoid the singularities and complexities introduced by local parameterizations or explicit constraints. We derive global expressions for the Euler-Lagrange equations on two-spheres which are more compact than existing equations written in terms of angles. Since the variational integrators are derived from Hamilton's principle, they preserve the geometric features of the dynamics such as symplecticity, momentum maps, or total energy, as well as the structure of the configuration manifold. Computational properties of the variational integrators are illustrated for several mechanical systems. In addition, we describe how Lie group variational integrators can be used to integrate Lagrangian flows on more general homogeneous spaces by lifting the discrete Hamilton's principle on homogeneous spaces to a constrained discrete variational principle on the Lie group by choosing a discrete connection. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: Lagrangian mechanics, geometric integrator, variational integrator, two-sphere, homogeneous manifold

*Correspondence to: Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109, Email: tylee@umich.edu

Contract/grant sponsor: National Science Foundation; contract/grant number: DMS-0504747, DMS-0726263, ECS-0244977, CMS-0555797

Received June 23, 2008

1. Introduction

The two-sphere $\mathbb{S}^2$ is the set of all points in the Euclidean space $\mathbb{R}^3$ which are a unit distance from the origin. It is a two dimensional manifold that is locally diffeomorphic to $\mathbb{R}^2$. Many classical and interesting mechanical systems, such as a spherical pendulum, a double spherical pendulum, and magnetic models, evolve on the two-sphere or on a product of two-spheres. In this paper, we derive Euler-Lagrange equations on configuration spaces of the form $(\mathbb{S}^2)^n$, for a positive integer n. We also develop geometric numerical integrators referred to as discrete Euler-Lagrange equations or variational integrators on $(\mathbb{S}^2)^n$.

In most of the literature that treats dynamic systems on $(\mathbb{S}^2)^n$, either $2n$ angles or n explicit equality constraints enforcing unit length are used to describe the configuration of the system [1, 2]. These descriptions involve complicated trigonometric expressions and introduce additional complexity in analysis and computations. In this paper, we focus on developing continuous equations of motion and discrete equations of motion directly on $(\mathbb{S}^2)^n$, without need of local parameterizations, constraints, or reprojections. This provides a remarkably compact form of the equations of motion.

Geometric numerical integrators are numerical integration algorithms that preserve the geometric structure of the continuous dynamics, such as invariants, symplecticity, and the configuration manifold [3]. Conventional numerical integrators construct a discrete approximation to the flow using only information about the vector field, and ignore the physical laws and the geometric properties inherent in the differential equations [4]. Consequently, they do not preserve important characteristics of the dynamics of the continuous equations of motion. In contrast, variational integrators are constructed by discretizing Hamilton's principle, rather than discretizing the continuous Euler-Lagrange equation [5, 6]. Since they are developed by using a discrete version of a physical principle, the resulting integrators have the desirable property that they are symplectic and momentum preserving, and they exhibit good energy behavior for exponentially long times.

Geometric numerical integration on $\mathbb{S}^2$ has been studied in [7, 8, 9]. The two-sphere is a homogeneous manifold; the special orthogonal group $\mathrm{SO}(3)$ acts transitively on $\mathbb{S}^2$, and Lie group methods [10] can be adapted to generate numerical flows on $\mathbb{S}^2$.

In this paper, we study Lagrangian mechanical systems on $(\mathbb{S}^2)^n$. Thus, it is desirable to preserve the geometric properties of the dynamics, such as momentum map, symplecticity, and total energy, in addition to the structure of the configuration manifold [11]. We combine the approaches of geometric integrators on homogeneous manifolds and variational integrators to obtain variational integrators on $(\mathbb{S}^2)^n$ that preserve the geometric properties of the dynamics as well as the homogeneous structure of the configuration manifold $(\mathbb{S}^2)^n$ concurrently.

The main contributions of this paper are as follows. (i) In the continuous time setting, the global Euler-Lagrange equations on $(\mathbb{S}^2)^n$ are developed in a compact form without local

parameterization or constraints. This provides insight into the global dynamics on $(\mathbb{S}^2)^n$, which is desirable for theoretical studies. (ii) As a geometric numerical integrator, the discrete Euler-Lagrange equations on $(\mathbb{S}^2)^n$ are unique in the sense that they conserve both the geometric properties of the dynamics and the manifold structure of $(\mathbb{S}^2)^n$ simultaneously. The exact geometric properties of the discrete flow not only generate improved qualitative behavior, but they also provide accurate and reliable computational results in long-time simulation. (iii) More generally, Lie group variational integrators can be applied to homogeneous spaces through a constrained lifted discrete variational principle. This allows the construction of homogeneous variational integrators for homogeneous spaces such as the Stiefel and Grassmannian manifolds, which are of particular interest in many scientific and engineering applications.

A substantial class of nontrivial configuration spaces that arise in mechanical systems, including robotic arms, satellites, and underwater vehicles, can be described as products of homogeneous spaces and Lie groups. By extending the applicability of Lie group variational integrators to Lagrangian flows on homogeneous spaces, this paper provides a systematic framework for modeling and simulating systems with complicated internal shape degrees of freedom that have not been extensively studied due to the limitations of existing techniques. As demonstrated in the numerous examples presented in this paper, the proposed approach yields compact expressions that clarifies the manner in which the interconnections between bodies enter into the resulting equations of motion.

This paper is organized as follows. Lagrangian mechanics on $(\mathbb{S}^2)^n$ is described in Section 2. Variational integrators on $(\mathbb{S}^2)^n$ are developed in Section 3. Computational properties are illustrated for several mechanical systems, namely a double spherical pendulum, an n-body problem on a sphere, an interconnected system of spherical pendula, pure bending of an elastic rod, a spatial array of magnetic dipoles, and molecular dynamics that evolves on a sphere. The general theory of Lagrangian mechanics on homogeneous spaces in both continuous and discrete time is described in Section 4.

2. Lagrangian mechanics on $(\mathbb{S}^2)^n$

In this section, continuous equations of motion for a mechanical system defined on $(\mathbb{S}^2)^n$ are developed in the context of Lagrangian mechanics. It is common in the published literature that the equations of motion are developed by using either two angles or a unit length constraint to characterize $\mathbb{S}^2$. Any description with two angles has singularities, and any trajectory near a singularity experiences numerical ill-conditioning. The unit length constraint leads to additional complexity in numerical computations. We develop global continuous equations of motion without resorting to local parameterizations or constraints. To achieve this, it is critical to understand the global characteristics of a mechanical system on $(\mathbb{S}^2)^n$. This section provides a good background for understanding the theory of discrete Lagrangian mechanics

on $(\mathbb{S}^2)^n$ to be introduced in the next section.

The two-sphere is the set of points that have the unit length from the origin of $\mathbb{R}^3$, i.e. $\mathbb{S}^2 = \{q \in \mathbb{R}^3 \,|\, q \cdot q = 1\}$. The tangent space $T_q\mathbb{S}^2$ for $q \in \mathbb{S}^2$ is a plane tangent to the two-sphere at the point q. Thus, a curve $q : \mathbb{R} \to \mathbb{S}^2$ and its time derivative satisfy $q \cdot \dot{q} = 0$. The time-derivative of a curve can be written as

$$\dot{q} = \omega \times q, \tag{1}$$

where the angular velocity $\omega \in \mathbb{R}^3$ is constrained to be orthogonal to q, i.e. $q \cdot \omega = 0$. The time derivative of the angular velocity is also orthogonal to q, i.e. $q \cdot \dot{\omega} = 0$.

2.1. Euler-Lagrange equations on $(\mathbb{S}^2)^n$

We consider a mechanical system on the configuration manifold $\mathbb{S}^2 \times \cdots \times \mathbb{S}^2 = (\mathbb{S}^2)^n$. We assume that the Lagrangian $L : T(\mathbb{S}^2)^n \to \mathbb{R}$ is given by the difference between a quadratic kinetic energy and a configuration-dependent potential energy as follows.

$$L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) = \frac{1}{2} \sum_{i,j=1}^{n} M_{ij} \dot{q}_i \cdot \dot{q}_j - V(q_1, \ldots, q_n), \tag{2}$$

where $(q_i, \dot{q}_i) \in T\mathbb{S}^2$ for $i \in \{1, \ldots, n\}$, and $M_{ij} \in \mathbb{R}$ is the i, j-th element of a symmetric positive definite inertia matrix $M \in \mathbb{R}^{n \times n}$ for $i, j \in \{1, \ldots, n\}$. The configuration dependent potential is denoted by $V : (\mathbb{S}^2)^n \to \mathbb{R}$.

The action integral is defined as the time integral of the Lagrangian, and the variation of the action integral leads to continuous equations of motion by applying Hamilton's principle. These are standard procedures to derive the Euler-Lagrange equations. The expression for the infinitesimal variation of $q_i \in \mathbb{S}^2$ should be carefully developed, since the configuration manifold is not a linear vector space. As in (1), the infinitesimal variation of q_i can be written as a vector cross product,

$$\delta q_i = \xi_i \times q_i, \tag{3}$$

where $\xi_i \in \mathbb{R}^3$ is constrained to be orthogonal to q_i, i.e. $\xi_i \cdot q_i = 0$. From this, the expression for the infinitesimal variation of $\dot{q}_i$ is given by

$$\delta \dot{q}_i = \dot{\xi}_i \times q_i + \xi_i \times \dot{q}_i. \tag{4}$$

These expressions are the key elements to obtaining global equations of motion on $(\mathbb{S}^2)^n$.

The variation of the Lagrangian can be written as

$$\delta L = \sum_{i,j=1}^{n} \delta \dot{q}_i \cdot M_{ij} \dot{q}_j - \sum_{i=1}^{n} \delta q_i \cdot \frac{\partial V}{\partial q_i},$$

where the symmetric property $M_{ij} = M_{ji}$ is used. Substituting (3) and (4) into this, and using the vector identity $(a \times b) \cdot c = a \cdot (b \times c)$ for any $a, b, c \in \mathbb{R}^3$, we obtain

$$\delta L = \sum_{i,j=1}^{n} \dot{\xi}_i \cdot (q_i \times M_{ij}\dot{q}_j) + \xi_i \cdot (\dot{q}_i \times M_{ij}\dot{q}_j) - \sum_{i=1}^{n} \xi_i \cdot \left(q_i \times \frac{\partial V}{\partial q_i} \right).$$

Let $\mathfrak{G}$ be the action integral defined as $\mathfrak{G} = \int_0^T L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n)\, dt$. Using the above equation and integrating by parts, the variation of the action integral is given by

$$\delta\mathfrak{G} = \sum_{i,j=1}^{n} \xi_i \cdot (q_i \times M_{ij}\dot{q}_j)\Big|_0^T - \sum_{i=1}^{n} \int_0^T \xi_i \cdot \left[(q_i \times \sum_{j=1}^{n} M_{ij}\ddot{q}_j) + q_i \times \frac{\partial V}{\partial q_i} \right].$$

From Hamilton's principle, $\delta\mathfrak{G} = 0$ for any ξ_i vanishing at $t = 0, T$. Since ξ_i is orthogonal to q_i, the continuous equations of motion satisfy

$$(q_i \times \sum_{j=1}^{n} M_{ij}\ddot{q}_j) + q_i \times \frac{\partial V}{\partial q_i} = c_i(t) q_i \tag{5}$$

for some scalar valued functions $c_i(t)$ for $i \in \{1, \ldots, n\}$. Taking the cross product of (5) and q_i yields

$$q_i \times (q_i \times \sum_{j=1}^{n} M_{ij}\ddot{q}_j) + q_i \times \left(q_i \times \frac{\partial V}{\partial q_i} \right) = 0. \tag{6}$$

From the vector identity $a \times (b \times c) = (a \cdot c)b - (a \cdot b)c$ for any $a, b, c \in \mathbb{R}^3$, we have

$$\begin{aligned} q_i \times (q_i \times \ddot{q}_i) &= (q_i \cdot \ddot{q}_i) q_i - (q_i \cdot q_i)\ddot{q}_i, \\ &= -(\dot{q}_i \cdot \dot{q}_i) q_i - \ddot{q}_i, \end{aligned}$$

where we use the properties $\frac{d}{dt}(q_i \cdot \dot{q}_i) = q_i \cdot \ddot{q}_i + \dot{q}_i \cdot \dot{q}_i = 0$ and $q_i \cdot q_i = 1$. Substituting these into (6), we obtain an expression for $\ddot{q}_i$, which is summarized as follows.

Proposition 1. *Consider a mechanical system on $(\mathbb{S}^2)^n$ whose Lagrangian is expressed as (2). The continuous equations of motion are given by*

$$M_{ii}\ddot{q}_i = q_i \times (q_i \times \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{ij}\ddot{q}_j) - (\dot{q}_i \cdot \dot{q}_i) M_{ii} q_i + q_i \times \left(q_i \times \frac{\partial V}{\partial q_i} \right) \tag{7}$$

for $i \in \{1, \ldots, n\}$. Equivalently, this can be written in matrix form as

$$\begin{bmatrix} M_{11}I_{3\times 3} & -M_{12}\hat{q}_1\hat{q}_1 & \cdots & -M_{1n}\hat{q}_1\hat{q}_1 \\ -M_{21}\hat{q}_2\hat{q}_2 & M_{22}I_{3\times 3} & \cdots & -M_{2n}\hat{q}_2\hat{q}_2 \\ \vdots & \vdots & & \vdots \\ -M_{n1}\hat{q}_n\hat{q}_n & -M_{n2}\hat{q}_n\hat{q}_n & \cdots & M_{nn}I_{3\times 3} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \vdots \\ \ddot{q}_n \end{bmatrix} = \begin{bmatrix} -(\dot{q}_1 \cdot \dot{q}_1) M_{11} q_1 + \hat{q}_1^2 \frac{\partial V}{\partial q_1} \\ -(\dot{q}_2 \cdot \dot{q}_2) M_{22} q_2 + \hat{q}_2^2 \frac{\partial V}{\partial q_2} \\ \vdots \\ -(\dot{q}_n \cdot \dot{q}_n) M_{nn} q_n + \hat{q}_n^2 \frac{\partial V}{\partial q_n} \end{bmatrix}, \tag{8}$$

where the hat map $\hat{\cdot} : \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ is defined such that $\hat{a}b = a \times b$ for any $a, b \in \mathbb{R}^3$.

Since $\dot{q}_i = \omega_i \times q_i$ for the angular velocity ω_i satisfying $q_i \cdot \omega_i = 0$, we have

$$\begin{aligned}\ddot{q}_i &= \dot{\omega}_i \times q_i + \omega_i \times (\omega_i \times q_i),\\ &= \dot{\omega}_i \times q_i - (\omega_i \cdot \omega_i) q_i.\end{aligned}$$

Substituting this into (5) and using the fact that $q_i \cdot \dot{\omega}_i = 0$, we obtain continuous equations of motion in terms of the angular velocity.

Corollary 1. *The continuous equations of motion given by (7) can be written in terms of the angular velocity as*

$$M_{ii}\dot{\omega}_i = \sum_{\substack{j=1\\ j\neq i}}^{n} (M_{ij} q_i \times (q_j \times \dot{\omega}_j) + M_{ij}(\omega_j \cdot \omega_j) q_i \times q_j) - q_i \times \frac{\partial V}{\partial q_i}, \tag{9}$$

$$\dot{q}_i = \omega_i \times q_i \tag{10}$$

for $i \in \{1, \ldots, n\}$. Equivalently, this can be written in matrix form as

$$\begin{bmatrix} M_{11}I_{3\times 3} & -M_{12}\hat{q}_1\hat{q}_2 & \cdots & -M_{1n}\hat{q}_1\hat{q}_n \\ -M_{21}\hat{q}_2\hat{q}_1 & M_{22}I_{3\times 3} & \cdots & -M_{2n}\hat{q}_2\hat{q}_n \\ \vdots & \vdots & & \vdots \\ -M_{n1}\hat{q}_n\hat{q}_1 & -M_{n2}\hat{q}_n\hat{q}_2 & \cdots & M_{nn}I_{3\times 3} \end{bmatrix} \begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \\ \vdots \\ \dot{\omega}_n \end{bmatrix} = \begin{bmatrix} \sum_{j=2}^{n} M_{1j}(\omega_j \cdot \omega_j)\hat{q}_1 q_j - \hat{q}_1 \frac{\partial V}{\partial q_1} \\ \sum_{j=1, j\neq 2}^{n} M_{2j}(\omega_j \cdot \omega_j)\hat{q}_2 q_j - \hat{q}_2 \frac{\partial V}{\partial q_2} \\ \vdots \\ \sum_{j=1}^{n-1} M_{nj}(\omega_j \cdot \omega_j)\hat{q}_n q_j - \hat{q}_n \frac{\partial V}{\partial q_n} \end{bmatrix}. \tag{11}$$

Equations (7)–(11) are global continuous equations of motion for a mechanical system on $(\mathbb{S}^2)^n$. They avoid singularities completely, and they preserve the structure of $T(\mathbb{S}^2)^n$ automatically, if an initial condition is chosen properly. These equations are useful for understanding global characteristics of the dynamics. In addition, these expressions are dramatically more compact than the equations of motion written in terms of any local parameterization.

We need to check that the $3n \times 3n$ matrices given by the first terms of (8) and (11) are nonsingular. This is a property of the mechanical system itself, rather than a consequence of the particular form of equations of motion. For example, when $n = 2$, it can be shown that

$$\begin{aligned}\det \begin{bmatrix} M_{11}I_{3\times 3} & -M_{12}\hat{q}_1\hat{q}_1 \\ -M_{12}\hat{q}_2\hat{q}_2 & M_{22}I_{3\times 3} \end{bmatrix} &= \det \begin{bmatrix} M_{11}I_{3\times 3} & -M_{12}\hat{q}_1\hat{q}_2 \\ -M_{12}\hat{q}_2\hat{q}_1 & M_{22}I_{3\times 3} \end{bmatrix},\\ &= M_{11}^2 M_{22}^2 (M_{11}M_{22} - M_{12}^2 (q_1 \cdot q_2)^2)(M_{11}M_{22} - M_{12}^2).\end{aligned}$$

Since the inertia matrix is symmetric positive definite, $M_{11}, M_{22} > 0$, $M_{11}M_{22} > M_{12}^2$, and from the Cauchy-Schwarz inequality, $(q_1 \cdot q_2)^2 \leq (q_1 \cdot q_1)(q_2 \cdot q_2) = 1$. Thus, the above matrices are non-singular. One may show a similar property for $n > 2$. Throughtout this paper, it is assumed that the $3n \times 3n$ matrices given at the first terms of (8) and (11) are nonsingular. Under this assumption, the Legendre transformation given in the next subsection is a diffeomorphism; the Lagrangian is hyperregular.

2.2. Legendre transformation

The Legendre transformation of the Lagrangian gives an equivalent Hamiltonian form of equations of motion in terms of conjugate momenta if the Lagrangian is hyperregular. Here, we find expressions for the conjugate momenta, which are used in the following section for the discrete equations of motion. For $q_i \in \mathbb{S}^2$, the corresponding conjugate momentum p_i lies in the dual space $T^*_{q_i}\mathbb{S}^2$. We identify the tangent space $T_{q_i}\mathbb{S}^2$ and its dual space $T^*_{q_i}\mathbb{S}^2$ by using the usual dot product in $\mathbb{R}^3$. The Legendre transformation is given by

$$\begin{aligned} p_i \cdot \delta q_i &= \mathbf{D}_{\dot{q}_i} L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) \cdot \delta q_i, \\ &= \sum_{j=1}^{n} M_{ij} \dot{q}_j \cdot \delta q_i, \end{aligned}$$

which is satisfied for any δq_i perpendicular to q_i. Here $\mathbf{D}_{\dot{q}_i} L$ denotes the derivative of the Lagrangian with respect to $\dot{q}_i$. The momentum p_i is an element of the dual space identified with the tangent space, and the component parallel to q_i has no effect since $\delta q_i \cdot q_i = 0$. As such, the vector representing p_i is perpendicular to q_i, and p_i is equal to the projection of $\sum_{j=1}^{n} M_{ij}\dot{q}_j$ onto the orthogonal complement to q_i,

$$\begin{aligned} p_i &= \sum_{j=1}^{n} (M_{ij}\dot{q}_j - (q_i \cdot M_{ij}\dot{q}_j) q_i) = \sum_{j=1}^{n} ((q_i \cdot q_i) M_{ij}\dot{q}_j - (q_i \cdot M_{ij}\dot{q}_j) q_i), \\ &= M_{ii}\dot{q}_i - q_i \times (q_i \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}\dot{q}_j). \end{aligned} \tag{12}$$

The time derivative of p_i is given by

$$\dot{p}_i = M_{ii}\ddot{q}_i - \dot{q}_i \times (q_i \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}\dot{q}_j) - q_i \times (\dot{q}_i \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}\dot{q}_j) - q_i \times (q_i \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}\ddot{q}_j)$$

Substituting (7), and using the vector identity $a \times (b \times c) = (a \cdot c)b - (a \cdot b)c$, we obtain the following Hamilton's equations.

Corollary 2. *Consider a mechanical system on $(\mathbb{S}^2)^n$ whose Lagrangian is expressed as (2). The Hamilton's equations are given by*

$$p_i = M_{ii}\dot{q}_i - q_i \times (q_i \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}\dot{q}_j), \tag{13}$$

$$\dot{p}_i = -\sum_{j=1}^{n} (\dot{q}_i \cdot M_{ij}\dot{q}_j) q_i - \sum_{\substack{j=1 \\ j\neq i}}^{n} (q_i \cdot M_{ij}\dot{q}_j)\dot{q}_i + q_i \times \left(q_i \times \frac{\partial U}{\partial q_i} \right) \tag{14}$$

for $i \in \{1, \ldots, n\}$. Equivalently, (13) can be written in a matrix form as

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \vdots \\ \dot{q}_n \end{bmatrix} = \begin{bmatrix} M_{11} I_{3\times 3} & -M_{12}\hat{q}_1\hat{q}_1 & \cdots & -M_{1n}\hat{q}_1\hat{q}_1 \\ -M_{21}\hat{q}_2\hat{q}_2 & M_{22} I_{3\times 3} & \cdots & -M_{2n}\hat{q}_2\hat{q}_2 \\ \vdots & \vdots & & \vdots \\ -M_{n1}\hat{q}_n\hat{q}_n & -M_{n2}\hat{q}_n\hat{q}_n & \cdots & M_{nn} I_{3\times 3} \end{bmatrix}^{-1} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}. \tag{15}$$

3. Variational integrators on $(\mathbb{S}^2)^n$

The dynamics of Lagrangian and Hamiltonian systems on $(\mathbb{S}^2)^n$ have unique geometric properties; the Hamiltonian flow is symplectic, the total energy is conserved in the absence of non-conservative forces, and the momentum map associated with a symmetry of the system is preserved. The configuration space is a homogeneous manifold. These geometric features determine the qualitative dynamics of the system, and serve as a basis for theoretical study.

Conventional numerical integrators construct a discrete approximation of the flow using only information about the vector field. Other than the direction specified by the vector field, they completely ignore the physical laws and the geometric properties inherent in the differential equations [4]. For example, if we integrate (11) by using an explicit Runge-Kutta method, the unit length of the vector q_i, and the total energy are not preserved numerically; we will see this later in this paper.

Numerical integration methods that preserve the simplecticity of a Hamiltonian system have been studied [12]. Coefficients of a Runge-Kutta method can be carefully chosen to satisfy a simplecticity criterion and order conditions to obtain a symplectic Runge-Kutta method. However, it can be difficult to construct such integrators, and it is not guaranteed that other invariants of the system, such as the momentum map, are preserved. Alternatively, variational integrators are constructed by discretizing Hamilton's principle, rather than by discretizing the continuous Euler-Lagrange equation [5, 6]. The key feature of variational integrators is that they are derived by a discrete version of a physical principle, so the resulting integrators satisfy the physical properties automatically in a discrete sense; they are symplectic and momentum preserving, and they exhibit good energy behavior for exponentially long times. Lie group methods are numerical integrators that preserve the Lie group structure of the configuration space [10]. Recently, these two approaches have been unified to obtain Lie group variational integrators that preserve the geometric properties of the dynamics as well as the Lie group structure of the configuration manifold [13].

The two-sphere is a homogeneous manifold. It does not have a Lie group structure by itself, but instead, the special orthogonal group, $\mathrm{SO}(3) = \{F \in \mathbb{R}^{3\times 3} \,|\, F^T F = I_{3\times 3}, \det F = 1\}$, acts on $\mathbb{S}^2$ in a transitive way; for any $q_1, q_2 \in \mathbb{S}^2$, there exist $F \in \mathrm{SO}(3)$ such that $q_2 = Fq_1$. If a group acts transitively on a manifold, a curve on the manifold can be represented as

the action of a curve in the Lie group on an initial point on the manifold. As such, Lie group methods can be applied to obtain numerical integration schemes for homogeneous manifolds [7, 9, 8]. However, it is not guaranteed that these methods preserve the geometric properties of the dynamics. In this paper, we focus on a Lagrangian mechanical system evolving on the homogeneous manifold, $(\mathbb{S}^2)^n$ by extending the method of Lie group variational integrators [11, 13]. The resulting integrator preserves the dynamic characteristics and the homogeneous manifold structure concurrently.

3.1. Discrete Euler-Lagrange equations on $(\mathbb{S}^2)^n$

The procedure to derive discrete Euler-Lagrange equations follows the development of the continuous time case; the tangent bundle is replaced by a cartesian product of the configuration manifold, a discrete Lagrangian is chosen to approximate the integral of the Lagrangian over a discrete time step, and the variation of the corresponding discrete action sum provides discrete Euler-Lagrange equations, referred to as a variational integrator. The discrete version of the Legendre transformation yields the discrete equations in Hamiltonian form.

Let the number of timesteps be N, with constant timesteps $h > 0$. A variable with subscript k denotes the value of variable at $t = kh$. Define a discrete Lagrangian $L_d : (\mathbb{S}^2)^n \times (\mathbb{S}^2)^n \to \mathbb{R}$ such that it approximates the integral of the Lagrangian given by (2) over a discrete time step

$$L_d(q_{1_k}, \ldots, q_{n_k}, q_{1_{k+1}}, \ldots, q_{n_{k+1}}) = \frac{1}{2h} \sum_{i,j=1}^{n} M_{ij}(q_{i_{k+1}} - q_{i_k}) \cdot (q_{j_{k+1}} - q_{j_k}) - \frac{h}{2} V_k - \frac{h}{2} V_{k+1}, \tag{16}$$

where V_k denotes the value of the potential at the k-th step, i.e. $V_k = V(q_{1_k}, \ldots, q_{n_k})$. As given in (3), the infinitesimal variation of q_{i_k} is written as

$$\delta q_{i_k} = \xi_{i_k} \times q_{i_k}, \tag{17}$$

where $\xi_{i_k} \in \mathbb{R}^3$ is constrained to be orthogonal to q_{i_k}, i.e. $\xi_{i_k} \cdot q_{i_k} = 0$. The variation of the discrete Lagrangian can be written as

$$\delta L_{d_k} = \frac{1}{h} \sum_{i,j=1}^{n} (\delta q_{i_{k+1}} - \delta q_{i_k}) \cdot M_{ij}(q_{j_{k+1}} - q_{j_k}) - \frac{h}{2} \sum_{i=1}^{n} \left(\delta q_{i_k} \cdot \frac{\partial V_k}{\partial q_{i_k}} + \delta q_{i_{k+1}} \cdot \frac{\partial V_{k+1}}{\partial q_{i_{k+1}}} \right). \tag{18}$$

Substituting (17) into (18), and using the vector identity $(a \times b) \cdot c = a \cdot (b \times c)$ for any $a, b, c \in \mathbb{R}^3$, we obtain

$$\begin{aligned} \delta L_{d_k} = \frac{1}{h} \sum_{i,j=1}^{n} & \left(\xi_{i_{k+1}} \cdot (q_{i_{k+1}} \times M_{ij}(q_{j_{k+1}} - q_{j_k})) - \xi_{i_k} \cdot (q_{i_k} \times M_{ij}(q_{j_{k+1}} - q_{j_k})) \right) \\ & - \frac{h}{2} \sum_{i=1}^{n} \left(\xi_{i_k} \cdot \left(q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}} \right) + \xi_{i_{k+1}} \cdot \left(q_{i_{k+1}} \times \frac{\partial V_{k+1}}{\partial q_{i_{k+1}}} \right) \right). \end{aligned} \tag{19}$$

Let $\mathfrak{G}_d$ be the discrete action sum defined as $\mathfrak{G}_d = \sum_{k=0}^{N-1} L_{d_k}$, which approximates the action integral as the discrete Lagrangian approximates a piece of the action integral over a discrete time step. The variation of the action sum is obtained by using (19). Using the fact that ξ_{i_k} vanish at $k = 0$ and $k = N$, we can reindex the summation, which is the discrete analog of integration by parts, to yield

$$\delta\mathfrak{G}_d = \sum_{k=1}^{N-1} \sum_{i=1}^{n} \xi_{i_k} \cdot \left[\frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(-q_{j_{k+1}} + 2q_{j_k} - q_{j_{k-1}})) - hq_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}} \right].$$

From discrete Hamilton's principle $\delta\mathfrak{G}_d = 0$ for any ξ_{i_k} perpendicular to q_{i_k}. Using the same argument given in (5), the discrete equations of motion are given by

$$\frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(-q_{j_{k+1}} + 2q_{j_k} - q_{j_{k-1}})) - hq_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}} = 0 \tag{20}$$

for $i \in \{1, \dots n\}$. In addition, we require that the unit length of the vector q_{i_k} is preserved. This is achieved by viewing $\mathbb{S}^2$ as a homogeneous manifold. Since the special orthogonal group SO(3) acts on $\mathbb{S}^2$ transitively, we can define a discrete update map for q_{i_k} as

$$q_{i_{k+1}} = F_{i_k} q_{i_k}$$

for $F_{i_k} \in \mathrm{SO}(3)$. Then, the unit length of the vector q_i is preserved through the discrete equations of motion, since $q_{i_{k+1}} \cdot q_{i_{k+1}} = q_{i_k}^T F_{i_k}^T F_{i_k} q_{i_k} = 1$. These results are summarized as follows.

Proposition 2. *Consider a mechanical system on $(\mathbb{S}^2)^n$ whose Lagrangian is expressed as (2). The discrete equations of motion are given by*

$$M_{ii} q_{i_k} \times F_{i_k} q_{i_k} + q_{i_k} \times \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{ij}(F_{j_k} - I_{3\times 3}) q_{j_k} = q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_k} - q_{j_{k-1}}) - h^2 q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}}, \tag{21}$$

$$q_{i_{k+1}} = F_{i_k} q_{i_k} \tag{22}$$

for $i \in \{1, \dots n\}$. For given $(q_{i_{k-1}}, q_{i_k})$, we solve (21) to obtain $F_{i_k} \in \mathrm{SO}(3)$. Then, $q_{i_{k+1}}$ is computed by (22). This yields a discrete flow map $(q_{i_{k-1}}, q_{i_k}) \mapsto (q_{i_k}, q_{i_{k+1}})$, and this process is repeated.

Remark 1. *It should be noted that the equations given in Proposition 2 do not uniquely define the rotation matrices F_{i_k} which are used to generate the updates on $(\mathbb{S}^2)^n$. This is since the Lie group action of $SO(3)$ on $\mathbb{S}^2$ has a non-trivial isotropy subgroup, corresponding to the fact that rotating a vector q_i about the q_i axis leaves the vector invariant. In order to make the equations well-defined, we need to choose a connection on $TSO(3)$ that is complementary to*

the local isotropy direction, and require that the rotation matrix used to update the solution is in the horizontal distribution associated with this connection. In Section 3.3, the connection is chosen by requiring that $F_i = \mathrm{Cay}(f_i)$, where f_i is constrained to satisfy $f_i \cdot q_i = 0$.

3.2. Discrete Legendre transformation

We find discrete equations of motion in terms of the angular velocity. The discrete Legendre transformation is given as follows [6].

$$\begin{aligned} p_{i_k} \cdot \delta q_{i_k} &= -\mathbf{D}_{q_{i_k}} L_{d_k} \cdot \delta q_{i_k}, \\ &= \left[\frac{1}{h} \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k}) + \frac{h}{2} \frac{\partial V_k}{\partial q_{i_k}} \right] \cdot \delta q_{i_k}, \end{aligned}$$

which can be directly obtained from (18). This is satisfied for any δq_{i_k} perpendicular to q_{i_k}. Using the same argument used to derive (12), the conjugate momenta p_{i_k} is the projection of the expression in brackets onto the orthogonal complement of q_{i_k}. Thus, we obtain

$$p_{i_k} = -\frac{1}{h} q_{i_k} \times (q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{i_k} \times \left(q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}} \right).$$

Similarly, we obtain

$$\begin{aligned} p_{i_{k+1}} \cdot \delta q_{i_{k+1}} &= \mathbf{D}_{q_{i_{k+1}}} L_{d_k} \cdot \delta q_{i_{k+1}} \\ &= \left[\frac{1}{h} \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k}) - \frac{h}{2} \frac{\partial V_{k+1}}{\partial q_{i_{k+1}}} \right] \cdot \delta q_{i_{k+1}}. \end{aligned}$$

Since $p_{i_{k+1}}$ is perpendicular to $q_{i_{k+1}}$, it is given by

$$p_{i_{k+1}} = -\frac{1}{h} q_{i_{k+1}} \times (q_{i_{k+1}} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) + \frac{h}{2} q_{i_{k+1}} \times \left(q_{i_{k+1}} \times \frac{\partial V_{k+1}}{\partial q_{i_{k+1}}} \right). \tag{23}$$

This yields the discrete-time Hamilton's equations as follows.

Corollary 3. *Consider a mechanical system on $(\mathbb{S}^2)^n$ whose Lagrangian is expressed as (2). The discrete-time Hamilton's equations are given by*

$$p_{i_k} = -\frac{1}{h} q_{i_k} \times (q_{i_k} \times \sum_{j=1}^{n} M_{ij}(F_{j_k} - I_{3\times 3}) q_{j_k}) - \frac{h}{2} q_{i_k} \times \left(q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}} \right), \tag{24}$$

$$q_{i_{k+1}} = F_{i_k} q_{i_k}, \tag{25}$$

$$p_{i_{k+1}} = -\frac{1}{h} q_{i_{k+1}} \times (q_{i_{k+1}} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) + \frac{h}{2} q_{i_{k+1}} \times \left(q_{i_{k+1}} \times \frac{\partial V_{k+1}}{\partial q_{i_{k+1}}} \right) \tag{26}$$

for $i \in \{1, \ldots, n\}$. For given (q_{i_k}, p_{i_k}), we solve (24) to obtain $F_{i_k} \in \mathrm{SO}(3)$. Then, $q_{i_{k+1}}$ and $p_{i_{k+1}}$ are computed by (25) and (26), respectively. This yields a discrete-time flow map $(q_{i_k}, p_{i_k}) \mapsto (q_{i_{k+1}}, p_{i_{k+1}})$, and this process is repeated.

This provides a discrete-time flow map in terms of the conjugate momenta. Now, we find a discrete-time flow map written in terms of the angular velocity. Comparing (24) to (12), substituting $\dot{q}_{i_k} = \omega_{i_k} \times q_{i_k}$, and rearranging, we obtain

$$q_{j_k} \times \left[M_{ii}\omega_{i_k} + (q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(\omega_{j_k} \times q_{j_k})) - \frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}} \right] = 0.$$

Since the expression in the brackets is orthogonal to q_{i_k}, the left hand side is equal to zero if and only if the expression in the brackets is zero. Thus,

$$M_{ii}\omega_{i_k} + (q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(\omega_{j_k} \times q_{j_k})) = \frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) + \frac{h}{2} q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}}. \quad (27)$$

This provides a relationship between (q_{i_k}, ω_{i_k}) and $(q_{i_k}, q_{i_{k+1}})$. Comparing this with (20), we obtain

$$M_{ii}\omega_{i_k} + (q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(\omega_{j_k} \times q_{j_k})) = \frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_k} - q_{j_{k-1}})) - \frac{h}{2} q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}}, \quad (28)$$

which provides a relationship between (q_{i_k}, ω_{i_k}) and $(q_{i_{k-1}}, q_{i_k})$. Equations (27) and (28) give a discrete flow map in terms of the angular velocity; for a given (q_{i_k}, ω_{i_k}), we find $(q_{i_k}, q_{i_{k+1}})$ by using (27). Substituting this into (28) expressed at the $(k+1)$th step, we obtain $(q_{i_{k+1}}, \omega_{i_{k+1}})$. This procedure is summarized as follows.

Corollary 4. *The discrete equations of motion given by (21) and (22) can be written in terms of the angular velocity as*

$$M_{ii}q_{i_k} \times F_{i_k}q_{i_k} + q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(F_{j_k} - I_{3\times 3})q_{j_k} = M_{ii}h\omega_{i_k} - (q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(q_{j_k} \times h\omega_{j_k})) - \frac{h^2}{2} q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}}, \quad (29)$$

$$q_{i_{k+1}} = F_{i_k}q_{i_k}, \quad (30)$$

$$M_{ii}\omega_{i_{k+1}} - (q_{i_{k+1}} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(q_{j_{k+1}} \times \omega_{j_{k+1}})) = \frac{1}{h}(q_{i_{k+1}} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{i_{k+1}} \times \frac{\partial V_{k+1}}{\partial q_{i_{k+1}}} \quad (31)$$

for $i \in \{1, \dots, n\}$. Equivalently, (31) can be written in a matrix form as

$$\begin{bmatrix} M_{11}I_{3\times 3} & -M_{12}\hat{q}_{1_{k+1}}\hat{q}_{2_{k+1}} & \cdots & -M_{1n}\hat{q}_1\hat{q}_{n_{k+1}} \\ -M_{21}\hat{q}_{2_{k+1}}\hat{q}_{1_{k+1}} & M_{22}I_{3\times 3} & \cdots & -M_{2n}\hat{q}_{2_{k+1}}\hat{q}_{n_{k+1}} \\ \vdots & \vdots & & \vdots \\ -M_{n1}\hat{q}_{n_{k+1}}\hat{q}_{1_{k+1}} & -M_{n2}\hat{q}_{n_{k+1}}\hat{q}_{2_{k+1}} & \cdots & M_{nn}I_{3\times 3} \end{bmatrix} \begin{bmatrix} \omega_{1_{k+1}} \\ \omega_{2_{k+1}} \\ \vdots \\ \omega_{n_{k+1}} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{h}(q_{1_{k+1}} \times \sum_{j=1}^{n} M_{1j}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{1_{k+1}} \times \frac{\partial V_{k+1}}{\partial q_{1_{k+1}}} \\ \frac{1}{h}(q_{2_{k+1}} \times \sum_{j=1}^{n} M_{2j}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{2_{k+1}} \times \frac{\partial V_{k+1}}{\partial q_{2_{k+1}}} \\ \vdots \\ \frac{1}{h}(q_{n_{k+1}} \times \sum_{j=1}^{n} M_{nj}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{n_{k+1}} \times \frac{\partial V_{k+1}}{\partial q_{n_{k+1}}} \end{bmatrix}. \tag{32}$$

For a given (q_{i_k}, ω_{i_k}), we solve (29) to obtain $F_{i_k} \in \mathrm{SO}(3)$. Then, $q_{i_{k+1}}$ and $\omega_{i_{k+1}}$ are computed by (30) and (32), respectively. This yields a discrete flow map in terms of the angular velocity $(q_{i_k}, \omega_{i_k}) \mapsto (q_{i_{k+1}}, \omega_{i_{k+1}})$, and this process is repeated.

3.3. Computational approach

For the discrete equations of motion, we need to solve (21) and (29) to obtain $F_{i_k} \in \mathrm{SO}(3)$. Here we present a computational approach. The implicit equations given by (21) and (29) have the following structure.

$$M_{ii} q_i \times F_i q_i + q_i \times \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{ij}(F_j - I_{3\times 3}) q_j = d_i \tag{33}$$

for $i \in \{1, \dots, n\}$, where $M_{ij} \in \mathbb{R}$, $q_i \in \mathbb{S}^2$, $d_i \in \mathbb{R}^3$ are known, and we need to find $F_i \in \mathrm{SO}(3)$. We derive an equivalent equation in terms of local coordinates for F_i. This is reasonable since F_i represents the relative update between two integration steps. Using the Cayley transformation [14], $F_i \in \mathrm{SO}(3)$ can be expressed in terms of $f_i \in \mathbb{R}^3$ as

$$\begin{aligned} F_i &= (I_{3\times 3} + \hat{f}_i)(I_{3\times 3} - \hat{f}_i)^{-1}, \\ &= \frac{1}{1 + f_i \cdot f_i}((1 - f_i \cdot f_i) I_{3\times 3} + 2 f_i f_i^T + 2\hat{f}_i). \end{aligned}$$

The operation $F_i q_i$ can be considered as a rotation of the vector q_i about the direction f_i with rotation angle $2\tan^{-1}\|f_i\|$. Since the rotation of the vector q_i about the direction q_i has no effect, we can assume that f_i is orthogonal to q_i, i.e. $f_i \cdot q_i = 0$. Under this assumption, $F_i q_i$ is given by

$$F_i q_i = \frac{1}{1 + f_i \cdot f_i}((1 - f_i \cdot f_i) q_i + 2\hat{f}_i q_i). \tag{34}$$

Thus, we obtain

$$q_i \times F_i q_i = \frac{2}{1 + f_i \cdot f_i} q_i \times (f_i \times q_i) = \frac{2}{1 + f_i \cdot f_i} f_i,$$
$$(F_j - I_{3\times 3}) q_j = -\frac{2}{1 + f_j \cdot f_j} (q_j f_j^T + \hat{q}_j) f_j,$$

where we use the property, $\hat{q}_i f_i = q_i \times f_i = -\hat{f}_i q_i$. Substituting these into (33), we obtain

$$\begin{bmatrix} \frac{2M_{11}I_{3\times 3}}{1+f_1\cdot f_1} & -\frac{2M_{12}\hat{q}_1(\hat{q}_2+q_2f_2^T)}{1+f_2\cdot f_2} & \cdots & -\frac{2M_{1n}\hat{q}_1(\hat{q}_n+q_nf_n^T)}{1+f_n\cdot f_n} \\ -\frac{2M_{21}\hat{q}_2(\hat{q}_1+q_1f_1^T)}{1+f_1\cdot f_1} & \frac{2M_{22}I_{3\times 3}}{1+f_2\cdot f_2} & \cdots & -\frac{2M_{2n}\hat{q}_2(\hat{q}_n+q_nf_n^T)}{1+f_n\cdot f_n} \\ \vdots & \vdots & & \vdots \\ -\frac{2M_{n1}\hat{q}_n(\hat{q}_1+q_1f_1^T)}{1+f_1\cdot f_1} & -\frac{2M_{n2}\hat{q}_n(\hat{q}_2+q_2f_2^T)}{1+f_2\cdot f_2} & \cdots & \frac{2M_{nn}I_{3\times 3}}{1+f_n\cdot f_n} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}, \tag{35}$$

which is an equation equivalent to (33), written in terms of local coordinates for F_i using the Cayley transformation. Any numerical method to solve nonlinear equations can be applied to find f_i. Then, $F_i q_i$ is computed by using (34). In particular, (35) is written in a form that can be readily applied to a fixed point iteration method [15].

If there are no coupling terms in the kinetic energy, we can obtain an explicit solution of (33). When $M_{ij} = 0$ for $i \neq j$, (35) reduces to

$$\frac{2M_{ii}}{1 + f_i \cdot f_i} f_i = d_i.$$

Using the identity, $\frac{2\tan\theta}{1+\tan^2\theta} = \sin 2\theta$ for any $\theta \in \mathbb{R}$, it can be shown that the solution of this equation is given by $f_i = \tan\left(\frac{1}{2}\sin^{-1}(\|d_i\|/M_{ii})\right)\frac{d_i}{\|d_i\|}$. Substituting this into (34) and rearranging, we obtain

$$F_i q_i = \frac{d_i}{M_{ii}} \times q_i + \left(1 - \left\|\frac{d_i}{M_{ii}}\right\|^2\right)^{1/2} q_i.$$

Using this expression, we can rewrite the discrete equations of motion given in (29)–(32) in an explicit form.

Corollary 5. *Consider a mechanical system on $(\mathbb{S}^2)^n$ whose Lagrangian is expressed as (2) where $M_{ij} = 0$ for $i \neq j$, i.e. the dynamics are coupled only though the potential energy. The explicit discrete equations of motion are given by*

$$q_{i_{k+1}} = \left(h\omega_{i_k} - \frac{h^2}{2M_{ii}} q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}}\right) \times q_{i_k} + \left(1 - \left\|h\omega_{i_k} - \frac{h^2}{2M_{ii}} q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}}\right\|^2\right)^{1/2} q_{i_k}, \tag{36}$$

$$\omega_{i_{k+1}} = \omega_{i_k} - \frac{h}{2M_{ii}} q_{i_k} \times \frac{\partial V_k}{\partial q_{i_k}} - \frac{h}{2M_{ii}} q_{i_{k+1}} \times \frac{\partial V_{k+1}}{\partial q_{i_{k+1}}} \tag{37}$$

for $i \in \{1, \ldots, n\}$.

3.4. Properties of variational integrators on $(\mathbb{S}^2)^n$

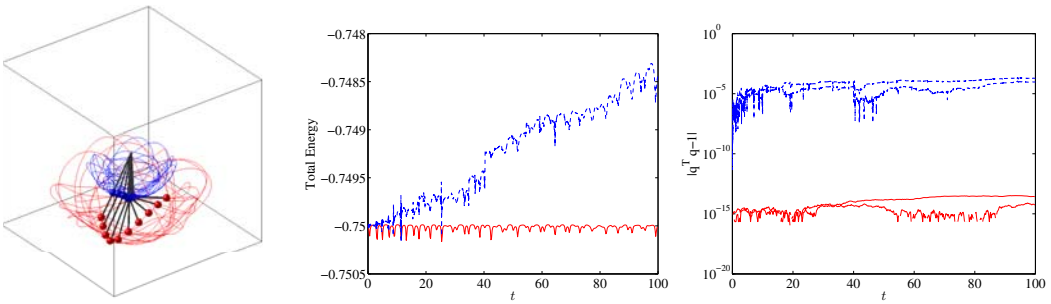
Since variational integrators are derived from the discrete Hamilton's principle, they are symplectic, and momentum preserving. The discrete action sum can be considered as a zero-form on $(\mathbb{S}^2)^n \times (\mathbb{S}^2)^n$ which maps the initial condition of a discrete flow satisfying the discrete Euler-Lagrange equation to the action sum for that trajectory. The simplecticity of the discrete flow follows from the fact the the iterated exterior derivative of any differential form is zero. If the discrete Lagrangian exhibits a symmetry, the corresponding momentum map is preserved since by symmetry, the variation of the discrete Lagrangian in the symmetry direction is zero, which in combination with the discrete Euler–Lagrange equations, implies a discrete version of Noether's theorem. Detailed proofs for the symplectic property and the momentum preserving property can be found in [6]. The total energy oscillates around its initial value with small bounds on a comparatively short timescale, but there is no tendency for the mean of the oscillation in the total energy to drift (increase or decrease) over exponentially long times [16].

The variational integrators presented in this paper preserve the structure of $(\mathbb{S}^2)^n$ without need of local parameterizations, explicit constraints or reprojection. Using the characteristics of the homogeneous manifold, the discrete update map is represented by a group action of SO(3), and a proper subspace is searched to obtain a compact, possibly explicit, form for the numerical integrator. As a result, the following numerical problems are avoided: (i) local parameterizations yield singularities; (ii) numerical trajectories in the vicinity of a singularity experiences numerical ill-conditioning; (iii) unit length constraints lead to additional computational complexity; (iv) reprojection corrupts the numerical accuracy of trajectories [3, 9].

It can be shown that these variational integrators have second-order accuracy as the discrete action sum is a second-order approximation of the action integral, which we have constructed through linear interpolation and the trapezoidal approximation of the integral. Indeed, the variational integrator obtained for $(\mathbb{S}^2)^n$ for this discrete Lagrangian is equivalent to the RATTLE discretization of the corresponding constrained Lagrangian system with the Lagrange multipliers explicitly eliminated. More formally, one can consider the construction we have presented in this paper to be a systematic means of constructing a symplectic integrator based on index-reduction of the constraints. While higher-order integrators can be easily constructed by applying a symmetric composition method [17], this work also provides a framework for directly obtaining higher-order variational integrators for Lagrangian flows on $(\mathbb{S}^2)^n$.

3.5. Numerical examples

The computational properties of variational integrators on $(\mathbb{S}^2)^n$ and explicit Runge-Kutta methods are compared for several mechanical systems taken from variety of scientific areas, namely a double spherical pendulum, an n-body problem on a sphere, an interconnected system



(a) Trajectory of pendulum

(b) Computed total energy

(c) Unit length error

Figure 1. Numerical simulation of a double spherical pendulum (RK45: blue, dotted, VI: red, solid)

of spherical pendula, pure bending of an elastic rod, a spatial array of magnetic dipoles, and molecular dynamics that evolves on a sphere.

Example 1 (Double Spherical Pendulum) A double spherical pendulum is defined by two mass particles serially connected to frictionless two degree-of-freedom pivots by rigid massless links acting under a uniform gravitational potential. The dynamics of a double spherical pendulum has been studied in [2], and a variational integrator is developed in [18] by explicitly using unit length constraints.

Let the mass and the length of the pendulum be $m_1, m_2, l_1, l_2 \in \mathbb{R}$, respectively, and let $e_3 = [0, 0, 1] \in \mathbb{R}^3$ be the direction of gravity. The vector $q_1 \in \mathbb{S}^2$ represents the direction from the pivot to the first mass, and the vector $q_2 \in \mathbb{S}^2$ represents the direction from the first mass to the second mass. The inertia matrix is given by $M_{11} = (m_1+m_2)l_1^2$, $M_{12} = m_2 l_1 l_2$, and $M_{22} = m_2 l_2^2$. The gravitational potential is written as $V(q_1, q_2) = -(m_1+m_2)gl_1 e_3 \cdot q_1 - m_2 g l_2 e_3 \cdot q_2$ for the gravitational acceleration $g \in \mathbb{R}$. Substituting these into (10)–(11), the continuous equations of motion for the double spherical pendulum are given by

$$\dot{q}_1 = \omega_1 \times q_1 \quad \dot{q}_2 = \omega_2 \times q_2, \tag{38}$$

$$\begin{bmatrix} (m_1+m_2)l_1^2 I_{3\times 3} & -m_2 l_1 l_2 \hat{q}_1 \hat{q}_2 \\ -m_2 l_1 l_2 \hat{q}_2 \hat{q}_1 & m_2 l_2^2 I_{3\times 3} \end{bmatrix} \begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \end{bmatrix} = \begin{bmatrix} m_2 l_1 l_2 (\omega_2 \cdot \omega_2) \hat{q}_1 q_2 + (m_1+m_2) g l_1 \hat{q}_1 e_3 \\ m_2 l_1 l_2 (\omega_1 \cdot \omega_1) \hat{q}_2 q_1 + m_2 g l_2 \hat{q}_2 e_3 \end{bmatrix}, \tag{39}$$

which are more compact than existing equations written in terms of angles. Another nice property is that the same structure for the equations of motion is maintained for $n > 2$. Thus, it is easy to generalize these equations of motion to a triple, or more generally, a multiple-link spherical pendulum.

We compare the computational properties of the discrete equations of motion given by (29)–(32) with a 4(5)-th order variable step size Runge-Kutta method for (38)–(39). We choose $m_1 = m_2 = 1\,\text{kg}$, $l_1 = l_2 = 9.81\,\text{m}$. The initial conditions are $q_{1_0} = [0.8660,\ 0,\ 0.5]$,

$q_{2_0} = [0,\, 0,\, 1]$, $\omega_{1_0} = [-0.4330,\, 0,\, 0.75]$, $\omega_{2_0} = [0,\, 1,\, 0]$ rad/sec. The simulation time is 100 sec, and the step-size of the discrete equations of motion is $h = 0.01$. Figure 1 shows the computed total energy and the configuration manifold errors. The variational integrator preserves the total energy and the structure of $(\mathbb{S}^2)^n$ well for this chaotic motion of the double spherical pendulum. The mean total energy variation is 2.1641×10^{-5} Nm, and the mean unit length error is 8.8893×10^{-15}. But, there is a notable increase of the computed total energy for the Runge-Kutta method, where the mean variation of the total energy is 7.8586×10^{-4} Nm. The Runge-Kutta method also fails to preserve the structure of $(\mathbb{S}^2)^n$. The mean unit length error is 6.2742×10^{-5}.

Example 2 (n-body Problem on Sphere) An n-body problem on the two-sphere deals with the motion of n mass particles constrained to lie on a two-sphere, acting under a mutual potential. Let $m_i \in \mathbb{R}$ and $q_i \in \mathbb{S}^2$ be the mass and the position vector of the i-th particle, respectively. The i, j-th element of the inertia matrix is $M_{ij} = m_i$ when $i = j$, and $M_{ij} = 0$ otherwise. In [19], the following expression for the potential is introduced as an analog of a gravitational potential,

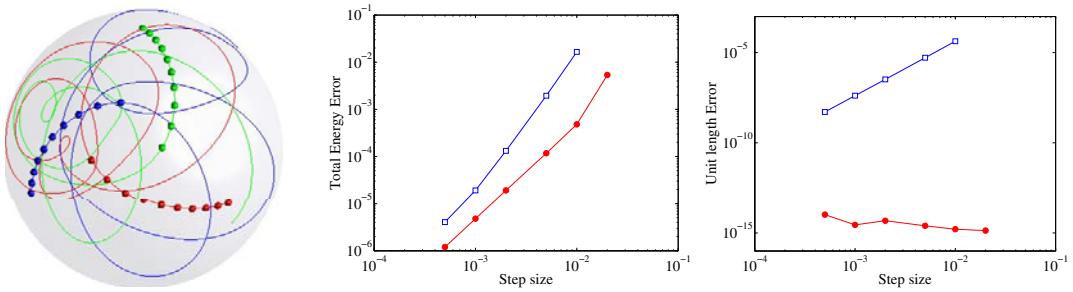
$$V(q_1, \ldots, q_n) = -\frac{\gamma}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \frac{q_i \cdot q_j}{\sqrt{1 - (q_i \cdot q_j)^2}}$$

for a constant γ. Substituting these into (7), the continuous equations of motion for the n-body problem on a sphere are given by

$$m_i \ddot{q}_i = -m_i (\dot{q}_i \cdot \dot{q}_i) q_i - q_i \times \Big(q_i \times \gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{q_j}{(1 - (q_i \cdot q_j)^2)^{3/2}} \Big) \tag{40}$$

for $i \in \{1, \ldots, n\}$.

A two-body problem on the two-sphere under this gravitational potential is studied in [20] by explicitly using unit length constraints. Here we study a three-body problem, $n = 3$. Since there are no coupling terms in the kinetic energy, we use the explicit form of the variational integrator. We compare the computational properties of the discrete equations of motion given by (36)–(37) with a 2-nd order fixed step size Runge-Kutta method for (40). We choose $m_1 = m_2 = m_3 = 1$, and $\gamma = 1$. The initial conditions are $q_{1_0} = [0,\, -1,\, 0]$, $q_{2_0} = [0,\, 0,\, 1]$, $q_{3_0} = [-1,\, 0,\, 0]$, $\omega_{1_0} = [0,\, 0,\, -1.1]$, $\omega_{2_0} = [1,\, 0,\, 0]$, and $\omega_{3_0} = [0,\, 1,\, 0]$. The simulation time is 10 sec. Figure 2 shows the computed total energy and the unit length errors for various step sizes. The total energy variations and the unit length errors for the variational integrator are smaller than those of the Runge-Kutta method for the same time step size by several orders of magnitude. For the variational integrator, the total energy error is reduced by almost 100 times from 1.1717×10^{-4} to 1.1986×10^{-6} when the step size is reduced by 10 times from 10^{-3} to 10^{-4}, which verifies the second order accuracy numerically.



(a) Trajectory of particles (b) Total energy error v.s. step size (c) Unit length error v.s. step size

Figure 2. Numerical simulation of a 3-body problem on sphere (RK2: blue, square, VI: red, circle)

Example 3 (Interconnection of Spherical Pendula) We study the dynamics of n spherical pendula connected by linear springs. Each pendulum is a mass particle connected to a frictionless two degree-of-freedom pivot by a rigid massless link acting under a uniform gravitational potential. It is assumed that all of the pivot points lie on a horizontal plane, and some pairs of pendulua are connected by linear springs at the centers of links.

Let the mass and the length of the i-th pendulum be $m_i, l_i \in \mathbb{R}$, respectively. The vector $q_i \in \mathbb{S}^2$ represents the direction from the i-th pivot to the i-th mass. The inertia matrix is given by $M_{ij} = m_i l_i^2$ when $i = j$, and $M_{ij} = 0$ otherwise. Let Ξ be a set defined such that $(i, j) \in \Xi$ if the i-th pendulum and the j-th pendulum are connected. For a connected pair $(i, j) \in \Xi$, define $\kappa_{ij} \in \mathbb{R}$ and $r_{ij} \in \mathbb{R}^3$ as the corresponding spring constant and the vector from the i-th pivot to the j-th pivot, respectively. The bases for the inertial frame are chosen such that the direction along gravity is denoted by $e_3 = [0, 0, 1] \in \mathbb{R}^3$, and the horizontal plane is spanned by $e_1 = [0, 0, 1], e_2 = [0, 1, 0] \in \mathbb{R}^3$. The potential energy is given by

$$V(q_1, \dots q_n) = -\sum_{i=1}^{n} m_i g l_i q_i \cdot e_3 + \sum_{(i,j)\in\Xi} \frac{1}{2}\kappa_{ij} \left(\left\| r_{ij} + \frac{1}{2} l_j q_j - \frac{1}{2} l_i q_i \right\| - \|r_{ij}\| \right)^2.$$

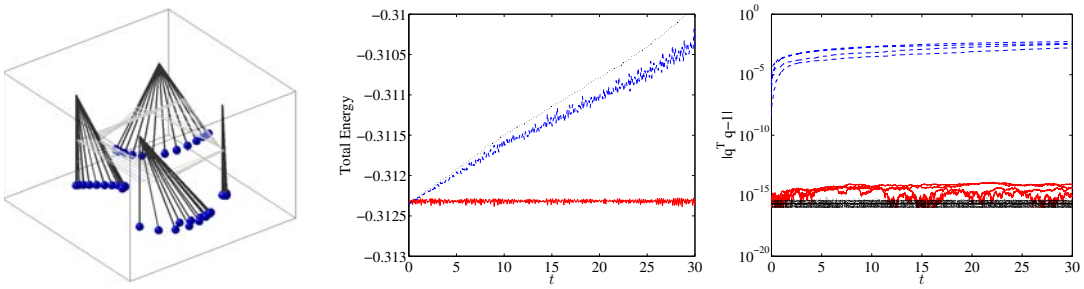
Substituting these into (9)–(10), the continuous equations of motion for the interconnection of spherical pendula are given by

$$m_i l_i^2 \dot{\omega}_i = -q_i \times \frac{\partial V}{\partial q_i} \tag{41}$$

$$\dot{q}_i = \omega_i \times q_i \tag{42}$$

for $i \in \{1, \dots, n\}$.

We compare the computational properties of the discrete equations of motion given by (36)–(37) with a 2-nd order fixed step size explicit Runge-Kutta method for (41)–(42), and the same Runge-Kutta method with reprojection; at each time step, the vectors q_{i_k} are projected onto $\mathbb{S}^2$ by using normalization.



(a) Motion of pendula (b) Computed total energy (c) Unit length error

Figure 3. Numerical simulation of a system of 4 spherical pendula (RK2: blue, dotted, RK2 with projection: black, dashed, VI: red, solid)

We choose four interconnected pendula, $n = 4$, and we assume each pendulum has the same mass and length; $m_i = 0.1\,\mathrm{kg}$, $l_i = 0.1\,\mathrm{m}$. The pendula are connected as $\Xi = \{(1,2),(2,3),(3,4),(4,1)\}$, and the corresponding spring constants and the relative vector between pivots are given by $\kappa_{12} = 10$, $\kappa_{12} = 20$, $\kappa_{12} = 30$, $\kappa_{12} = 40\,\mathrm{N/m}$, $r_{12} = -r_{34} = l_i e_1$, and $r_{23} = -r_{41} = -l_i e_2$. The initial conditions are chosen as $q_{1_0} = q_{2_0} = q_{4_0} = e_3$, $q_{3_0} = [0.4698, 0.1710, 0.8660]$, $\omega_{1_0} = [-10, 4, 0]$, and $\omega_{2_0} = \omega_{3_0} = \omega_{4_0} = 0\,\mathrm{rad/sec}$

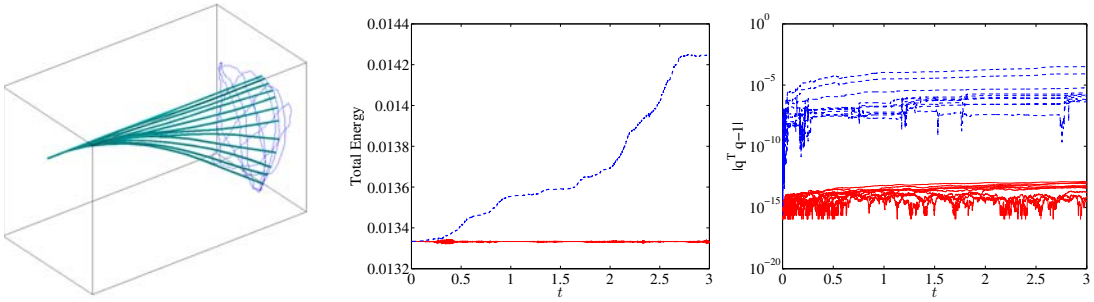
Figure 3 shows the computed total energy and the unit length errors. The variational integrator preserves the total energy and the structure of $(\mathbb{S}^2)^n$ well. The mean total energy variation is $3.6171\times10^{-5}\,\mathrm{Nm}$, and the mean unit length error is 4.2712×10^{-15}. For both Runge-Kutta methods, there is a notable increase of the computed total energy. It is interesting to see that the reprojection approach makes the total energy error worse, even though it preserves the structure of $(\mathbb{S}^2)^n$ accurately. This shows that a standard reprojection method can corrupt numerical trajectories [3, 9].

Example 4 (Pure Bending of Elastic Rod) We study the dynamics of $(n+1)$ rigid rod elements that are serially connected by rotational springs, where the 'zeroth' rod is assumed to be fixed to a wall. Thus, the configuration space is $(\mathbb{S}^2)^n$. This can be considered as a simplified dynamics model for pure non-planar bending of a thin elastic rod that is clamped at one end and free at the other end. Notably, this approach is geometrically exact, and preserves the length of the elastic rod in the presence of large displacements.

The mass and the length of the i-th rod element are denoted by $m_i, l_i \in \mathbb{R}$, respectively. The inertia matrix is given by

$$M_{ii} = \frac{1}{3} m_i l_i^2 + \sum_{k=i+1}^{n} m_i l_i^2, \quad M_{ij} = \sum_{k=\max\{i,j\}}^{n} \frac{1}{2} m_k l_k^2$$

for $i, j \in \{1, \dots n\}$ and $i \neq j$. The potential energy is composed of gravitational terms and



(a) Deformation of rod (b) Computed total energy (c) Unit length error

Figure 4. Numerical simulation of an elastic rod (RK45: blue, dotted, VI: red, solid)

elastic bending terms given by

$$V(q_1,\ldots,q_n) = -\sum_{i=1}^{n} m_i g\Big(\sum_{j=1}^{i-1} l_j q_j + \frac{1}{2} l_i q_i\Big)\cdot e_3 + \frac{1}{2}\kappa_i (1 - q_{i-1}\cdot q_i)^2,$$

where a constant vector $q_0 \in \mathbb{S}^2$ denotes the direction of the zeroth rod element fixed to a wall, and $\kappa_i \in \mathbb{R}$ denotes spring constants. The bases for the inertial frame are chosen such that the gravity direction is denoted by $e_3 = [0,0,1] \in \mathbb{R}^3$, and the horizontal plane is spanned by $e_1 = [0,0,1], e_2 = [0,1,0] \in \mathbb{R}^3$. Suppose that the total mass and length of rod are given by m, l, and each rod element has the same mass and length, i.e. $m_i = \frac{m}{n+1}$, $l_i = \frac{l}{n+1}$ for $i \in \{0,\ldots,n\}$. Substituting these into (8), the continuous equations of motion for the pure bending of an elastic rod are given by

$$\begin{bmatrix} \frac{n-2/3}{(n+1)^3} ml^2 I_{3\times 3} & -\frac{n-1}{2(n+1)^3} ml^2 \hat{q}_1\hat{q}_1 & \cdots & -\frac{1}{2(n+1)^3} ml^2 \hat{q}_1\hat{q}_1 \\ -\frac{n-1}{2(n+1)^3} ml^2 \hat{q}_2\hat{q}_2 & \frac{n-5/3}{(n+1)^3} ml^2 I_{3\times 3} & \cdots & -\frac{1}{2(n+1)^3} ml^2 \hat{q}_2\hat{q}_2 \\ \vdots & \vdots & & \vdots \\ -\frac{1}{2(n+1)^3} ml^2 \hat{q}_n\hat{q}_n & -\frac{1}{2(n+1)^3} ml^2 \hat{q}_n\hat{q}_n & \cdots & \frac{1/3}{(n+1)^3} ml^2 I_{3\times 3} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \vdots \\ \ddot{q}_n \end{bmatrix} = \begin{bmatrix} -\frac{n-2/3}{(n+1)^3} ml^2 (\dot{q}_1\cdot\dot{q}_1) q_1 + \hat{q}_1^2 \frac{\partial V}{\partial q_1} \\ -\frac{n-5/3}{(n+1)^3} ml^2 (\dot{q}_2\cdot\dot{q}_2) q_2 + \hat{q}_2^2 \frac{\partial V}{\partial q_2} \\ \vdots \\ -\frac{1/3}{(n+1)^3} ml^2 (\dot{q}_n\cdot\dot{q}_n) q_n + \hat{q}_n^2 \frac{\partial V}{\partial q_n} \end{bmatrix}, \tag{43}$$

We compare the computational properties of the discrete equations of motion given by (29)–(32) with a 4(5)-th order variable step size Runge-Kutta method for (43). We choose 10 rod elements, $n = 10$, and the total mass and the total length are $m = 55\,\mathrm{g}$, $l = 1.1,\mathrm{m}$. The spring constants are chosen as $\kappa_i = 1000\,\mathrm{Nm}$. Initially, the rod is aligned horizontally; $q_{i_0} = e_1$ for all $i \in 1,\ldots n$. The initial angular velocity for each rod element is zero except $\omega_{5_0} = [0,0,10]\,\mathrm{rad/sec}$. This represents the dynamics of the rod after an initial impact. The simulation time is 3 sec, and the step-size of the discrete equations of motion is $h = 0.0001$.

Figure 4 shows the computed total energy and the unit length errors. The variational integrator preserves the total energy and the structure of $(\mathbb{S}^2)^n$. The mean total energy

variation is 1.4310×10^{-6} Nm, and the mean unit length error is 2.9747×10^{-14}. There is a notable dissipation of the computed total energy for the Runge-Kutta method, where the mean variation of the total energy is 3.5244×10^{-4} Nm. The Runge-Kutta method also fail to preserve the structure of $(\mathbb{S}^2)^n$. The mean unit length error is 1.8725×10^{-5}.

Example 5 (Spatial Array of Magnetic Dipoles) We study dynamics of n magnetic dipoles uniformly distributed on a plane. Each magnetic dipole is modeled as a spherical compass; a thin rod magnet supported by a frictionless, two degree-of-freedom pivot acting under their mutual magnetic field. This can be considered as a simplified model for the dynamics of micromagnetic particles [21].

The mass and the length of the i-th magnet are denoted by $m_i, l_i \in \mathbb{R}$, respectively. The magnetic dipole moment of the i-th magnet is denoted by $\nu_i q_i$, where $\nu_i \in \mathbb{R}$ is the constant magnitude of the magnetic moment measured in ampere square-meters, and $q_i \in \mathbb{S}^2$ is the direction of the north pole from the pivot point. Thus, the configuration space is $(\mathbb{S}^2)^n$. The inertia matrix is given by $M_{ij} = \frac{1}{12} m_i l_i^2$ when $i = j$, and $M_{ij} = 0$ otherwise. Let $r_{ij} \in \mathbb{R}^3$ be the vector from the i-pivot point to the j-th pivot point. The mutual potential energy of the array of magnetic dipoles are given by

$$V(q_1, \ldots, q_n) = \frac{1}{2} \sum_{\substack{i,j=1 \\ j \neq i}}^{n} \frac{\mu \, \nu_i \nu_j}{4\pi \|r_{ij}\|^3} \left[(q_i \cdot q_j) - \frac{3}{\|r_{ij}\|^2} (q_i \cdot r_{ij})(q_j \cdot r_{ij}) \right],$$

where $\mu = 4\pi \times 10^{-7} \, \mathrm{N} \cdot \mathrm{A}^{-2}$ is the permeability constant. Substituting these into (9)–(10), the continuous equations of motion for the spatial array of magnetic dipoles are given by

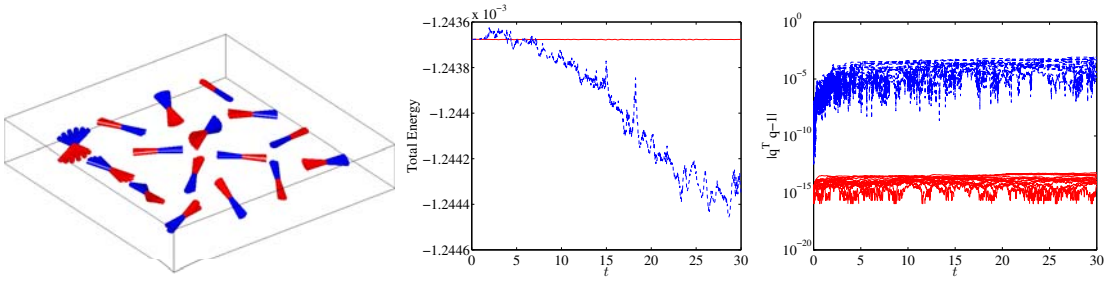
$$\frac{1}{12} m_i l_i^2 \dot{\omega}_i = -q_i \times \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\mu \, \nu_i \nu_j}{4\pi \|r_{ij}\|^3} \left[q_j - \frac{3}{\|r_{ij}\|^2} r_{ij} (q_j \cdot r_{ij}) \right], \tag{44}$$

$$\dot{q}_i = \omega_i \times q_i \tag{45}$$

for $i \in \{1, \ldots, n\}$.

We compare the computational properties of the discrete equations of motion given by (36)–(37) with a 4(5)-th order variable step size Runge-Kutta method for (44)–(45). We choose 16 magnetic dipoles, $n = 16$, and we assume each magnetic dipole has the same mass, length, and magnitude of magnetic moment; $m_i = 0.05\,\mathrm{kg}$, $l_i = 0.02\,\mathrm{m}$, $\nu_i = 0.1\,\mathrm{A} \cdot \mathrm{m}^2$. The magnetic dipoles are located at vertices of a 4×4 square grid in which the edge of a unit square has the length of $1.2 l_i$. The initial conditions are chosen as $q_{i_0} = [1, 0, 0]$, $\omega_{i_0} = [0, 0, 0]$ for all $i \in \{1, \ldots, 16\}$ except $q_{16_0} = [0.3536, 0.3536, -0.8660]$ and $\omega_{1_0} = [0, 0.5, 0]\,\mathrm{rad/sec}$.

Figure 5 shows the computed total energy and the unit length errors. The variational integrator preserves the total energy and the structure of $(\mathbb{S}^2)^n$ well. The mean total energy variation is 8.5403×10^{-10} Nm, and the mean unit length error is 1.6140×10^{-14}. There is a notable dissipation of the computed total energy for the Runge-Kutta method, where the



(a) Motion of magnetic dipoles (b) Computed total energy (c) Unit length error

Figure 5. Numerical simulation of an array of magnetic dipoles (RK45: blue, dotted, VI: red, solid)

mean variation of the total energy is 2.9989×10^{-7} Nm. The Runge-Kutta method also fail to preserve the structure of $(\mathbb{S}^2)^n$. The mean unit length error is 1.7594×10^{-4}.

Example 6 (Molecular Dynamics on a Sphere) We study molecular dynamics on $\mathbb{S}^2$. Each molecule is modeled as a particle moving on $\mathbb{S}^2$. Molecules are subject to two distinct forces: an attractive force at long range and a repulsive force at short range. Let $m_i \in \mathbb{R}$ and $q_i \in \mathbb{S}^2$ be the mass and the position vector of the i-th molecule, respectively. The i, j-th element of the inertia matrix is $M_{ij} = m_i$ when $i = j$, and $M_{ij} = 0$ otherwise. The Lennard-Jones potential is a simple mathematical model that represents the behavior of molecules [22]

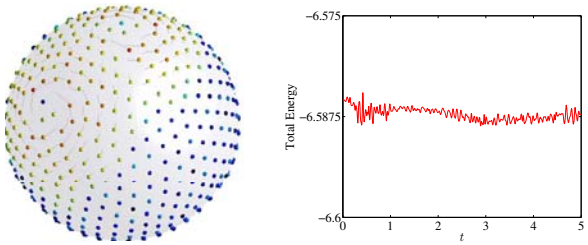
$$V(q_1, \ldots, q_n) = \frac{1}{2} \sum_{\substack{i,j=1 \\ j \neq i}}^{n} 4\epsilon \left[\left(\frac{\sigma}{\|q_i - q_j\|} \right)^{12} - \left(\frac{\sigma}{\|q_i - q_j\|} \right)^{6} \right],$$

where the first term models repulsion between molecules at short distance according to the Pauli principle, and the second term models attraction at long distance generated by van der Walls forces. The constant ϵ and σ are molecular constants; ϵ is proportional to the strength of the mutual potential, and σ characterize inter-molecular force. Substituting these into (7), the continuous equations of motion for the molecular dynamics on a sphere are given by

$$m_i \ddot{q}_i = -m_i (\dot{q}_i \cdot \dot{q}_i) q_i - q_i \times \Big(q_i \times \sum_{\substack{j=1 \\ j \neq i}}^{n} 4\epsilon \frac{q_i - q_j}{\|q_i - q_j\|} \left[\frac{12\sigma^{12}}{\|q_i - q_j\|^{13}} - \frac{6\sigma^6}{\|q_i - q_j\|^7} \right] \Big) \tag{46}$$

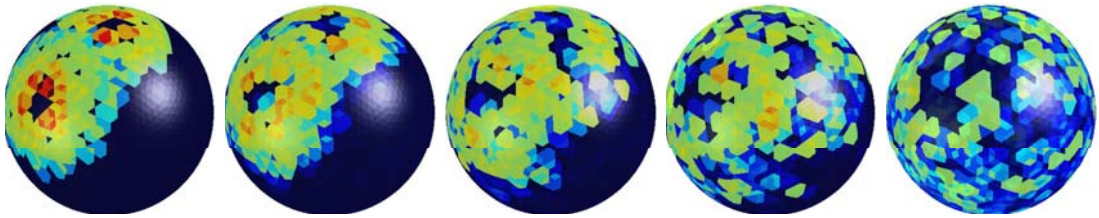
for $i \in \{1, \ldots, n\}$.

We choose 642 molecules, $n = 642$, and we assume each molecule has the same mass, $m_i = 1$. Initially, molecules are uniformly distributed on a sphere. The strength of the potential is chosen as $\epsilon = 0.01$, and the constant σ is chosen such that the inter-molecular force between neighboring molecules is close to zero. The initial velocities are modeled as two vortices separated by 30 degrees. The simulation time is 5 sec, and the step size is $h = 0.005$.



(a) Initial trajectories (b) Computed total energy

Figure 6. Numerical simulation of molecular dynamics on a sphere



(a) $t = 0$ (b) $t = 0.25$ (c) $t = 0.5$ (d) $t = 0.75$ (e) $t = 5$

Figure 7. Kinetic energy distributions over time

Trajectories of molecules and the computed total energy is shown at Figure 6. The mean deviation of the total energy is 1.8893×10^{-3}, and the mean unit length error is 5.2623×10^{-15}. In molecular dynamics simulations, macroscopic quantities such as temperature and pressure are more useful than trajectories of molecules. Figure 7 shows the change of kinetic energy distributions over time, which measures the temperature [23]; the sphere is discretized by an icosahedron with 5120 triangular faces, and the color of a face is determined by the average kinetic energy for molecules that lie within the face and within its neighboring faces. The local kinetic energy is represented by color shading of blue, green, yellow, and red colors in ascending order.

4. General Theory of Lagrangian Mechanics on Homogeneous Spaces

In this section, we will describe briefly how the approach adopted in this paper generalizes to arbitrary homogeneous spaces in both continuous and discrete time.

4.1. Continuous Time Formulation

Let G be a Lie group that acts transitively on the homogeneous space Q, with an Ad-invariant isotropy subgroup $H_q = \{h \in G | h \cdot q = q\}$. This induces an isotropy subalgebra $\mathfrak{h} \subset \mathfrak{g}$. We

then choose a principle bundle connection $\mathcal{A} : TG \to \mathfrak{h}$.

Given a Lagrangian system on the homogeneous space described by a Lagrangian $L : TQ \to \mathbb{R}$, and a distinguished point $q_0 \in Q$, we consider the horizontal lift of the second-order curve $q : [0, T] \to Q$, with respect to the initial point q_0, which we denote by $g : [0, T] \to G$. This is defined by the following system of equations,

$$q(\cdot) = g(\cdot)q_0, \qquad \dot{q}(\cdot) = \dot{g}(\cdot)q_0, \qquad \mathcal{A}(g, \dot{g}) = 0, \qquad g(0) = e.$$

We construct the lifted Lagrangian which, by an abuse of notation, is denoted $L : TG \to \mathbb{R}$, and is defined by,

$$L(g, \dot{g}) = L(gq_0, \dot{g}q_0).$$

Then, given a solution of Hamilton's principle $(q, \dot{q})$, with initial point q_0, the horizontal lift of the solution satisfies the following constrained variational principle,

$$\delta \int L(g, \dot{g})dt = 0,$$

subject to the constraint that $\mathcal{A}(g, \dot{g}) = 0$, and $g(0) = e$. The equivalence to Hamilton's principle follows by construction of the lifted Lagrangian, and the horizontal lift. It is important to note that while the lifted curve $g(\cdot)$ on the Lie group depends on the choice of connection, it induces a curve on the homogeneous space through the relation $q(\cdot) = g(\cdot)q_0$ that is independent of the choice of connection.

4.2. Discrete Time Formulation

The essential issue in the extension of the prior construction to discrete time is discretizing the constraint induced by the connection. It is possible to formally introduce a discrete connection [24], $\mathcal{A}_d : G \times G \to H_{g_1}$, and imposing the discrete horizontal curve condition which is given by $\mathcal{A}_d(g_k, g_{k+1}) = e$.

Then, the discrete horizontal lift of a discrete curve $\{q_k\}_{k=0}^N$, is given by, $\{g_k\}_{k=0}^N$, where,

$$q_k = g_k q_0, \qquad \mathcal{A}_d(g_k, g_{k+1}) = e, \qquad g_0 = e.$$

Given a discrete Lagrangian system described by a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$, we construct the lifted discrete Lagrangian $L_d : G \times G \to \mathbb{R}$ by,

$$L_d(g_k, g_{k+1}) = L_d(g_k q_0, g_{k+1} q_0).$$

Then, given the solution of the discrete Hamilton's principle, $\{q_k\}_{k=0}^N$, the horizontal lift of the discrete solution satisfies the following constrained discrete variational principle,

$$\delta \sum_{k=0}^{N-1} L_d(g_k, g_{k+1}) = 0,$$

subject to the constraint that $\mathcal{A}_d(g_k, g_{k+1}) = e$, and $g_0 = e$.

In practice, one can obtain an appropriate discretization of the constraint by introducing a retraction $\varphi_g : T_gG \to G$, which satisfies $\varphi_g(v) = g$ iff $v = 0 \in T_gG$, and $\varphi_g'\big|_0 = \mathbf{1}_{T_gG}$, and then requiring that,

$$\mathcal{A}(\varphi_{g_k}^{-1}(g_{k+1})) = 0.$$

This is equivalent to the condition in terms of a discrete connection constructed by,

$$\mathcal{A}_d(g_k, g_{k+1}) = \varphi_{g_k} \circ \mathcal{A} \circ \varphi_{g_k}^{-1}(g_{k+1}).$$

As with the continuous time case, while the lifted constrained variational principle on G yields a discrete curve $\{g_k\}_{k=0}^N$ that depends on the choice of the discrete connection, it induces a discrete curve on Q, through the relation $q_k = g_k q_0$, that is independent of the choice of connection. This is in contrast to integration algorithms for homogeneous spaces that are based on traditional Lie group methods, wherein the numerical properties of the algorithm depend critically on the choice of connection.

5. Conclusions

Euler-Lagrange equations and variational integrators are developed for Lagrangian mechanical systems evolving on $(\mathbb{S}^2)^n$ where the Lagrangian is written in a particular form given by (2). The structure of $\mathbb{S}^2$ is carefully considered to obtain global equations of motion on $(\mathbb{S}^2)^n$ without local parameterizations or explicit constraints.

In the continuous time setting, this provides a remarkably compact form of the equations of motion compared to the popular angular description. For example, it is not practical to study a triple spherical pendulum by using angles due to the complexity of the trigonometric expressions involved. The global Euler-Lagrange equations on $(\mathbb{S}^2)^n$ maintain the same compact structure for arbitrary n. In addition to being useful for the theoretical study of global dynamical properties, it is also possible to use them as a finite element model for a continuum problem as shown in Example 4. More generally, combining Lie group or homogeneous variational integrators with noncommutative harmonic analysis techniques will yield multisymplectic variational integrators for problems where either the base space or the fiber is a Lie group or homogeneous space.

The variational integrators on $(\mathbb{S}^2)^n$ preserve the geometric properties of the dynamics as well as the structure of the configuration manifold concurrently. They are symplectic, momentum preserving, and they exhibit good energy behavior for exponentially long time as they are derived from discrete Hamilton's principle. Using the characteristics of the homogeneous manifold $(\mathbb{S}^2)^n$, the discrete update map is represented by a group action of SO(3) to obtain compactly represented, and possibly explicit, numerical integrators. In particular, variational integrators on $(\mathbb{S}^2)^n$ completely avoid the singularities and complexity introduced by local parameterizations and explicit constraints.

While the numerical method presented in the paper is equivalent to the RATTLE algorithm with the Lagrange multipliers explicitly eliminated, the methodology proposed is quite general, and can be readily applied to directly construct higher-order variational integrators based on an index-reduced formulation of Lagrangian flows on $(\mathbb{S}^2)^n$.

At a more abstract level, the procedure described in this paper corresponds to considering a variational principle on a homogeneous space, and lifting this to a variational principle at the level of the Lie group. Due to the isotropy subgroup associated with the action of the Lie group on the homogeneous space, in order to uniquely define the lift of a curve on the homogeneous space to the Lie group, we need to introduce a connection that specifies a horizontal space that is complementary to tangent space of the isotropy subgroup.

This means that the lifted variational principle at the level of the Lie group is a constrained variational principle which uses a connection to specify the admissible set of variations. While the solution of the constrained variational principle depends on the choice of connection, we obtain the solution curve on the homogeneous space by the action of the solution curve on the Lie group on the initial point on the homogeneous space, and the homogeneous space curve is independent of the choice of connection.

While the construction of variational integrators on $\mathbb{S}^2$ in this paper is performed explicitly, the general methodology of lifting a variational principle on a homogeneous space to a constrained variational principle on the Lie group is quite general, and serves as the basis for extending the techniques of Lie group variational integrators to the setting of homogeneous spaces. Furthermore, the general construction is naturally related to continuous and discrete reduction, and particularly reduction by stages.

REFERENCES

1. Bendersky S, Sandler B. Investigation of spatial double pendulum: an engineering approach. *Discrete Dynamics in Nature and Society* 2006; **2006**:1–12.
2. Marsden JE, Scheurle E, Wendlandt J. Visualization of orbits and pattern evocation for the double spherical pendulum. *International Congress on Industrial and Applied Mathematics*, vol. 87, 1995.
3. Hairer E, Lubich C, Wanner G. *Geometric Numerical Integration*. Springer, 2000.
4. McLachlan R, Quispel R. *Six Lectures on The Geometric Integration of ODEs, in Foundations of Computational Mathematics*. London Mathematical Society Lecture Note, 284, Cambridge University Press, 2001; 155–210.
5. Moser J, Veselov AP. Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Communications in Mathematical Physics* 1991; **139**:217–243.
6. Marsden JE, West M. Discrete mechanics and variational integrators. *Acta Numerica*, vol. 10. Cambridge University Press, 2001; 317–514.
7. Munthe-Kaas H, Zanna A. *Numerical Integration of Differential Equations on Homogeneous Manifolds, in Foundations of Computational Mathematics*. Springer, 1997; 305–315.
8. Lewis D, Olver PJ. Geometric integration algorithms on homogeneous manifolds. *Foundations of Computational Mathematics* 2002; **2**(4):363–392, doi:10.1007/s102080010028.

9. Lewis D, Nigam N. Geometric integration on spheres and some interesting applications. *Journal of Computational and Applied Mathematics* 2003; **151**(1):141–170, doi:10.1016/S0377-0427(02)00743-4.
10. Iserles A, Munthe-Kaas H, Nørsett SP, Zanna A. Lie-group methods. *Acta Numerica*, vol. 9. Cambridge University Press, 2000; 215–365.
11. Lee T, Leok M, McClamroch NH. Lie group variational integrators for the full body problem in orbital mechanics. *Celestial Mechanics and Dynamical Astronomy* June 2007; **98**(2):121–144, doi:10.1007/s10569-007-9073-x.
12. Sanz-Serna JM. Symplectic integrators for hamiltonian problems: an overview. *Acta Numerica* 1992; :243–286.
13. Lee T, Leok M, McClamroch NH. Lie group variational integrators for the full body problem. *Computer Methods in Applied Mechanics and Engineering* May 2007; **196**:2907–2924, doi:10.1016/j.cma.2007.01.017.
14. Shuster MD. Survey of attitude representations. *Journal of the Astronautical Sciences* 1993; **41**:439–517.
15. Kelley CT. *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.
16. Hairer E. Backward analysis of numerical integrators and symplectic methods. *Annals of Numerical Mathematics* 1994; **1**(1-4):107–132. Scientific computation and differential equations (Auckland, 1993).
17. Yoshida H. Construction of high order symplectic integrators. *Physics Letters A* 1990; **150**:262–268.
18. Wendlandt JM, Marsden JE. Mechanical integrators derived from a discrete variational principle. *Physica D* 1997; **106**(3-4):223–246.
19. Kozlov VV, Harin AO. Kepler's problem in constant curvature spaces. *Celestial Mechanics and Dynamical Astronomy* 1992; **54**:393–399.
20. Hairer E, Lubich C, Wanner G. Geometric numerical integration illustrated by the Störmer-Verlet method. *Acta Numer.* 2003; **12**:399–450.
21. Cheng XZ, Jalil MBA, Lee HK. Time-quantified monte carlo algorithm for interacting spin array micromagnetic dynamics. *Physical Review B* 2006; **73**:224 438.
22. Lennard-Jones JE. Cohesion. *The Proceedings of the Physical Society* 1931; **43**:461–482.
23. Allen MP, Tildesley DJ. *Computer Simulation of Liquids*. Clarendon Press, 1987.
24. Leok M, Marsden JE, Weinstein AD. A discrete theory of connections on principal bundles 2005; ArXiv:math.DG/0508338.

NONLINEAR DYNAMICS OF THE 3D PENDULUM*

NALIN A. CHATURVEDI†, TAEYOUNG LEE‡, MELVIN LEOK§, AND N. HARRIS MCCLAMROCH¶

Abstract. A 3D pendulum consists of a rigid body, supported at a fixed pivot, with three rotational degrees of freedom. The pendulum is acted on by a gravitational force. 3D pendulum dynamics have been much studied in integrable cases that arise when certain physical symmetry assumptions are made. This paper treats the non-integrable case of the 3D pendulum dynamics when the rigid body is asymmetric and the center of mass is distinct from the pivot location. Full and reduced models of the 3D pendulum are introduced and used to study important features of the nonlinear dynamics: conserved quantities, equilibria, relative equilibria, invariant manifolds, local dynamics, and presence of chaotic motions. The paper provides a unified treatment of the 3D pendulum dynamics that includes prior results and new results expressed in the framework of geometric mechanics. These results demonstrate the rich and complex dynamics of the 3D pendulum.

Key words. Pendulum, rigid body, nonlinear dynamics, attitude, equilibria, relative equilibria, stability, chaos

AMS subject classifications. 70E17, 70K20, 70K42, 65P20

1. Introduction. Pendulum models have been a rich source of examples in nonlinear dynamics and in recent decades, in nonlinear control. The most common rigid pendulum model consists of a mass particle that is attached to one end of a massless, rigid link; the other end of the link is fixed to a pivot point that provides a rotational joint for the link and mass particle. If the link and mass particle are constrained to move within a fixed plane, the system is referred to as a planar 1D pendulum. If the link and mass particle are unconstrained, it is referred to as a spherical 2D pendulum. Planar and spherical pendulum models have been studied, for example, in [2, 13]. Spinning tops, such as the Lagrange top, also constitute another special category of pendulum models [7, 12, 18].

Pendulum models are useful for both pedagogical and research reasons. They represent physical mechanisms that can be viewed as simplified academic versions of mechanical systems that arise in, for example, robotics and spacecraft. In addition to their important role in illustrating the fundamental techniques of nonlinear dynamics, pendulum models have motivated new research directions and applications in nonlinear dynamics.

This paper treats 3D pendulum models, some of which were studied by Euler; see [1, 7] and references therein for historical background. Physically, these models describe the dynamics of a rigid body, supported at a fixed, frictionless, pivot point that has three rotational degrees of freedom; it is acted on by a uniform gravity force. Control and disturbance forces and moments are ignored in this development.

This paper arose out of our continuing research on a laboratory facility, referred to as the Triaxial Attitude Control Testbed (TACT). The TACT was constructed to provide a testbed for physical experiments on attitude dynamics and attitude control. The most important feature of the TACT design is that it is supported by a three-dimensional air bearing that serves as an ideal frictionless pivot, allowing nearly unrestricted three degrees of rotation. The TACT has been described in several prior publications [5, 11]. Issues of nonlinear dynamics for the TACT have been treated in [10, 11]; issues of nonlinear control for the

*NC, TL and NHM have been supported in part by NSF Grant ECS-0244977 and CMS-0555797. TL and ML have been supported in part by NSF Grant DMS-0504747 and DMS-0726263.

†Robert Bosch, LLC, Palo Alto, California (nalin.chaturvedi@us.bosch.com).

‡Department of Aerospace Engineering, The University of Michigan, Ann Arbor, Michigan 48109-2140 (tylee@umich.edu).

§Department of Mathematics, Purdue University, West Lafayette, IN 47907-2067 (mleok@math.purdue.edu).

¶Department of Aerospace Engineering, The University of Michigan, Ann Arbor, Michigan 48109-2140 (nhm@umich.edu).

TACT have been treated in [21]. The present paper is partly motivated by the realization that the TACT is, in fact, a physical realization of a 3D pendulum.

2. Description of the 3D Pendulum. A 3D pendulum is a rigid body supported by a fixed, frictionless pivot, acted on by gravitational forces. The supporting pivot allows three degrees of rotational freedom of the pendulum. Uniform, constant gravity is assumed. The terminology 3D pendulum refers to the fact that the pendulum is a rigid body with three spatial dimensions and the pendulum has three rotational degrees of freedom.

Two reference frames are introduced. An inertial reference frame has its origin at the pivot; the first two axes lie in the horizontal plane and the third axis is vertical in the direction of gravity. A reference frame fixed to the pendulum body is also introduced. The origin of this body-fixed frame is located at the pivot. In the body-fixed frame, the inertia tensor of the pendulum is constant. This inertia tensor can be computed from the traditional inertia tensor of a translated frame whose origin is located at the center of mass of the pendulum using the parallel axis theorem. Since the origin of the body-fixed frame is located at the pivot, principal axes with respect to this frame can be defined for which the inertia tensor is diagonal. Note that the center of mass of the 3D pendulum does not necessarily lie on one of the principal axes. Throughout this paper, we assume that the 3D pendulum is asymmetric, that is its three principal moments of inertia are distinct, and the center of mass of the 3D pendulum is not at the pivot location.

Rotation matrices are used to describe the attitude of the rigid 3D pendulum. A rotation matrix maps a representation of vectors expressed in the body-fixed frame to a representation expressed in the inertial frame. Rotation matrices provide global representations of the attitude of the pendulum, which is why they are utilized here. Other attitude representations, such as exponential coordinates, quaternions, or Euler angles, can also be used following standard descriptions, but each of the representations has a disadvantage of introducing an ambiguity or coordinate singularity. In this paper, the attitude configuration of the pendulum is a rotation matrix R in the special orthogonal group $SO(3)$ defined as

$$SO(3) \triangleq \{R \in \mathbb{R}^{3\times 3} : RR^{\mathrm{T}} = I_{3\times 3}, \det(R) = 1\}.$$

The associated angular velocity of the pendulum, expressed in the body-fixed frame, is denoted by ω in $\mathbb{R}^3$. The constant inertia tensor of the rigid body pendulum, in the body-fixed frame, is denoted by the symbol J. The constant body-fixed vector from the pivot to the center of mass of the pendulum is denoted by $\rho = [\rho_1\ \rho_2\ \rho_3]^{\mathrm{T}}$. We assume without loss of generality that the inertia tensor is diagonal, i.e. $J = \mathrm{diag}(J_1, J_2, J_3)$, where $J_1 > J_2 > J_3 > 0$ and $\rho \neq 0$. The symbol g denotes the constant acceleration due to gravity.

Three categories of 3D pendulum models are subsequently introduced and studied. The "full" dynamics of the 3D pendulum are based on Euler's equations that include the gravity moment and the rotational kinematics, expressed in terms of the angular velocity and a rotation matrix; this model describes the dynamics that evolves on $TSO(3)$. Since the gravity moment depends solely on the direction of gravity in the body-fixed frame, it is possible to obtain a reduced model expressed in terms of the angular velocity and a unit vector that defines the direction of gravity in the body-fixed frame; this model describes the dynamics that evolve on $TSO(3)/S^1$, and corresponds to the case of Lagrange–Poincaré reduction [8]. Since there is a symmetry action given by a rotation about the gravity direction, Lagrange–Routh reduction [20] leads to a reduced model expressed in terms of the unit vector that defines the direction of gravity in the body-fixed frame and its derivative; this model describes the dynamics that evolve on TS^2. Each of these 3D pendulum models provides special insight into the nonlinear dynamics. We develop each of these models in this paper,

and we investigate the features of the nonlinear dynamics, namely invariants, equilibria, and stability, for each model.

3. 3D Pendulum Dynamics on $TSO(3)$. The dynamics of the 3D pendulum are given by the Euler equation that includes the moment due to gravity:

$$J\dot{\omega} = J\omega \times \omega + mg\rho \times R^{\mathrm{T}} e_3. \tag{3.1}$$

The rotational kinematics equations are

$$\dot{R} = R\widehat{\omega}. \tag{3.2}$$

Equations (3.1) and (3.2) define the full dynamics of a rigid pendulum on the tangent bundle $TSO(3)$. Throughout the paper, $e_1 = (1,0,0)^{\mathrm{T}}$, $e_2 = (0,1,0)^{\mathrm{T}}$, $e_3 = (0,0,1)^{\mathrm{T}}$. In the inertial frame, e_1 and e_2 are assumed to lie in a horizontal plane, while e_3 is assumed to be in the direction of gravity; consequently $R^{\mathrm{T}}e_3$ is the direction of gravity in the body-fixed frame. The cross product notation $a \times b$ for vectors a and b in $\mathbb{R}^3$ is

$$a \times b = [a_2b_3 - a_3b_2,\ a_3b_1 - a_1b_3,\ a_1b_2 - a_2b_1] = \widehat{a}b, \tag{3.3}$$

where, the skew-symmetric matrix $\widehat{a}$ is defined as

$$\widehat{a} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \tag{3.4}$$

3.1. Integrals of the 3D Pendulum Dynamics. There are two conserved quantities, or integrals of motion, for the 3D pendulum. First, the total energy, which is the sum of the rotational kinetic energy and the gravitational potential energy, is conserved. In addition, there is a symmetry corresponding to rotations about the gravity direction through the pivot. This symmetry leads to conservation of the component of angular momentum about the gravity direction. These two well known results are summarized as follows.

Proposition 1. *The total energy*

$$E = \frac{1}{2}\,\omega^{\mathrm{T}} J\omega - mg\rho^{\mathrm{T}} R^{\mathrm{T}} e_3,$$

and the component of the angular momentum vector about the vertical axis through the pivot

$$h = \omega^{\mathrm{T}} J R^{\mathrm{T}} e_3.$$

are each constant along motions of the 3D pendulum.

Proof. The proof follows by showing that the time derivative of the total energy and the time derivative of the angular momentum component about the vertical axis are each identically zero. We use (3.1) and (3.2) to compute the derivatives, yielding

$$\dot{E} = \omega^{\mathrm{T}} J\dot{\omega} - mg\rho^{\mathrm{T}}\dot{R}^{\mathrm{T}} e_3 = mg\,\omega^{\mathrm{T}}(\rho \times R^{\mathrm{T}} e_3) + mg\rho^{\mathrm{T}}(\omega \times R^{\mathrm{T}} e_3) = 0,$$

and similarly,

$$\begin{aligned} \dot{h} &= \dot{\omega}^{\mathrm{T}} J R^{\mathrm{T}} e_3 + \omega^{\mathrm{T}} J \dot{R}^{\mathrm{T}} e_3, \\ &= (J\omega \times \omega)^{\mathrm{T}}(R^{\mathrm{T}} e_3) + (mg\rho \times R^{\mathrm{T}} e_3)^{\mathrm{T}}(R^{\mathrm{T}} e_3) - \omega^{\mathrm{T}} J(\omega \times R^{\mathrm{T}} e_3), \\ &= (J\omega)^{\mathrm{T}}(\omega \times R^{\mathrm{T}} e_3) - \omega^{\mathrm{T}} J(\omega \times R^{\mathrm{T}} e_3) = 0. \end{aligned}$$

□

The assumption that the 3D pendulum is asymmetric and the center of mass is not located at the pivot implies that there are no nontrivial integrals of motion other than those given in the above proposition. That is, the 3D pendulum dynamics are not integrable [12].

We do not further consider the integrable cases of the 3D pendulum dynamics, which include the free rigid body, the Lagrange top, the Kovalevskaya top, and the Goryachev-Chaplygin top. These integrable case have been extensively treated in the existing literature [1, 7, 12, 16, 18, 19].

Constant values of the total energy of the 3D pendulum and constant values of the component of angular momentum of the 3D pendulum in the direction of gravity define invariant manifolds of the 3D pendulum. These invariant manifolds are important characterizations of the 3D pendulum dynamics.

3.2. Equilibria of the 3D Pendulum. To further understand the dynamics of the 3D pendulum, we study its equilibria. Equating the RHS of equations (3.1) and (3.2) to zero yields conditions for an equilibrium $(R_\mathrm{e}, \omega_\mathrm{e})$:

$$J\omega_\mathrm{e} \times \omega_\mathrm{e} + mg\rho \times R_\mathrm{e}^\mathrm{T} e_3 = 0, \tag{3.5}$$

$$R_\mathrm{e}\widehat{\omega}_\mathrm{e} = 0. \tag{3.6}$$

Since $R_\mathrm{e} \in SO(3)$ is non-singular, and $\widehat{\cdot}\colon \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ is a linear injection, $R_\mathrm{e}\widehat{\omega}_\mathrm{e} = 0$ if and only if $\omega_\mathrm{e} = 0$. Substituting $\omega_\mathrm{e} = 0$ in (3.5), we obtain

$$\rho \times R_\mathrm{e}^\mathrm{T} e_3 = 0. \tag{3.7}$$

Hence,

$$R_\mathrm{e}^\mathrm{T} e_3 = \frac{\rho}{\|\rho\|}, \tag{3.8}$$

or

$$R_\mathrm{e}^\mathrm{T} e_3 = -\frac{\rho}{\|\rho\|}. \tag{3.9}$$

An attitude R_e is an equilibrium attitude if and only if the direction of gravity resolved in the body-fixed frame, $R_\mathrm{e}^\mathrm{T} e_3$, is collinear with the vector ρ. If $R_\mathrm{e}^\mathrm{T} e_3$ is in the same direction as the vector ρ, then $(R_\mathrm{e}, 0)$ is a hanging equilibrium of the 3D pendulum; if $R_\mathrm{e}^\mathrm{T} e_3$ is in the opposite direction to the vector ρ, then $(R_\mathrm{e}, 0)$ is an inverted equilibrium of the 3D pendulum.

Thus, if R_e defines an equilibrium attitude for the 3D pendulum, then a rotation of the 3D pendulum about the gravity vector by an arbitrary angle is also an equilibrium. Consequently, in $TSO(3)$ there are two disjoint equilibrium manifolds of the 3D pendulum. The manifold corresponding to the first case in the above description is referred to as the hanging equilibrium manifold, since the center of mass is always below the pivot. The manifold corresponding to the second case in the above description is referred to as the inverted equilibrium manifold, since the center of mass is always above the pivot.

Following equations (3.8) and (3.9) and the discussion above, we define

$$[R]_h \triangleq \left\{ R \in SO(3) : R^\mathrm{T} e_3 = \frac{\rho}{\|\rho\|} \right\}, \tag{3.10}$$

$$[R]_i \triangleq \left\{ R \in SO(3) : R^\mathrm{T} e_3 = -\frac{\rho}{\|\rho\|} \right\}, \tag{3.11}$$

as the *hanging attitude manifold* and the *inverted attitude manifold*, respectively.

From (3.8) and (3.9),

$$\mathsf{H} \triangleq \Big\{(R,0) \in TSO(3) : R \in [R]_h\Big\}, \tag{3.12}$$

is the manifold of hanging equilibria and

$$\mathtt{I} \triangleq \Big\{(R,0) \in TSO(3) : R \in [R]_i\Big\}, \tag{3.13}$$

is the manifold of inverted equilibria, and these two equilibrium manifolds are clearly distinct.

3.3. Local Analysis of the 3D Pendulum near an Equilibrium. A perturbation from a hanging equilibrium $(R_{\mathrm{e}}, 0)$ of the 3D pendulum can be expressed using an exponential representation and a perturbation parameter ε. Let $R^\varepsilon(t)$ and $\omega^\varepsilon(t)$ represent the perturbed solution, corresponding to initial conditions $R^\varepsilon(0) = R_{\mathrm{e}} \exp \varepsilon\widehat{\delta\Theta}$ and $\omega^\varepsilon(0) = \varepsilon\delta\omega$, where $\delta\Theta, \delta\omega \in \mathbb{R}^3$ are constant vectors. Note that if $\varepsilon = 0$ then, $(R^0(0), \omega^0(0)) = (R_{\mathrm{e}}, 0)$ and hence

$$(R^0(t), \omega^0(t)) \equiv (R_{\mathrm{e}}, 0) \tag{3.14}$$

for all time $t \in \mathbb{R}$.

The perturbed solution satisfies the perturbed equations of motion for the 3D pendulum:

$$J\dot{\omega}^\varepsilon = J\omega^\varepsilon \times \omega^\varepsilon + mg\rho \times (R^\varepsilon)^{\mathrm{T}} e_3, \tag{3.15}$$

$$\dot{R}^\varepsilon = R^\varepsilon \widehat{\omega^\varepsilon}. \tag{3.16}$$

Next, we differentiate both sides with respect to ε and substitute $\varepsilon = 0$, yielding

$$J\dot{\omega}^0_\varepsilon = J\omega^0_\varepsilon \times \omega^0 + J\omega^0 \times \omega^0_\varepsilon + mg\rho \times (R^0_\varepsilon)^{\mathrm{T}} e_3, \tag{3.17}$$

$$\dot{R}^0_\varepsilon = R^0_\varepsilon \widehat{\omega}^0 + R^0 \widehat{\omega}^0_\varepsilon, \tag{3.18}$$

where the subscripts denote derivatives. Substituting $R^0 = R_{\mathrm{e}}$ and $\omega^0 = 0$ from equation (3.14) into equations (3.17) and (3.18) yields

$$J\dot{\omega}^0_\varepsilon = mg\rho \times (R^0_\varepsilon)^{\mathrm{T}} e_3, \tag{3.19}$$

$$\dot{R}^0_\varepsilon = R_{\mathrm{e}} \widehat{\omega}^0_\varepsilon. \tag{3.20}$$

Now define perturbation variables $\Delta\omega \triangleq \omega^0_\varepsilon$ and $\widehat{\Delta\Theta} \triangleq R_{\mathrm{e}}^{\mathrm{T}} R^0_\varepsilon$. It can be shown that $\Delta\omega = \Delta\dot{\Theta}$. Thus, (3.19) and (3.20) can be written as

$$J\Delta\ddot{\Theta} - \frac{mg\widehat{\rho}^{\,2}}{\|\rho\|}\Delta\Theta = 0. \tag{3.21}$$

Note that equation (3.21) can be interpreted as defining a mechanical system with mass matrix J, stiffness matrix $-\dfrac{mg\widehat{\rho}^{\,2}}{\|\rho\|}$, but no damping. Since $\widehat{\rho}^{\,2}$ is a negative-semidefinite matrix with two negative eigenvalues and one zero eigenvalue, the stiffness matrix is positive-semidefinite with two positive eigenvalues and one zero eigenvalue. The zero eigenvalue corresponds to rotations about the vertical axis, for which gravity has no influence. To see this more explicitly, we next perform a change of variables.

Since $\widehat{\rho}^{\,2}$ is a rank 2, symmetric, negative-semidefinite matrix, it follows from [3, 4] that one can simultaneously diagonalize J and $\widehat{\rho}^{\,2}$. Thus, there exists a non-singular matrix M such that $J = MM^{\mathrm{T}}$ and

$-\dfrac{mg}{\|\rho\|}\widehat{\rho}^{\,2} = M\Lambda M^{\mathrm{T}}$, where Λ is a diagonal matrix. Denote $\Lambda = \mathrm{diag}(mgl_1, mgl_2, 0)$, where l_1 and l_2 are positive. Define $x \triangleq M^{\mathrm{T}}\Delta\Theta$. Then expressing $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, equation (3.21) can be written as

$$\ddot{x}_1 + mgl_1 x_1 = 0, \tag{3.22}$$

$$\ddot{x}_2 + mgl_2 x_2 = 0, \tag{3.23}$$

$$\ddot{x}_3 = 0. \tag{3.24}$$

Thus, the variable x_3 represents a perturbation in attitude of the 3D pendulum that corresponds to a rotation about the vertical axis.

Due to the presence of imaginary and zero eigenvalues of the linearized equations, no conclusion can be made about the stability of the hanging equilibrium or the hanging equilibrium manifold of the 3D pendulum. Indeed, the local structure of trajectories in an open neighborhood of the equilibrium is that of a center manifold; there are no stable or unstable manifolds. We next show that the hanging equilibrium manifold of the 3D pendulum is Lyapunov stable.

Proposition 2. *Consider the 3D pendulum model described by equations (3.1) and (3.2). Then, the hanging equilibrium manifold* H *given by (3.12) is stable in the sense of Lyapunov.*

Proof. Consider the following positive-semidefinite function on $TSO(3)$

$$V(R, \omega) = \frac{1}{2}\,\omega^{\mathrm{T}} J\omega + mg(\|\rho\| - \rho^{\mathrm{T}} R^{\mathrm{T}} e_3). \tag{3.25}$$

Note that $V(R, 0) = 0$ for all $(R, \omega) \in \mathsf{H}$ and $V(R, \omega) > 0$ elsewhere. Furthermore, the derivative along a solution of equations (3.1) and (3.2) is given by

$$\begin{aligned}
\dot{V}(R, \omega) &= \omega^{\mathrm{T}} J\dot{\omega} - mg\rho^{\mathrm{T}}\dot{R}^{\mathrm{T}} e_3, \\
&= \omega^{\mathrm{T}}(J\omega \times \omega + mg\rho \times R^{\mathrm{T}} e_3) - mg\rho^{\mathrm{T}}(-\widehat{\omega} R^{\mathrm{T}} e_3), \\
&= mg\Big[\omega^{\mathrm{T}}(\rho \times R^{\mathrm{T}} e_3) + \rho^{\mathrm{T}}(\omega \times R^{\mathrm{T}} e_3)\Big] = 0.
\end{aligned}$$

Thus, $\dot{V}$ is negative-semidefinite on $TSO(3)$. Also, every sublevel set of the function V is compact. Therefore, the hanging equilibrium manifold H is Lyapunov stable. □

Similarly, one can linearize the 3D pendulum dynamics about an equilibrium in the inverted equilibrium manifold. Expressing this linearization in terms of $(x_1, x_2, x_3) \in \mathbb{R}^3$ as in equations (3.22)–(3.24), it can be shown that the linearization of the 3D pendulum about an inverted equilibrium can be written as

$$\ddot{x}_1 - mgl_1 x_1 = 0, \tag{3.26}$$

$$\ddot{x}_2 - mgl_2 x_2 = 0, \tag{3.27}$$

$$\ddot{x}_3 = 0. \tag{3.28}$$

The linearization of the 3D pendulum about an inverted equilibrium results in a system that has two positive eigenvalues, two negative eigenvalues and two zero eigenvalues. Thus, the inverted equilibrium has a two-dimensional stable manifold, a two-dimensional unstable manifold and a two-dimensional center manifold. It is clear that due to the presence of the two positive eigenvalues, the inverted equilibrium is unstable.

Proposition 3. *Consider the 3D pendulum model described by equations (3.1) and (3.2). Then, each equilibrium in the inverted equilibrium manifold* $\mathtt{I}$ *given by (3.13) is unstable.*

4. Lagrange–Poincaré Reduced 3D Pendulum Dynamics on $TSO(3)/S^1$. The equations of motion (3.1) and (3.2) for the 3D pendulum are viewed as a model for the dynamics on the tangent bundle $TSO(3)$ [6]; these are referred to as the full equations of motion since they characterize the full attitude dynamics of the 3D pendulum. Since there is a rotational symmetry corresponding to the group of rotations about the vertical axis through the pivot and an associated conserved angular momentum component, it is possible to obtain a lower dimensional reduced model for the rigid pendulum. This Lagrange–Poincaré reduction is based on the fact that the dynamics and kinematics equations can be written in terms of the reduced attitude vector $\Gamma = R^{\mathrm{T}} e_3 \in S^2$, which is the unit vector that expresses the gravity direction in the body-fixed frame [18].

Specifically, let Φ_θ denote the group action of S^1 on $SO(3)$, given by $\Phi_\theta : S^1 \times SO(3) \to SO(3)$, $\Phi_\theta(R) = \exp(\theta \widehat{e}_3) R$. This induces an equivalence class by identifying elements of $SO(3)$ that belong to the same orbit; explicitly, for $R_1, R_2 \in SO(3)$, we write $R_1 \sim R_2$ if there exists a $\theta \in S^1$ such that $\Phi_\theta(R_1) = R_2$. The *orbit space* $SO(3)/S^1$ is the set of equivalence classes,

$$[R] \triangleq \{\Phi_\theta(R) \in SO(3) : \theta \in S^1\}. \tag{4.1}$$

For this equivalence relation, it is easy to see that $R_1 \sim R_2$ if and only if $R_1^{\mathrm{T}} e_3 = R_2^{\mathrm{T}} e_3$ and hence the equivalence class in (4.1) can alternately, be expressed as

$$[R] \triangleq \{R_s \in SO(3) : R_s^{\mathrm{T}} e_3 = R^{\mathrm{T}} e_3\}. \tag{4.2}$$

Thus, for each $R \in SO(3)$, $[R]$ can be identified with $\Gamma = R^{\mathrm{T}} e_3 \in S^2$ and hence $SO(3)/S^1 \cong S^2$. This group action induces a projection $\Pi : SO(3) \to SO(3)/S^1 \cong S^2$ given by $\Pi(R) = R^T e_3$.

Proposition 4 ([9]). *The dynamics of the 3D pendulum given by equations (3.1) and (3.2) induce a flow on the quotient space $TSO(3)/S^1$, through the projection $\pi : TSO(3) \to TSO(3)/S^1$ defined as $\pi(R, \Omega) = (R^T e_3, \Omega)$, given by the dynamics*

$$J\dot{\omega} = J\omega \times \omega + mg\rho \times \Gamma, \tag{4.3}$$

and the kinematics for the reduced attitude

$$\dot{\Gamma} = \Gamma \times \omega. \tag{4.4}$$

Furthermore, $TSO(3)/S^1$ is identified with $S^2 \times \mathbb{R}^3$.

Equations (4.3) and (4.4) are expressed in a non-canonical form; they are referred to as the Lagrange–Poincaré reduced attitude dynamics of the 3D pendulum on $S^2 \times \mathbb{R}^3$.

4.1. Integrals of the Lagrange–Poincaré Reduced Model. In a previous section, we presented two integrals of motion for the full model of the 3D pendulum. In this section we summarize similar well known results on integrals for the Lagrange–Poincaré reduced model of the 3D pendulum.

Proposition 5. *The total energy*

$$E = \frac{1}{2}\, \omega^{\mathrm{T}} J \omega - mg\rho^{\mathrm{T}} \Gamma, \tag{4.5}$$

and the component of the angular momentum vector about the vertical axis through the pivot

$$h = \omega^{\mathrm{T}} J\, \Gamma.$$

are each constant along trajectories of the Lagrange–Poincaré reduced model of the 3D pendulum given by equations (4.3) and (4.4).

Constant values of the total energy of the 3D pendulum and constant values of the component of angular momentum of the 3D pendulum in the direction of gravity define invariant manifolds of the Lagrange–Poincaré reduced model of the 3D pendulum. These invariant manifolds are important characterizations of the 3D pendulum dynamics.

4.2. Equilibria of the Lagrange–Poincaré Reduced Model. We study the equilibria of the Lagrange–Poincaré reduced equations of motion of the 3D pendulum given by equations (4.3) and (4.4). Equating the RHS of equations (4.3) and (4.4) to zero yields conditions for an equilibrium $(\Gamma_{\mathrm{e}}, \omega_{\mathrm{e}})$:

$$J\omega_{\mathrm{e}} \times \omega_{\mathrm{e}} + mg\rho \times \Gamma_{\mathrm{e}} = 0, \tag{4.6}$$

$$\Gamma_{\mathrm{e}} \times \omega_{\mathrm{e}} = 0. \tag{4.7}$$

Equation (4.7) implies that $\omega_{\mathrm{e}} = k\,\Gamma_{\mathrm{e}}$, where $k \in \mathbb{R}$. Substituting this into (4.6) yields

$$k^2\, J\Gamma_{\mathrm{e}} \times \Gamma_{\mathrm{e}} + mg\rho \times \Gamma_{\mathrm{e}} = 0. \tag{4.8}$$

If $k = 0$, then $\omega_{\mathrm{e}} = 0$ gives two *static* equilibria of the 3D pendulum; if $k \neq 0$ then relative equilibria [14] of the 3D pendulum are obtained.

We assume without loss of generality that the inertia tensor is diagonal, i.e. $J = \mathrm{diag}(J_1, J_2, J_3)$, where $J_1 > J_2 > J_3 > 0$ and $\rho \neq 0$. The following result describes the generic equilibria structure of the Lagrange–Poincaré reduced equations without further assumptions.

Proposition 6. *Consider the Lagrange-Poincaré model of the 3D pendulum given by equations (4.3) and (4.4). The Lagrange-Poincaré model on $S^2 \times \mathbb{R}^3$ has the following equilibria and relative equilibria:*

1. *There is a hanging equilibrium: $\left(\frac{\rho}{\|\rho\|}, 0\right) \in S^2 \times \mathbb{R}^3$,*
2. *There is an inverted equilibrium: $\left(-\frac{\rho}{\|\rho\|}, 0\right) \in S^2 \times \mathbb{R}^3$,*
3. *There are two relative equilibria in $S^2 \times \mathbb{R}^3$:*

$$\left(-\frac{J^{-1}\rho}{\|J^{-1}\rho\|}, \sqrt{\frac{mg}{\|J^{-1}\rho\|}} J^{-1}\rho\right), \left(-\frac{J^{-1}\rho}{\|J^{-1}\rho\|}, -\sqrt{\frac{mg}{\|J^{-1}\rho\|}} J^{-1}\rho\right), \tag{4.9}$$

4. *There are one-dimensional relative equilibria manifolds in $S^2 \times \mathbb{R}^3$ described by the parameterizations:*

$$\left(-\frac{n_\alpha}{\|n_\alpha\|}, \sqrt{\frac{mg}{\|n_\alpha\|}} n_\alpha\right), \left(-\frac{n_\alpha}{\|n_\alpha\|}, -\sqrt{\frac{mg}{\|n_\alpha\|}} n_\alpha\right), \quad \textit{for } \alpha \in \mathcal{L}_i, i \in \{1,2,3,4,5\}, \tag{4.10}$$

where $n_\alpha = (J - \frac{1}{\alpha} I_{3\times 3})^{-1}\rho \in \mathbb{R}^3$ and the intervals of the reals are defined by

$$\mathcal{L}_1 = (-\infty, 0),\ \mathcal{L}_2 = (0, \tfrac{1}{J_1}),\ \mathcal{L}_3 = (\tfrac{1}{J_1}, \tfrac{1}{J_2}),\ \mathcal{L}_4 = (\tfrac{1}{J_2}, \tfrac{1}{J_3}),\ \mathcal{L}_5 = (\tfrac{1}{J_3}, \infty).$$

Proof. From equation (4.8), an equilibrium $(\Gamma_{\mathrm{e}}, \omega_{\mathrm{e}})$ satisfies

$$k^2 J\Gamma_{\mathrm{e}} + mg\rho = k_1\Gamma_{\mathrm{e}}, \tag{4.11}$$

for some constant $k_1 \in \mathbb{R}$. We solve this equation to obtain the expression for an equilibrium attitude Γ_{e} for two cases; when $k_1 = 0$ and when $k_1 \neq 0$. The corresponding value of the constant k yields the expression

for the equilibrium angular velocity as $\omega_{\mathrm{e}} = k\Gamma_{\mathrm{e}}$. Recall that the inertia tensor J is diagonal with distinct diagonal entries and $\rho \neq 0$.

Equilibria 3: Suppose $k_1 = 0$. It follows that $k \neq 0$ from (4.11). Thus, we have $\Gamma_{\mathrm{e}} = -\frac{mg}{k^2} J^{-1}\rho$. Since $\|\Gamma_{\mathrm{e}}\| = 1$, we obtain $k^2 = mg\|J^{-1}\rho\|$, which gives (4.9).

Equilibria 1, 2, 4: Suppose $k_1 \neq 0$. If $k = 0$, (4.11) yields the hanging and the inverted equilibrium. Suppose $k \neq 0$. Define $\alpha = \frac{k^2}{k_1} \in \mathbb{R}\backslash\{0\}$, and $v = k_1\Gamma_{\mathrm{e}} \in \mathbb{R}^3$. Then, (4.11) can be written as

$$(\alpha J - I_{3\times 3})v = -mg\rho. \tag{4.12}$$

Note that for $\alpha \in \mathbb{R}\backslash\{0, \frac{1}{J_1}, \frac{1}{J_2}, \frac{1}{J_3}\}$ the matrix $(J - \frac{1}{\alpha} I_{3\times 3})$ is invertible. Then, (4.12) can be solved to obtain $v = -\frac{mg}{\alpha}(J - \frac{1}{\alpha} I_{3\times 3})^{-1}\rho$. Since $\|\Gamma_{\mathrm{e}}\| = 1$, we have $\|v\| = \|k_1\Gamma_{\mathrm{e}}\| = |k_1|$. We consider two sub-cases; when $k_1 > 0$, and $k_1 < 0$.

If $k_1 > 0$, we have $k_1 = \|v\|$ and $\alpha > 0$. Thus, we obtain the expression for values of the equilibria attitudes as

$$\Gamma_{\mathrm{e}} = \frac{v}{\|v\|} = \frac{-\frac{1}{\alpha} n_\alpha}{|-\frac{1}{\alpha}|\,\|n_\alpha\|} = -\frac{n_\alpha}{\|n_\alpha\|}. \tag{4.13}$$

where $n_\alpha = (J - \frac{1}{\alpha} I_{3\times 3})^{-1}\rho \in \mathbb{R}^3$. Since $k^2 = \alpha k_1 = \alpha\,\|v\|$, we obtain the expression for values of the equilibria angular velocities as

$$\omega_{\mathrm{e}} = k\Gamma_{\mathrm{e}} = \mp\sqrt{\alpha\,\|v\|}\,\frac{n_\alpha}{\|n_\alpha\|} = \mp\sqrt{\frac{mg}{\|n_\alpha\|}}\, n_\alpha. \tag{4.14}$$

Thus, (4.13) and (4.14) correspond to the families of equilibria given by (4.10) for $\alpha > 0$.

Similarly, if $k_1 < 0$, we have $k_1 = -\,\|v\|$, $\alpha < 0$, and $k^2 = -\alpha\,\|v\|$. Thus, we obtain the expression for the values of the relative equilibria as

$$\Gamma_{\mathrm{e}} = -\frac{v}{\|v\|} = \frac{\frac{1}{\alpha} n_\alpha}{|-\frac{1}{\alpha}|\,\|n_\alpha\|} = -\frac{n_\alpha}{\|n_\alpha\|}, \quad \omega_{\mathrm{e}} = \mp\sqrt{\frac{mg}{\|n_\alpha\|}}\, n_\alpha, \tag{4.15}$$

which corresponds to the families of equilibria given by (4.10) for $\alpha < 0$.

Equation (4.12) for $\alpha \in \mathbb{R}$ characterizes the equilibria of (4.3) and (4.4). Condition (4.10) presents solutions of (4.12) for all $\alpha \in \mathbb{R}\backslash\{0, \frac{1}{J_1}, \frac{1}{J_2}, \frac{1}{J_3}\}$. The parameter value $\alpha = 0$ yields the hanging and the inverted equilibria. □

The first three statements in Proposition 6 are self-explanatory. The fourth statement describes parameterizations for one-dimensional relative equilibria manifolds; these parameterizations are expressed in terms of the real parameter α that lies in one of the four defined intervals. The intervals exclude only the real values $0, \frac{1}{J_1}, \frac{1}{J_2}, \frac{1}{J_3}$ at which the parameterizations are not continuous.

Denote

$$\Gamma_h \triangleq \frac{\rho}{\|\rho\|}, \quad \Gamma_i \triangleq -\frac{\rho}{\|\rho\|}, \quad \text{and} \quad \Gamma_\infty \triangleq -\frac{J^{-1}\rho}{\|J^{-1}\rho\|}.$$

Then it follows from Proposition 6 that $(\Gamma_h, 0)$ and $(\Gamma_i, 0)$ are equilibria of the Lagrange–Poincaré reduced model of the 3D pendulum. These are called the hanging equilibrium and the inverted equilibrium of the Lagrange–Poincaré reduced model, respectively. We refer to Γ_h and to Γ_i as the hanging equilibrium attitude and the inverted equilibrium attitude, respectively. As shown subsequently, Γ_∞ is a vector that is used to define the limit of two of the relative equilibrium manifolds in (4.10).

Convergence properties of the relative equilibria as the parameter α tends to each of the distinguished real values $0, \frac{1}{J_1}, \frac{1}{J_2}, \frac{1}{J_3}$ are addressed in the following proposition.

Proposition 7. *The equilibria and relative equilibria of the 3D pendulum given in the parameterizations (4.10) have the following convergence properties:*

1. $\lim\limits_{\alpha\to\infty} \Gamma_e = \Gamma_\infty, \quad \lim\limits_{\alpha\to\infty} \omega_e = \mp\sqrt{\frac{mg}{\|J^{-1}\rho\|}} J^{-1}\rho;$
2. $\lim\limits_{\alpha\to 0^-} \Gamma_e = \Gamma_i, \quad \lim\limits_{\alpha\to 0^-} \omega_e = 0;$
3. $\lim\limits_{\alpha\to 0^+} \Gamma_e = \Gamma_h, \quad \lim\limits_{\alpha\to 0^+} \omega_e = 0;$
4. *for* $i \in \{1,2,3\}$: $\lim\limits_{\alpha\to \frac{1}{J_i}^-} \Gamma_e = \mathrm{sgn}(\rho_i) e_i, \quad \lim\limits_{\alpha\to \frac{1}{J_i}^-} \omega_e = \pm\infty\, e_i;$
5. *for* $i \in \{1,2,3\}$: $\lim\limits_{\alpha\to \frac{1}{J_i}^+} \Gamma_e = -\mathrm{sgn}(\rho_i) e_i, \quad \lim\limits_{\alpha\to \frac{1}{J_i}^+} \omega_e = \pm\infty\, e_i.$

Proof. Consider the limiting case when $\alpha \to \infty$. Since $\alpha > 0$, we have, from (4.13) and (4.14),

$$\lim_{\alpha\to\infty} \Gamma_{\mathrm{e}} = \lim_{\alpha\to\infty} -\frac{n_\alpha}{\|n_\alpha\|} = \lim_{\alpha\to\infty} -\frac{(J - I_{3\times 3}/\alpha)^{-1}\rho}{\|(J - I_{3\times 3}/\alpha)^{-1}\rho\|} = -\frac{J^{-1}\rho}{\|J^{-1}\rho\|} = \Gamma_\infty.$$

Similarly,

$$\lim_{\alpha\to\infty} \omega_{\mathrm{e}} = \lim_{\alpha\to\infty} \sqrt{\frac{mg}{\|n_\alpha\|}} n_\alpha = \sqrt{\frac{mg}{\|J^{-1}\rho\|}} J^{-1}\rho.$$

Thus, as $\alpha \to \infty$, the relative equilibria converge to the relative equilibria given in (4.9). It can be similarly shown that the relative equilibria also converge to the relative equilibria given in (4.9) when $\alpha \to -\infty$.

Next, consider (4.10) as $\alpha \to 0$. Expressing $n_\alpha = \alpha(\alpha J - I_{3\times 3})^{-1}\rho$, we obtain

$$\lim_{\alpha\to 0^-} \Gamma_{\mathrm{e}} = \lim_{\alpha\to 0^-} -\frac{n_\alpha}{\|n_\alpha\|} = \lim_{\alpha\to 0^-} -\frac{\alpha(\alpha J - I_{3\times 3})^{-1}\rho}{\|\alpha(\alpha J - I_{3\times 3})^{-1}\rho\|} = -\frac{\rho}{\|\rho\|} = \Gamma_{\mathrm{i}},$$

which corresponds to the inverted attitude. Similarly,

$$\lim_{\alpha\to 0^+} \Gamma_{\mathrm{e}} = \lim_{\alpha\to 0^+} -\frac{n_\alpha}{\|n_\alpha\|} = \lim_{\alpha\to 0^+} -\frac{\alpha(\alpha J - I_{3\times 3})^{-1}\rho}{\|\alpha(\alpha J - I_{3\times 3})^{-1}\rho\|} = \frac{\rho}{\|\rho\|} = \Gamma_{\mathrm{h}},$$

which corresponds to the hanging attitude. Also,

$$\lim_{\alpha\to 0^-} \omega_{\mathrm{e}} = \lim_{\alpha\to 0^-} \sqrt{\frac{mg}{\|n_\alpha\|}} n_\alpha = \lim_{\alpha\to 0^-} \sqrt{mg\|n_\alpha\|}\frac{n_\alpha}{\|n_\alpha\|} = -\frac{\rho}{\|\rho\|} \lim_{\alpha\to 0^-} \sqrt{mg\|\alpha\rho\|} = 0.$$

Similarly,

$$\lim_{\alpha\to 0^+} \omega_{\mathrm{e}} = 0.$$

Now, it follows that

$$\lim_{\alpha\to 1/J_1{}^-} \Gamma_e = \lim_{\alpha\to 1/J_1{}^-} -\frac{\left[\frac{\rho_1}{J_1 - 1/\alpha},\ \frac{\rho_2}{J_2 - J_1},\ \frac{\rho_3}{J_3 - J_1}\right]^T}{\left\|\left[\frac{\rho_1}{J_1 - 1/\alpha},\ \frac{\rho_2}{J_2 - J_1},\ \frac{\rho_3}{J_3 - J_1}\right]\right\|} = \mathrm{sgn}(\rho_1) e_1,$$

and

$$\lim_{\alpha\to 1/J_1{}^-} \omega_{\mathrm{e}} = \lim_{\alpha\to 1/J_1{}^-} \sqrt{mg}\frac{\left[\frac{\rho_1}{J_1 - 1/\alpha},\ \frac{\rho_2}{J_2 - J_1},\ \frac{\rho_3}{J_3 - J_1}\right]^T}{\sqrt{\left\|\left[\frac{\rho_1}{J_1 - 1/\alpha},\ \frac{\rho_2}{J_2 - J_1},\ \frac{\rho_3}{J_3 - J_1}\right]^T\right\|}} = [-\mathrm{sgn}(\rho_1)\,\infty,\ 0,\ 0]^T.$$

The remaining cases are analyzed in a similar way. □

If additional assumptions are made about the location of the center of mass, then additional relative equilibria exist. The following results summarizes this situation.

Proposition 8. *Consider the 3D pendulum and assume that the location of the center of mass vector ρ satisfies the indicated property.*

1. *Assume there is exactly one index $i \in \{1,2,3\}$ for which $\rho_i = 0$. There are one-dimensional relative equilibria manifolds in $S^2 \times \mathbb{R}^3$ described by the parameterizations:*

$$\left(-\frac{p_i}{\|p_i\|}, \sqrt{\frac{mg}{\|p_i\|}}p_i\right), \left(-\frac{p_i}{\|p_i\|}, -\sqrt{\frac{mg}{\|p_i\|}}p_i\right) \text{ for any } \gamma \in \mathbb{R},$$

   where $p_1 = \left(\gamma, \frac{\rho_2}{J_2-J_1}, \frac{\rho_3}{J_3-J_1}\right)$, $p_2 = \left(\frac{\rho_1}{J_1-J_2}, \gamma, \frac{\rho_3}{J_3-J_2}\right)$, $p_3 = \left(\frac{\rho_1}{J_1-J_3}, \frac{\rho_2}{J_2-J_3}, \gamma\right)$.
2. *Assume $\rho_i = 0$ for exactly two indices $i \in \{1,2,3\}$. There are one-dimensional relative equilibria manifolds in $S^2 \times \mathbb{R}^3$ described by the parameterizations:*

$$(e_i, \gamma e_i), (-e_i, \gamma e_i), \text{ for any } \gamma \in \mathbb{R}.$$

Proof. Consider the first statement. It is easy to see that for $\alpha = 1/J_1$, (4.12) has a solution iff $\rho_1 = 0$. In this case, (4.12) can be written as

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & J_2 - J_1 & 0 \\ 0 & 0 & J_3 - J_1 \end{bmatrix} v = -mgJ_1 \begin{bmatrix} 0 \\ \rho_2 \\ \rho_3 \end{bmatrix}.$$

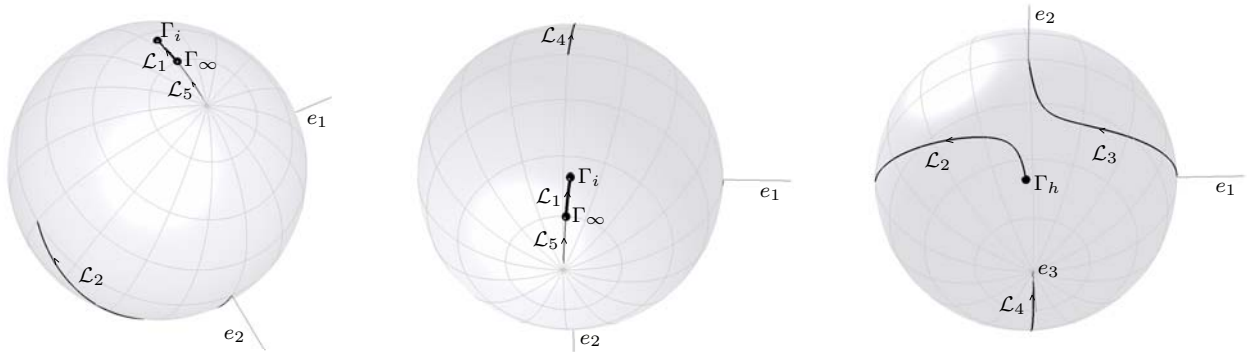
Since $\alpha > 0$, it can be shown as in (4.13) that $\Gamma_e = -\frac{p_1}{\|p_1\|}$ and $\omega_e = \pm\sqrt{\frac{mg}{\|p_1\|}}p_1$, where $p_1 = \left(\gamma, \frac{\rho_2}{J_2-J_1}, \frac{\rho_3}{J_3-J_1}\right)$ and $\gamma \in \mathbb{R}$. Similarly, one can obtain solutions of (4.12) for the case where $\alpha = 1/J_2$ and $\alpha = 1/J_3$ iff $\rho_2 = 0$ and $\rho_3 = 0$, respectively.

Now consider the second statement. The assumption guarantees that the vector ρ can be expressed as $\rho = se_i$, for some $s \in \mathbb{R}$. Then for all $\alpha \in \mathbb{R}\backslash\{0, \frac{1}{J_1}, \frac{1}{J_2}, \frac{1}{J_3}\}$ it follows that $n_\alpha = \frac{1}{\alpha}(\alpha J - I_{3\times 3})^{-1}\rho = \frac{s}{\alpha(\alpha J_i - 1)}e_i$. Thus the description of the relative equilibria given in (4.10) can be parameterized as $\{(e_i, \gamma e_i), (-e_i, \gamma e_i)\}$ for $\gamma \in \mathbb{R}$. □

This Proposition states the well known result for an asymmetric 3D pendulum that if the center of mass lies on a principal axis, then there exist relative equilibria defined by a constant angular velocity vector with arbitrary magnitude and direction that is aligned with that principal axis.

The geometric description of relative equilibria of the Lagrange-Poincaré model given in Proposition 6 and Proposition 8 provides important insight into the equilibria structure of the 3D pendulum. This geometric description is consistent with the implicit characterization of relative equilibria given in [19].

We now relate the equilibria structure for the Lagrange–Poincaré reduced model, as described in this section, to the equilibria structure for the full model, as described in the previous section. Let $(R_e, 0)$ denote an equilibrium in either the hanging equilibrium manifold or the inverted equilibrium manifold of the full equations (3.1) and (3.2) and $\pi : TSO(3) \to TSO(3)/S^1$ be the projection as in Proposition 4. Then, it can be shown that either $\pi(R_e, 0) = (\Gamma_h, 0)$ or $\pi(R_e, 0) = (\Gamma_i, 0)$. Thus, we obtain the following.



(a) 3D view (b) Top view (c) Bottom view

FIG. 4.1. *Relative equilibria attitudes for an elliptic cylinder 3D pendulum model*

Proposition 9 ([9]). *The hanging equilibrium manifold and the inverted equilibrium manifold of the 3D pendulum given by the full equations (3.1) and (3.2) are identified with the hanging equilibrium* $(\Gamma_h, 0)$ *and the inverted equilibrium* $(\Gamma_i, 0)$ *of the Lagrange–Poincaré reduced equations given by (4.3) and (4.4).*

4.3. Visualization of Equilibria and Relative Equilibria. We examine the equilibrium structure of a particular 3D pendulum model, demonstrating how this equilibrium structure can be visualized. We choose an elliptic cylinder with its semimajor axis $a = 0.8\,m$, semiminor axis $b = 0.2\,m$, and height $0.6\,m$. The pivot point is located at the surface of the upper ellipse, and it is offset from the center by $[-\frac{a}{6}, \frac{b}{2}, 0]$. The inertia tensor is given by $J = \text{diag}(0.3061, 0.2136, 0.1159)\,\text{kgm}^2$ and the vector from the pivot to the mass center is $\rho = [-0.0160,\ 0.2077,\ 0.2727]\,\text{m}$.

Figures 4.1(a)–4.1(c) show the equilibrium and relative equilibrium attitudes on S^2. Figure (a) provides a 3D perspective; Figure (b) provides a top view, with the inverted equilibrium attitude located at the center and Figure (c) provides a bottom view, with the hanging equilibrium attitude located at the center. The inverted equilibrium attitude Γ_i and the limiting relative equilibrium attitude Γ_∞ described by (4.9) are denoted by circles in Fig. 4.1(b); the hanging equilibrium attitude Γ_h is denoted by a circle in Fig. 4.1(c). The one-dimensional manifolds of relative equilibrium attitudes described in (4.10) are shown by five curve segments illustrated by the solid lines corresponding to $\alpha \in \mathcal{L}_i$ for $i \in \{1, 2, 3, 4, 5\}$, where the increasing value of α is denoted by arrows on each segment.

The relative equilibrium attitudes for $\alpha \in \mathcal{L}_1 = (-\infty, 0)$ are shown by a segment of a thick line in Fig. 4.1(b), which starts from Γ_∞, and converges to the inverted equilibrium attitude Γ_i as α increases, according to Proposition 7. For $\alpha \in \mathcal{L}_2 = (0, 1/J_1)$, the line of relative equilibrium attitudes starts from Γ_h, and ends at $-e_1$ in Fig. 4.1(c), since $\rho_1 < 0$ for the given pendulum model. Similarly, for $\alpha \in \mathcal{L}_3 = (1/J_1, 1/J_2)$, the line of relative equilibrium attitudes begins at e_1 and tends to e_2, and for $\alpha \in \mathcal{L}_4 = (1/J_2, 1/J_3)$ the line of relative equilibrium attitudes begins from $-e_2$ and ends at e_3, as α increases. The relative equilibrium attitudes for $\alpha \in \mathcal{L}_5 = (1/J_3, \infty)$ are shown by a segment of a thin line in Fig. 4.1(b), which begins at $-e_3$, and ends at Γ_∞. Therefore, the line of relative equilibrium attitudes for $\alpha \in \mathcal{L}_1$ and the line of relative equilibrium attitudes for $\alpha \in \mathcal{L}_5$ are connected. Since no component of the center of mass vector vanishes, there are no additional relative equilibria.

In summary, we have provided a graphical illustration of the hanging equilibrium attitude, the inverted equilibrium attitude, the relative equilibrium attitude given by (4.9), and the four mutually disjoint one–

dimensional relative equilibrium attitude curve segments.

4.4. Local Analysis of the Lagrange–Poincaré Reduced Model near an Equilibrium. In the last section, we showed that the Lagrange–Poincaré reduced model of the 3D pendulum has exactly two *static* equilibria, namely the hanging equilibrium and the inverted equilibrium. As stated in Proposition 9, these equilibria correspond to the disjoint equilibrium manifolds of the full equations of the 3D pendulum.

We focus on these static equilibria of the Lagrange–Poincaré reduced equations. The identification mentioned in Proposition 9 relates properties of the equilibrium manifolds of the full equations and the equilibria of the Lagrange–Poincaré reduced equations. We can deduce the stability of the hanging and the inverted equilibrium manifolds of the full equations by studying the stability property of the hanging equilibrium and the inverted equilibrium of the Lagrange–Poincaré reduced equations.

Consider the linearization of equations (4.3)–(4.4) about an equilibrium $(\Gamma_h, 0) = (R_{\mathrm{e}}^{\mathrm{T}} e_3, 0)$, where $(R_{\mathrm{e}}, 0)$ is an equilibrium of the hanging equilibrium manifold H.

Proposition 10. *The linearization of the Lagrange–Poincaré reduced equations for the 3D pendulum, about the equilibrium* $(\Gamma_h, 0) = (R_{\mathrm{e}}^{\mathrm{T}} e_3, 0)$ *described by equations (4.3)–(4.4) can be expressed using* $(x_1, x_2, \dot{x}_1, \dot{x}_2, \dot{x}_3) \in \mathbb{R}^5$ *according to equations (3.22)–(3.24).*

Proof. Consider a perturbation in terms of the perturbation parameter $\varepsilon \in \mathbb{R}$ as before. Let $(R^\varepsilon, \omega^\varepsilon)$ denote a perturbed solution of the Lagrange–Poincaré reduced equations (3.1)–(3.2). Since $\Gamma = R^{\mathrm{T}} e_3$, the perturbed solution of equations (4.3)–(4.4) is given by $(\Gamma^\varepsilon, \omega^\varepsilon)$ where $\Gamma^\varepsilon = (R^\varepsilon)^{\mathrm{T}} e_3$. Define the perturbation variables $\Delta\omega \triangleq \omega_\varepsilon^0$ and $\Delta\Gamma \triangleq \Gamma_\varepsilon^0 = (R_\varepsilon^0)^{\mathrm{T}} e_3$. From definition of $\Delta\Theta$ in Subsection 3.3, note that

$$\Delta\Gamma = -\widehat{\Delta\Theta} R_{\mathrm{e}}^{\mathrm{T}} e_3 = \widehat{\Gamma}_h \Delta\Theta \in T_{\Gamma_h} S^2.$$

Then from equation (3.19) and the definition of $\Delta\Gamma$, it can be shown that the linearization of equations (4.3)–(4.4) is given by

$$J\Delta\dot{\omega} = mg\widehat{\rho}\,\Delta\Gamma, \tag{4.16}$$

$$\Delta\dot{\Gamma} = \widehat{\Gamma}_h\,\Delta\omega. \tag{4.17}$$

Following the development in Subsection 3.3, we give an orthogonal decomposition of the vector $\Delta\Theta = M^{-\mathrm{T}} x$ into a component along the vector ρ and a component normal to the vector ρ. This decomposition is

$$\Delta\Theta = M^{-\mathrm{T}} x = -\frac{\widehat{\rho}^{\,2}}{\|\rho\|^2}(M^{-\mathrm{T}} x) + \frac{1}{\|\rho\|^2}\Big[\rho^{\mathrm{T}}(M^{-\mathrm{T}} x)\Big]\rho,$$

where $\dfrac{1}{\|\rho\|^2}\Big[\rho^{\mathrm{T}}(M^{-\mathrm{T}} x)\Big]\rho \in \mathrm{span}\{\rho\}$ and $-\dfrac{\widehat{\rho}^{\,2}}{\|\rho\|^2}(M^{-\mathrm{T}} x) \in \mathrm{span}\{\rho\}^{\perp}$. Thus,

$$\Delta\Gamma = \widehat{\Gamma}_h\,\Delta\Theta = \frac{\widehat{\rho}}{\|\rho\|} M^{-\mathrm{T}} x = \frac{1}{mg\|\rho\|^2}\widehat{\rho}\, M\Lambda\, x,$$

does not depend on x_3 since $\Lambda = \mathrm{diag}(mgl_1, mgl_2, 0)$. Thus, we can express the linearization of equations (4.3)–(4.4) at $(\Gamma_h, 0) = (R_{\mathrm{e}}^{\mathrm{T}} e_3, 0)$ in terms of the variables $(x_1, x_2, \dot{x}_1, \dot{x}_2, \dot{x}_3)$ according to equations (3.22)–(3.24). ∎

Due to our careful choice of variables, one can discard x_3 from equations (3.22)–(3.24) when studying the stability properties of the inverted equilibrium manifold. However, the angular velocity expression $\dot{x}_3$ is retained.

Summarizing the above, the linearization of equations (4.3)–(4.4) about the hanging equilibrium $(\Gamma_h, 0)$ is expressed as

$$\ddot{x}_1 + mgl_1 x_1 = 0, \tag{4.18}$$

$$\ddot{x}_2 + mgl_2 x_2 = 0, \tag{4.19}$$

$$\dot{x}_3 = 0. \tag{4.20}$$

It is clear that due to the presence of zero and imaginary eigenvalues, one cannot arrive at a conclusion about the stability of the hanging equilibrium $(\Gamma_h, 0)$ from this linear analysis. Therefore we next consider a Lyapunov analysis.

Proposition 11. *The hanging equilibrium* $(\Gamma_h, 0) = \left(\frac{\rho}{\|\rho\|}, 0\right)$*, of the Lagrange–Poincaré reduced dynamics of the 3D pendulum described by equations (4.3) and (4.4) is stable in the sense of Lyapunov.*

Proof. Consider the Lyapunov function

$$V(\Gamma, \omega) = \frac{1}{2}\, \omega^{\mathrm{T}} J \omega + mg(\|\rho\| - \rho^{\mathrm{T}}\Gamma). \tag{4.21}$$

Note that $V(\Gamma_h, 0) = 0$ and $V(\Gamma, \omega) > 0$ elsewhere. Furthermore, the derivative along a solution of (4.3) and (4.4) is given by

$$\begin{aligned} \dot{V}(\Gamma, \omega) &= \omega^{\mathrm{T}} J \dot{\omega} - mg\rho^{\mathrm{T}}\dot{\Gamma}, \\ &= \omega^{\mathrm{T}}(J\omega \times \omega + mg\rho \times \Gamma) - mg\rho^{\mathrm{T}}(\Gamma \times \omega), \\ &= \omega^{\mathrm{T}} mg\rho \times \Gamma - mg\rho^{\mathrm{T}}\Gamma \times \omega = 0. \end{aligned}$$

Thus, the hanging equilibrium is Lyapunov stable. □

Note that combining Proposition 11 with Proposition 9 immediately confirms the stability result for the hanging equilibrium manifold in Proposition 2.

We next examine the local properties of the Lagrange–Poincaré reduced equations of the 3D pendulum near the inverted equilibrium $(\Gamma_i, 0)$. Consider the linearization of equations (4.3)–(4.4) about an equilibrium $(\Gamma_i, 0) = (R_{\mathrm{e}}^{\mathrm{T}} e_3, 0)$, where $(R_{\mathrm{e}}, 0)$ is an equilibrium in the inverted equilibrium manifold $\mathtt{I}$. A result similar to Proposition 10 follows.

Proposition 12. *The linearization of the Lagrange–Poincaré reduced equations for the 3D pendulum, about the equilibrium* $(\Gamma_i, 0) = (R_{\mathrm{e}}^{\mathrm{T}} e_3, 0)$ *described by equations (4.3)–(4.4) can be expressed using* $(x_1, x_2, \dot{x}_1, \dot{x}_2, \dot{x}_3) \in \mathbb{R}^5$ *according to (3.26)–(3.28).*

The linearization of equations (4.3)–(4.4) at the inverted equilibrium $(\Gamma_i, 0)$ is expressed as

$$\ddot{x}_1 - mgl_1 x_1 = 0, \tag{4.22}$$

$$\ddot{x}_2 - mgl_2 x_2 = 0, \tag{4.23}$$

$$\dot{x}_3 = 0. \tag{4.24}$$

Note that the inverted equilibrium of the Lagrange–Poincaré reduced equations has two negative eigenvalues, two positive eigenvalues and a zero eigenvalue. Thus, the inverted equilibrium $(\Gamma_i, 0)$ is unstable and locally there exists a two-dimensional stable manifold, a two-dimensional unstable manifold and a one-dimensional center manifold.

Proposition 13. *The inverted equilibrium* $(\Gamma_i, 0) = \left(-\frac{\rho}{\|\rho\|}, 0\right)$ *of the Lagrange–Poincaré reduced model of the 3D pendulum described by equations (4.3) and (4.4) is unstable.*

Note that combining Proposition 13 with Proposition 9 immediately recovers the result that the inverted equilibrium manifold I of the full equations for the 3D pendulum given by (3.1)–(3.2) is unstable.

We have analyzed the local stability properties of the hanging equilibrium and of the inverted equilibrium of the Lagrange-Poincaré model. We have not analyzed local stability properties of other equilibrium solutions, namely the relative equilibria of of the of the Lagrange-Poincaré model. Such analysis can easily be carried out using constrained variations that respect the Lie group structure of the attitude configurations following the methods introduced in this paper. Alternatively, an analysis of the stability of the relative equilibria of the Lagrange-Poincaré model has been provided in [19] using constrained second variation methods that enforce the Lie group constraints using Lagrange multipliers.

5. Lagrange–Routh Reduced 3D Pendulum Dynamics on TS^2. In the previous sections we studied the full dynamics and the Lagrange–Poincaré reduced dynamics of the 3D pendulum. These involved the study of the dynamics of the 3D pendulum on $TSO(3)$ and on $S^2 \times \mathbb{R}^3$, respectively, using variables (R, ω) and (Γ, ω) to express the equations of motion. In this section, we present Lagrange–Routh reduction of the 3D pendulum, and we study the equations of motion that describe the evolution of $(\Gamma, \dot{\Gamma}) \in TS^2$.

5.1. Lagrange–Routh Reduction of the 3D pendulum. Lagrange–Routh reduction involves identifying trajectories that are related by the symmetry group action, and further restricting the dynamics to a level set of the associated momentum map. Since the symmetry group is abelian, the dynamics on the configuration manifold $SO(3)$ can be reduced to the shape manifold, which is the quotient manifold associated with the symmetry action. The resulting equations of motion on the quotient manifold are described not in terms of the Lagrangian but in terms of the Routhian [15, 18, 20].

The 3D pendulum has a S^1 symmetry given by a rotation about the vertical axis. The symmetry action $\Phi_\theta : S^1 \times SO(3) \to SO(3)$ is given by

$$\Phi_\theta(R) = \exp(\theta \widehat{e}_3) R,$$

for $\theta \in S^1$ and $R \in SO(3)$. It can be shown that the Lagrangian of the 3D pendulum is invariant under this symmetry action. Thus, the configuration manifold $SO(3)$ is reduced to the shape manifold $SO(3)/S^1 \cong S^2$, and the reduced dynamics of the 3D pendulum are described on the tangent bundle TS^2. This reduction procedure is interesting and challenging, since the projection $\Pi : SO(3) \to S^2$ given by $\Pi(R) = R^T e_3$ together with the symmetry action has a nontrivial principal bundle structure. In other words, the angle of the rotation about the vertical axis is not a global cyclic variable.

Here we present expressions for the Routhian and the reduced equations of motion. The detailed description and development can be found in the Appendix.

Proposition 14 ([20]). *We identify the Lie algebra of S^1 with $\mathbb{R}$. For $(R, \omega) \in T_R SO(3)$, the momentum map $\mathbf{J} : TSO(3) \to \mathbb{R}^*$, the locked inertia tensor $\mathbb{I}(R) : \mathbb{R} \to \mathbb{R}^*$, and the mechanical connection $\mathcal{A} : TSO(3) \to \mathbb{R}$ for the 3D pendulum are given as follows*

$$\mathbf{J}(R, \hat{\omega}) = e_3^T R J \omega, \tag{5.1}$$

$$\mathbb{I}(R) = e_3^T R J R^T e_3, \tag{5.2}$$

$$\mathcal{A}(R, \hat{\omega}) = \frac{e_3^T R J \omega}{e_3^T R J R^T e_3}. \tag{5.3}$$

The value of the momentum map $\mu = \mathbf{J}(R, \hat{\omega})$ corresponds to the vertical component of the angular momentum. Noether's theorem states that the symmetry of the Lagrangian implies conservation of the corresponding momentum map. This is an alternative method of establishing the invariance properties of the 3D pendulum dynamics, as opposed to the direct computation in Section 3.1.

Based on the above expressions, Lagrange–Routh reduction is carried out to obtain the following result.

Proposition 15. *For a given value of the momentum map μ, the Routhian of the 3D pendulum is given by*

$$R^{\mu}(\Gamma, \dot{\Gamma}) = \frac{1}{2}(\dot{\Gamma} \times \Gamma) \cdot J(\dot{\Gamma} \times \Gamma) - \frac{1}{2}(b^2 + \nu^2)(\Gamma \cdot J\Gamma) + mg\Gamma \cdot \rho, \tag{5.4}$$

where $b = \frac{J\Gamma \cdot (\dot{\Gamma} \times \Gamma)}{\Gamma \cdot J\Gamma}$, $\nu = \frac{\mu}{\Gamma \cdot J\Gamma}$, and the magnetic two-form can be written as

$$\beta_{\mu}(\Gamma \times \eta, \Gamma \times \zeta) = -\frac{\mu}{(\Gamma \cdot J\Gamma)^2}\left[-(\Gamma \cdot J\Gamma)\mathrm{tr}[J] + 2\left\|J\Gamma\right\|^2\right]\Gamma \cdot (\eta \times \zeta). \tag{5.5}$$

The Routhian satisfies the Euler-Lagrange equation, with the magnetic term, given by

$$\delta \int_0^T R^{\mu}(\Gamma, \dot{\Gamma}) dt = \int_0^T \mathbf{i}_{\dot{\Gamma}} \beta_{\mu}(\delta\Gamma) dt. \tag{5.6}$$

This yields the reduced equation of motion on TS^2:

$$\ddot{\Gamma} = -\|\dot{\Gamma}\|^2\Gamma + \Gamma \times \Sigma(\Gamma, \dot{\Gamma}), \tag{5.7}$$

where

$$\Sigma(\Gamma, \dot{\Gamma}) = b\dot{\Gamma} + J^{-1}\left[(J(\dot{\Gamma} \times \Gamma) - bJ\Gamma) \times ((\dot{\Gamma} \times \Gamma) - b\Gamma) + \nu^2 J\Gamma \times \Gamma - mg\Gamma \times \rho - c\dot{\Gamma}\right], \tag{5.8}$$

$$c = \nu\left\{\mathrm{tr}[J] - 2\frac{\|J\Gamma\|^2}{\Gamma \cdot J\Gamma}\right\}, \quad b = \frac{J\Gamma \cdot (\dot{\Gamma} \times \Gamma)}{\Gamma \cdot J\Gamma}, \quad \nu = \frac{\mu}{\Gamma \cdot J\Gamma}. \tag{5.9}$$

Proof. See the Appendix. □

The function $\Sigma(\Gamma, \dot{\Gamma})$ is an exceedingly complicated function of its arguments. This makes direct analysis of equation (5.7) a challenge.

5.2. Lagrange–Routh Reconstruction of the 3D pendulum. For a given value of the momentum map μ, let $\Gamma(t) \in S^2$ be a curve in the reduced space S^2 satisfying the Lagrange–Routh reduced model given by equation (5.7). The reconstruction procedure is to find the curve $R(t) \in SO(3)$ that satisfies $\Pi(R(t)) = \Gamma(t)$ and $\mathbf{J}(R(t), R(t)^T\dot{R}(t)) = \mu$.

This can be achieved in two steps. First, we choose any curve $\tilde{R}(t) \in SO(3)$ such that its projection is equal to the reduced curve, i.e. $\Pi(\tilde{R}(t)) = \Gamma(t)$. Then, the curve $R(t)$ can be written as $R(t) = \Phi_{\theta(t)}(\tilde{R}(t))$ for some $\theta(t) \in S^1$. We find a differential equation for $\theta(t)$ so that the value of the momentum map for the reconstructed curve is conserved. In order to simplify the differential equation for $\theta(t)$, we restrict our discussion to the horizontal lift $R_{\mathrm{hor}}(t)$, which is a particular reconstruction curve that has the additional property that $\mathcal{A}(\dot{\tilde{R}}(t)) = 0$, where $\mathcal{A}$ is the mechanical connection in equation (5.3).

Proposition 16. *Suppose that the integral curve of the Lagrange–Routh reduced equation (5.7) is given by $(\Gamma(t), \dot{\Gamma}(t)) \in TS^2$ and the value of the momentum map μ is known. The following procedure*

reconstructs the motion of the 3D pendulum to obtain $(R(t), \omega(t)) \in TSO(3)$ *such that* $\Pi(R(t)) = \Gamma(t)$ *and* $\mathbf{J}(R(t), \omega(t)) = \mu$.

1. *Horizontally lift* $\Gamma(t)$ *to obtain* $R_{\text{hor}}(t)$ *by integrating the following equation with* $R_{\text{hor}}(0) = R(0)$*:*

$$\dot{R}_{\text{hor}}(t) = R_{\text{hor}}(t)\hat{\omega}_{\text{hor}}(t), \tag{5.10}$$

where

$$\omega_{\text{hor}}(t) = \dot{\Gamma}(t) \times \Gamma(t) - b(t)\Gamma(t). \tag{5.11}$$

2. *Determine* $\theta_{\text{dyn}}(t) \in S^1$ *from:*

$$\theta_{\text{dyn}}(t) = \int_0^t \frac{\mu}{\Gamma(s) \cdot J\Gamma(s)}\, ds. \tag{5.12}$$

3. *Reconstruct the desired curve in* $TSO(3)$ *from:*

$$R(t) = \Phi_{\theta_{\text{dyn}}(t)}(R_{\text{hor}}(t)) = \exp[\theta_{\text{dyn}}(t)\hat{e}_3]R_{\text{hor}}(t), \tag{5.13}$$

$$\omega(t) = \omega_{\text{hor}}(t) + \nu(t)\Gamma(t). \tag{5.14}$$

Proof. See the Appendix. ∎

This leads to the geometric phase formula that expresses the rotation angle about the vertical axis along a closed integral curve of the Lagrange-Routh reduced equations.

Proposition 17. *Assume the value of the momentum map* $\mu = 0$*. Let* $\Gamma(t)$*,* $t \in R$*, define a closed curve in* S^2*, i.e.* $\Gamma(0) = \Gamma(T)$ *for some* T*. The geometric phase* $\theta_{\text{geo}}(T) \in S^1$ *of the 3D pendulum is defined by the relationship* $R(T) = \Phi_{\theta_{\text{geo}}(T)}(R(0))$ *where*

$$\theta_{\text{geo}}(T) = \int_{\mathcal{B}} \frac{2\,\|J\Gamma(t)\|^2 - \text{tr}[J]\,(\Gamma(t) \cdot J\Gamma(t))}{(\Gamma(t) \cdot J\Gamma(t))^2}\, dA, \tag{5.15}$$

and $\mathcal{B}$ *is a surface in* S^2 *with boundary* $\Gamma(t)$*.*

5.3. Integral of the Lagrange–Routh Reduced Model. In this section we find an integral of motion for the Lagrange–Routh reduced model of the 3D pendulum, namely the total energy of the system. Note that the Lagrange–Routh reduced equations of motion are derived by eliminating the conserved vertical component of the body-fixed angular momentum. In a later section, we make use of the constant energy surfaces to visualize the dynamics of the 3D pendulum.

Proposition 18. *Assume the constant value of the momentum map is* μ*. The total energy*

$$E = \frac{1}{2}(\dot{\Gamma} \times \Gamma + (\nu - b)\Gamma)^T J(\dot{\Gamma} \times \Gamma + (\nu - b)\Gamma) - mg\rho^T\Gamma \tag{5.16}$$

is constant along solutions of the Lagrange–Routh reduced equations for the 3D pendulum given by equation (5.7).

Proof. Substituting the reconstruction equations for the angular velocity (5.11), (5.14) into the total energy expression (4.5), we obtain equation (5.16). The time derivative expression for the total energy is given by

$$\dot{E} = (\dot{\Gamma} \times \Gamma + (\nu - b)\Gamma)^T J(\ddot{\Gamma} \times \Gamma + (\dot{\nu} - \dot{b})\Gamma + (\nu - b)\dot{\Gamma}) - mg\rho^T\dot{\Gamma}.$$

Substituting the reduced equation of motion (5.7) into the above equation and rearranging, we can show that $\dot{E} = 0$. ∎

5.4. Equilibria of the Lagrange–Routh Reduced Model. The Lagrange–Routh reduced model is related to the Lagrange-Poincaré reduced model through a projection of $TSO(3)/S^1$ onto TS^2. Thus, the equilibria structure of the Lagrange-Routh reduced model is equivalent to the Lagrange-Poincaré reduced model, but it is represented in terms of the reduced attitude Γ_e and the value of the momentum map μ instead of (Γ_e, ω_e).

We summarize the equilibria structure of the Lagrange-Routh reduced model using equation (5.7), showing its equivalence to the equilibria structure presented in Proposition 6.

Proposition 19. *Consider the Lagrange–Routh reduced model of the 3D pendulum given by equation (5.7). For the indicated values of the momentum map μ, the Lagrange–Routh model on TS^2 has the following equilibria and relative equilibria:*

1. *If $\mu = 0$ there is a hanging equilibrium: $\left(\frac{\rho}{\|\rho\|}, 0\right) \in TS^2$,*
2. *If $\mu = 0$ there is an inverted equilibrium: $\left(-\frac{\rho}{\|\rho\|}, 0\right) \in TS^2$,*
3. *If the momentum map has one of the values*

$$\mu = \pm\sqrt{\frac{mg}{\|J^{-1}\rho\|^3}}\rho^T J^{-1}\rho, \tag{5.17}$$

there is a relative equilibria in TS^2:

$$\left(-\frac{J^{-1}\rho}{\|J^{-1}\rho\|}, 0\right) \tag{5.18}$$

4. *If the momentum map has one of the values*

$$\mu = \pm\sqrt{\frac{mg}{\|n_\alpha\|^3}}n_\alpha^T J n_\alpha, \tag{5.19}$$

there are one-dimensional relative equilibrium manifolds in TS^2 described by the parameterization:

$$\left(-\frac{n_\alpha}{\|n_\alpha\|}, 0\right) \quad \text{for } \alpha \in \mathcal{L}_i, i \in \{1,2,3,4,5\}, \tag{5.20}$$

where $n_\alpha = (J - \frac{1}{\alpha}I_{3\times 3})^{-1}\rho \in \mathbb{R}^3$ and the intervals of the reals are defined by

$$\mathcal{L}_1 = (-\infty, 0),\ \mathcal{L}_2 = (0, \tfrac{1}{J_1}),\ \mathcal{L}_3 = (\tfrac{1}{J_1}, \tfrac{1}{J_2}),\ \mathcal{L}_4 = (\tfrac{1}{J_2}, \tfrac{1}{J_3}),\ \mathcal{L}_5 = (\tfrac{1}{J_3}, \infty).$$

Proof. Substituting $\dot{\Gamma}_e = 0$ into equations (5.7)–(5.9), we obtain a condition for an equilibrium Γ_e:

$$\Gamma_e \times J^{-1}\left[\nu^2 J\Gamma_e \times \Gamma_e - mg\Gamma_e \times \rho\right] = 0.$$

This is equivalent to

$$\left[\nu^2 J\Gamma_e \times \Gamma_e - mg\Gamma_e \times \rho\right] = k_2 J\Gamma_e \tag{5.21}$$

for some constant $k_2 \in \mathbb{R}$. Taking the dot product of this and Γ_e implies that $0 = k_2\Gamma_e^T J\Gamma_e$. Since $\Gamma_e^T J\Gamma_e > 0$ as the inertia tensor J is positive-definite and $\Gamma_e \in S^2$, it follows that $k_2 = 0$. Thus, equation (5.21) is equivalent to

$$\nu^2 J\Gamma_e + mg\rho = k_1\Gamma_e \tag{5.22}$$

for some constant $k_1 \in \mathbb{R}$. Note that this is equivalent to the equilibrium condition for the Lagrange-Poincaré reduced model given by equation (4.11): for any solution (Γ_e, k, k_1) of equation (4.11), we can choose μ such that $k^2 = \nu^2 = \frac{\mu^2}{(\Gamma_e^T J \Gamma_e)^2}$, which gives a solution of equation (5.22), and vice versa. Thus, the equilibria structure of the Lagrange-Routh reduced model is equivalent to the equilibria of the Lagrange-Poincaré reduced model. For an equilibrium (Γ_e, ω_e) of the Lagrange-Poincaré reduced model, the value of the momentum map at the corresponding equilibrium of the Lagrange-Routh model is given by

$$\mu = k(\Gamma_e^T J \Gamma_e) = \omega_e^T \Gamma_e (\Gamma_e^T J \Gamma_e) \tag{5.23}$$

Substituting this into the equilibria expressions in Proposition 6, we obtain the stated results. □

According to Proposition 7, the following results hold for the Lagrange-Routh reduced model of the 3D pendulum.

Proposition 20. *The equilibria of the 3D pendulum given in the parameterizations (5.20) have the following convergence properties:*

1. $\lim_{\alpha \to \infty} \Gamma_e = \Gamma_\infty$;
2. $\lim_{\alpha \to 0^-} \Gamma_e = \Gamma_i$;
3. $\lim_{\alpha \to 0^+} \Gamma_e = \Gamma_h$;
4. *for* $i \in \{1,2,3\}$: $\lim_{\alpha \to \frac{1}{J_i}^-} \Gamma_e = \mathrm{sgn}(\rho_i) e_i$;
5. *for* $i \in \{1,2,3\}$: $\lim_{\alpha \to \frac{1}{J_i}^+} \Gamma_e = -\mathrm{sgn}(\rho_i) e_i$.

As previously, if additional assumptions are made about the location of the center of mass, then additional relative equilibria exist. The following results summarizes this situation.

Proposition 21.

Consider the 3D pendulum and assume that the location of the center of mass vector ρ satisfies the indicated property.

1. *Assume there is exactly one index $i \in \{1,2,3\}$ for which $\rho_i = 0$. If the momentum map has one of the values*

$$\mu = \pm \sqrt{\frac{mg}{\|p_i\|^3}} p_i^T J p_i,$$

there are one-dimensional relative equilibrium manifolds in TS^2 described by the parameterizations:

$$\left(-\frac{p_i}{\|p_i\|}, 0 \right) \quad \text{for any } \gamma \in \mathbb{R},$$

where $p_1 = \left(\gamma, \frac{\rho_2}{J_2 - J_1}, \frac{\rho_3}{J_3 - J_1}\right)$, $p_2 = \left(\frac{\rho_1}{J_1 - J_2}, \gamma, \frac{\rho_3}{J_3 - J_2}\right)$, $p_3 = \left(\frac{\rho_1}{J_1 - J_3}, \frac{\rho_2}{J_2 - J_3}, \gamma\right)$.

2. *Assume $\rho_i = 0$ for exactly two indices $i \in \{1,2,3\}$. There are one-dimensional relative equilibrium manifolds in TS^2 described by the parameterizations:*

$$(e_i, 0), (-e_i, 0) \quad \text{for any } \mu \in \mathbb{R}.$$

This Proposition states the well known result for an asymmetric 3D pendulum that if the center of mass lies on a principal axis, then for any value of the angular momentum map there exist relative equilibria defined by a constant angular velocity vector that is aligned with that principal axis.

The geometric description of relative equilibria, based on the Lagrange-Routh model, provides additional insight into the equilibria structure of the 3D pendulum.

5.5. Local Analysis of the Lagrange–Routh Reduced Model on TS^2. We showed that when the angular momentum map $\mu = 0$, the Lagrange–Routh reduced model of the 3D pendulum has two isolated equilibria, namely the hanging equilibrium and the inverted equilibrium. These equilibria correspond to the disjoint equilibrium manifolds of the full equations of the 3D pendulum.

We next focus on these isolated equilibria of the Lagrange–Routh reduced equations. Using Proposition 19, the stability properties of the equilibrium manifolds of the 3D pendulum can be deduced by studying the Lagrange–Routh reduced equilibria for the case $\mu = 0$. Compared to the Lagrange–Poincaré reduced model, the Lagrange–Routh reduction procedure results in a set of complicated equations that are a challenge to analyze.

Consider the equations (3.22)–(3.24) representing the linearization of the full equations of motion of the 3D pendulum at the hanging equilibrium. It was shown before that the Lagrange–Poincaré reduced equations of motion can be written in terms of $(x_1, x_2, \dot{x}_1, \dot{x}_2, \dot{x}_3)$. As shown in Proposition 10, this result follows from the fact that any perturbation in $\Gamma \in S^2$ at Γ_h can be expressed in terms of $(x_1, x_2) \in \mathbb{R}^2$. In a similar fashion, one obtains the following result.

Proposition 22. *Assume that the momentum map $\mu = 0$. The linearization of the Lagrange–Routh reduced model of the 3D pendulum, at the equilibrium $(\Gamma_h, 0) = (R_e^{\mathrm{T}} e_3, 0)$, described by equation (5.7) can be expressed using $(x_1, x_2, \dot{x}_1, \dot{x}_2) \in \mathbb{R}^4$ according to equations (3.22) and (3.23).*

Proof. In Proposition 10, it was shown that the perturbations in Γ at Γ_h can be described in terms of $(x_1, x_2) \in \mathbb{R}^2$. The result then follows by noting that the equations of motion of the Lagrange–Routh reduced 3D pendulum are described in terms of $(\Gamma, \dot{\Gamma}) \in TS^2$. Thus the linearization of the Lagrange–Routh reduced 3D pendulum model can be described using using $(x_1, x_2, \dot{x}_1, \dot{x}_2) \in \mathbb{R}^4$ according to equations (3.22) and (3.23). □

The linearization of the Lagrange–Routh reduced dynamics of the 3D pendulum, about the equilibrium $(0, \Gamma_h)$ is obtained from the linearized model of the full 3D pendulum dynamics by neglecting the dynamics corresponding to x_3:

$$\ddot{x}_1 + mgl_1 x_1 = 0, \tag{5.24}$$

$$\ddot{x}_2 + mgl_2 x_2 = 0. \tag{5.25}$$

It is clear that due to the presence of imaginary eigenvalues, stability of the hanging equilibrium $(\Gamma_h, 0)$ cannot be concluded. Therefore we next consider Lyapunov analysis.

Proposition 23. *Assume that the momentum map $\mu = 0$. The hanging equilibrium $(\Gamma_h, 0) = \left(\frac{\rho}{\|\rho\|}, 0\right)$ of the Lagrange–Routh reduced dynamics of the 3D pendulum described by equation (5.7) is stable in the sense of Lyapunov.*

Proof. Consider the Lyapunov function

$$V(\Gamma, \dot{\Gamma}) = \frac{1}{2}(\dot{\Gamma} \times \Gamma + (\nu - b)\Gamma)^{\mathrm{T}} J(\dot{\Gamma} \times \Gamma + (\nu - b)\Gamma) + mg(\|\rho\| - \rho^{\mathrm{T}}\Gamma). \tag{5.26}$$

Note that $V(\Gamma_h, 0) = 0$ and $V(\Gamma, \dot{\Gamma}) > 0$ elsewhere. Furthermore, the derivative along a solution of (4.3) and (4.4) is given by

$$\dot{V}(\Gamma, \dot{\Gamma}) = (\dot{\Gamma} \times \Gamma + (\nu - b)\Gamma)^{\mathrm{T}} J(\ddot{\Gamma} \times \Gamma + (\dot{\nu} - \dot{b})\Gamma + (\nu - b)\dot{\Gamma}) - mg\rho^{\mathrm{T}}\dot{\Gamma}.$$

Substituting the reduced equation of motion (5.7) into the above equation and rearranging, we can show that $\dot{V}(\Gamma, \dot{\Gamma}) = 0$. Thus, the hanging equilibrium of (5.7) is Lyapunov stable. □

Note that combining Proposition 23 with Proposition 19 immediately yields the result in Proposition 2.

We next study the local properties of the Lagrange–Routh reduced dynamics of the 3D pendulum near the inverted equilibrium $(\Gamma_i, 0)$. Consider the linearization of equation (5.7) at an equilibrium $(\Gamma_i, 0) = (R_{\mathrm{e}}^{\mathrm{T}} e_3, 0)$, where $(R_{\mathrm{e}}, 0)$ is an equilibrium of the inverted equilibrium manifold $\mathtt{I}$. A result similar to Proposition 22 follows.

Proposition 24. *Assume that the momentum map $\mu = 0$. The linearization of the Lagrange-Routh reduced dynamics of the 3D pendulum, at the equilibrium $(\Gamma_i, 0) = (R_{\mathrm{e}}^{\mathrm{T}} e_3, 0)$, described by equation (5.7) can be expressed using $(x_1, x_2, \dot{x}_1, \dot{x}_2) \in \mathbb{R}^4$ according to equations (3.26) and (3.27).*

Summarizing, the linearization of equation (5.7) at the inverted equilibrium $(\Gamma_i, 0)$ is expressed as

$$\ddot{x}_1 - mgl_1 x_1 = 0, \tag{5.27}$$

$$\ddot{x}_2 - mgl_2 x_2 = 0. \tag{5.28}$$

Note that the linearization of equation (5.7) at the inverted equilibrium has two negative eigenvalues and two positive eigenvalues. Thus, the inverted equilibrium $(\Gamma_i, 0)$ of the Lagrange–Routh reduced model is unstable and locally there exists a two-dimensional stable manifold and a two-dimensional unstable manifold.

Proposition 25. *Assume that the momentum map $\mu = 0$. The inverted equilibrium $(\Gamma_i, 0) = \left(-\frac{\rho}{\|\rho\|}, 0\right)$ of the Lagrange–Routh reduced dynamics of the 3D pendulum described by equations (5.7) is unstable.*

Note that combining Proposition 25 with Proposition 19 immediately yields the result that the inverted equilibrium manifold $\mathtt{I}$ of the 3D pendulum given by (3.1)–(3.2) is unstable.

5.6. Poincaré Map on the Lagrange–Routh Reduced Model. A Poincaré map describes the evolution of successive intersection points of a trajectory with a transversal hypersurface of codimension one. Typically, one chooses a hyperplane, and considers a trajectory with initial conditions on the hyperplane. The points at which this trajectory returns to the hyperplane are then observed, which provides insight into the stability of periodic orbits or the global characteristics of the dynamics.

The Lagrange–Routh reduced equations for the 3D pendulum on TS^2 are a particularly suitable choice for analysis using a Poincaré map, since it has dimension four. On the manifold defined by a constant total energy as in equation (5.16) a Poincaré section on TS^2 defines a three-dimensional subspace of TS^2 on which the corresponding Poincaré map evolves. We define a Poincaré section on TS^2 for the Lagrange-Routh dynamics of the 3D pendulum given by equation (5.7) as follows.

$$\mathcal{P} = \left\{ (\Gamma, \dot{\Gamma}) \in TS^2 \,\middle|\, e_3^T \dot{\Gamma} = 0,\ e_3^T (\Gamma \times \dot{\Gamma}) > 0,\ \text{and } E(\Gamma, \dot{\Gamma}) = \text{constant} \right\}.$$

Suppose $\Gamma \in \mathcal{P}$ is given. The tangent space $T_\Gamma S^2$ is a plane that is tangential to S^2 and perpendicular to Γ. The first condition of the Poincaré section, $e_3^T \dot{\Gamma} = 0$ determines a line in which the tangent vector $\dot{\Gamma} \in T_\Gamma S^2$ should lie, and the constraint of total energy conservation fixes the magnitude of the tangent vector in that line. Thus, the tangent vector is uniquely determined up to sign. The second condition of the Poincaré section resolves this ambiguity. It also excludes two reduced attitudes $\Gamma = \pm e_3$ for which the first condition is trivial; $e_3^T \dot{\Gamma} = 0$ for any $\dot{\Gamma} \in T_{e_3} S^2 \bigcup T_{-e_3} S^2$. Thus, $\mathcal{P}$ can be equivalently identified as

$$\mathcal{P} = \left\{ \Gamma \in S^2 \,\middle|\, e_3^T \dot{\Gamma} = 0,\ e_3^T (\Gamma \times \dot{\Gamma}) > 0,\ \text{and } E(\Gamma, \dot{\Gamma}) = \text{constant} \right\},$$

(a) $E = -2.65$ (b) $E = 0$ (c) $E = 2.03$ (d) $E = 8.83$ (e) $E = 11.95$

FIG. 5.1. *Poincaré maps for 3D pendulum with varying total energy*

where $(\Gamma, \dot{\Gamma})$ satisfies (5.7).

This Poincaré section in TS^2 is well-defined in the sense that for each element, the corresponding tangent vector is uniquely determined. The attitude and the angular velocity in $TSO(3)$ can be obtained by using the reconstruction procedure for the given value of the momentum map μ.

Although the Poincaré map is best explained in terms of the Lagrange–Routh reduced model as above, successive intersection points of a trajectory need not be computed directly from the complicated Lagrange–Routh reduced model. In particular, trajectories can be computed directly from the full 3D pendulum model given by equations (3.1) and (3.2) or from the the Lagrange-Poincaré reduced model given by equations (4.3) and (4.4); the Poincaré map is obtained by a projection onto S^2.

5.7. Visualization of Poincaré Map. We examine the Poincaré map that describes dynamics features of a particular 3D pendulum model, demonstrating how this Poincaré map can be visualized. The 3D pendulum is chosen as an elliptic cylinder with properties $m = 1\,\text{kg}$, $J = \text{diag}[0.13, 0.28, 0.17]\,\text{kgm}^2$, $\rho = [0, 0, 0.3]\,\text{m}$. The initial conditions are given by $R_0 = I_{3\times 3}$ and $\omega_0 = c[1, 1, 1]\,\text{rad/s}$, where the constant c is varied to give different total energy levels as specified in Fig. 5.1; the specific values of the momentum map are not required since the Lagrange–Poincaré equations need not be formed. The Lie group variational integrator introduced in [17] is used to numerically integrate the full 3D pendulum equations (3.1) and (3.2), thereby obtaining the Poincaré map on S^2 numerically.

Since the center of mass lies on the third principal axis of the 3D pendulum, there is a relative equilibrium in which the angular velocity vector is aligned with the body-fixed principal axis e_3 axis and the value of the angular momentum map is arbitrary. Fig. 5.1 shows particular examples of the Poincaré map on S^2 corresponding to five different trajectories with various values of the total energy; the values of successive values of Γ in S^2 are shown. Each of these trajectories can be viewed as a perturbation of a relative equilibrium whose angular velocity vector is along the principal axis e_3. In Fig. 5.1 the center of each Poincaré map is defined by the body-fixed principal axis e_3 representing this relative equilibrium. That is, the origin of the body-fixed axes in Fig. 5.1 is located at the center of S^2 with the e_1 axis pointing to the right in the plane of the page, the e_2 axis pointing to the top in the plane of the page, and e_3 is perpendicular to the plane of the page pointing outward.

It is interesting to see the transition of the Poincaré maps with varying total energy levels. The attitude dynamics of the 3D pendulum is periodic in Fig. 5.1(a), but it exhibits chaotic behavior with increased energy level in Fig. 5.1(b) and 5.1(c). If the total energy is increased further, the attitude dynamics become periodic again in Fig. 5.1(e). This demonstrates the highly-nonlinear, and possibly chaotic, characteristics of the 3D pendulum dynamics.

6. Conclusions. The asymmetric 3D pendulum, assuming the center of mass is distinct from the pivot location, exhibits rich dynamics with nontrivial geometric structure; these dynamics are much richer and more complex than the dynamics of a 1D planar pendulum, a 2D spherical pendulum, or any of the integrable cases such as the Lagrange top. This paper has demonstrated that the methods of geometric mechanics and the methods of nonlinear dynamics can be combined to obtain insight into the complex, nonintegrable dynamics of the 3D pendulum.

The main contribution of the paper is that we have introduced three different models for the asymmetric 3D pendulum, including the full model defined on $TSO(3)$, the Lagrange–Poincaré reduced model on $TSO(3)/S^1$ obtained by identifying configurations in the same group orbit, and the Lagrange–Routh reduced model on TS^2 where one additionally utilizes the fact that the dynamics evolve on a constant momentum level set. Relationships between the various representations are discussed in the context of conservation properties, equilibria and their stability properties, and invariant manifolds.

In addition, we illustrate that the use of the Lagrange–Routh reduced equations of motion, together with the energy conservation properties, allow the construction of a Poincaré map that can be readily visualized, thereby providing a graphical tool for obtaining insight into the rich nonlinear dynamical properties of the 3D pendulum.

Appendix.

In this appendix, we summarize Lagrange-Routh reduction and reconstruction procedures for the 3D pendulum.

A.1. Reduction. A description of Lagrange-Routh reduction can be found in [20] including expressions for the mechanical connection and the Routhian of the 3D pendulum given by equation (5.3) and equation (5.4), respectively. Here we derive the reduced equation of motion (5.7) using the Euler-Lagrange equation with magnetic terms for the given Routhian (5.4).

Variation of Routhian. The Routhian satisfies the variational Euler-Lagrange equation with the magnetic term given by (5.6). We use a constrained variation of $\Gamma \in S^2$:

$$\delta\Gamma = \Gamma \times \eta, \tag{A.1}$$

$$\delta\dot{\Gamma} = \dot{\Gamma} \times \eta + \Gamma \times \dot{\eta}. \tag{A.2}$$

Here we assume that $\eta \cdot \Gamma = 0$, since the component of η parallel to Γ has no effect on $\delta\Gamma$. These expressions are essential for developing the reduced equation of motion.

Using (A.1), (A.2), and the properties $\Gamma \cdot \dot{\Gamma} = 0$, $\Gamma \cdot \eta = 0$, the variation of the Routhian is given by

$$\delta R^\mu = \dot{\eta} \cdot J(\dot{\Gamma} \times \Gamma - b\Gamma) - \eta \cdot \Gamma \times \left[-\dot{\Gamma} \times J(\dot{\Gamma} \times \Gamma) + (b^2 + \nu^2)J\Gamma - bJ(\dot{\Gamma} \times \Gamma) + b(\dot{\Gamma} \times J\Gamma) + mg\rho\right]. \tag{A.3}$$

Magnetic two-form. From the given mechanical connection $\mathcal{A}$ and a value of the momentum map $\mu \in \mathbb{R}^*$, define a one-form $\mathcal{A}_\mu$ on $TSO(3)$ by

$$\mathcal{A}_\mu(R) \cdot (R, \hat{\omega}) = \langle \mu, \mathcal{A}(R, \hat{\omega}) \rangle = \mu \frac{e_3^T RJ\omega}{e_3^T RJR^T e_3}$$

The magnetic two-form β_μ in (5.5) is the exterior derivative of $\mathcal{A}_\mu$, which can be obtained by using the identity $\mathbf{d}\mathcal{A}_\mu(X, Y) = X[\mathcal{A}_\mu(Y)] - Y[\mathcal{A}_\mu(X)] - \mathcal{A}_\mu([X, Y])$ for $X = R\hat{\eta}, Y = R\hat{\zeta} \in T_R SO(3)$. Suppose that $\dot{\Gamma} = \Gamma \times \omega$. Since $\Gamma \cdot (\omega \times \eta) = \eta \cdot (\Gamma \times \omega) = \eta \cdot \dot{\Gamma}$, the interior product of the magnetic two-form is given by

$$\mathbf{i}_{\dot{\Gamma}} \beta_\mu(\delta\Gamma) = \beta_\mu(\Gamma \times \omega, \Gamma \times \eta) = \nu \left\{ \mathrm{tr}[J] - 2\frac{\|J\Gamma\|^2}{\Gamma \cdot J\Gamma} \right\} \dot{\Gamma} \cdot \eta, \tag{A.4}$$

where $\nu = \frac{\mu}{\Gamma \cdot J\Gamma}$.

Euler-Lagrange equation with magnetic two-form. Substituting (A.3) and (A.4) into (5.6), and integrating by parts, the Euler-Lagrange equation for the reduced Routhian (5.4) is written as

$$-\int_0^T \eta \cdot \left[J(\ddot{\Gamma} \times \Gamma - b\dot{\Gamma} - \dot{b}\Gamma) + \Gamma \times X + c\dot{\Gamma} \right] dt = 0, \tag{A.5}$$

where

$$X = -\dot{\Gamma} \times J(\dot{\Gamma} \times \Gamma) + (b^2 + \nu^2) J\Gamma - bJ(\dot{\Gamma} \times \Gamma) + b(\dot{\Gamma} \times J\Gamma) + mg\rho, \tag{A.6}$$

and c is given by (5.9). Since (A.5) is satisfied for all η with $\Gamma \cdot \eta = 0$, we obtain

$$J(\ddot{\Gamma} \times \Gamma - b\dot{\Gamma} - \dot{b}\Gamma) + \Gamma \times X + c\dot{\Gamma} = \lambda\Gamma, \tag{A.7}$$

for $\lambda \in \mathbb{R}$. This is the reduced equation of motion. However, this equation has an ambiguity since the value of λ is unknown; this equation is implicit for $\ddot{\Gamma}$ since the term $\dot{b}$ is expressed in terms of $\ddot{\Gamma}$. The next step is to determine expressions for λ and $\dot{b}$ using the definition of b and some vector identities.

We first find an expression for λ in terms of $\Gamma, \dot{\Gamma}$. Taking the dot product of (A.7) with Γ, we obtain

$$\Gamma \cdot J(\ddot{\Gamma} \times \Gamma - b\dot{\Gamma} - \dot{b}\Gamma) = \lambda. \tag{A.8}$$

From the definition of b, we obtain the following identity: $\Gamma \cdot J(\dot{\Gamma} \times \Gamma - b\Gamma) = 0$. Differentiating this with respect to time and substituting into (A.8), we find an expression for λ in terms of $\Gamma, \dot{\Gamma}$ as

$$\lambda = -\dot{\Gamma} \cdot J(\dot{\Gamma} \times \Gamma - b\Gamma). \tag{A.9}$$

Substituting (A.9) into (A.7), and taking the dot product of the result with Γ, we obtain an expression for $\dot{b}$ in terms of $\Gamma, \dot{\Gamma}$ as

$$\dot{b} = \Gamma \cdot J^{-1} \left\{ \Gamma \times X + c\dot{\Gamma} + (\dot{\Gamma} \cdot J(\dot{\Gamma} \times \Gamma - b\Gamma))\Gamma \right\}. \tag{A.10}$$

Substituting (A.10) into (A.7), and using the vector identity $Y - (\Gamma \cdot Y)\Gamma = (\Gamma \cdot \Gamma)Y - (\Gamma \cdot Y)\Gamma = -\Gamma \times (\Gamma \times Y)$ for any $Y \in \mathbb{R}^3$, we obtain the following from for the reduced equation of motion

$$\ddot{\Gamma} \times \Gamma - b\dot{\Gamma} - \Gamma \times \left[\Gamma \times J^{-1} \left\{ \Gamma \times X + c\dot{\Gamma} + (\dot{\Gamma} \cdot J(\dot{\Gamma} \times \Gamma - b\Gamma))\Gamma \right\} \right] = 0.$$

Reduced equation of motion. This equation has no ambiguity. Now, we simplify this equation. The above expression is equivalent to the following equation

$$\Gamma \times \left[\ddot{\Gamma} \times \Gamma - b\dot{\Gamma} - \Gamma \times \left[\Gamma \times J^{-1} \left\{ \Gamma \times X + c\dot{\Gamma} + (\dot{\Gamma} \cdot J(\dot{\Gamma} \times \Gamma - b\Gamma))\Gamma \right\} \right] \right] = 0.$$

Since $\Gamma \cdot \ddot{\Gamma} = -\|\dot{\Gamma}\|^2$, the first term is given by

$$\Gamma \times (\ddot{\Gamma} \times \Gamma) = (\Gamma \cdot \Gamma)\ddot{\Gamma} - (\Gamma \cdot \ddot{\Gamma})\Gamma = \ddot{\Gamma} + \|\dot{\Gamma}\|^2 \Gamma.$$

Using the property $\Gamma \times (\Gamma \times (\Gamma \times Y)) = -(\Gamma \cdot \Gamma)\Gamma \times Y = -\Gamma \times Y$ for $Y \in \mathbb{R}^3$, the third term of the above equation can be simplified. Substituting (A.6) and rearranging, the reduced equation of motion for the 3D pendulum is given by

$$\ddot{\Gamma} = -\|\dot{\Gamma}\|^2 \Gamma + \Gamma \times \Sigma, \tag{A.11}$$

where $\Sigma = b\dot{\Gamma} + J^{-1} \left[(J(\dot{\Gamma} \times \Gamma) - bJ\Gamma) \times ((\dot{\Gamma} \times \Gamma) - b\Gamma) + \nu^2 J\Gamma \times \Gamma - mg\Gamma \times \rho - c\dot{\Gamma} \right]$.

A.2. Reconstruction. For a given integral curve of the reduced equation $(\Gamma(t), \dot{\Gamma}(t)) \in TS^2$, we find a curve $\tilde{R}(t) \in SO(3)$ that is projected into the reduced curve, i.e. $\Pi(\tilde{R}(t) = \Gamma(t)$. The reconstructed curve can be written as $R(t) = \Phi_{\theta(t)}(\tilde{R}(t))$ for some $\theta(t) \in S^1$. The conservation of the momentum map yields the following reconstruction equation [20].

$$\theta(t)^{-1}\dot{\theta}(t) = \mathbb{I}^{-1}(\tilde{R}(t))\mu - \mathcal{A}(\dot{\tilde{R}}(t)).$$

The particular choice of $\tilde{R}(t)$, the horizontal lift given by (5.10), simplifies the above equation, since the horizontal part of the tangent vector is annihilated by the mechanical connection, and as such the second term in the above equation vanishes. Furthermore, since the group S^1 is abelian, the solution reduces to a quadrature as in (5.12). The reconstructed curve is given by (5.13).

REFERENCES

[1] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, *Dynamical Systems III*, Springer, 1988.

[2] K. J. Astrom and K. Furuta, *Swinging up a pendulum by energy control*, Automatica, 36 (2000), pp. 287–295.

[3] R. E. Bellman, *Introduction to Matrix Analysis*, Society for Industrial and Applied Mathematics, 2nd ed., 1997.

[4] D. S. Bernstein, *Matrix Mathematics, Theory, Facts, and Formulas with Applications to Linear System Theory*, Princeton University Press, 2005.

[5] D. S. Bernstein, N. H. McClamroch, and A. M. Bloch, *Development of air spindle and triaxial air bearing testbeds for spacecraft dynamics and control experiments*, Proceedings of the American Control Conference, (2001), pp. 3967–3972.

[6] A. M. Bloch, *Nonholonomic Mechanics and Control*, Springer-Verlag, 2003.

[7] A. D. Bruno, *Analysis of the Euler-Poisson equations by methods of power geometry and normal form*, Journal of Applied Mathematics and Mechanics, 71 (2007), pp. 168–199.

[8] H. Cendra, J. E. Marsden, and T. S. Ratiu, *Lagrangian reduction by stages*, Mem. Amer. Math. Soc., 152 (2001).

[9] N. A. Chaturvedi and N. H. McClamroch, *Asymptotic Stabilization of the hanging equilibrium manifold of the 3D pendulum*, International Journal of Robust and Nonlinear Control, 17 (2007), pp. 1435–1454.

[10] S. Cho, J. Shen, and N. H. McClamroch, *Mathematical models for the triaxial attitude control testbed*, Mathematical and Computer Modeling of Dynamical Systems, 9 (2003), pp. 165–192.

[11] S. Cho, J. Shen, N. H. McClamroch, and D. S. Bernstein, *Equations of motion of the triaxial attitude control testbed*, Proceedings of the IEEE Conference on Decision and Control, (2001), pp. 3429–3434.

[12] R. H. Cushman and L. M. Bates, *Global Aspects of Classical Integrable Systems*, Birkhauser, 1997.

[13] K. Furuta, *Control of pendulum: From super mechano-system to human adaptive mechatronics*, Proceedings of the IEEE Conference on Decision and Control, (2003), pp. 1498–1507.

[14] A. Hernández-Garduño, J. K. Lawson, and J. E. Marsden, *Relative equilibria for the generalized rigid body*, Journal of Geometry and Physics, 53 (2005), pp. 259–274.

[15] P. J. Holmes and J. E. Marsden, *Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups*, Indiana U. Math. Journal, 32 (1983), pp. 962–967.

[16] A. V. Karapetyan, *Invariant sets in the Goryachev-Chaplygin problem: Existence, stability, and branching*, Journal of Applied Mathematics and Mechanics, 70 (2006), pp. 195–198.

[17] T. Lee, M. Leok, and N. H. McClamroch, *A Lie group variational integrator for the attitude dynamics of a rigid body with application to the 3D pendulum*, Proceedings of the IEEE Conference on Decision and Control, (2005), pp. 962–967.

[18] D. Lewis, T. S. Ratiu, J. C. Simo, and J. E. Marsden, *The heavy top: A geometric treatment*, Nonlinearity, 5 (1992), pp. 1–48.

[19] J. H. Maddocks, *Stability of relative equilibria*, IMA Journal of Applied Mathematics, 46 (1991), pp. 71–99.

[20] J. E. Marsden, T. S. Ratiu, and J. Scheurle, *Reduction theory and the Lagrange-Routh equations*, Journal of Mathematical Physics, 41 (2000), pp. 3379–3429.

[21] J. Shen, A. K. Sanyal, N. A. Chaturvedi, D. S. Bernstein, and N. H. McClamroch, *Dynamics and control of a 3D pendulum*, Proceedings of the IEEE Conference on Decision and Control, (2004), pp. 323–328.

Controlled Lagrangians and Stabilization of Discrete Mechanical Systems I

Anthony M. Bloch, Melvin Leok, Jerrold E. Marsden, and Dmitry V. Zenkov

Abstract— **Controlled Lagrangian and matching techniques are developed for the stabilization of relative equilibria and equilibria of discrete mechanical systems with symmetry as well as broken symmetry. Interesting new phenomena arise in the controlled Lagrangian approach in the discrete context that are not present in the continuous theory. In particular, to make the discrete theory effective, one can make an appropriate selection of momentum levels or, alternatively, introduce a new parameter into the controlled Lagrangian to complete the kinetic matching procedure. Specifically, new terms in the controlled shape equation that are necessary for potential matching in the discrete setting are introduced. The theory is illustrated with the problem of stabilization of the cart-pendulum system on an incline. The paper also discusses digital and model predictive controlers.**

I. Introduction

The method of controlled Lagrangians for stabilization of relative equilibria (steady state motions) originated in Bloch, Leonard, and Marsden [5] and was then developed in Auckly [1], Bloch, Leonard, and Marsden [6], [7], [8], Bloch, Chang, Leonard, and Marsden [9], and Hamberg [12], [13]. A similar approach for Hamiltonian controlled systems was introduced and further studied in the work of Blankenstein, Ortega, van der Schaft, Maschke, Spong, and their collaborators (see, e.g., [21], [22] and related references). The two methods were shown to be equivalent in [10] and a nonholonomic version was developed in [25], [26], and [2].

In the controlled Lagrangian approach, one considers a mechanical system with an uncontrolled (free) Lagrangian equal to kinetic energy minus potential energy. To start with, one considers the case in which the Lagrangian is invariant with respect to the action of a Lie group G on the configuration space. To stabilize a relative equilibrium of interest, the kinetic energy is modified to produce a *controlled Lagrangian* which describes the dynamics of the controlled closed-loop system. The equations corresponding to this controlled Lagrangian are the closed-loop equations and the new terms appearing in those equations corresponding to the directly controlled variables correspond to control inputs. The modifications to the Lagrangian are chosen so that no new terms appear in the equations corresponding to the variables that are not directly controlled. This process of obtaining controlled Euler–Lagrange equations by modifying the original Lagrangian is referred to as *kinetic matching*.

One advantage of this approach is that once the form of the control law is derived using the controlled Lagrangian, the stability of a relative equilibrium of the closed-loop system can be determined by energy methods, using any available freedom in the choice of the parameters of the controlled Lagrangian. To obtain asymptotic stabilization, dissipation-emulating terms are added to the control input.

The method is extended in [9] to the class of Lagrangian mechanical systems with potential energy that may break symmetry, *i.e.*, there is still a symmetry group G for the kinetic energy of the system but one may now have a potential energy that need not be G-invariant. Further, in order to define the controlled Lagrangian, a modification to the potential energy is introduced that also breaks symmetry in the group variables. After adding the dissipation-emulating terms to the control input, this procedure allows one to achieve complete state-space asymptotic stabilization of an equilibrium of interest.

The main objective of this paper is to develop the method of controlled Lagrangians for discrete mechanical systems. The discretization is done in the spirit of discrete variational mechanics, as in [20]. In particular, as the closed loop dynamics of a controlled Lagrangian system is itself Lagrangian, it is natural to adopt a variational discretization that exhibits good long-time numerical stability. This study is also motivated by the recent development of structure-preserving algorithms for the numerical simulation of discrete controlled systems, such as recent work on discrete optimization, such as in [11], [15], [16].

The matching procedure is carried out explicitly for discrete systems with one shape and one group degree of freedom to avoid technical issues and to concentrate on the new phenomena that emerge in the discrete setting that have not been observed in the continuous-time theory. In particular, it leads one to either carefully select the momentum levels or introduce a new term in the controlled Lagrangian to perform the discrete kinetic matching. Further, when the potential shaping is carried out, it is necessary to introduce non-conservative forcing in the shape equation associated with the controlled Lagrangian.

It is also shown that once energetically stabilized, the (relative) equilibria of interest can be asymptotically stabilized by adding dissipation emulating terms. The separation of controlled dissipation from physical dissipation remains an interesting topic for future research; even in the continuous theory there are interesting questions remaining, as discussed in [24].

The theoretical analysis is validated by simulating the discrete cart-pendulum system on an incline. When dissipation is added, the inverted pendulum configuration is seen to be asymptotically stabilized, as predicted.

The discrete controlled dynamics is used to construct a real-time model predictive controller with piecewise constant control inputs. This serves to illustrate how discrete mechanics can

be naturally applied to yield digital controllers for mechanical systems.

The paper is organized as follows: In Sections II and III we review discrete mechanics and the method of controlled Lagrangians for stabilization of equilibria of mechanical systems. The discrete version of the potential shaping procedure and related stability analysis are discussed in Section IV. The theory is illustrated with the discrete cart-pendulum system in Section V. Simulations and the construction of the digital controller are presented in Sections VI and VII.

In a future publication we intend to treat discrete systems with nonabelian symmetries as well as systems with nonholonomic constraints.

II. An Overview of Discrete Mechanics

A discrete analogue of Lagrangian mechanics can be obtained by considering a discretization of Hamilton's principle; this approach underlies the construction of variational integrators. See Marsden and West [20], and references therein, for a more detailed discussion of discrete mechanics.

Consider a Lagrangian mechanical system with configuration manifold Q and Lagrangian $L : TQ \to \mathbb{R}$. A key notion is that of a *discrete Lagrangian*, which is a map $L^d : Q \times Q \to \mathbb{R}$ that approximates the action integral along an exact solution of the Euler–Lagrange equations joining the configurations $q_k, q_{k+1} \in Q$,

$$L^d(q_k, q_{k+1}) \approx \operatorname*{ext}_{q \in \mathcal{C}([0,h],Q)} \int_0^h L(q, \dot{q})\, dt, \tag{1}$$

where $\mathcal{C}([0, h], Q)$ is the space of curves $q : [0, h] \to Q$ with $q(0) = q_k$, $q(h) = q_{k+1}$, and ext denotes extremum.

In the discrete setting, the action integral of Lagrangian mechanics is replaced by an action sum

$$S^d(q_0, q_1, \ldots, q_N) = \sum_{k=0}^{N-1} L^d(q_k, q_{k+1}),$$

where $q_k \in Q$, $k = 0, 1, \ldots, N$, is a finite sequence of points in the configuration space. The equations are obtained by the discrete Hamilton principle, which extremizes the discrete action given fixed endpoints q_0 and q_N. Taking the extremum over $q_1, \ldots, q_{N-1}$ gives the *discrete Euler–Lagrange equations*

$$D_1 L^d(q_k, q_{k+1}) + D_2 L^d(q_{k-1}, q_k) = 0,$$

for $k = 1, \ldots, N - 1$. This implicitly defines the update map $\Phi : Q \times Q \to Q \times Q$, where $\Phi(q_{k-1}, q_k) = (q_k, q_{k+1})$ and $Q \times Q$ replaces the phase space TQ of Lagrangian mechanics.

Since we are concerned with control, we need to consider the effect of external forces on Lagrangian systems. In the context of discrete mechanics, this is addressed by introducing the *discrete Lagrange–d'Alembert principle* (see Kane, Marsden, Ortiz, and West [17]), which states that

$$\delta \sum_{k=0}^{n-1} L^d\left(q_k, q_{k+1}\right) + \sum_{k=0}^{n-1} F^d\left(q_k, q_{k+1}\right) \cdot \left(\delta q_k, \delta q_{k+1}\right) = 0$$

for all variations $\boldsymbol{\delta q}$ of $\boldsymbol{q}$ that vanish at the endpoints. Here, $\boldsymbol{q}$ denotes the vector of positions $(q_0, q_1, \ldots, q_N)$, and $\boldsymbol{\delta q} = (\delta q_0, \delta q_1, \ldots, \delta q_N)$, where $\delta q_k \in T_{q_k} Q$. The discrete one-form F^d on $Q \times Q$ approximates the impulse integral between the points q_k and q_{k+1}, just as the discrete Lagrangian L^d approximates the action integral. We define the maps $F_1^d, F_2^d : Q \times Q \to T^*Q$ by the relations

$$F_2^d\left(q_0, q_1\right) \delta q_1 := F^d\left(q_0, q_1\right) \cdot \left(0, \delta q_1\right),$$
$$F_1^d\left(q_0, q_1\right) \delta q_0 := F^d\left(q_0, q_1\right) \cdot \left(\delta q_0, 0\right).$$

The discrete Lagrange–d'Alembert principle may then be rewritten as

$$\delta \sum_{k=0}^{n-1} L^d\left(q_k, q_{k+1}\right) + \sum_{k=0}^{n-1} \left[F_1^d\left(q_k, q_{k+1}\right) \delta q_k + F_2^d\left(q_k, q_{k+1}\right) \delta q_{k+1}\right] = 0$$

for all variations $\boldsymbol{\delta q}$ of $\boldsymbol{q}$ that vanish at the endpoints. This is equivalent to the *forced discrete Euler–Lagrange equations*

$$D_1 L^d\left(q_k, q_{k+1}\right) + D_2 L^d\left(q_{k-1}, q_k\right) + F_1^d\left(q_k, q_{k+1}\right) + F_2^d\left(q_{k-1}, q_k\right) = 0.$$

III. Matching and Controlled Lagrangians

A. Controlled Euler–Lagrange Equations

This paper focuses on systems with one shape and one group degree of freedom. It is further assumed that the configuration space Q is the direct product of a one-dimensional shape space S and a one-dimensional Lie group G.

The configuration variables are written as $q = (\phi, s)$, with $\phi \in S$, and $s \in G$. The velocity phase space, TQ, has coordinates $(\phi, s, \dot{\phi}, \dot{s})$. The Lagrangian is the kinetic minus potential energy

$$L(q, \dot{q}) = \tfrac{1}{2}\left[\alpha \dot{\phi}^2 + 2\beta(\phi)\dot{\phi}\dot{s} + \gamma \dot{s}^2\right] - V(q), \tag{2}$$

with G-invariant kinetic energy. The corresponding controlled Euler–Lagrange dynamics is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = 0, \tag{3}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{s}} = u, \tag{4}$$

where u is the control input.

B. Continuous-Time Kinetic Shaping

Assume that the potential energy is G-invariant, *i.e.,* $V(q) = V(\phi)$, and that the *relative equilibria* $\phi = \phi_e$, $\dot{s} =$ const are unstable and given by non-degenerate critical points of $V(\phi)$. To stabilize the relative equilibria $\phi = \phi_e$, $\dot{s} =$ const with respect to ϕ, kinetic shaping is used. The controlled Lagrangian in this case is defined by

$$L_{\tau,\sigma}(q, \dot{q}) = L(\phi, \dot{\phi}, \dot{s} + \tau(\phi)\dot{\phi}) + \tfrac{1}{2}\sigma\gamma(\tau(\phi)\dot{\phi})^2, \tag{5}$$

where $\tau(\phi) = \kappa\beta(\phi)$. This velocity shift corresponds to a new choice of the horizontal space (see [8] for details). The

dynamics is just the Euler–Lagrange dynamics for controlled Lagrangian (5),

$$\frac{d}{dt}\frac{\partial L_{\tau,\sigma}}{\partial \dot\phi} - \frac{\partial L_{\tau,\sigma}}{\partial \phi} = 0, \tag{6}$$

$$\frac{d}{dt}\frac{\partial L_{\tau,\sigma}}{\partial \dot s} = 0. \tag{7}$$

Lagrangian (5) satisfies the simplified matching conditions of [9] when the kinetic energy metric coefficient γ in (2) is constant.

Setting $u = -d\big(\gamma\tau(\phi)\dot\phi\big)/dt$ defines the control input, makes equations (4) and (7) identical, and results in controlled momentum conservation by dynamics (3) and (4). Setting $\sigma = -1/\gamma\kappa$ makes equations (3) and (6) reduced on the controlled momentum level identical.

A very interesting feature of systems (3), (4) and (6), (7) is that the *reduced* dynamics are the same on all momentum levels, which follows from the independence of equations (3) and (6) of the group velocity $\dot s$. We will see in Section IV that this property does not hold in the discrete setting, and one has to carefully select the momentum levels when performing discrete kinetic shaping.

C. Continuous-Time Potential Shaping

Now, consider the case when the kinetic energy is group invariant, but the potential is not. Consider the special case when the potential energy is $V(q) = V_1(\phi)+V_2(s)$ with $V_1(\phi)$ having a local non-degenerate maximum at ϕ_e, and the goal is to stabilize the *equilibrium* $\phi = \phi_e$, $s = s_e$. As it becomes necessary to shape the potential energy as well, the controlled Lagrangian is defined by the formula

$$L_{\tau,\sigma,\rho,\varepsilon}(\phi, s, \dot\phi, \dot s) = L(\phi, s, \dot\phi, \dot s + \tau(\phi)\dot\phi) + \tfrac{1}{2}\sigma\gamma(\tau(\phi)\dot\phi)^2 + \tfrac{1}{2}(\rho-1)\gamma(\dot s + (\sigma-1)\tau(\phi)\dot\phi)^2 + V_2(s) - V_\varepsilon(y), \tag{8}$$

where

$$y = s - \int_{\phi_e}^{\phi} \frac{1}{\gamma}\left(\frac{1}{\sigma} - \frac{\rho-1}{\rho}\right)\beta(z)\,dz, \tag{9}$$

$V_\varepsilon(y)$ is an arbitrary negative-definite function, and (ϕ_e, s_e) is the equilibrium of interest.

Below we assume that $\phi_e = 0$, which can be always accomplished by an appropriate choice of local coordinates for each (relative) equilibrium.

IV. Discrete Shaping

A. Discrete Controlled Dynamics

In discretizing the method of controlled Lagrangians, we combine formulae (1), (2), and (5). In the rest of this paper, we will adopt the notations

$$q_{k+1/2} = \frac{q_k + q_{k+1}}{2}, \quad \Delta q_k = q_{k+1} - q_k, \quad q_k = (\phi_k, s_k).$$

This allows us to construct a *second-order accurate* discrete Lagrangian

$$L^d(q_k, q_{k+1}) = hL(q_{k+1/2}, \Delta q_k/h).$$

Thus, for a system with one shape and one group degree of freedom the discrete Lagrangian is given by the formula

$$L^d(q_k, q_{k+1}) = \frac{h}{2}\Big[\alpha\Big(\frac{\Delta\phi_k}{h}\Big)^2 + 2\beta(\phi_{k+1/2})\frac{\Delta\phi_k}{h}\frac{\Delta s_k}{h} + \gamma\Big(\frac{\Delta s_k}{h}\Big)^2\Big] - hV(q_{k+1/2}). \tag{10}$$

The discrete dynamics is governed by the equations

$$\frac{\partial L^d(q_k, q_{k+1})}{\partial \phi_k} + \frac{\partial L^d(q_{k-1}, q_k)}{\partial \phi_k} = 0, \tag{11}$$

$$\frac{\partial L^d(q_k, q_{k+1})}{\partial s_k} + \frac{\partial L^d(q_{k-1}, q_k)}{\partial s_k} = -u_k, \tag{12}$$

where u_k is the control input.

B. Kinetic Shaping

At first, it will be assumed that the potential energy is G-invariant, *i.e.,* $V(q) = V(\phi)$, and that relative equilibria $\phi_k = 0, \Delta s_k = \text{const}$ of (11) and (12) in the absence of control input are unstable. We will see that one needs to either appropriately select the momentum levels or introduce a new parameter into the controlled Lagrangian to complete the matching procedure.

Motivated by the continuous-time matching procedure (see Section III), we define the discrete controlled Lagrangian by the formula

$$L^d_{\tau,\sigma}(q_k, q_{k+1}) = hL_{\tau,\sigma}(q_{k+1/2}, \Delta q_k/h) = h\Big[L\big(\phi_{k+\frac{1}{2}}, \Delta\phi_k/h, \Delta s_k/h + \kappa\beta(\phi_{k+\frac{1}{2}})\Delta\phi_k/h\big) + \frac{\sigma\gamma}{2}\Big(\kappa\beta(\phi_{k+\frac{1}{2}})\Delta\phi_k/h\Big)^2\Big]. \tag{13}$$

where $L_{\tau,\sigma}(q,\dot q)$ is the continuous-time controlled Lagrangian. The dynamics associated with (13) is

$$\frac{\partial L^d_{\tau,\sigma}(q_k, q_{k+1})}{\partial \phi_k} + \frac{\partial L^d_{\tau,\sigma}(q_{k-1}, q_k)}{\partial \phi_k} = 0, \tag{14}$$

$$\frac{\partial L^d_{\tau,\sigma}(q_k, q_{k+1})}{\partial s_k} + \frac{\partial L^d_{\tau,\sigma}(q_{k-1}, q_k)}{\partial s_k} = 0. \tag{15}$$

Equation (15) is equivalent to the *discrete controlled momentum conservation:*

$$p_k = \mu, \tag{16}$$

where

$$p_k = -\frac{\partial}{\partial s_k}L^d_{\tau,\sigma}(q_k, q_{k+1}) = \frac{(1+\gamma\kappa)\beta(\phi_{k+1/2})\Delta\phi_k + \gamma\Delta s_k}{h}. \tag{17}$$

Setting

$$u_k = -\frac{\gamma\Delta\phi_k\tau(\phi_{k+1/2}) - \gamma\Delta\phi_{k-1}\tau(\phi_{k-1/2})}{h} \tag{18}$$

makes equations (12) and (15) identical and allows one to represent the discrete momentum equation (12) as the discrete momentum conservation law

$$p_k = p. \tag{19}$$

Theorem 1: The dynamics determined by equations (11) *and* (12) *restricted to the momentum level $p_k = p$ is equivalent to the dynamics of equations* (14) *and* (15) *restricted to the momentum level $p_k = \mu$ if and only if the matching conditions*

$$\sigma = -\frac{1}{\gamma\kappa}, \qquad \mu = \frac{p}{1+\gamma\kappa}. \tag{20}$$

hold.

Proof: Solve equations (16) and (19) for Δs_k and substitute the solutions in equations (11) and (14), respectively. This process is a simple version of discrete reduction [14]. A computation shows that the equations obtained this way are equivalent if and only if

$$h\Bigg[\frac{\mu - p + \gamma\kappa\mu}{\gamma}\frac{\partial}{\partial\phi_k}\Big(\beta(\phi_{k+1/2})\frac{\Delta\phi_k}{h} + \beta(\phi_{k-1/2})\frac{\Delta\phi_{k-1}}{h}\Big) + \frac{\kappa+\gamma\sigma\kappa^2}{2}\frac{\partial}{\partial\phi_k}\Big(\beta^2(\phi_{k+1/2})\Big(\frac{\Delta\phi_k}{h}\Big)^2 + \beta^2(\phi_{k-1/2})\Big(\frac{\Delta\phi_{k-1}}{h}\Big)^2\Big)\Bigg] = 0. \tag{21}$$

Since $\beta(\phi) \neq 0$ and $\Delta\phi_k \neq 0$ generically, equations (11) and (12) are equivalent if and only if

$$\mu - p + \gamma\kappa\mu = 0, \qquad \kappa + \gamma\sigma\kappa^2,$$

which is equivalent to (20) Note that the momentum levels p and μ *are not* the same. ∎

Remark. As $h \to 0$, formulae (18) and (21) become

$$u = -\frac{d}{dt}\big(\gamma\tau(\phi)\dot\phi\big)$$

and

$$-(\kappa+\gamma\sigma\kappa^2)(\beta^2(\phi)\ddot\phi + \beta(\phi)\beta'(\phi)\dot\phi^2) = 0,$$

respectively. That is, as $h \to 0$, one recovers the continuous-time control input and the continuous-time matching condition, $\sigma = -1/\gamma\kappa$. Condition $\mu = p/(1+\gamma\kappa)$ becomes redundant after taking the limit, *i.e.,* the reduced dynamics can be matched on arbitrary momentum levels in the continuous-time case, which agrees with observations made in Section III.

We now discuss an alternative matching procedure. Define the discrete controlled Lagrangian $\Lambda^d_{\tau,\sigma,\lambda}(q_k, q_{k+1})$ by the formula

$$h\Big[L\big(\phi_{k+\frac12}, \Delta\phi_k/h, \Delta s_k/h + \kappa\beta(\phi_{k+\frac12})\Delta\phi_k/h\big) + \frac{\sigma\gamma}{2}\Big(\kappa\beta(\phi_{k+\frac12})\Delta\phi_k/h\Big)^2 + \lambda\kappa\beta(\phi_{k+\frac12})\Delta\phi_k/h\Big].$$

The discrete dynamics associated with this Lagrangian is

$$\frac{\partial\Lambda^d_{\tau,\sigma,\lambda}(q_k,q_{k+1})}{\partial\phi_k} + \frac{\partial\Lambda^d_{\tau,\sigma,\lambda}(q_{k-1},q_k)}{\partial\phi_k} = 0, \tag{22}$$

$$\frac{\partial\Lambda^d_{\tau,\sigma}(q_k,q_{k+1})}{\partial s_k} + \frac{\partial\Lambda^d_{\tau,\sigma}(q_{k-1},q_k)}{\partial s_k} = 0. \tag{23}$$

The discrete controlled momentum is given by formula

$$p_k = -\frac{\partial}{\partial s_k}\Lambda^d_{\tau,\sigma,\lambda}(q_k,q_{k+1}) = \frac{(1+\gamma\kappa)\beta(\phi_{k+1/2})\Delta\phi_k + \gamma\Delta s_k}{h} \tag{24}$$

and equation (23) is equivalent to the discrete momentum conservation (19).

Theorem 2: The dynamics (11) *and* (12) *restricted to the momentum level $p_k = p$ is equivalent to the dynamics* (22) *and* (23) *restricted to the same momentum level if and only if the matching conditions*

$$\sigma = -\frac{1}{\gamma\kappa}, \qquad \lambda = -p. \tag{25}$$

hold.

Proof: Similar to the proof of Theorem 1, solve equation (19) for Δs_k and substitute the solution in equations (11) and (22), respectively. A computation shows that the equations obtained this way are equivalent if and only if

$$h\Bigg[(\kappa p + \kappa\lambda)\frac{\partial}{\partial\phi_k}\Big(\beta(\phi_{k+1/2})\frac{\Delta\phi_k}{h} + \beta(\phi_{k-1/2})\frac{\Delta\phi_{k-1}}{h}\Big) + \frac{\kappa+\gamma\sigma\kappa^2}{2}\frac{\partial}{\partial\phi_k}\Big(\beta^2(\phi_{k+1/2})\Big(\frac{\Delta\phi_k}{h}\Big)^2 + \beta^2(\phi_{k-1/2})\Big(\frac{\Delta\phi_{k-1}}{h}\Big)^2\Big)\Bigg] = 0, \tag{26}$$

which implies (25). Note that in this case we add an extra term to the controlled Lagrangian which eliminates the need for adjusting the momentum level. ∎

Remark. The ratio $\Lambda^d_{\tau,\sigma,\lambda}/h$ becomes $L^d_{\tau,\sigma} + \lambda\kappa\beta(\varphi)\dot\varphi$ as $h \to 0$. That is, as we let the time step go to 0, we obtain the continuous-time controlled Lagrangian modified by a term which is a derivative of the function $\lambda\kappa\int\beta(\phi)\,d\phi$ with respect to time. It is well-known that adding such a derivative term to a Lagrangian does not change the dynamics associated with this Lagrangian.

The stability properties of the relative equilibria $\phi_k = 0$, $s_k = \text{const}$ of equations (11) and (12) are now investigated.

Theorem 3: The relative equilibria $\phi_k = 0$, $\Delta s_k = const$ of equations (11) *and* (12)*, with u_k defined by* (18)*, are* ***spectrally stable*** *if*

$$\kappa > \frac{\alpha\gamma - \beta^2(0)}{\beta^2(0)\,\gamma}. \tag{27}$$

Proof: Let $V''(0) = -C$, where $C > 0$ (see Section III). The linearization of the reduced dynamics (11) and (12) at $\phi = 0$ is computed to be

$$\frac{\alpha\gamma - \beta^2(0) - \beta^2(0)\gamma\kappa}{h^2\gamma}(\Delta\phi_{k-1} - \Delta\phi_k) + \frac{C}{4}(\phi_{k-1} + 2\phi_k + \phi_{k+1}) = 0. \tag{28}$$

Observe that the value of p does not affect the linearized dynamics.

The linearized dynamics preserves the quadratic approximation of the discrete energy

$$\frac{\alpha\gamma - \beta^2(0) - \beta^2(0)\gamma\kappa}{2\gamma}\Big(\frac{\Delta\phi_k}{h^2}\Big)^2 - \frac{C}{2}\phi^2_{k+1/2}. \tag{29}$$

The equilibrium $\phi_k = 0$ of (28) is stable if and only if the function (29) is negative-definite at $\phi_k = \phi_{k+1} = 0$. The latter requirement is equivalent to condition (27). ∎

Remark. The stability condition (27) is identical to the stability condition of the continuous-time cart-pendulum system, and it can be rewritten as

$$-\frac{\beta^2(0)}{\alpha\gamma-\beta^2(0)}<\sigma<0.$$

The spectrum of the linear map $(\phi_{k-1},\phi_k)\mapsto(\phi_k,\phi_{k+1})$ defined by (28) belongs to the unit circle. Spectral stability in this situation is not sufficient to conclude nonlinear stability.

We now modify the control input (18) by adding the *kinetic discrete dissipation-emulating term*

$$\frac{D(\Delta\phi_{k-1}+\Delta\phi_k)}{2h}$$

in order to achieve the asymptotic stabilization of the upward position of the pendulum. In the above, D is a positive constant. The discrete momentum conservation law becomes

$$p_k-\frac{D\,\phi_{k+1/2}}{h}=p.$$

Straightforward calculation shows that the spectrum of the matrix of the linear map $(\phi_{k-1},\phi_k)\mapsto(\phi_k,\phi_{k+1})$ defined by the reduced discrete dynamics belongs to the open unit disc. This implies that the equilibrium $\phi=0$ is asymptotically stable.

C. Potential Shaping

Recall that the the discrete dynamics associated with discrete Lagrangian (10) is governed by equations (11) and (12), where u_k is the control input. The goal of the procedure developed in this section is to stabilize the equilibrium $(\phi,s)=(0,0)$ of (11) and (12).

Motivated by (8), we define the second-order accurate discrete controlled Lagrangian by the formula

$$L^d_{\tau,\sigma,\rho,\varepsilon}(q_k,q_{k+1})=hL_{\tau,\sigma,\rho,\varepsilon}(q_{k+1/2},\Delta q_k/h),\qquad(30)$$

where $q_k=(\phi_k,s_k)$.

The dynamics associated with (30) is amended by the term w_k in the discrete shape equation:

$$\frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_k,q_{k+1})}{\partial\phi_k}+\frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1},q_k)}{\partial\phi_k}=-w_k,\qquad(31)$$

$$\frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_k,q_{k+1})}{\partial s_k}+\frac{\partial L^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1},q_k)}{\partial s_k}=0.\qquad(32)$$

This term w_k is important for matching systems (11), (12) and (31), (32). *The presence of the terms w_k represents an interesting (but manageable) departure from the continuous theory.* Let

$$J_k=\rho\gamma\big(\Delta s_k/h-(\sigma-1)\tau(\phi_{k+\frac12})\Delta\phi_k/h\big).$$

The following statement is proved by a straightforward calculation:

Theorem 4: The dynamics (11)*,* (12) *is equivalent to the dynamics* (31)*,* (32) *if and only if u_k and w_k are given by*

$$\begin{aligned}u_k=&\frac{h}{2}\Big[V_2'(s_{k+\frac12})+V_2'(s_{k-\frac12})\Big]\\&-\frac{h}{2\rho}\Big[V_\varepsilon'(s_{k+\frac12})+V_\varepsilon'(s_{k-\frac12})\Big]\\&-\frac{\gamma\Delta\phi_k\tau(\phi_{k+1/2})-\gamma\Delta\phi_{k-1}\tau(\phi_{k-1/2})}{h},\end{aligned}\qquad(33)$$

and

$$\begin{aligned}w_k=&\Big(1-\sigma+\frac{\sigma}{\rho}\Big)\Big(\tau(\phi_{k+\frac12})\Big[-\gamma\rho J_k+\frac{h}{2}V_\varepsilon'(y_{k+\frac12})\Big]\\&+\tau(\phi_{k-\frac12})\Big[\gamma\rho J_{k-1}+\frac{h}{2}V_\varepsilon'(y_{k-\frac12})\Big]\\&-\tau'(\phi_{k+\frac12})J_k\Delta\phi_k-\tau'(\phi_{k-\frac12})J_{k-1}\Delta\phi_{k-1}\Big),\end{aligned}$$

where y_k is obtained by substituting ϕ_k and s_k in formula (9).

Remark. Equations (11), (12) define closed-loop dynamics when u_k is given by formula (33). The terms w_k vanish when $\beta(\phi)=$ const as they become proportional to the left-hand side of equation (32).

As in the case of kinetic shaping, the stability analysis is done by means of an analysis of the spectrum of the linearized discrete equations. We assume that the equilibrium to be stabilized is $(\phi_k,s_k)=(0,0)$.

Theorem 5: The equilibrium $(\phi_k,s_k)=(0,0)$ of equations (31) *and* (32) *is* ***spectrally stable*** *if*

$$-\frac{\beta^2(0)}{\alpha\gamma-\beta^2(0)}<\sigma<0,\quad\rho<0,\quad\textit{and}\quad V_\varepsilon''(0)<0.\qquad(34)$$

Proof: The linearized discrete equations are

$$\frac{\partial\mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k,q_{k+1})}{\partial\phi_k}+\frac{\partial\mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1},q_k)}{\partial\phi_k}=0,\qquad(35)$$

$$\frac{\partial\mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k,q_{k+1})}{\partial s_k}+\frac{\partial\mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1},q_k)}{\partial s_k}=0,\qquad(36)$$

where $\mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k,q_{k+1})$ is the quadratic approximation of $L^d_{\tau,\sigma,\rho,\varepsilon}$ at the equilibrium (*i.e.*, $\beta(\phi)$, $V_1(\phi)$, and $V_\varepsilon(y)$ in $L^d_{\tau,\sigma,\rho,\varepsilon}$ are replaced by $\beta(0)$, $\frac12V_1''(0)\phi^2$, and $\frac12V_\varepsilon''(0)y^2$, respectively). *Note the absence of the term w_k in equation* (35)*.*

The linearized dynamics preserves the quadratic approximation of the discrete energy $E_{k,k+1}$ defined by

$$\begin{aligned}&\frac{\alpha\gamma\sigma^2-\beta(0)^2(\sigma-1)(\rho(\sigma-1)-\sigma)}{2\gamma\sigma^2h}\Delta\phi_k^2\\&\quad+\frac{\beta(0)\rho(\sigma-1)}{\sigma h}\Delta\phi_k\Delta s_k+\frac{\gamma\rho}{2h}\Delta s_k^2\\&\qquad+\frac{h}{2}V_1''(0)\phi_{k+\frac12}^2+\frac{h}{2}V_\varepsilon''(0)x_{k+\frac12}^2,\end{aligned}\qquad(37)$$

where

$$x=s+\left(\frac{\rho-1}{\rho}-\frac{1}{\sigma}\right)\frac{\beta(0)}{\gamma}\,\phi.\qquad(38)$$

Since $V_1''(0)$ is negative, the equilibrium $(\phi_k,s_k)=(0,0)$ of equations (35) and (36) is stable if the quadratic approximation of the discrete controlled energy (37) is negative-definite. The latter requirement is equivalent to conditions (34). The spectrum of the linearized discrete dynamics in this case belongs to the unit circle. ■

Remarks. Spectral stability in this situation is not sufficient to conclude nonlinear stability. The stability conditions (34) are identical to the stability conditions of the corresponding continuous-time system.

Following [9], we now modify the control input (33) by adding the *potential discrete dissipation-emulating term*

$$\frac{D(\Delta y_{k-1}+\Delta y_k)}{2h}\qquad(39)$$

in order to achieve the asymptotic stabilization of the equilibrium $(\phi_k, s_k) = (0, 0)$. In the above, D is a constant. The linearized discrete dynamics becomes

$$\frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})}{\partial \phi_k} + \frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1}, q_k)}{\partial \phi_k} = -\left(\frac{\rho - 1}{\rho} - \frac{1}{\sigma}\right)\frac{\beta(0)}{\gamma}\frac{D(\Delta x_{k-1} + \Delta x_k)}{2h}, \quad (40)$$

$$\frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_k, q_{k+1})}{\partial s_k} + \frac{\partial \mathcal{L}^d_{\tau,\sigma,\rho,\varepsilon}(q_{k-1}, q_k)}{\partial s_k} = -\frac{D(\Delta x_{k-1} + \Delta x_k)}{2h}, \quad (41)$$

where x_k is obtained by substituting ϕ_k and s_k in formula (38).

Theorem 6: The equilibrium $(\phi_k, s_k) = (0, 0)$ *of equations* (40) *and* (41) *is asymptotically stable if conditions* (27) *are satisfied and* D *is positive.*

Proof: Multiplying equations (40) and (41) by $(\Delta\phi_{k-1} + \Delta\phi_k)/2$ and $(\Delta s_{k-1} + \Delta s_k)/2$, respectively, we obtain

$$E_{k,k+1} = E_{k-1,k} + \frac{Dh}{2}\left(\frac{\Delta x_{k-1} + \Delta x_k}{2h}\right)^2,$$

where $E_{k,k+1}$ is the quadratic approximation of the discrete energy (29). Recall that $E_{k,k+1}$ is negative-definite (see the proof of Theorem 5). It is possible to show that, in some neighborhood of $(\phi_k, s_k) = (0, 0)$, the quantity $\Delta x_{k-1} + \Delta x_k \not\equiv 0$ along a solution of equations (40) and (41) unless this solution is the equilibrium $(\phi_k, s_k) = (0, 0)$. Therefore, $E_{k,k+1}$ increases along non-equilibrium solutions of (40) and (41). Since equations (40) and (41) are linear, this is only possible if the spectrum of (40) and (41) is inside the open unit disk, which implies asymptotic stability of the equilibrium of both linear system (40) and (41) and nonlinear system (11) and (12) with potenital discrete dissipation-emulating term (39) added to u_k. ■

V. Stabilization of the Discrete Pendulum on the Cart

A basic example treated in earlier papers in the smooth setting is the *pendulum on a cart*. Let s denote the position of the cart on the s-axis, ϕ denote the angle of the pendulum with the upright vertical, and ψ denote the elevation angle of the incline, as in Figure 1. The configuration space for this system is $Q = S \times G = S^1 \times \mathbb{R}$, with the first factor being the pendulum angle ϕ and the second factor being the cart position s. The symmetry group G of the kinetic energy of the pendulum-cart system is that of translation in the s variable, so $G = \mathbb{R}$.

The length of the pendulum is l, the mass of the pendulum is m and that of the cart is M.

For the cart-pendulum system, α, $\beta(\phi)$, γ, are given by

$$\alpha = ml^2, \quad \beta(\phi) = ml\cos(\phi - \psi), \quad \gamma = M + m. \quad (42)$$

The potential energy is $V(q) = V_1(\phi) + V_2(s)$, where

$$V_1(\phi) = -mgl\cos\phi, \quad V_2(s) = -\gamma g s \sin\psi.$$

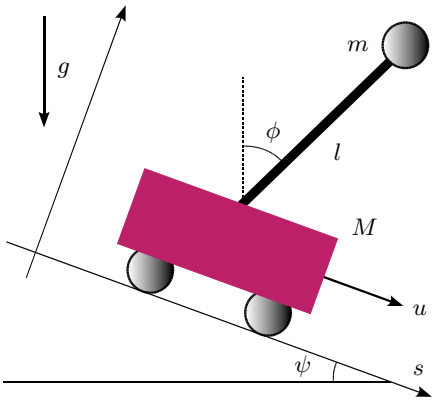


Fig. 1. The pendulum on a cart going down an inclined plane under gravity. The control force is in the direction s, the overall motion of the cart.

Note that $\alpha\gamma - \beta^2(\phi) > 0$, and that the potential energy becomes G-invariant when the plane is horizontal, *i.e.,* when $\psi = 0$.

Since the Lagrangian for the cart-pendulum system is of the form (2), the discrete control laws (18) and (33) stabilize the upward vertical equilibrium of the *pendulum*. As in the continuous-time setting, the *cart* is stabilized by symmetry-breaking controller (33) and is not stabilized by symmetry-preserving controller (18).

Simulations of the discrete cart-pendulum system are shown in the next section.

VI. Simulations

Simulating the behavior of the discrete controlled Lagrangian system involves viewing equations (11) and (32) as an implict update map $\Phi : (q_{k-2}, q_{k-1}) \mapsto (q_{k-1}, q_k)$. This presupposes that the initial conditions are given in the form (q_0, q_1); however it is generally preferable to specify the initial conditions as $(q_0, \dot{q}_0)$. This is achieved by solving the boundary condition

$$\frac{\partial L}{\partial \dot{q}}(q_0, \dot{q}_0) + D_1 L^d(q_0, q_1) + F_1^d(q_0, q_1) = 0$$

for q_1. Once the initial conditions are expressed in the form (q_0, q_1), the discrete evolution can be obtained using the implicit update map Φ.

We first consider the case of kinetic shaping on a level surface (with $\psi = 0$), when κ is twice the critical value, and without dissipation. Here, $h = 0.05\,\text{sec}$, $m = 0.14\,\text{kg}$, $M = 0.44\,\text{kg}$, and $l = 0.215\,\text{m}$. As shown in Figure 2, the ϕ dynamics is stabilized, but since there is no dissipation, the oscillations are sustained. The s dynamics exhibits both a drift and oscillations, as potential shaping is necessary to stabilize the translational dynamics.

When dissipation is added, the ϕ dynamics is asymptotically stabilized, as shown in Figure 3. However, even though the oscillations are damped, the s dynamics retains a drift motion, as expected.

We next consider the case of potential shaping on an inclined surface (with $\psi = \frac{\pi}{9}$ radians) without dissipation, with the other physical parameters as before. Here, our goal

is to regulate the cart at $s = 0$ and the pendulum at $\phi = 0$. We set $V_\varepsilon = -\frac{\varepsilon}{2}y^2$. The control gains are chosen to be $\kappa = 20$, $\rho = -0.02$, and $\varepsilon = 0.00001$. It is worth noting that the discrete dynamics remain bounded near the desired equilibrium, and this behavior persists even for significantly longer simulation runs involving 10^6 time-steps. To more clearly visualize the dynamics, we only include a 4000 time-step segment of this computation in Figure 4. The exceptional stability of the discrete controlled trajectory can presumably be understood in terms of the bounded energy oscillations characteristic of symplectic and variational integrators.

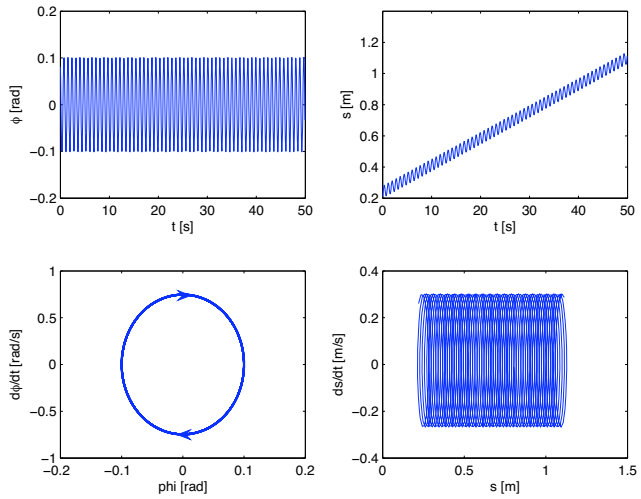


Fig. 2. Discrete controlled dynamics with kinetic shaping and without dissipation. The discrete controlled system stabilizes the ϕ motion about the equilibrium, but the s dynamics is not stabilized; since there is no dissipation, the oscillations are sustained.

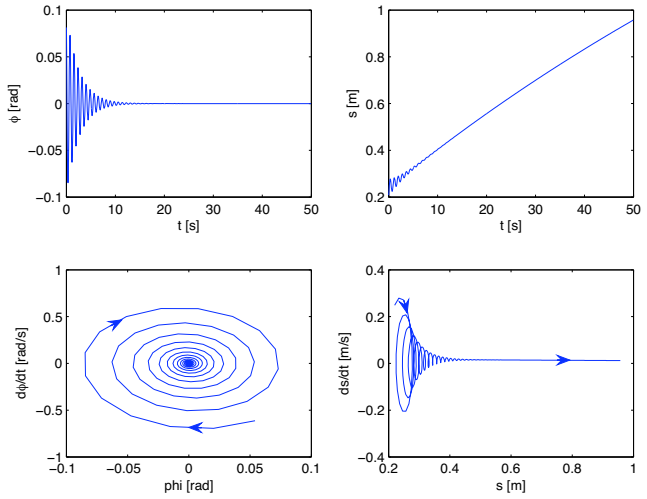


Fig. 3. Discrete controlled dynamics with kinetic shaping and dissipation. The discrete controlled system asymptotically stabilizes the ϕ motion about the equilibrium; since there is no potential shaping, the s dynamics is not stabilized, and there is a slow drift in s.

When dissipation is added, we obtain an asymptotically

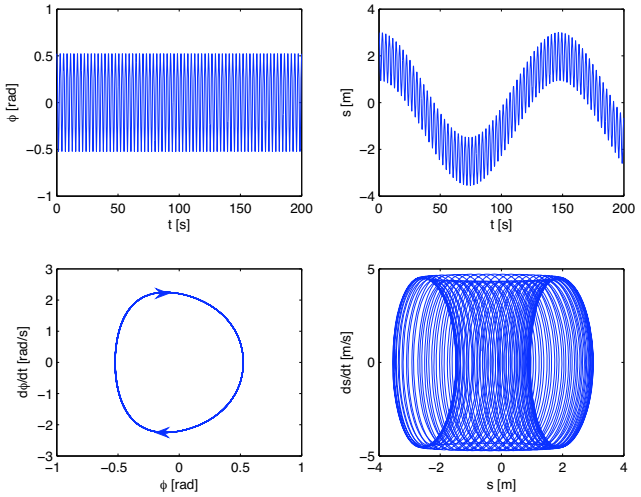


Fig. 4. Discrete controlled dynamics with potential shaping and without dissipation. The discrete controlled system stabilizes the motion about the equilibrium; since there is no dissipation, the oscillations are sustained.

stabilizing control law, as illustrated in Figure 5. This is consistent with the stability analysis of Section V.

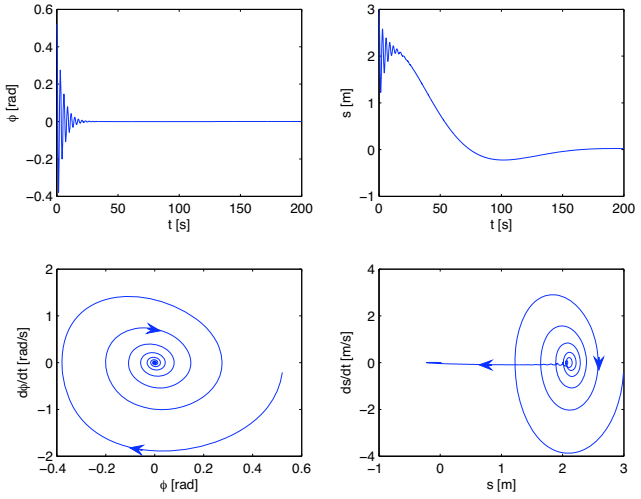


Fig. 5. Discrete controlled dynamics with potential shaping and dissipation. Here the oscillations die out and both the cart position and the pendulum angle converge to their desired values $s = 0$ and $\phi = 0$.

VII. Model Predictive Controller

We now explore the use of the forced discrete Euler–Lagrange equations as the model in a real-time model predictive controller, with piecewise constant control forces. Algorithm 1 below describes the details of the procedure.

The digital controller uses the position information it senses for $t = -2h, -h$ to estimate the positions at $t = 0, h$ during the time interval $t = [-h, 0]$. This allows it to compute a symmetric finite difference approximation to the continuous

Algorithm 1 DIGITAL CONTROLLER ($q(\cdot), T_f, h$)

$q_0 \leftarrow$ **sense** $q(0)$
$q_1 \leftarrow$ **sense** $q(h)$
$\bar{q}_2 \leftarrow$ **solve** $D_2L^d(q_0, q_1) + D_1L^d(q_1, \bar{q}_2) = 0$
$\bar{q}_3 \leftarrow$ **solve** $D_2L^d(q_1, \bar{q}_2) + D_1L^d(\bar{q}_2, \bar{q}_3) + F_1^d(\bar{q}_2, \bar{q}_3) = 0$

$u_{2+1/2} \leftarrow u\left(\frac{\bar{q}_2+\bar{q}_3}{2}, \frac{\bar{q}_3-\bar{q}_2}{h}\right)$
actuate $u = u_{2+1/2}$ for $t \in [2h, 3h]$
$q_2 \leftarrow$ **sense** $q(2h)$
$\bar{q}_3 \leftarrow$ **solve** $D_2L^d(q_1, q_2) + D_1L^d(q_2, \bar{q}_3) + F_1^d(q_2, \bar{q}_3) = 0$

$\bar{q}_4 \leftarrow$ **solve** $D_2L^d(q_2, \bar{q}_3) + D_1L^d(\bar{q}_3, \bar{q}_4)$
$\quad + F_2^d(q_2, \bar{q}_3) + F_1^d(\bar{q}_3, \bar{q}_4) = 0$
$u_{3+1/2} \leftarrow u\left(\frac{\bar{q}_3+\bar{q}_4}{2}, \frac{\bar{q}_4-\bar{q}_3}{h}\right)$
actuate $u = u_{3+1/2}$ for $t \in [3h, 4h]$
for $k = 4$ to $(T_f/h - 1)$ **do**
  $q_{k-1} \leftarrow$ **sense** $q((k-1)h)$
  $\bar{q}_k \leftarrow$ **solve** $D_2L^d(q_{k-2}, q_{k-1}) + D_1L^d(q_{k-1}, \bar{q}_k)$
  $\quad + F_2^d(q_{k-2}, q_{k-1}) + F_1^d(q_{k-1}, \bar{q}_k) = 0$
  $\bar{q}_{k+1} \leftarrow$ **solve** $D_2L^d(q_{k-1}, \bar{q}_k) + D_1L^d(\bar{q}_k, \bar{q}_{k+1})$
  $\quad + F_2^d(q_{k-1}, \bar{q}_k) + F_1^d(\bar{q}_k, \bar{q}_{k+1}) = 0$
  $u_{k+1/2} \leftarrow u\left(\frac{\bar{q}_k+\bar{q}_{k+1}}{2}, \frac{\bar{q}_{k+1}-\bar{q}_k}{h}\right)$
  actuate $u = u_{k+1/2}$ for $t \in [kh, (k+1)h]$
end for

control force $u(\phi, s, \dot{\phi}, \dot{s})$ at $t = h/2$ using the approximation

$$u_{1/2} = u\left(\frac{\bar{\phi}_0 + \bar{\phi}_1}{2}, \frac{\bar{s}_0 + \bar{s}_1}{2}, \frac{\bar{\phi}_1 - \bar{\phi}_0}{h}, \frac{\bar{s}_1 - \bar{s}_0}{h}\right),$$

where the overbar indicates that the position variable is being estimated by the numerical model. This control is then applied as a constant control input for the time interval $[0, h]$. This algorithm can be implemented in real-time if the two forward solves can be computed within the time interval h. On a 2.5 GHz PowerPC G5 running MATLAB, two forward solves take 631.2 μsec, which is sufficiently fast to drive a digital controller with a frequency in excess of 1.5 kHz. In our simulation, the digital controller had a frequency of 20 Hz, which involves a computational load that is easily accommodated by an embedded controller.

The initialization of the discrete controller is somewhat involved, since the system is unforced during the time interval $[0, 2h]$ while the controller senses the initial states, and computes the appropriate control forces. Consequently, a combination of the forced and unforced discrete Euler–Lagrange equations are used to predict the initial evolution of the system.

We present the numerical simulation results for the digital controller in both the case of kinetic shaping (Figure 6) and potential shaping (Figure 7). We see that in the case of kinetic shaping, the system is asymptotically stabilized in only the ϕ variable, and the s dynamics exhibits a drift, whereas in the case of potential shaping, the system is asymptotically stabilized in both the ϕ and s variables. Notice that the use of a piecewise constant control introduces dissipation-like effects, which are reduced as the time-step is decreased.

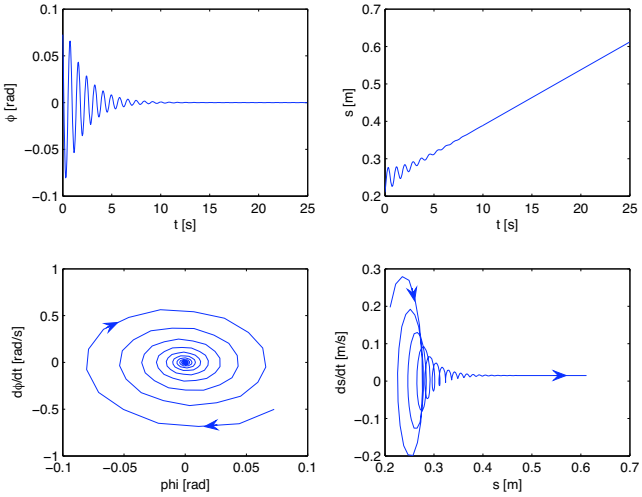


Fig. 6. The discrete real-time piecewise constant model predictive controller with kinetic shaping stabilizes ϕ to zero, but not s.

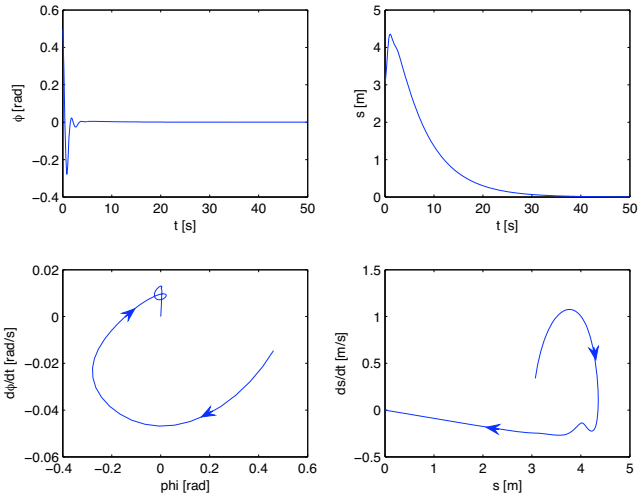


Fig. 7. The discrete real-time piecewise constant model predictive controller with potential shaping stabilizes ϕ and s to zero.

VIII. CONCLUSIONS

In this paper we have introduced potential shaping techniques for discrete systems and have shown that these lead to an effective numerical implementation for stabilization in the case of the discrete cart-pendulum model. The method in this paper is related to other discrete methods in control that have a long history; recent papers that use discrete mechanics in the context of optimal control and celestial navigation are [11], [15], [16], and [23]. The method of discrete controlled Lagrangians for systems with higher-dimensional configuration space and with non-commutative symmetry will be developed in a forthcoming paper.

IX. Acknowledgments

The research of AMB was supported by NSF grants DMS-0305837, DMS-0604307, and CMS-0408542. The research of ML was partially supported by NSF grant DMS-0504747 and a University of Michigan Rackham faculty grant. The research of JEM was partially supported by AFOSR Contract FA9550-05-1-0343. The research of DVZ was partially supported by NSF grants DMS-0306017 and DMS-0604108.

References

[1] Auckly, D., L. Kapitanski, & W. White, Control of Nonlinear Underactuated Systems, *Commun. Pure Appl. Math.* **53**, 2000, 354–369.

[2] Bloch, A. M., *Nonholonomic Mechanics and Control*, Interdisciplinary Appl. Math. **24**, Springer-Verlag, 2003.

[3] Bloch, A. M., M. Leok, J. E. Marsden, and D. V. Zenkov [2005], Controlled Lagrangians and Stabilization of the Discrete Cart-Pendulum System, *Proc. CDC* **44**, 6579–6584.

[4] Bloch, A. M., M. Leok, J. E. Marsden, and D. V. Zenkov [2006], Controlled Lagrangians and Potential Shaping for Stabilization of Discrete Mechanical Systems, *Proc. CDC* **45**, ?–?.

[5] Bloch, A. M., N. Leonard, & J. E. Marsden, Stabilization of Mechanical Systems Using Controlled Lagrangians, *Proc. CDC* **36**, 1997, 2356–2361.

[6] Bloch, A. M., N. Leonard, & J. E. Marsden, Matching and Stabilization by the Method of Controlled Lagrangians, *Proc. CDC* **37**, 1998, 1446–1451.

[7] Bloch, A. M., N. Leonard, & J. E. Marsden, Potential Shaping and the Method of Controlled Lagrangians, *Proc. CDC* **38**, 1999, 1652–1657.

[8] Bloch, A. M., N. E. Leonard, & J. E. Marsden, Controlled Lagrangians and the Stabilization of Mechanical Systems I: The First Matching Theorem, *IEEE Trans. on Systems and Control* **45**, 2000, 2253–2270.

[9] Bloch, A. M., D-E. Chang, N. E. Leonard, & J. E. Marsden, Controlled Lagrangians and the Stabilization of Mechanical Systems II: Potential Shaping, *Trans. IEEE on Autom. Contr.* **46**, 2001, 1556–1571.

[10] Chang, D-E., A. M. Bloch, N. E. Leonard, J. E. Marsden, & C. Woolsey, The Equivalence of Controlled Lagrangian and Controlled Hamiltonian Systems, *Control and the Calculus of Variations (special issue dedicated to J. L. Lions)* **8**, 2002, 393–422.

[11] Guibout, V. & A. M. Bloch, A Discrete Maximum Principle for Solving Optimal Control Problems, *Proc. CDC* **43**, 2004, 1806–1811.

[12] Hamberg, J., General Matching Conditions in the Theory of Controlled Lagrangians, *Proc. CDC* **38**, 1999, 2519–2523.

[13] Hamberg, J., Controlled Lagrangians, Symmetries and Conditions for Strong Matching, *In: Lagrangian and Hamiltonian Methods for Nonlinear Control,* Elsevier, 2000.

[14] Jalnapurkar, S. M., M. Leok, J. E. Marsden & M. West, Discrete Routh Reduction, 2005, arXiv:math.NA/0508330.

[15] Junge, O., J. Marsden, & S. Ober-Blöbaum, Discrete Mechanics and Optimal Control, *Proc. of the 16th IFAC World Congress*, Prague, 2005.

[16] Junge, O., J. Marsden, & S. Ober-Blöbaum, Optimal reconfiguration of formation flying spacecraft–a decentralized approach, *Proc. 45th IEEE Conf. on Decision and Control*, 2006, 5210-5215

[17] Kane, C., J. E. Marsden, M. Ortiz, & M. West, Variational Integrators and the Newmark Algorithm for Conservative and Dissipative Mechanical Systems, *Int. J. Numer. Math. Eng.* **49**, 2000, 1295–1325.

[18] Marsden, J. E., *Lectures on Mechanics,* London Mathematical Society Lecture Note Series **174**, Cambridge University Press, 1992.

[19] Marsden, J. E. & T. S. Ratiu, *An Introduction to Mechanics and Symmetry.* Texts in Appl. Math. **17**, Springer-Verlag, 1999.

[20] Marsden, J. E. & M. West, Discrete Mechanics and Variational Integrators, *Acta Numerica* **10**, 2001, 357–514.

[21] Maschke, B., R. Ortega, & A. van der Schaft, Energy-Based Lyapunov Functions for Forced Hamiltonian Systems with Dissipation, *IEEE Trans. Automat. Control* **45**, 2001, 1498–1502.

[22] Ortega, R., M.W. Spong, F. Gómez-Estern, & G. Blankenstein, Stabilization of a Class of Underactuated Mechanical Systems via Interconnection and Damping Assignment, *IEEE Trans. Aut. Control* **47**, 2002, 1218–1233.

[23] Sanyal, A., J. Shen, N. H. McClamroch, & A. M. Bloch, Stability and Stabilization of Relative Equilibria of the Dumbbell Satellite in Central Gravity, 2006, *Journal of the American Institute of Aeronautics and Astronautics,* (to appear).

[24] Woolsey, C., C. K. Reddy, A. M. Bloch, D. E. Chang, N. E. Leonard and J. E. Marsden, Controlled Lagrangian systems with gyroscopic forcing and dissipation, *European Journal of Control*, **10**, number 5, 2004.

[25] Zenkov, D. V., A. M. Bloch, N. E. Leonard, & J. E. Marsden, Matching and Stabilization of Low-Dimensional Nonholonomic Systems, *Proc. CDC* **39**, 2000, 1289–1295.

[26] Zenkov, D. V., A. M. Bloch, & J. E. Marsden, Flat Nonholonomic Matching, *Proc. ACC*, 2002, 2812–2817.

GENERALIZED GALERKIN VARIATIONAL INTEGRATORS

MELVIN LEOK

Abstract. We introduce generalized Galerkin variational integrators, which are a natural generalization of discrete variational mechanics, whereby the discrete action, as opposed to the discrete Lagrangian, is the fundamental object. This is achieved by approximating the action integral with appropriate choices of a finite-dimensional function space that approximate sections of the configuration bundle and numerical quadrature to approximate the integral. We discuss how this general framework allows us to recover higher-order Galerkin variational integrators, asynchronous variational integrators, and symplectic-energy-momentum integrators. In addition, we will consider function spaces that are not parameterized by field values evaluated at nodal points, which allows the construction of Lie group, multiscale, and pseudospectral variational integrators. The construction of pseudospectral variational integrators is illustrated by applying it to the (linear) Schrödinger equation. G-invariant discrete Lagrangians are constructed in the context of Lie group methods through the use of natural charts and interpolation at the level of the Lie algebra. The reduction of these G-invariant Lagrangians yield a higher-order analogue of discrete Euler–Poincaré reduction. By considering nonlinear approximation spaces, spatio-temporally adaptive variational integrators can be introduced as well.

Contents

1. Introduction

We will review some of the previous work on discrete mechanics and their multisymplectic generalizations (see, for example, Marsden et al. [1998, 2001]), before introducing a general formulation of discrete mechanics that recovers higher-order variational integrators (see, for example, Marsden and West [2001]), asynchronous variational integrators (see, for example, Lew et al. [2003]), as well symplectic-energy-momentum integrators (see, for example, Kane et al. [1999]).

While discrete variational integrators exhibit desirable properties such as symplecticity, momentum preservation, and good energy behavior, it does not address other important issues in numerical analysis, such as adaptivity and approximability. Generalized variational integrators are introduced with a view towards addressing such issues in the context of discrete variational mechanics.

By formulating the construction of a generalized variational integrator in terms of the choice of a finite-dimensional function space and a numerical quadrature scheme, we are able to draw upon the extensive literature on approximation theory and numerical quadrature to construct variational schemes that are appropriate for a larger class of problems. Within this framework, we will introduce multiscale, spatio-temporally adaptive, Lie group, and pseudospectral variational integrators.

1.1. **Standard Formulation of Discrete Mechanics.** The standard formulation of discrete variational mechanics (see, for example, Marsden and West [2001]) is to consider the ***discrete Hamilton's principle***,

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

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

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

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

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

where $q_{01}(0) = q_0$, $q_{01}(h) = q_1$, and q_{01} satisfies the Euler–Lagrange equation in the time interval $(0, h)$. The exact discrete Lagrangian is related to the Jacobi solution of the Hamilton–Jacobi equation. The discrete variational principle then yields the ***discrete Euler–Lagrange (DEL)*** equation,

$$D_2 L_d(q_0, q_1) + D_1 L_d(q_1, q_2) = 0,$$

which yields an implicit update map $(q_0, q_1) \mapsto (q_1, q_2)$ that is valid for initial conditions (q_0, q_1) that as sufficiently close to the diagonal of $Q \times Q$.

1.2. **Multisymplectic Geometry.** The generalization of the variational principle to the setting of partial differential equations involves a multisymplectic formulation (see, for example, Marsden et al. [1998, 2001]). Here, the ***base space*** $\mathcal{X}$ consists of the independent variables, which are denoted by $(x^0, \ldots, x^n)$, where x^0 is time, and $x^1, \ldots, x^n$ are space variables.

The independent or field variables, denoted $(q^1, \ldots, q^m)$, form a fiber over each space-time basepoint. The set of independent variables, together with the field variables over them, form a fiber bundle, $\pi : Y \to \mathcal{X}$, referred to as the ***configuration bundle***. The configuration of the system is specified by giving the field values at each space-time point. More precisely, this can be represented as a section of Y over $\mathcal{X}$, which is a continuous map $q : \mathcal{X} \to Y$, such that $\pi \circ q = 1_{\mathcal{X}}$. This means that for every $x \in \mathcal{X}$, $q(x)$ is in the fiber over x, which is $\pi^{-1}(x)$.

In the case of ordinary differential equations, the Lagrangian is dependent on the position variable, and its time derivative, and the action integral is obtained by integrating the Lagrangian in time. In the multisymplectic case, the Lagrangian density is dependent on the field variables, and the derivatives of the field variables with respect to the space-time variables, and the action integral is obtained by integrating the Lagrangian density over a region of space-time.

The analogue of the tangent bundle TQ in the multisymplectic setting is referred to as the ***first jet bundle*** J^1Y, which consists of the configuration bundle Y, together with the first derivatives of the field variables with respect to independent variables. We denote these as,

$$v^i{}_j = q^i{}_{,j} = \frac{\partial q^i}{\partial x^j},$$

for $i = 1, \ldots, m$, and $j = 0, \ldots, n$.

We can think of J^1Y as a fiber bundle over $\mathcal{X}$. Given a section $q : \mathcal{X} \to Y$, we obtain its ***first jet extension***, $j^1q : \mathcal{X} \to J^1Y$, that is given by

$$j^1q(x^0, \dots, x^n) = \left(x^0, \dots, x^n, q^1, \dots, q^m, q^1{}_{,1}, \dots, q^m{}_{,n}\right),$$

which is a section of the fiber bundle J^1Y over $\mathcal{X}$.

The Lagrangian density is a map $L : J^1Y \to \Omega^{n+1}(\mathcal{X})$, and the action integral is given by

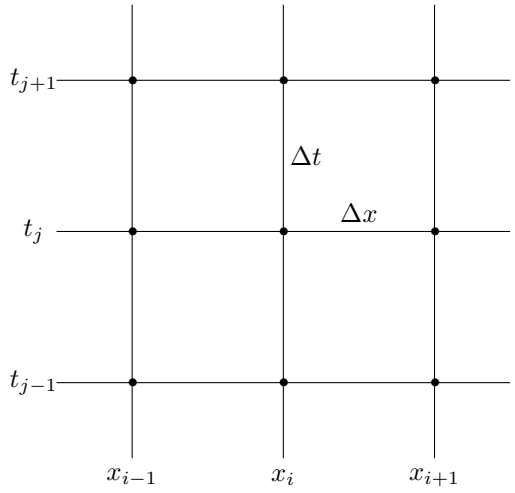
$$\mathcal{S}(q) = \int_{\mathcal{X}} L(j^1q),$$

and then Hamilton's principle states that

$$\delta\mathcal{S} = 0.$$

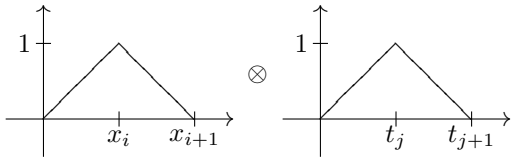
We will see in the next subsection how this allows us to construct multisymplectic variational integrators.

1.3. **Multisymplectic Variational Integrator.** We introduce a multisymplectic variational integrator through the use of a simple but illustrative example. We consider a tensor product discretization of $(1+1)$-space-time, given by



and tensor product shape functions given by

$$\varphi_{i,j}(x,t) =$$



We construct the discrete Lagrangian as follows,

$$L_d(q_{i,j}, q_{i+1,j}, q_{i,j+1}, q_{i+1,j+1}) = \int_{[x_i,x_{i+1}]} \int_{[t_j,t_{j+1}]} L\Big(j^1\Big(\sum\nolimits_{a=i}^{i+1} \sum\nolimits_{b=j}^{j+1} q_{a,b}\varphi_{a,b}\Big)\Big),$$

where $q_{i,j} \simeq q(i\Delta x, j\Delta t)$.

Consider a variation that varies the value of $q_{i,j}$, and leaves the other degrees of freedom fixed, then we obtain the following discrete Euler–Lagrange equation,

$$\begin{aligned} D_1L_d(q_{i,j}, q_{i+1,j}, q_{i,j+1}, q_{i+1,j+1}) &+ D_2L_d(q_{i-1,j}, q_{i,j}, q_{i-1,j+1}, q_{i,j+1}) \\ &+ D_3L_d(q_{i,j-1}, q_{i+1,j-1}, q_{i,j}, q_{i+1,j}) \\ &+ D_4L_d(q_{i-1,j-1}, q_{i,j-1}, q_{i-1,j}, q_{i,j}) = 0\,, \end{aligned}$$

In general, when we take variations in a degree of freedom, we will have a discrete Euler–Lagrange equation that involves all terms in the discrete action sum that are associated with regions in space-time that overlap with the support of the shape function associated with that degree of freedom.

2. Generalized Galerkin Variational Integrators

There are a few essential observations that go into constructing a general framework that encompasses the prior work on variational integrators, asynchronous variational integrators, and symplectic-energy-momentum integrators, while yielding generalizations that allow the construction of multiscale, spatio-temporally adaptive, Lie group, and pseudospectral variational integrators.

The first is that a generalized Galerkin variational integrator involves the choice of a finite-dimensional function space that discretizes a section of the configuration bundle, and the second is that we approximate the action integral though a numerical quadrature scheme to yield a discrete action sum. The discrete variational equations we obtain from this discrete variational principle are simply the Karush–Kuhn–Tucker (KKT) conditions (see, for example, Nocedal and Wright [1999]) with respect to the degrees of freedom that generate the finite-dimensional function space.

To recap, the choices which are made in discretizing a variational problem are:

(1) A finite-dimensional function space to represent sections of the configuration bundle.
(2) A numerical quadrature scheme to evaluate the action integral.

Given these two choices, we obtain an expression for the discrete action in terms of the degrees of freedom, and the KKT conditions for the discrete action to be stationary with respect to variations in the degrees of freedom yield the generalized discrete Euler–Lagrange equations.

Current variational integrators are based on piecewise polynomial interpolation, with function spaces that are parameterized by the value of the field variables at nodal points, and a set of internal points. By relaxing the condition that the interpolation is piecewise, we will be able to consider pseudospectral discretizations, and by relaxing the condition that the parameterization is in terms of field values, we will be able to consider Lie group variational integrators. By considering shape functions motivated by multiscale finite elements (see, for example, Hou and Wu [1999], Efendiev et al. [2000], Chen and Hou [2003]), we will obtain multiscale variational integrators. And by generalizing the approach used in symplectic-energy-momentum integrators, and considering nonlinear approximation spaces (see, for example, DeVore [1998]), we will be able to introduce spatio-temporally adaptive variational integrators.

As we will see, there is nothing canonical about the form of the discrete Euler–Lagrange equations, or the notion that the discrete Lagrangian is a map $L_d : Q \times Q \to \mathbb{R}$. These expressions arise because of the finite-dimensional function space which has been chosen.

2.1. Special Cases of Generalized Galerkin Variational Integrators.

In this subsection, we will show how higher-order Galerkin variational integrators, multisymplectic variational integrators, and symplectic-energy-momentum integrators are all special cases of generalized Galerkin variational integrators.

Higher-Order Galerkin Variational Integrators. In the case of higher-order Galerkin variational integrators, we have chosen a piecewise interpolation for each time interval $[0, h]$, that is parameterized by control points $q_0^1, \dots q_0^s$, corresponding to the value of the curve at a set of control times $0 = d_0 < d_1 < \dots < d_{s-1} < d_s = 1$. The interpolation within each interval $[0, h]$ is given by the unique degree s polynomial $q_d(t; q_0^\nu, h)$, such that $q_d(d_\nu h) = g_0^\nu$, for $\nu = 0, \dots, s$.

By an appropriate choice of quadrature scheme, we can break up the action integral into pieces, which we denote by

$$\mathbb{S}_d^i(q_i^\nu) \approx \int_0^h L(q_d(t; q_i^\nu, h), \dot{q}_d(t; q_i^\nu, h))dt.$$

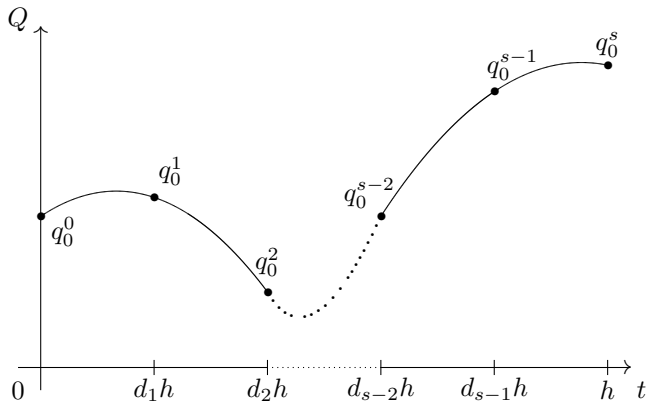


Figure 1: Polynomial interpolation used in higher-order Galerkin variational integrators.

If we further require that the piecewise defined curve is continuous at the node points, we obtain the following augmented discrete action,

$$\mathbb{S}_d\left(\{q_i^\nu\}_{\substack{i=0,\dots,N-1\\ \nu=0,\dots,s}}\right) = \sum_{i=0}^{N-1} \mathbb{S}_d^i\left(\{q_i^\nu\}_{\nu=0}^s\right) - \sum_{i=0}^{N-2} \lambda_i(q_i^s - q_{i+1}^0).$$

The discrete action is stationary when

$$\begin{aligned}\frac{\partial \mathbb{S}_d^i}{\partial q_i^s}(q_i^\nu) &= \lambda_i,\\ \frac{\partial \mathbb{S}_d^i}{\partial q_{i+1}^0}(q_{i+1}^\nu) &= -\lambda_i,\\ \frac{\partial \mathbb{S}_d^i}{\partial q_i^j}(q_i^\nu) &= 0.\end{aligned}$$

We can identify the pieces of the discrete action with the discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$, by setting

$$L_d(q_i, q_{i+1}) = \mathbb{S}_d^i(q_i^\nu), \tag{2.1}$$

where, $q_i^0 = q_i$, $q_i^s = q_{i+1}$, and the other q_i^j's are defined implicitly by the system of equations

$$\frac{\partial \mathbb{S}_d^i}{\partial q_i^j}(q_i^\nu) = 0, \tag{2.2}$$

for $\nu = 1, \dots, s-1$. Once we have made this identification, we have that

$$\begin{aligned}-D_1 L_d(q_{i+1}, q_{i+2}) &= -\frac{\mathbb{S}_d^{i+1}}{\partial q_{i+1}^0}(q_{i+1}^\nu)\\ &= \lambda_i\\ &= \frac{\partial \mathbb{S}_d^i}{\partial q_i^s}(q_i^\nu)\\ &= D_2 L_d(q_i, q_{i+1}),\end{aligned}$$

from which we recover the discrete Euler–Lagrange equation,

$$D_1 L_d(q_{i+1}, q_{i+2}) + D_2 L_d(q_i, q_{i+1}) = 0.$$

The DEL equations, together with the definition of the discrete Lagrangian, given in Equations 2.1 and 2.2, yield a higher-order Galerkin variational integrator. As we have shown, it is only because the interpolation was piecewise that we were able to decompose the equations into a DEL equation, and a set of equations that define the discrete Lagrangian in terms of conditions on the internal control points.

Multisymplectic Variational Integrators. Recall the example of a multisymplectic variational integrator we introduced in §1.3, where we used tensor product linear shape functions with localized supports to discretize the configuration bundle. The discrete action is then given by

$$\begin{aligned}\mathbb{S}_d\left(\{q_{a,b}\}_{\substack{a=0,\dots,M-1\\b=0,\dots,N-1}}\right) &= \int_{[x_0,x_M]}\int_{[t_0,t_N]} L\Big(j^1\Big(\sum_{a=0}^{N}\sum_{b=0}^{N} q_{a,b}\varphi_{a,b}\Big)\Big)\\ &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\int_{[x_i,x_{i+1}]}\int_{[t_j,t_{j+1}]} L\Big(j^1\Big(\sum_{a=0}^{N}\sum_{b=0}^{N} q_{a,b}\varphi_{a,b}\Big)\Big)\\ &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\int_{[x_i,x_{i+1}]}\int_{[t_j,t_{j+1}]} L\Big(j^1\Big(\sum_{a=i}^{i+1}\sum_{b=j}^{j+1} q_{a,b}\varphi_{a,b}\Big)\Big),\end{aligned}$$

where we first decomposed the integral into pieces, and then used the local support of the shape functions to simplify the inner sums over $q_{a,b}\varphi_{a,b}$. As before, it is due to the local support of the shape functions that we can express the discrete action as

$$\mathbb{S}_d = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} L_d(q_{i,j}, q_{i+1,j}, q_{i,j+1}, q_{i+1,j+1}),$$

where

$$L_d(q_{i,j}, q_{i+1,j}, q_{i,j+1}, q_{i+1,j+1}) = \int_{[x_i,x_{i+1}]}\int_{[t_j,t_{j+1}]} L\Big(j^1\Big(\sum\nolimits_{a=i}^{i+1}\sum\nolimits_{b=j}^{j+1} q_{a,b}\varphi_{a,b}\Big)\Big),$$

This localized support is also the reason why the discrete Euler–Lagrange equation in this case consists of four terms, since to each degree of freedom, there are four other degrees of freedom that have shape functions with overlapping support. In the case of ordinary differential equations, this was two. In general, for tensor product meshes of $(n+1)$-space-times, with tent function shape functions, the number of terms in the discrete Euler–Lagrange equation will be 2^{n+1}. In contrast, for pseudospectral variational integrators with m spatial degrees of freedom per time level, and k degrees of freedom per piecewise polynomial in time, each of the $m(k-1)$ discrete Euler–Lagrange equations will involve $m(k-1)$ terms.

This is simply a reflection of the fact that shape functions with compact support yield schemes with banded matrix structure, whereas pseudospectral and spectral methods tend to yield fuller matrices. The payoff in using pseudospectral and spectral methods for problems with smooth or analytic solutions is due to the approximation theoretic property that these solutions are approximated at an exponential rate of accuracy by spectral expansions.

Symplectic-Energy-Momentum Integrators. In the case of symplectic-energy-momentum integrators, the degrees of freedom involve both the base variables and the field variables. We will first derive a second-order symplectic-energy momentum integrator, and in the next section, we will derive a higher-order generalization. We choose a piecewise linear interpolation for our configuration bundle. Each piece is parameterized by the endpoint values of the field variable q_i^0, q_i^1, and the endpoint times h_i^0, h_i^1, and we approximate the action integral using the midpoint rule. To ensure continuity, we require that, $q_i^1 = q_{i+1}^0$, and $h_i^1 = h_{i+1}^0$.

Remark 2.1. *The approach of allowing each piecewise defined curve to float around freely, and imposing the continuity conditions using Lagrange multipliers was used in Lall and West [2003] to unify the formulation*

of discrete variational mechanics and optimal control. Through the use of a primal-dual formalism, discrete analogues of Hamiltonian mechanics and the Hamilton–Jacobi equation were also introduced.

This yields the following discrete action,

$$\mathbb{S}_d = \sum_{i=0}^{N-1} (h_i^1 - h_i^0) L\left(\frac{q_i^0 + q_i^1}{2}, \frac{q_i^1 - q_i^0}{h_i^1 - h_1^0}\right) - \sum_{i=0}^{N-2} \lambda_i (q_i^1 - q_{i+1}^0) - \sum_{i=0}^{N-2} \omega_i (h_i^1 - h_{i+1}^0).$$

To simplify the expressions, we define

$$\begin{aligned} h_i &\equiv h_i^1 - h_i^0, \\ q_{i+\frac{1}{2}} &\equiv \frac{q_i^0 + q_i^1}{2}, \\ \dot{q}_{i+\frac{1}{2}} &\equiv \frac{q_i^1 - q_i^0}{h_i}. \end{aligned}$$

Then, the variational equations are given by

$$\begin{aligned}
0 =& L\left(q_{i+\frac{1}{2}}, \dot{q}_{i+\frac{1}{2}}\right) - h_i \frac{\partial L}{\partial \dot{q}}\left(q_{i+\frac{1}{2}}, \dot{q}_{i+\frac{1}{2}}\right) \frac{1}{h_i} \dot{q}_{i+\frac{1}{2}} - \omega_i, && \text{for } i = 0, \ldots, N-2, \\
0 =& -L\left(q_{i+\frac{1}{2}}, \dot{q}_{i+\frac{1}{2}}\right) + h_i \frac{\partial L}{\partial \dot{q}}\left(q_{i+\frac{1}{2}}, \dot{q}_{i+\frac{1}{2}}\right) \frac{1}{h_i} \dot{q}_{i+\frac{1}{2}} + \omega_{i-1}, && \text{for } i = 1, \ldots, N-1, \\
0 =& h_i \left[\frac{\partial L}{\partial q}\left(q_{i+\frac{1}{2}}, \dot{q}_{i+\frac{1}{2}}\right) \frac{1}{2} + \frac{\partial L}{\partial \dot{q}}\left(q_{i+\frac{1}{2}}, \dot{q}_{i+\frac{1}{2}}\right) \frac{1}{h_i}\right] - \lambda_i, && \text{for } i = 0, \ldots, N-2, \\
0 =& h_i \left[\frac{\partial L}{\partial q}\left(q_{i+\frac{1}{2}}, \dot{q}_{i+\frac{1}{2}}\right) \frac{1}{2} - \frac{\partial L}{\partial \dot{q}}\left(q_{i+\frac{1}{2}}, \dot{q}_{i+\frac{1}{2}}\right) \frac{1}{h_i}\right] + \lambda_{i-1}, && \text{for } i = 1, \ldots, N-1, \\
q_i^1 =& q_{i+1}^0, && \text{for } i = 0, \ldots, N-2, \\
h_i^1 =& h_{i+1}^0, && \text{for } i = 0, \ldots, N-2.
\end{aligned}$$

If we define the discrete Lagrangian to be

$$L_d(q_i, q_{i+1}, h_i) \equiv h_i L\left(\frac{q_i + q_{i+1}}{2}, \frac{q_{i+1} - q_i}{h_i}\right),$$

and the discrete energy to be

$$\begin{aligned} E_d(q_i, q_{i+1}, h_i) &\equiv -D_3 L_d(q_i, q_{i+1}, h_i) \\ &= -L\left(\frac{q_i + q_{i+1}}{2}, \frac{q_{i+1} - q_i}{h_i}\right) + h_i \frac{\partial L}{\partial \dot{q}}\left(\frac{q_i + q_{i+1}}{2}, \frac{q_{i+1} - q_i}{h_i}\right) \frac{1}{h_i} \frac{q_{i+1} - q_i}{h_i}, \end{aligned}$$

and identify points as follows,

$$\begin{aligned} q_i &= q_{i-1}^1 = q_i^0, \\ h_i &= h_{i-1}^1 = h_i^0, \end{aligned}$$

we obtain

$$\begin{aligned}
0 =& E_d(q_i, q_{i+1}, h_i) - \omega_i, && \text{for } i = 0, \ldots, N-2, \\
0 =& -E_d(q_i, q_{i+1}, h_i) + \omega_{i-1}, && \text{for } i = 1, \ldots, N-1, \\
0 =& D_2 L_d(q_i, q_{i+1}, h_i) - \lambda_i, && \text{for } i = 0, \ldots, N-2, \\
0 =& D_1 L_d(q_i, q_{i+1}, h_i) + \lambda_{i-1}, && \text{for } i = 1, \ldots, N-1.
\end{aligned}$$

After eliminating the Lagrange multipliers, we obtain the conservation of discrete energy equation,

$$E_d(q_i, q_{i+1}, h_i) = E_d(q_{i+1}, q_{i+2}, h_{i+1}),$$

and the discrete Euler–Lagrange equation,

$$D_2L_d(q_i, q_{i+1}, h_i) + D_1L_d(q_{i+1}, q_{i+2}, h_{i+1}) = 0.$$

This recovers the results obtained in Kane et al. [1999], but the derivation is new. In §5, we will see how this derivation can be generalized to yield a higher-order scheme.

Discrete Action as the Fundamental Object. The message of this section is that the discrete action is the fundamental object in discrete mechanics, as opposed to the discrete Lagrangian. In instances whereby the shape function associated with individual degrees of freedom have localized supports, it is possible to decompose the discrete action into terms that can be identified with discrete Lagrangians. While this approach might seem artificial at first, we will find that in discussing pseudospectral variational integrators, it does not make sense to break up the discrete action into individual pieces.

3. Lie Group Variational Integrators

In this section, we will introduce higher-order Lie group variational integrators. The basic idea behind all Lie group techniques is to express the update map of the numerical scheme in terms of the exponential map,

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

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

For a more in depth review of Lie group methods, please refer to Iserles et al. [2000]. In the case of variational Lie group methods, we will express the variational problem in terms of finding Lie algebra elements, such that the discrete action is stationary.

As we will consider the reduction of these higher-order Lie group integrators in the next section, we will chose a construction that yields a G-invariant discrete Lagrangian whenever the continuous Lagrangian is G-invariant. This is achieved through the use of G-equivariant interpolatory functions, and in particular, natural charts on G.

3.1. Galerkin Variational Integrators.

We first recall the construction of higher-order Galerkin variational integrators, as originally described in Marsden and West [2001]. Given a Lie group G, the associated ***state space*** is given by the tangent bundle TG. In addition, the dynamics on G is described by a ***Lagrangian***, $L : TG \to \mathbb{R}$. Given a time interval $[0, h]$, the ***path space*** is defined to be

$$\mathcal{C}(G) = \mathcal{C}([0, h], G) = \{g : [0, h] \to G \mid g \text{ is a } C^2 \text{ curve}\},$$

and the ***action map***, $\mathfrak{S} : \mathcal{C}(G) \to \mathbb{R}$, is given by

$$\mathfrak{S}(g) \equiv \int_0^h L(g(t), \dot{g}(t))dt.$$

We approximate the action map, by numerical quadrature, to yield $\mathfrak{S}^s : \mathcal{C}([0, h], G) \to \mathbb{R}$,

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

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

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

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

If we restrict the extremization procedure to the subspace spanned by the interpolatory function that is parameterized by $s+1$ internal points, $\varphi : G^{s+1} \to \mathcal{C}([0,h],G)$, we obtain the following discrete Lagrangian,

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

The interpolatory function is G-equivariant if

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

Lemma 3.1. *If the interpolatory function $\varphi(g^\nu;t)$ is G-equivariant, and the Lagrangian, $L : TG \to \mathbb{R}$, is G-invariant, then the discrete Lagrangian, $L_d : G \times G \to \mathbb{R}$, given by*

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

is G-invariant.

Proof.

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

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

Remark 3.1. *While G-equivariant interpolatory functions provide a computationally efficient method of constructing G-invariant discrete Lagrangians, we can construct a G-invariant discrete Lagrangian (when G is compact) by averaging an arbitrary discrete Lagrangian. In particular, given a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$, the averaged discrete Lagrangian, given by*

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

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

3.2. **Natural Charts.** Following the construction in Marsden et al. [1999], we use the group exponential map at the identity, $\exp_e : \mathfrak{g} \to G$, to construct a G-equivariant interpolatory function, and a higher-order discrete Lagrangian. As shown in Lemma 3.1, this construction yields a G-invariant discrete Lagrangian if the Lagrangian itself is G-invariant.

In a finite-dimensional Lie group G, $\exp_e$ is a local diffeomorphism, and thus there is an open neighborhood $U \subset G$ of e such that $\exp_e^{-1} : U \to \mathfrak{u} \subset \mathfrak{g}$. When the group acts on the left, we obtain a chart $\psi_g : L_g U \to \mathfrak{u}$

at $g \in G$ by

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

Lemma 3.2. *The interpolatory function given by*

$$\varphi(g^\nu; \tau h) = \psi_{g^0}^{-1}\Big(\sum\nolimits_{\nu=0}^{s} \psi_{g^0}(g^\nu) \tilde{l}_{\nu,s}(\tau) \Big),$$

is G-equivariant.

Proof.

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

□

Remark 3.2. *In the proof that φ is G-equivariant, it was important that the base point for the chart should transform in the same way as the internal points g^ν. As such, the interpolatory function will be G-equivariant for a chart that it based at any one of the internal points g^ν that parameterize the function, but will not be G-equivariant if the chart is based at a fixed $g \in G$. Without loss of generality, we will consider the case when the chart is based at the first point g_0.*

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

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

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

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

and the interpolated curve in the algebra is given by

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

which is related to the curve in the group,

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

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

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

Using the standard formula for the derivative of the exponential,

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

where

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

we obtain the following expression for discrete Lagrangian,

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

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

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

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

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

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

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

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

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

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

4. Higher-Order Discrete Euler–Poincaré Equations

In this section, we will apply discrete Euler–Poincaré reduction (see, for example, Marsden et al. [1999]) to the Lie group variational integrator we derived previously, to construct a higher-order generalization of discrete Euler–Poincaré reduction.

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

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

Setting $\xi^0 = 0$, and $\xi^s = \log(g_0^{-1}g_1)$, we can solve the stationarity conditions for the other Lie algebra elements $\{\xi^\nu\}_{\nu=1}^{s-1}$ using the following implicit system of equations,

$$0 = h\sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial g}(c_i h) T_e L_{\exp(\xi(c_i h))} \cdot \operatorname{dexp}_{\operatorname{ad}_{\xi(c_i h)}} \tilde{l}_{\nu,s}(c_i) \right.$$
$$\left. + \frac{1}{h}\frac{\partial L}{\partial \dot{g}}(c_i h) T_e^2 L_{\exp(\xi(c_i h))} \cdot \operatorname{ddexp}_{\operatorname{ad}_{\xi(c_i h)}} \dot{\tilde{l}}_{\nu,s}(c_i) \right]$$

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

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

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

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

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

can be replaced by

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

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

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

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

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

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

As before, we set $\xi^0 = 0$, and $\xi^s = \log(g_0^{-1}g_1)$, and solve the stationarity conditions for the other Lie algebra elements $\{\xi^\nu\}_{\nu=1}^{s-1}$ using the following implicit system of equations,

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

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

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

$$f_{kk+1} \equiv g_k g_{k+1}^{-1}.$$

We will now recall the derivation of the discrete Euler–Poincaré equations, introduced in Marsden et al. [1999]. The variations in f_{kk+1} induced by variations in g_k, g_{k+1} are computed as follows,

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

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

$$\delta \mathbb{S} = \sum_{k=0}^{N-1} l_d'(f_{kk+1}) \delta f_{kk+1}$$

$$= \sum_{k=0}^{N-1} l_d'(f_{kk+1}) TR_{f_{kk+1}}(-g_k^{-1}\delta g_k + \mathrm{Ad}_{f_{kk+1}}\, g_{k+1}\delta g_{k+1})$$
$$= \sum_{k=1}^{N-1} \left[l_d'(f_{k-1k}) TR_{f_{k-1k}}\, \mathrm{Ad}_{f_{k-1k}} - l_d'(f_{kk+1}) TR_{f_{kk+1}} \right] \vartheta_k \,,$$

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

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

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

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

where

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

and

$$\xi^0 = 0 \,,$$
$$\xi^s = \log(f_{kk+1}) \,,$$

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

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

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

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

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

5. Higher-Order Symplectic-Energy-Momentum Variational Integrators

In this section, we will generalize our new derivation of the symplectic-energy-momentum preserving variational integrators (see, Kane et al. [1999]) to yield integrators with higher-order accuracy.

As before, we consider a piecewise interpolation, with both the control points in the field variables, q_i^ν, and the endpoints of the interval, h_i^0, h_i^1, as degrees of freedom. The continuity conditions for this function space are $q_i^s = q_{i+1}^0$, and $h_i^1 = h_{i+1}^0$. Then, we have that the discrete action is given by

$$\mathbb{S}_d = \sum_{i=0}^{N-1} (h_i^1 - h_i^0) \sum_{j=1}^{s} b_j L(q_i(c_j(h_i^1 - h_i^0); q_i^\nu), \dot{q}_i(c_j(h_i^1 - h_i^0); q_i^\nu))$$
$$- \sum_{i=0}^{N-2} \lambda_i (q_i^s - q_{i+1}^0) - \sum_{i=0}^{N-2} \omega_i (h_i^1 - h_{i+1}^0),$$

where

$$q_i(\tau(h_i^1 - h_i^0); q_i^\nu) = \sum_{\kappa=0}^{s} q_i^\kappa \tilde{l}_{\kappa,s}(\tau),$$
$$\dot{q}_i(\tau(h_i^1 - h_i^0); q_i^\nu) = \frac{1}{h_i^1 - h_i^0} \sum_{\kappa=0}^{s} q_i^\kappa \dot{\tilde{l}}_{\kappa,s}(\tau).$$

To simplify the expressions, we define $h_i \equiv h_i^1 - h_i^0$. Then, the variational equations are given by

$$\begin{aligned}
0 =& \sum_{j=1}^{s} b_j L(q_i(c_j h_i), \dot{q}_i(c_j h_i)) - h_i \sum_{j=1}^{s} b_j \frac{\partial L}{\partial \dot{q}}(c_j h_i) \frac{1}{h_i} \dot{q}_i(c_j h_i) - \omega_i, && \text{for } i = 0, \dots, N-2,\\
0 =& -\sum_{j=1}^{s} b_j L(q_i(c_j h_i), \dot{q}_i(c_j h_i)) + h_i \sum_{j=1}^{s} b_j \frac{\partial L}{\partial \dot{q}}(c_j h_i) \frac{1}{h_i} \dot{q}_i(c_j h_i) + \omega_{i-1}, && \text{for } i = 1, \dots, N-1,\\
0 =& h_i \sum_{j=1}^{s} b_j \left[\frac{\partial L}{\partial q}(c_j h_i) \tilde{l}_{s,s}(c_j) + \frac{1}{h_i} \frac{\partial L}{\partial \dot{q}}(c_j h_i) \dot{\tilde{l}}_{s,s}(c_j) \right] - \lambda_i, && \text{for } i = 0, \dots, N-2,\\
0 =& h_i \sum_{j=1}^{s} b_j \left[\frac{\partial L}{\partial q}(c_j h_i) \tilde{l}_{0,s}(c_j) + \frac{1}{h_i} \frac{\partial L}{\partial \dot{q}}(c_j h_i) \dot{\tilde{l}}_{0,s}(c_j) \right] + \lambda_{i-1}, && \text{for } i = 1, \dots, N-1,\\
0 =& h_i \sum_{j=1}^{s} b_j \left[\frac{\partial L}{\partial q}(c_j h_i) \tilde{l}_{\nu,s}(c_j) + \frac{1}{h_i} \frac{\partial L}{\partial \dot{q}}(c_j h_i) \dot{\tilde{l}}_{\nu,s}(c_j) \right], && \begin{array}{r}\text{for } i = 0, \dots, N-1,\\ \nu = 1, \dots, s-1,\end{array}\\
q_i^s =& q_{i+1}^0, && \text{for } i = 0, \dots, N-2,\\
h_i^1 =& h_{i+1}^0, && \text{for } i = 0, \dots, N-2.
\end{aligned}$$

We can eliminate the Lagrange multipliers, to yield

$$\begin{aligned}
0 =& \sum_{j=1}^{s} b_j L(q_i(c_j h_i), \dot{q}_i(c_j h_i)) - \sum_{j=1}^{s} b_j \frac{\partial L}{\partial \dot{q}}(c_j h_i) \dot{q}_i(c_j h_i)\\
&+ \sum_{j=1}^{s} b_j L(q_{i-1}(c_j h_{i-1}), \dot{q}_{i-1}(c_j h_{i-1}))\\
&- \sum_{j=1}^{s} b_j \frac{\partial L}{\partial \dot{q}}(c_j h_{i-1}) \dot{q}_{i-1}(c_j h_{i-1}), && \text{for } i = 1, \dots, N-1,\\
0 =& h_i \sum_{j=1}^{s} b_j \left[\frac{\partial L}{\partial q}(c_j h_i) \tilde{l}_{s,s}(c_j) + \frac{1}{h_i} \frac{\partial L}{\partial \dot{q}}(c_j h_i) \dot{\tilde{l}}_{s,s}(c_j) \right]\\
&+ h_{i-1} \sum_{j=1}^{s} b_j \left[\frac{\partial L}{\partial q}(c_j h_{i-1}) \tilde{l}_{0,s}(c_j) + \frac{1}{h_{i-1}} \frac{\partial L}{\partial \dot{q}}(c_j h_{i-1}) \dot{\tilde{l}}_{0,s}(c_j) \right], && \text{for } i = 1, \dots, N-1,\\
0 =& h_i \sum_{j=1}^{s} b_j \left[\frac{\partial L}{\partial q}(c_j h_i) \tilde{l}_{\nu,s}(c_j) + \frac{1}{h_i} \frac{\partial L}{\partial \dot{q}}(c_j h_i) \dot{\tilde{l}}_{\nu,s}(c_j) \right], && \begin{array}{r}\text{for } i = 0, \dots, N-1,\\ \nu = 1, \dots, s-1,\end{array}\\
q_i^s =& q_{i+1}^0, && \text{for } i = 0, \dots, N-2,\\
h_i^1 =& h_{i+1}^0, && \text{for } i = 0, \dots, N-2.
\end{aligned}$$

If we define the discrete Lagrangian as follows,

$$L_d(q_i, q_{i+1}, h_i) \equiv h_i \sum_{j=1}^{s} b_j L(q_i(c_j h_i), \dot{q}_i(c_j h_i)),$$

where

$$q_i(\tau h_i; q_i^\nu) = \sum_{\kappa=0}^{s} q_i^\kappa \tilde{l}_{\kappa,s}(\tau),$$
$$\dot{q}_i(\tau h_i; q_i^\nu) = \frac{1}{h_i} \sum_{\kappa=0}^{s} q_i^\kappa \dot{\tilde{l}}_{\kappa,s}(\tau),$$

and $q_i^0 = q_i$, $q_i^s = q_1$, and the remaining terms were defined implicitly by

$$0 = h_i \sum_{j=1}^{s} b_j \left[\frac{\partial L}{\partial q}(c_j h_i) \tilde{l}_{\nu,s}(c_j) + \frac{1}{h_i} \frac{\partial L}{\partial \dot{q}}(c_j h_i) \dot{\tilde{l}}_{\nu,s}(c_j) \right],$$

then the equations reduce to the following,

$$E_d(q_i, q_{i+1}, h_i) = -\frac{\partial}{\partial h_i}[L_d(q_i, q_{i+1}, h_i)],$$
$$E_d(q_i, q_{i+1}, h_i) = E_d(q_{i+1}, q_{i+2}, h_{i+1}),$$
$$0 = D_2 L_d(q_i, q_{i+1}, h_i) + D_1 L_d(q_{i+1}, q_{i+2}, h_{i+1}).$$

which is a ***higher-order symplectic-energy-momentum variational integrator***.

Solvability of the Energy Equation. It should be noted that the discrete energy conservation equation is not necessarily solvable, in general, particularly near stationary points. This issue is discussed in Kane et al. [1999], Lew et al. [2004], and can be addressed by reformulating the discrete energy conservation equation as an optimization problem that chooses the time step by minimizing the discrete energy error squared. Clearly, reformulating the discrete energy conservation equation yields the desired behavior whenever the discrete energy conservation equation can be solved, while allowing the computation to proceed when discrete energy conservation cannot be achieved, albeit with a slight energy error in that case. This does not degrade performance significantly, since instances in which discrete energy conservation cannot be achieved are rare.

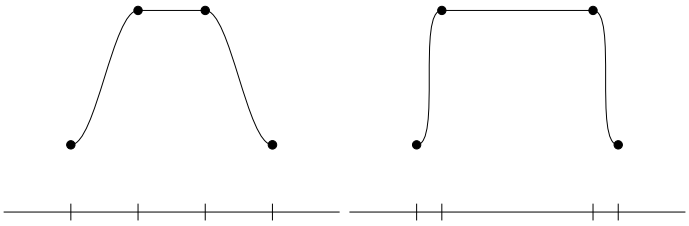
6. Spatio-Temporally Adaptive Variational Integrators

As is the case with all inner approximation techniques in numerical analysis, the quality of the numerical solution we obtain is dependent on the rate at which the sequence of finite-dimensional function spaces approximates the actual solution as the number of degrees of freedom is increased.

For problems that exhibit shocks, nonlinear approximation spaces (see, for example, DeVore [1998]), as opposed to linear approximation spaces, are clearly preferable. Adaptive techniques have been developed in the context of finite elements under the name of r-adaptivity and moving finite elements (see, for example, Baines [1995]), and has been developed in a variational context for elasticity in Thoutireddy and Ortiz [2003]. The standard motivation in discrete mechanics to introduce function spaces that have degrees of freedom associated with the base space is to achieve energy or momentum conservation, as discussed in §5, or Kane et al. [1999], Oliver et al. [2004]. However, if the solution to be approximated exhibits shocks, nonlinear approximation techniques achieve better results for a given number of degrees of freedom.

In this section, we will sketch the use of regularizing transformations of the base space, to achieve a computational representation of sections of the configuration bundle that will yield more accurate numerical results.

Consider the situation when we are representing a characteristic function using piecewise spline interpolation. We show in Figure 2, the difference between linear and nonlinear approximation of the characteristic function.



(a) Using equispaced nodes (b) Using adaptive nodes

Figure 2: Linear and nonlinear approximation of a characteristic function.

When the derivatives of the solution vary substantially in a spatially distributed manner, we obtain additional accuracy, for a fixed representation cost, if we allow nodal points to cluster near regions of high curvature. It is therefore desirable to consider variational integrators based on function spaces that are parameterized by both the position of the nodal points on the base space, as well as the field values over the nodes.

This is represented by having a regular grid for the computational domain $\mathcal{R}$, which is then mapped to the physical base space $\mathcal{X}$, as shown in Figure 3.

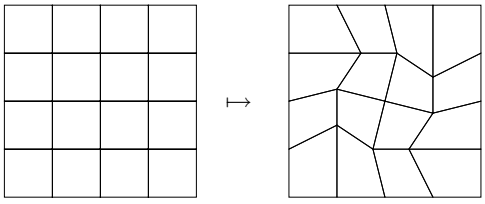


Figure 3: Mapping of the base space from the computational to the physical domain.

The sections of the configuration bundle factor as follows,

$$\begin{array}{ccc} & & Y \\ & \overset{\tilde{q}}{\nearrow} & \uparrow{\scriptstyle q} \\ \mathcal{R} & \xrightarrow[\varphi]{} & \mathcal{X} \end{array}$$

The mapping $\varphi : \mathcal{R} \to \mathcal{X}$ results in a regularized computational representation $\tilde{q} : \mathcal{R} \to Y$ of the original section $q : \mathcal{X} \to Y$. The relationship between the discrete section of the configuration bundle and its computational representation is illustrated in Figure 4.

The action integral is then given by

$$\mathcal{S}(q) = \int_{\mathcal{X}} L(j^1 q) = \int_{\mathcal{R}} L(j^1 \tilde{q}) |\mathbf{D}\varphi|.$$

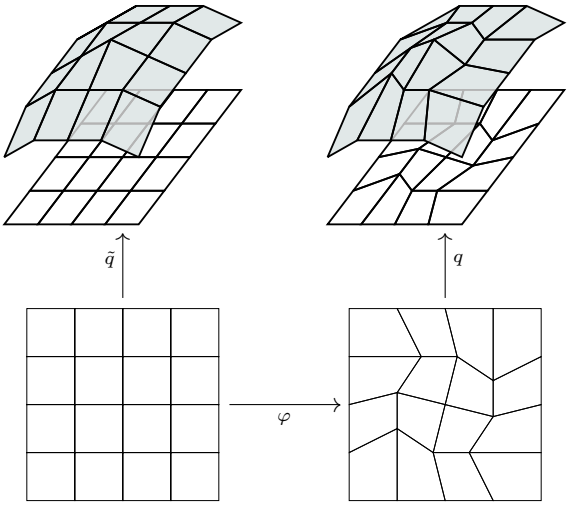


Figure 4: Factoring the discrete section.

Thus, even though $q : \mathcal{X} \to Y$ may exhibit shocks, the computational representation we work with, $\tilde{q} : \mathcal{R} \to Y$, is substantially more regular, and consequently, a numerical quadrature scheme in $\mathcal{R}$ applied to

$$\int_{\mathcal{R}} L(j^1\tilde{q})|\mathbf{D}\varphi|$$

is significantly more accurate than the corresponding numerical quadrature scheme in $\mathcal{X}$ applied to

$$\int_{\mathcal{X}} L(j^1 q).$$

As such, spatio-temporally adaptive variational integrators achieve increased accuracy by allowing the accurate representation of shock solutions using an adapted free knot representation, while using a smooth computational representation to compute the action integral.

7. Multiscale Variational Integrators

In the work on multiscale finite elements (MsFEM), introduced and developed in Hou and Wu [1999], Efendiev et al. [2000], Chen and Hou [2003], shape functions that are solutions of the fast dynamics in the absence of slow forces are constructed to yield finite element schemes that achieve convergence rates that are independent of the ratio of fast to slow scales.

In constructing a multiscale variational integrator, we need to choose finite-dimensional function spaces that do a good job of approximating the fast dynamics of the problem, when the slow variables are frozen. In addition, we require an appropriate choice of numerical quadrature scheme to be able to evaluate the action, which involves integrating a highly-oscillatory Lagrangian. In this section, we will discuss how to go about making such choices of function spaces and quadrature methods.

We will start with a discussion of the multiscale finite element method, to illustrate the importance of a good choice of shape functions in computing solutions to problems with multiple scales. After that, we will walk through the construction of a multiscale variational integrator for the case of a planar pendulum with a stiff spring. Finally, we will discuss how we might proceed if we do not possess knowledge of which variables, or forces, are fast or slow.

7.1. **Multiscale Shape Functions.** We will illustrate the idea of constructing shape functions that are solutions of the fast dynamics by introducing a model multiscale second-order elliptic partial differential equation given by

$$\nabla \cdot a(x/\epsilon)\nabla u^\epsilon(x) = f(x),$$

with homogeneous boundary conditions. In the one-dimensional case, we can solve for the solution analytically, and it has the form

$$u^\epsilon(x) = \int_0^x \frac{F(y)}{a(y/\epsilon)}dy - \frac{\int_0^1 \frac{F(y)}{a(y/\epsilon)}dy}{\int_0^1 \frac{dy}{a(y/\epsilon)}} \int_0^x \frac{dy}{a(y/\epsilon)},$$

where $F(x) = \int_0^x f(y)dy$. If we have nodal points at $\{x_i\}_{i=0}^N$, then the appropriate multiscale shape functions to adopt in this example is to use shape functions that are solutions of the homogeneous problem at the element level. These shape functions φ_i^ϵ satisfy

$$\begin{cases} \frac{\partial}{\partial x}\left(a(x/\epsilon)\frac{\partial}{\partial x}\varphi_i^\epsilon\right) = 0, & \text{for } x_{i-1} < x < x_{i+1}; \\ \varphi_i^\epsilon(x_{i-1}) = 0; \quad \varphi_i^\epsilon(x_{i+1}) = 0; \quad \varphi_i^\epsilon(x_i) = 1. \end{cases}$$

And they have the explicit form given by

$$\varphi_i^\epsilon(x) = \begin{cases} \left[\int_{x_{i-1}}^{x_i} \frac{ds}{a(s/\epsilon)}\right]^{-1} \left[\int_{x_{i-1}}^{x} \frac{ds}{a(s/\epsilon)}\right], & x \in [x_{i-1}, x_i]\,; \\ \left[\int_{x_i}^{x_{i+1}} \frac{ds}{a(s/\epsilon)}\right]^{-1} \left[\int_{x}^{x_{i+1}} \frac{ds}{a(s/\epsilon)}\right], & x \in (x_i, x_{i+1}]\,; \\ 0, & \text{otherwise}\,. \end{cases}$$

It can be shown that this will yield a numerical scheme that solves exactly for the solution at the nodal points.

As an example, we will compute the analytical solution for $a(x) = \frac{10}{1+0.95\sin(2\pi x)}$, $f(x) = x^2$, and $\epsilon = 0.025$. This is illustrated in Figure 5(a). What is particularly interesting is to compare the zoomed plot of the exact solution and the multiscale shape function over the same interval, shown in Figures 5(b) and 5(c), respectively. The multiscale finite element method is able to achieve excellent results because the multiscale shape functions are able to capture the fast dynamics well. In the next subsection, we will discuss how this insight is relevant in the construction of multiscale variational integrators.

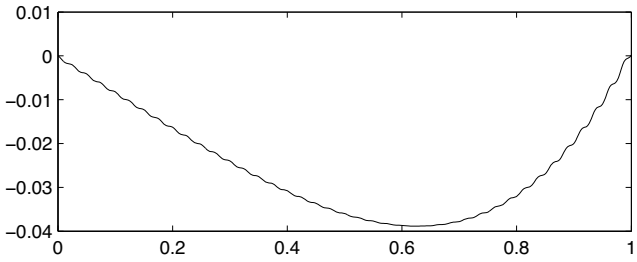
7.2. **Multiscale Variational Integrator for the Planar Pendulum with a Stiff Spring.** As was shown previously, a shape function that captures the fast dynamics of a multiscale problem is able to achieve superior accuracy when used for computation. While this idea has primarily been used for problems with multiple spatial scales, it is natural to consider its application to a problem with multiple temporal scales, such as the problem of the planar pendulum with a stiff spring, as illustrated in Figure 6.

We will use this example to illustrate the issues that arise in constructing a multiscale variational integrator. The variables are $q = (a, \theta)$, where a is the spring extension, and θ is the angle from the vertical. The Lagrangian is given by

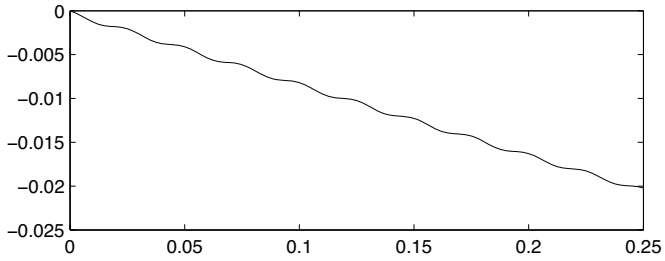
$$L(a, \theta, \dot{a}, \dot{\theta}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - mgy - \frac{k}{2}a^2,$$

where

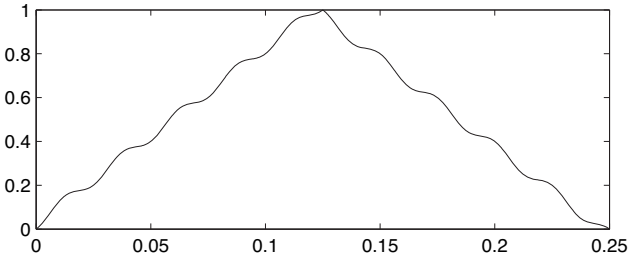
$$\begin{aligned} x &= (l + a)\sin\theta, \\ y &= -(l + a)\cos\theta, \\ \dot{x} &= \dot{a}\sin\theta + \dot{\theta}(l + a)\cos\theta, \\ \dot{y} &= -\dot{a}\cos\theta + \dot{\theta}(l + a)\sin\theta. \end{aligned}$$



(a) Exact solution



(b) Exact solution (zoomed)



(c) Multiscale shape function

Figure 5: Comparison of the multiscale shape function and the exact solution for the elliptic problem.

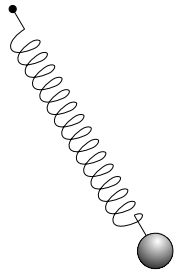


Figure 6: Planar pendulum with a stiff spring

The Hamilton's equations for the planar pendulum with a stiff spring are

$$\begin{aligned}
\dot{a} &= \frac{p_a}{m},\\
\dot{\theta} &= \frac{p_\theta}{m(l+a)^2},\\
\dot{p}_a &= -ka + gm\cos\theta + m(l+a)\dot{\theta}^2,\\
\dot{p}_\theta &= -gm(l+a)\sin\theta.
\end{aligned}$$

The timescale arising from the mass-spring system is $2\pi\sqrt{m/k}$. The timescale arising from the planar pendulum system is $2\pi\sqrt{l/g}$. The ratio of timescales is given by $\epsilon = \sqrt{mg/kl}$.

Multiscale Shape Function. In this problem, the fast scale is associated with the stiff spring, and if we set the slow variable $\theta = 0$, we obtain the equation

$$\ddot{a} = \frac{\dot{p}_a}{m} = -\frac{k}{m}a,$$

which has solutions of the form

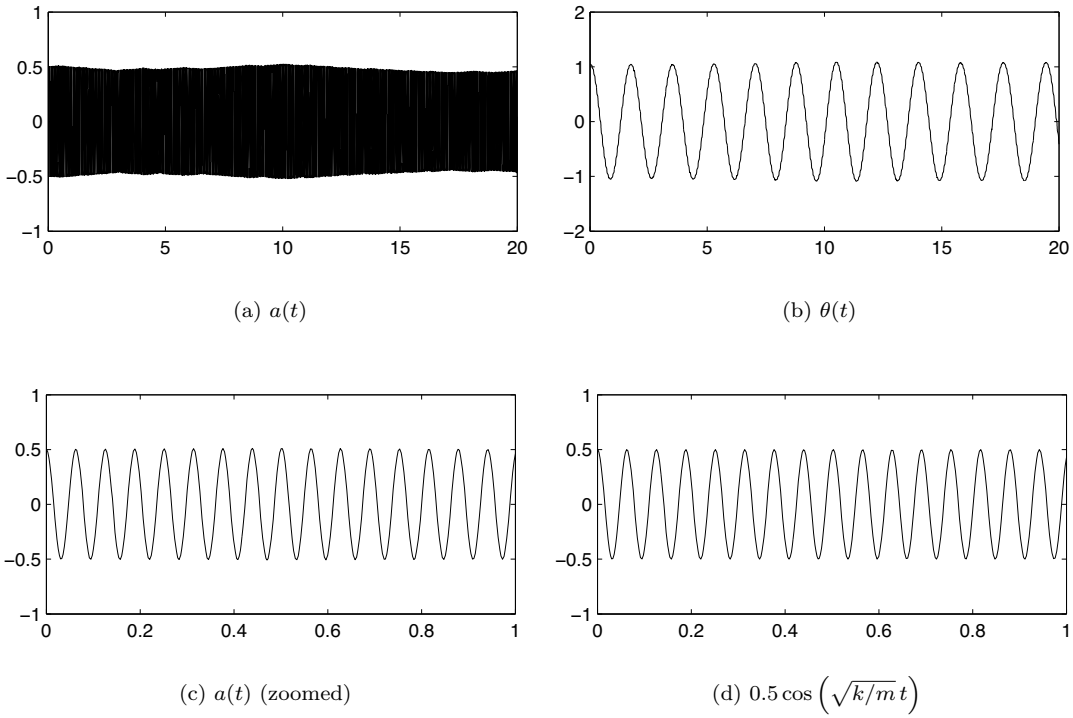
$$a(t) = a_0 \sin\left(\sqrt{k/m}\,t\right) + a_1 \cos\left(\sqrt{k/m}\,t\right).$$

We will now consider a well-resolved simulation of this system using the `ode15s` stiff solver from MATLAB, with parameters $m = 1$, $g = 9.81$, $k = 10000$, $l = 1$, giving a scale separation of $\epsilon = 0.0313$. The simulation results are show in Figure 7.

Clearly, if we wish to choose time steps that do not resolve the fast oscillations in a, but do resolve the slow oscillations in θ, it would be desirable to include $\sin\left(\sqrt{k/m}\,t\right)$, and $\cos\left(\sqrt{k/m}\,t\right)$ in the finite-dimensional function space used to interpolate $a(t)$.

Evaluating the Discrete Lagrangian. Since we have chosen time steps that do not resolve the fast oscillations, it follows that over the interval $[0, h]$, the Lagrangian will oscillate rapidly as well. In computing the discrete Lagrangian, it is therefore necessary to ensure that this highly-oscillatory integral is well-approximated.

It is conventional wisdom, in numerical analysis, that the numerical quadrature of highly-oscillatory integrals is a challenging problem requiring the use of many function evaluations. Recently, there has been a series of papers, Iserles [2003a,b, 2004], Iserles and Nørsett [2004], that provides an analysis of Filon-type quadrature schemes that provide an efficient and accurate method of evaluating such integrals. The method is applicable to weighted integrals as well, but we will summarize the results from §3 of Iserles [2003a] restricted to unweighted integrals, and refer the reader to the original reference for an in-depth discussion and analysis.



(a) $a(t)$ (b) $\theta(t)$

(c) $a(t)$ (zoomed) (d) $0.5\cos\left(\sqrt{k/m}\,t\right)$

Figure 7: Comparison of the multiscale shape function and the exact solution for the planar pendulum with a stiff spring.

The Filon-type method aims to evaluate an integral of the form

$$I_h[f] = \int_0^h f(x)e^{i\omega x}dx = h\int_0^h f(hx)e^{i\omega x}dx.$$

Given a set of distinct quadrature points, $c_1 < c_2 < \cdots < c_\nu$ in $[0,1]$, the Filon-type quadrature method is given by

$$Q_h^F[f] = h\sum_{i=1}^{\nu} b_i(ih\omega)f(c_i h),$$

where

$$b_i(ih\omega) = \int_0^1 l_i(x)e^{ih\omega x}dx,$$

and l_i are the Legendre polynomials. Here, we draw attention to the fact that the quadrature weights are dependent on $h\omega$.

If the quadrature points correspond to Gauss-Christoffel quadrature of order p, then the error for the Filon-type method is given as follows.

$$\begin{aligned} \mathcal{O}(h^{p+1}), &\quad \text{if} \quad h\omega \ll 1 \\ \mathcal{O}(h^{\nu}), &\quad \text{if} \quad h\omega = \mathcal{O}(1) \\ \mathcal{O}(h^{\nu+1}/(h\omega)), &\quad \text{if} \quad h\omega \gg 1 \end{aligned}$$

$$\mathcal{O}(h^{\nu+1}/(h\omega)^2), \quad \text{if} \quad h\omega \gg 1, c_1 = 0, c_\nu = 1.$$

Clearly, for highly-oscillatory functions, the last case, which corresponds to the Lobatto quadrature points, is most desirable. Thus, it is appropriate to use the Filon-Lobatto method to evaluate the discrete Lagrangian in our case.

Discrete Variational Equations. As we discussed previously, instead of looking for stationary solutions to the discrete Hamilton's principle in polynomial spaces, we will consider solutions that are piecewise of the form

$$q(t; \{p_j\}, \omega, a_0, a_1) = \left(\sum\nolimits_{j=0}^{n} p_j t^j\right) (1 + a_0 \sin(\omega t) + a_1 \cos(\omega t)) .$$

This function space approximates the highly-oscillatory nature of the solution well, in contrast to a polynomial function space, thereby avoiding approximation-theoretic errors. The degrees of freedom in this function space are $\{p_j\}$, ω, a_0, and a_1. Since there are no distinguished degrees of freedom that are responsible for the endpoint values of the curve, we need to impose continuity at the nodes using a Lagrange multiplier. The augmented discrete action is given by

$$\begin{aligned}\mathbb{S}_d = \sum_{i=0}^{N-1} \int_0^h & L(j^1 q_i(t; \{p_j^i\}, \omega^i, a_0^i, a_1^i))dt \\ & + \sum_{i=0}^{N-2} \lambda_i (q_i(h; \{p_j^i\}, \omega^i, a_0^i, a_1^i) - q_{i+1}(0; \{p_j^{i+1}\}, \omega^{i+1}, a_0^{i+1}, a_1^{i+1})) ,\end{aligned}$$

where each of the integrals are evaluated using the Filon-Lobatto method. Taking variations with respect to the degrees of freedom yields an update map,

$$\left(\{p_j^i\}, \omega^i, a_0^i, a_1^i\right) \mapsto \left(\{p_j^{i+1}\}, \omega^{i+1}, a_0^{i+1}, a_1^{i+1}\right) ,$$

which gives the ***multiscale variational integrator*** for the planar pendulum with a stiff spring.

7.3. **Computational Aspects.** Multiscale variational integrators have the advantage of directly accounting for the contribution of the fast dynamics, thereby allowing the scheme to use significantly larger time-steps, while maintaining accuracy and stability. It is possible to take advantage of knowledge about which of the variables, or forces, are fast or slow, by using a low degree polynomial and oscillatory functions for the fast variables, and a higher-order polynomial for the slow variables. In the absence of such information, it is appropriate to use a function space with both polynomials and oscillatory functions, and apply it to all the variables.

Recall that the Filon-type method has quadrature coefficients that depend on the frequency. As such, the initial fast frequency has to be estimated numerically using a fully resolved computation for a short period of time. Since both the function space and the quadrature weights depend on the fast frequency ω, the resulting scheme is implicit and fairly nonlinear, and as such, it may be expensive for large systems.

8. Pseudospectral Variational Integrators

The use of spectral expansions of the solution in space are particularly appropriate for highly accurate simulations of the evolution of smooth solutions, such as those arising from quantum mechanics. We will introduce pseudospectral variational integrators, and consider the Schrödinger equation as an example.

In particular we will adopt the tensor product of a spectral expansion in space, and a polynomial expansion in time. For example, we could have an interpolatory function of the form

$$\psi(x, (\tau + l)\Delta t) = \frac{1}{2\pi} {\sum_{k=-N/2}^{N/2}}' e^{ikx} \left((1-\tau)\hat{v}_k^l + \tau \hat{v}_k^{l+1}\right) ,$$

which is the tensor product of a discrete Fourier expansion in space, and linear interpolation in time. Here, the $\sum'$ notation denotes a weighted sum where the terms with indices $\pm N/2$ are weighted by $1/2$, and the other terms are weighted by 1. See page 19 of Trefethen [2000] for a discussion of why this is necessary to fix an issue with derivatives of the interpolant.

The degrees of freedom are given by $\hat{v}_k^l$, which are the discrete Fourier coefficients. We will later see how such an interpolation can be applied to the Schrödinger equation. The action integral can be exactly evaluated for this class of shape functions, as we will see below.

It is straightforward to generalize the pseudospectral approach we present in this section to a spectral variational integrator, with discrete Fourier expansions in space for periodic domains, or Chebyshev expansions in space for non-periodic domains, and Chebyshev expansions in time. This will however result in all the degrees of freedom on the space-time mesh being coupled, and is therefore substantially more expensive computationally than the pseudospectral method. The payoff for adopting the spectral approach is spectral accuracy, which is accuracy beyond all orders.

8.1. **Variational Derivation of the Schrödinger Equation.** Let $\mathcal{H}$ be a complex Hilbert space, for example, the space of complex-valued functions ψ on $\mathbb{R}^3$ with the Hermitian inner product,

$$\langle \psi_1, \psi_2 \rangle = \int \psi_1(x) \overline{\psi}_2(x) d^3x \,,$$

where the overbar denotes complex conjugation. We will present a Lagrangian derivation of the Schrödinger equation, following worked example 9.1 on pages 568–569 of José and Saletan [1998].

Consider the Lagrangian density $\mathcal{L}$ given by

$$\mathcal{L}(j^1\psi) = \frac{i\hbar}{2}\{\dot{\psi}\overline{\psi} - \psi\dot{\overline{\psi}}\} - \hat{H}\psi\overline{\psi},$$

where $\hat{H} : \mathcal{H} \to \mathcal{H}$ is given by

$$\hat{H}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi,$$

which yields

$$\mathcal{L}(j^1\psi) = \frac{i\hbar}{2}\{\dot{\psi}\overline{\psi} - \psi\dot{\overline{\psi}}\} - \frac{\hbar^2}{2m}\nabla\psi\cdot\nabla\overline{\psi} - V\psi\overline{\psi}.$$

We take $\psi, \overline{\psi}$ as independent variables, and compute,

$$\begin{aligned}
\delta\int L dt &= \int \left[\left(\frac{\partial \mathcal{L}}{\partial \overline{\psi}}\delta\overline{\psi} + \frac{\partial \mathcal{L}}{\partial \dot{\overline{\psi}}}\delta\dot{\overline{\psi}} + \frac{\partial \mathcal{L}}{\partial \nabla\overline{\psi}}\delta\nabla\overline{\psi} \right) + \left(\frac{\partial \mathcal{L}}{\partial \psi}\delta\psi + \frac{\partial \mathcal{L}}{\partial \dot{\psi}}\delta\dot{\psi} + \frac{\partial \mathcal{L}}{\partial \nabla\psi}\delta\nabla\psi \right) \right] d^3x dt \\
&= \int \left[\left(\frac{\partial \mathcal{L}}{\partial \overline{\psi}} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\overline{\psi}}} - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\overline{\psi}} \right)\delta\overline{\psi} + \left(\frac{\partial \mathcal{L}}{\partial \psi} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\psi}} - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\psi} \right)\delta\psi \right] d^3x dt \\
&= \int \left[\left(\frac{i\hbar}{2}\dot{\psi} - V\psi + \frac{i\hbar}{2}\dot{\psi} + \frac{\hbar^2}{2m}\nabla^2\psi \right)\delta\overline{\psi} + \left(\frac{i\hbar}{2}\dot{\overline{\psi}} - V\overline{\psi} + \frac{i\hbar}{2}\dot{\overline{\psi}} + \frac{\hbar^2}{2m}\nabla^2\overline{\psi} \right)\delta\psi \right] d^3x dt,
\end{aligned}$$

where we integrated by parts, and neglected boundary terms as the variations vanish at the boundary of the space-time region. Since the variations are arbitrary, we obtain the nonrelativistic (linear) Schrödinger equation as a result,

$$i\hbar\dot{\psi} = \left\{ -\frac{\hbar^2}{2m}\nabla^2 + V \right\}\psi \,.$$

We note that the Lagrangian density is invariant under the internal phase shift given by

$$\psi \mapsto e^{i\epsilon}\psi, \qquad \overline{\psi} \mapsto e^{-i\epsilon}\overline{\psi}.$$

The space part of the multi-momentum map is given by

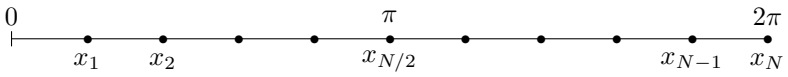
$$j^k = \frac{\partial \mathcal{L}}{\partial(\partial_k \psi)} i\psi + \frac{\partial \mathcal{L}}{\partial(\partial_k \overline{\psi})}(-i\overline{\psi}) = \frac{i\hbar^2}{2m}(\psi \partial_k \overline{\psi} - \overline{\psi}\partial_k \psi),$$

and the time part is given by

$$j^0 = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} i\psi - \frac{\partial \mathcal{L}}{\partial \dot{\overline{\psi}}}(-i\overline{\psi}) = -\frac{\hbar}{2}(\dot{\overline{\psi}}\psi - \dot{\psi}\overline{\psi}).$$

The norm of the wavefunction is automatically preserved by variational integrators, since the norm is a quadratic invariant.

8.2. **Pseudospectral Variational Integrator for the Schrödinger Equation.** Consider a periodic domain $[0, 2\pi]$, discretized with a discrete Fourier series expansion in space, and a linear interpolation in time. Let N be an even integer, then, our computation is done on the following mesh,



This implies that the grid spacing is given by

$$h = \frac{2\pi}{N}\,.$$

The interpolation is given by

$$\psi(x,(\tau+l)\Delta t) = \frac{1}{2\pi}\sum_{k=-N/2}^{N/2}{}' e^{ikx}\left((1-\tau)\hat{v}_k^l + \tau \hat{v}_k^{l+1}\right),$$

$$\dot{\psi}(x,(\tau+l)\Delta t) = \frac{1}{2\pi\Delta t}\sum_{k=-N/2}^{N/2}{}' e^{ikx}\left(\hat{v}_k^{l+1} - \hat{v}_k^l\right),$$

$$\bar{\psi}(x,(\tau+l)\Delta t) = \frac{1}{2\pi}\sum_{k=-N/2}^{N/2}{}' e^{-ikx}\left((1-\tau)\bar{\hat{v}}_k^l + \tau \bar{\hat{v}}_k^{l+1}\right),$$

$$\dot{\bar{\psi}}(x,(\tau+l)\Delta t) = \frac{1}{2\pi\Delta t}\sum_{k=-N/2}^{N/2}{}' e^{-ikx}\left(\bar{\hat{v}}_k^{l+1} - \bar{\hat{v}}_k^l\right),$$

and the discrete Fourier transformation is given by

$$\hat{v}_j = \frac{1}{2\pi} h \sum_{j=1}^{N} e^{-ikx_j} v_j\,,$$

for $k = -N/2+1, \ldots, N/2$, and $\hat{v}_{-N/2} \equiv \hat{v}_{N/2}$. Recall that

$$\mathcal{L}(j^1\psi) = \frac{i\hbar}{2}\{\dot{\psi}\overline{\psi} - \psi\dot{\overline{\psi}}\} - \hat{H}\psi\overline{\psi},$$

where $\hat{H} : \mathcal{H} \to \mathcal{H}$ is given by

$$\hat{H}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi.$$

Furthermore, the potential V is expressed using a discrete Fourier expansion,

$$V(x) = \frac{1}{2\pi}\sum_{k=-N/2}^{N/2}{}' e^{ikx}\hat{V}_k\,.$$

In addition, we will need to introduce a normalization condition, so as to eliminate trivial solutions of the partial differential equation. The normalization condition is

$$1 = \langle \psi_l, \psi_l \rangle = \int_0^{2\pi} \left(\frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime} e^{ikx} \hat{v}_k^l \right) \left(\frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime} e^{-ikx} \bar{\hat{v}}_k^l \right) dx = \frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime\prime} \hat{v}_k^l \bar{\hat{v}}_k^l,$$

which is enforced using a Lagrange multiplier. **Discrete Action for the Schrödinger Equation.** The discrete action in the space-time region $[0, 2\pi] \times [l\Delta t, (l+1)\Delta t]$ is given by

$$\begin{aligned}
\mathbb{S}_d &= \int_{l\Delta t}^{(l+1)\Delta t} \int_0^{2\pi} \mathcal{L}(j^1\psi) dx dt + \lambda_l (1 - \langle \psi_l, \psi_l \rangle) \\
&= \int_{l\Delta t}^{(l+1)\Delta t} \int_0^{2\pi} \left[\frac{i\hbar}{2} \{ \dot{\psi}\bar{\psi} - \psi\dot{\bar{\psi}} \} + \frac{\hbar^2}{2m} \nabla^2 \psi \bar{\psi} - V \psi \bar{\psi} \right] dx dt + \lambda_l \left(1 - \frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime\prime} \hat{v}_k^l \bar{\hat{v}}_k^l \right) \\
&= \int_0^1 \int_0^{2\pi} \frac{i\hbar}{2} \left[\left(\frac{1}{2\pi\Delta t} \sum_{k=-N/2}^{N/2\,\prime} e^{ikx} \left(\hat{v}_k^{l+1} - \hat{v}_k^l \right) \right) \left(\frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime} e^{-ikx} \left((1-\tau) \bar{\hat{v}}_k^l + \tau \bar{\hat{v}}_k^{l+1} \right) \right) \right. \\
&\qquad \left. - \left(\frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime} e^{ikx} \left((1-\tau) \hat{v}_k^l + \tau \hat{v}_k^{l+1} \right) \right) \left(\frac{1}{2\pi\Delta t} \sum_{k=-N/2}^{N/2\,\prime} e^{-ikx} \left(\bar{\hat{v}}_k^{l+1} - \bar{\hat{v}}_k^l \right) \right) \right] \Delta t \, dx d\tau \\
&\quad + \int_0^1 \int_0^{2\pi} \frac{\hbar^2}{2m} \left(\frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime} (-k^2) e^{ikx} \left((1-\tau) \hat{v}_k^l + \tau \hat{v}_k^{l+1} \right) \right) \\
&\qquad \cdot \left(\frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime} e^{-ikx} \left((1-\tau) \bar{\hat{v}}_k^l + \tau \bar{\hat{v}}_k^{l+1} \right) \right) \Delta t \, dx d\tau \\
&\quad - \int_0^1 \int_0^{2\pi} \left[\left(\frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime} e^{ikx} \hat{V}_k \right) \left(\frac{1}{2\pi} \sum_{m=-N/2}^{N/2\,\prime} e^{imx} \left((1-\tau) \hat{v}_m^l + \tau \hat{v}_m^{l+1} \right) \right) \right. \\
&\qquad \left. \cdot \left(\frac{1}{2\pi} \sum_{n=-N/2}^{N/2\,\prime} e^{-inx} \left((1-\tau) \bar{\hat{v}}_n^l + \tau \bar{\hat{v}}_n^{l+1} \right) \right) \right] \Delta t \, dx dt \\
&\quad + \lambda_l \left(1 - \frac{1}{2\pi} \sum_{k=-N/2}^{N/2\,\prime\prime} \hat{v}_k^l \bar{\hat{v}}_k^l \right)
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 \frac{i\hbar}{2}\left[\frac{1}{2\pi}\sum_{k=-N/2}^{N/2}{}''\left((\hat{v}_k^{l+1}-\hat{v}_k^l)((1-\tau)\bar{v}_k^l+\tau\bar{v}_k^{l+1})-((1-\tau)\hat{v}_k^l+\tau\hat{v}_k^{l+1})(\bar{v}_k^{l+1}-\bar{v}_k^l)\right)\right]d\tau\\
&\quad-\int_0^1\left[\frac{\hbar^2}{2\pi}\frac{k^2}{2\pi}\sum_{k=-N/2}^{N/2}{}''((1-\tau)\hat{v}_k^l+\tau\hat{v}_k^{l+1})((1-\tau)\bar{v}_k^l+\tau\bar{v}_k^{l+1})\right]\Delta t\,d\tau\\
&\quad-\int_0^1\left(\frac{1}{2\pi}\right)^2\left[\sum_{n=-N/2}^{-1}{}'\sum_{m=-N/2}^{N/2+n}{}'\left(\hat{V}_{n-m}((1-\tau)\hat{v}_m^l+\tau\hat{v}_m^{l+1})((1-\tau)\bar{v}_n^l+\tau\bar{v}_n^{l+1})\right)\right.\\
&\qquad\left.+\sum_{n=0}^{N/2}{}'\sum_{m=n-N/2}^{N/2}{}'\left(\hat{V}_{n-m}((1-\tau)\hat{v}_m^l+\tau\hat{v}_m^{l+1})((1-\tau)\bar{v}_n^l+\tau\bar{v}_n^{l+1})\right)\right]\Delta t\,d\tau\\
&\quad+\lambda_l\left(1-\frac{1}{2\pi}\sum_{k=-N/2}^{N/2}{}''\hat{v}_k^l\bar{v}_k^l\right)\\
&=\frac{i\hbar}{4\pi}\sum_{k=-N/2}^{N/2}{}''\left[\hat{v}_k^{l+1}\bar{v}_k^l-\hat{v}_k^l\bar{v}_k^{l+1}\right]-\frac{\hbar^2k^2\Delta t}{24\pi^2}\sum_{k=-N/2}^{N/2}{}''\left[\hat{v}_k^l(2\bar{v}_k^l+\bar{v}_k^{l+1})+\hat{v}_k^{l+1}(\bar{v}_k^l+2\bar{v}_k^{l+1})\right]\\
&\quad-\frac{\Delta t}{24\pi^2}\left(\sum_{n=-N/2}^{-1}{}'\sum_{m=-N/2}^{N/2+n}{}'+\sum_{n=0}^{N/2}{}'\sum_{m=n-N/2}^{N/2}{}'\right)\hat{V}_{n-m}\left[\hat{v}_m^l(2\bar{v}_n^l+\bar{v}_n^{l+1})+\hat{v}_m^{l+1}(\bar{v}_n^l+2\bar{v}_n^{l+1})\right]\\
&\quad+\lambda_l\left(1-\frac{1}{2\pi}\sum_{k=-N/2}^{N/2}{}''\hat{v}_k^l\bar{v}_k^l\right),
\end{aligned}
$$

where we used the fact that

$$\int_0^{2\pi}e^{ikx}dx=2\pi\delta_0^i\,,$$

for $k\in\mathbb{Z}$, and we define $\sum'$ as a weighted sum where the terms with indices $\pm N/2$ are weighted by $1/2$, and $\sum''$ as a weighted sum where the terms with indices $\pm N/2$ are weighted by $1/4$. We should note that using the same approach, it would be possible to exactly evaluate the action integral for the class of tensor product shape functions with a discrete Fourier expansion in space, and a polynomial expansion in time. In particular, a similar approach is valid in exactly evaluating the action integral when we use shape functions that are spectral in both space and time.

Discrete Euler–Lagrange Equations. We are now in a position to compute the discrete Euler–Lagrange equations associated with the Schrödinger equation when using a tensor product of a discrete Fourier expansion in space, and a linear interpolation in time.

The discrete variational equations are given by

$$
\begin{aligned}
0&=\frac{i\hbar}{4\pi}\left[\bar{v}_j^{l-1}-\bar{v}_j^{l+1}\right]-\frac{\hbar^2k^2\Delta t}{24\pi^2}\left[\bar{v}_j^{l-1}+4\bar{v}_j^l+\bar{v}_j^{l+1}\right]\\
&\quad-\frac{\Delta t}{24\pi^2}\sum_{n=-N/2}^{N/2+j}{}'\hat{V}_{n-j}\left[\bar{v}_n^{l-1}+4\bar{v}_n^l+\bar{v}_n^{l+1}\right]-\frac{\lambda_l}{2\pi}\bar{v}_j^l\,,\qquad\text{for } j=-N/2+1,\ldots,-1,
\end{aligned}
$$

$$
\begin{aligned}
0 &= \frac{i\hbar}{4\pi}\left[\hat{v}_j^{l+1} - \hat{v}_j^{l-1}\right] - \frac{\hbar^2 k^2 \Delta t}{24\pi^2}\left[\hat{v}_j^{l-1} + 4\hat{v}_j^{l} + \hat{v}_j^{l+1}\right] \\
&\quad - \frac{\Delta t}{24\pi^2}{\sum_{n=-N/2}^{N/2+j}}' \hat{V}_{j-n}\left[\hat{v}_n^{l-1} + 4\hat{v}_n^{l} + \hat{v}_n^{l+1}\right] - \frac{\lambda_l}{2\pi}\hat{v}_j^{l}\,, \qquad \text{for } j = -N/2+1,\dots,-1, \\
0 &= \frac{i\hbar}{4\pi}\left[\bar{\hat{v}}_j^{l-1} - \bar{\hat{v}}_j^{l+1}\right] - \frac{\hbar^2 k^2 \Delta t}{24\pi^2}\left[\bar{\hat{v}}_j^{l-1} + 4\bar{\hat{v}}_j^{l} + \bar{\hat{v}}_j^{l+1}\right] \\
&\quad - \frac{\Delta t}{24\pi^2}{\sum_{n=j-N/2}^{N/2}}' \hat{V}_{n-j}\left[\bar{\hat{v}}_n^{l-1} + 4\bar{\hat{v}}_n^{l} + \bar{\hat{v}}_n^{l+1}\right] - \frac{\lambda_l}{2\pi}\bar{\hat{v}}_j^{l}\,, \qquad \text{for } j = 0,\dots,N/2-1, \\
0 &= \frac{i\hbar}{4\pi}\left[\hat{v}_j^{l+1} - \hat{v}_j^{l-1}\right] - \frac{\hbar^2 k^2 \Delta t}{24\pi^2}\left[\hat{v}_j^{l-1} + 4\hat{v}_j^{l} + \hat{v}_j^{l+1}\right] \\
&\quad - \frac{\Delta t}{24\pi^2}{\sum_{n=j-N/2}^{N/2}}' \hat{V}_{j-n}\left[\hat{v}_n^{l-1} + 4\hat{v}_n^{l} + \hat{v}_n^{l+1}\right] - \frac{\lambda_l}{2\pi}\hat{v}_j^{l}\,, \qquad \text{for } j = 0,\dots,N/2-1, \\
0 &= \frac{i\hbar}{16\pi}\left[\bar{\hat{v}}_{N/2}^{l-1} - \bar{\hat{v}}_{N/2}^{l+1}\right] - \frac{\hbar^2 k^2 \Delta t}{96\pi^2}\left[\bar{\hat{v}}_{N/2}^{l-1} + 4\bar{\hat{v}}_{N/2}^{l} + \bar{\hat{v}}_{N/2}^{l+1}\right] \\
&\quad - \frac{\Delta t}{48\pi^2}{\sum_{n=0}^{N/2}}' \hat{V}_{n-N/2}\left[\bar{\hat{v}}_n^{l-1} + 4\bar{\hat{v}}_n^{l} + \bar{\hat{v}}_n^{l+1}\right] - \frac{\lambda_l}{2\pi}\bar{\hat{v}}_{N/2}^{l}\,, \\
0 &= \frac{i\hbar}{16\pi}\left[\hat{v}_{N/2}^{l+1} - \hat{v}_{N/2}^{l-1}\right] - \frac{\hbar^2 k^2 \Delta t}{96\pi^2}\left[\hat{v}_{N/2}^{l-1} + 4\hat{v}_{N/2}^{l} + \hat{v}_{N/2}^{l+1}\right] \\
&\quad - \frac{\Delta t}{48\pi^2}{\sum_{n=0}^{N/2}}' \hat{V}_{N/2-n}\left[\hat{v}_n^{l-1} + 4\hat{v}_n^{l} + \hat{v}_n^{l+1}\right] - \frac{\lambda_l}{2\pi}\hat{v}_{N/2}^{l}\,, \\
1 &= \frac{1}{2\pi}{\sum_{k=-N/2}^{N/2}}'' \hat{v}_k^{l}\bar{\hat{v}}_k^{l}\,, \\
0 &= \hat{v}_{-N/2}^{l} - \hat{v}_{N/2}^{l}\,, \\
0 &= \bar{\hat{v}}_{-N/2}^{l} - \bar{\hat{v}}_{N/2}^{l}\,.
\end{aligned}
$$

This system of $(2N+3)$-equations, allow us to solve for $\{\hat{v}_k^{l+1}, \bar{\hat{v}}_k^{l+1}\}_{k=-N/2}^{N/2}$ and λ_l from initial data, $\{\hat{v}_k^{l-1}, \bar{\hat{v}}_k^{l-1}\}_{k=-N/2}^{N/2}$ and $\{\hat{v}_k^{l}, \bar{\hat{v}}_k^{l}\}_{k=-N/2}^{N/2}$. As such, this system of equations are an example of a spectral in space, second-order in time, ***pseudospectral variational integrator*** for the time-dependent Schrödinger equation. The expressions for the variational integrator for the time-independent Schrödinger equation, which has spectral accuracy in space, are given by

$$
\begin{aligned}
\hbar^2 k^2 \bar{\hat{v}}_j &= -{\sum_{n=-N/2}^{N/2+j}}' \hat{V}_{n-j}\bar{\hat{v}}_n - \lambda\bar{\hat{v}}_j\,, \qquad \text{for } j = -N/2+1,\dots,-1, \\
\hbar^2 k^2 \hat{v}_j &= -{\sum_{n=-N/2}^{N/2+j}}' \hat{V}_{j-n}\hat{v}_n - \lambda\hat{v}_j\,, \qquad \text{for } j = -N/2+1,\dots,-1,
\end{aligned}
$$

$$\hbar^2 k^2 \bar{\hat{v}}_j = -\sum_{n=j-N/2}^{N/2}{}' \hat{V}_{n-j} \bar{\hat{v}}_n - \lambda \bar{\hat{v}}_j \,, \qquad \text{for } j = 0, \ldots, N/2-1,$$

$$\hbar^2 k^2 \hat{v}_j = -\sum_{n=j-N/2}^{N/2}{}' \hat{V}_{j-n} \hat{v}_n - \lambda \hat{v}_j \,, \qquad \text{for } j = 0, \ldots, N/2-1,$$

$$\frac{\hbar^2 k^2}{2} \bar{\hat{v}}_{N/2} = -\sum_{n=0}^{N/2}{}' \hat{V}_{n-N/2} \bar{\hat{v}}_n - \lambda \bar{\hat{v}}_{N/2} \,,$$

$$\frac{\hbar^2 k^2}{2} \hat{v}_{N/2} = -\sum_{n=0}^{N/2}{}' \hat{V}_{N/2-n} \hat{v}_n - \lambda \hat{v}_{N/2},$$

$$1 = \frac{1}{2\pi} \sum_{k=-N/2}^{N/2}{}'' \hat{v}_k \bar{\hat{v}}_k \,,$$

$$\hat{v}^l_{-N/2} = \hat{v}^l_{N/2} \,,$$

$$\bar{\hat{v}}^l_{-N/2} = \bar{\hat{v}}^l_{N/2} \,.$$

As mentioned previously, it is possible generalize this approach to construct a fully spectral variational integrator in space-time, using Chebyshev polynomials to interpolate in time the coefficients of the discrete Fourier expansion used in the spatial interpolation. The computational cost of implementing such a scheme would be significantly higher, since this would require all the spatio-temporal degrees of freedom to be solved for simultaneously.

9. Conclusions and Future Work

We have introduced the notion of a generalized Galerkin variational integrator, which is based on the idea of appropriately choosing a finite-dimensional approximation of the section of the configuration bundle, and approximating the action integral by a numerical quadrature scheme.

In contrast to standard variational methods, that are typically formulated in terms of interpolatory schemes parameterized by values of field variables at nodal and internal points, generalized Galerkin methods utilize function spaces that can be generated by arbitrary degrees of freedom. This allows the introduction of Lie group methods, and their symmetry reduction using discrete Euler–Poincaré reduction, as well as multiscale, and pseudospectral methods. Nonlinear approximation spaces allow the construction of spatio-temporally adaptive methods, which are better able to resolve shocks and other kinds of localized discontinuities in the solution.

It would be interesting to compare the performance of pseudospectral variational integrators with traditional pseudospectral schemes to see if any additional benefits arise from constructing pseudospectral schemes using a variational approach. More interesting still would be the comparison for fully spectral methods, since both variational and non-variational methods would achieve spectral accuracy, and it would make a particularly compelling case for variational integrators if their advantages persist even when compared to numerical methods with spectral accuracy.

Most mesh adaptive methods use the principle of equipartitioning the error of the numerical scheme over the mesh elements to obtain moving mesh equations. These methods rely on *a posteriori* error estimators that are related to the norm in which the accuracy of the numerical method is measured. While adaptive variational integrators exhibit an equipartitioning principle, in the sense that the discrete conjugate momentum associated with the horizontal variations are preserved from element to element in each connected

component of the domain, it would be interesting to carefully explore the question of whether this can be understood as arising from error equipartitioning with respect to a geometrically motivated error estimator.

While we have only discussed the application of multiscale variational integrators to the case of ordinary differential equations, it would be natural to consider their generalizations to partial differential equations, whereby the multiscale shape functions are obtained through well-resolved solutions of the cell problem, as in the case with multiscale finite elements (see, for example, Hou and Wu [1999]). In general, short-term simulations at the fine scale can be used to construct appropriate shape functions to obtain generalized Galerkin variational integrators at a coarser level, through the use of principal orthogonal decomposition and balanced truncation, for example. This is consistent with the coarse-fine computational approach proposed in Theodoropoulos et al. [2000], or the framework of heterogeneous multiscale methods as proposed in E and Engquist [2003].

A natural generalization would be to consider wavelet based variational integrators, as well as schemes based on conforming, hierarchical, adaptive refinement methods (CHARMS) introduced in Grinspun et al. [2002] and further developed in Krysl et al. [2003].

References

M. J. Baines. *Moving Finite Elements*. Numerical Mathematics and Scientific Computation. Oxford University Press, 1995.

Z. Chen and T. Y. Hou. A mixed multiscale finite element method for elliptic problems with oscillating coefficients. *Math. Comp.*, 72(242):541–576 (electronic), 2003.

R. A. DeVore. Nonlinear approximation. In *Acta Numerica*, volume 7, pages 51–150. Cambridge University Press, 1998.

W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.

Y. R. Efendiev, T. Y. Hou, and X. -H. Wu. Convergence of a nonconforming multiscale finite element method. *SIAM J. Numer. Anal.*, 37(3):888–910 (electronic), 2000.

E. Grinspun, P. Krysl, and P. Schröder. CHARMS: A simple framework for adaptive simulation. *ACM Transactions on Graphics (SIGGRAPH)*, 21(21):281–290, July 2002.

T. Y. Hou and X. -H. Wu. A multiscale finite element method for PDEs with oscillatory coefficients. In *Numerical treatment of multi-scale problems (Kiel, 1997)*, volume 70 of *Notes Numer. Fluid Mech.*, pages 58–69. Vieweg, 1999.

A. Iserles. On the numerical quadrature of highly-oscillating integrals I: Fourier transforms. Technical Report 2003/NA05, DAMTP, Cambridge, 2003a. (to appear in IMA J. Num. Anal.).

A. Iserles. On the numerical quadrature of highly-oscillating integrals II: Irregular oscillators. Technical Report 2003/NA09, DAMTP, Cambridge, 2003b.

A. Iserles. On the method of Neumann series for highly oscillatory equations. Technical Report 2004/NA02, DAMTP, Cambridge, 2004.

A. Iserles, H. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. In *Acta Numerica*, volume 9, pages 215–365. Cambridge University Press, 2000.

A. Iserles and S. P. Nørsett. Efficient quadrature of highly oscillatory integrals using derivatives. Technical Report 2004/NA03, DAMTP, Cambridge, 2004.

J. V. José and E. J. Saletan. *Classical Dynamics: A Contemporary Approach*. Cambridge University Press, 1998.

C. Kane, J. E. Marsden, and M. Ortiz. Symplectic-energy-momentum preserving variational integrators. *J. Math. Phys.*, 40(7):3353–3371, 1999.

P. Krysl, A. Trivedi, and B . Zhu. Object-oriented hierarchical mesh refinement with CHARMS. *Int. J. Numer. Meth. Eng.*, 2003. (to appear).

S. Lall and M. West. Discrete variational mechanics and duality. (in preparation), 2003.

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

A. Lew, J. E. Marsden, M. Ortiz, and M. West. Variational time integrators. *Int. J. Numer. Meth. Eng.*, 2004. (to appear).

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

J. E. Marsden, S. Pekarsky, and S. Shkoller. Discrete Euler–Poincaré and Lie–Poisson equations. *Nonlinearity*, 12(6):1647–1662, 1999.

J. E. Marsden, S. Pekarsky, S. Shkoller, and M. West. Variational methods, multisymplectic geometry and continuum mechanics. *J. Geom. Phys.*, 38(3-4):253–284, 2001.

J. E. Marsden and M. West. Discrete mechanics and variational integrators. In *Acta Numerica*, volume 10, pages 317–514. Cambridge University Press, 2001.

J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, 1999.

M. Oliver, M. West, and C. Wulff. Approximate momentum conservation for spatial semidiscretizations of nonlinear wave equations. *Numer. Math.*, 2004. (to appear).

K. Theodoropoulos, Y. -H. Qian, and I. G. Kevrekidis. "Coarse" stability and bifurcation analysis using timesteppers: a reaction diffusion example. *Proc. Natl. Acad. Sci.*, 97(18):9840–9843, 2000.

P. Thoutireddy and M. Ortiz. A variational r-adaptation and shape-optimization method for finite-deformation elasticity. *Int. J. Numer. Meth. Eng.*, 2003. (to appear).

L. N. Trefethen. *Spectral methods in MATLAB*. Software, Environments, and Tools. Society for Industrial and Applied Mathematics (SIAM), 2000.

107-81, Control and Dynamical Systems, Caltech, Pasadena, CA 91125.
E-mail address: mleok@cds.caltech.edu

A DISCRETE THEORY OF CONNECTIONS ON PRINCIPAL BUNDLES

MELVIN LEOK, JERROLD E. MARSDEN, AND ALAN D. WEINSTEIN

Abstract. Connections on principal bundles play a fundamental role in expressing the equations of motion for mechanical systems with symmetry in an intrinsic fashion. A discrete theory of connections on principal bundles is constructed by introducing the discrete analogue of the Atiyah sequence, with a connection corresponding to the choice of a splitting of the short exact sequence. Equivalent representations of a discrete connection are considered, and an extension of the pair groupoid composition, that takes into account the principal bundle structure, is introduced. Computational issues, such as the order of approximation, are also addressed. Discrete connections provide an intrinsic method for introducing coordinates on the reduced space for discrete mechanics, and provide the necessary discrete geometry to introduce more general discrete symmetry reduction. In addition, discrete analogues of the Levi-Civita connection, and its curvature, are introduced by using the machinery of discrete exterior calculus, and discrete connections.

Contents

1. Introduction

One of the major goals of geometric mechanics is the study of symmetry, and its consequences. An important tool in this regard is the non-singular reduction of mechanical systems under the action of free and proper symmetries, which is naturally formulated in the setting of principal bundles.

The reduction procedure results in the decomposition of the equations of motion into terms involving the shape and group variables, and the coupling between these are represented in terms of a connection on the principal bundle.

Connections and their associated curvature play an important role in the phenomena of geometric phases. A discussion of the history of geometric phases can be found in Berry [1990]. Shapere and Wilczek [1989] is a collection of papers on the theory and application of geometric phases to physics. In the rest of this section, we will survey some of the applications of geometric phases and connections to geometric mechanics and control, some of which were drawn from Marsden [1994, 1997], Marsden and Ratiu [1999].

The simulation of these phenomena requires the construction of a discrete notion of connections on principal bundles that is compatible with the approach of discrete variational mechanics, and it towards this end that this chapter is dedicated.

Falling Cat. Geometric phases arise in nature, and perhaps the most striking example of this is the falling cat, which is able to reorient itself by 180°, while remaining at *zero* angular momentum, as show in Figure 1.

The key to reconciling this with the constancy of the angular momentum is that angular momentum depends on the moment of inertia, which in turn depends on the shape of the cat. When the cat changes it shape by curling up and twisting, its moment of inertia changes, which is in turn compensated by its overall orientation changing to maintain the zero angular momentum condition. The zero angular momentum condition induces a connection on the principal bundle, and the curvature of this connection is what allows the cat to reorient itself.

A similar experiment can be tried on Earth, as described on page 10 of Vedral [2003]. This involves standing on a swivel chair, lifting your arms, and rotating them over your head, which will result in the chair swivelling around slowly.

Holonomy. The sense in which curvature is related to geometric phases is most clearly illustrated by considering the parallel transport of a vector around a curve on the sphere, as shown in Figure 2.

Think of the point on the sphere as representing the shape of the cat, and the vector as representing its orientation. The fact that the vector experiences a phase shift when parallel transported around the sphere is an example of ***holonomy***. In general, holonomy refers to a situation in geometry wherein an orthonormal frame that is parallel transported around a closed loop, back to its original position, is rotated with respect to its original orientation.

Curvature of a space is critically related to the presence of holonomy. Indeed, curvature should be thought of as being an infinitesimal version of holonomy, and this interpretation will resurface when considering the discrete analogue of curvature in the context of a discrete exterior calculus.

Foucault Pendulum. Another example relating geometric phases and holonomy is that of the Foucault pendulum. As the Earth rotates about the Sun, the Foucault pendulum exhibits a phase shift of $\Delta\theta = 2\pi\cos\alpha$ (where α is the co-latitude). This phase shift is geometric in nature, and is a consequence of holonomy. If one parallel transports an orthonormal frame around the line of constant latitude, it exhibits a phase shift that is identical to that of the Foucault pendulum, as illustrated in Figure 3.

True Polar Wander. A particular striking example of the consequences of geometric phases and the conservation of angular momentum is the phenomena of true polar wander, that was studied by Goldreich and Toomre [1969], and more recently by Leok [1998]. It is thought that some 500 to 600 million years ago, during the Vendian–Cambrian transition, the Earth, over a 15-million-year period, experienced an inertial interchange true polar wander event. This occurred when the intermediate and maximum moments of inertia crossed due to the redistribution of mass anomalies, associated with continental drift and mantle convection, thereby causing a catastrophic shift in the axis of rotation.

This phenomena is illustrated in Figure 4, wherein the places corresponding to the North and South poles of the Earth migrate towards the equator as the axis of rotation changes.

Geometric Control Theory. Geometric phases also have interesting applications and consequences in geometric control theory, and allow, for example, astronauts in free space to reorient themselves by changing their shape. By holding one of their legs straight, swivelling at the hip, and moving their foot in a circle, they are able to change their orientation. Since the reorientation only occurs as the shape is being changed, this allows the reorientation to be done with extremely high precision. Such ideas have been applied to the

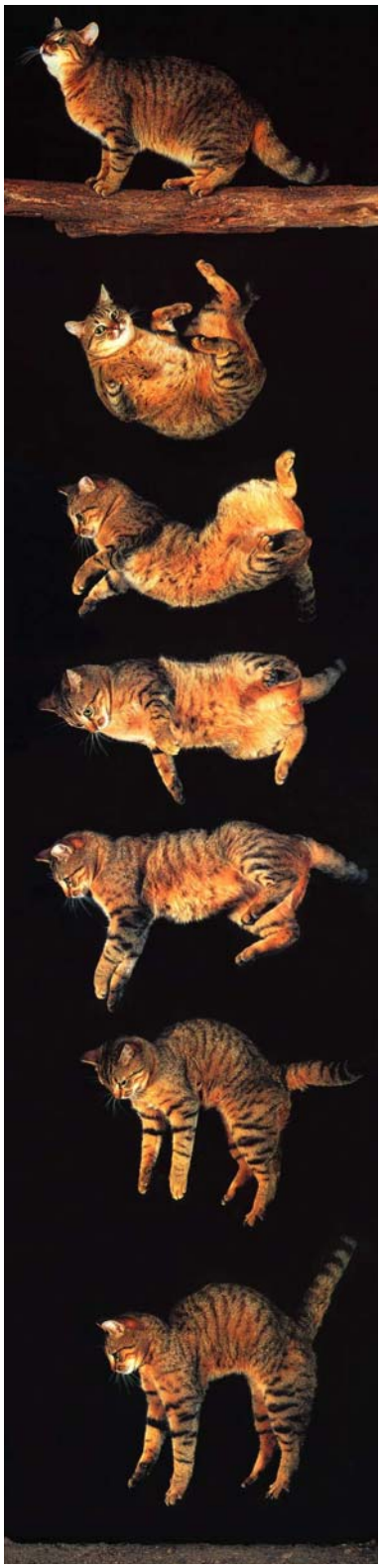


Figure 1: Reorientation of a falling cat at zero angular momentum.

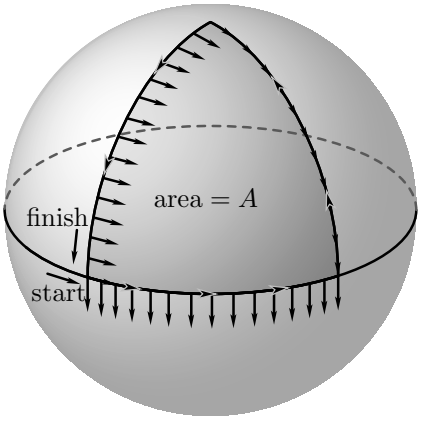


Figure 2: A parallel transport of a vector around a spherical triangle produces a phase shift.

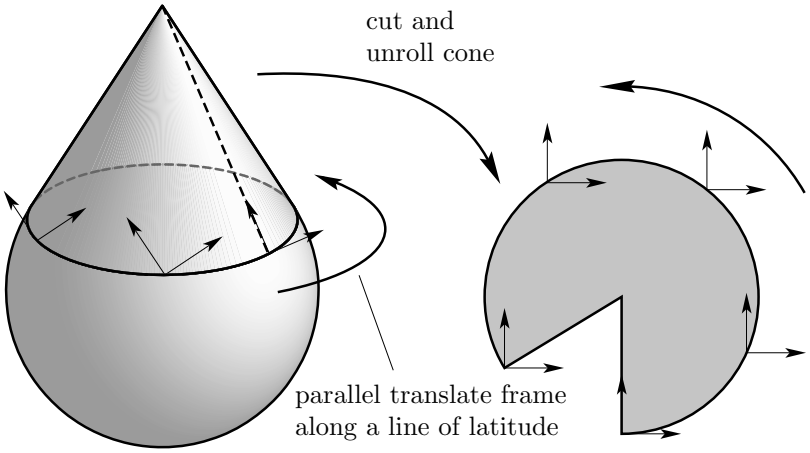


Figure 3: Geometric phase of the Foucault pendulum.

control of robots and spacecrafts; see, for example, Walsh and Sastry [1993]. The role of connections in geometric control is also addressed in-depth in Marsden [1994, 1997].

One of the theoretical underpinnings of the application of geometric phases to geometric control was developed in Montgomery [1991] and Marsden et al. [1990], in the form of the ***rigid-body phase formula***,

$$\Delta\theta = \frac{1}{\|\mu\|}\left\{\int_D \omega_\mu + 2H_\mu T\right\} = -\Lambda + \frac{2H_\mu T}{\|\mu\|},$$

the geometry of which is illustrated in Figure 5.

An example that has been studied extensively is that of the satellite with internal rotors, with a configuration space given by $Q = \mathrm{SE}(3) \times S^1 \times S^1 \times S^1$, and illustrated in Figure 6.

The generalization of the rigid-body phase formula in the presence of feedback control is particularly useful in the study and design of attitude control algorithms.

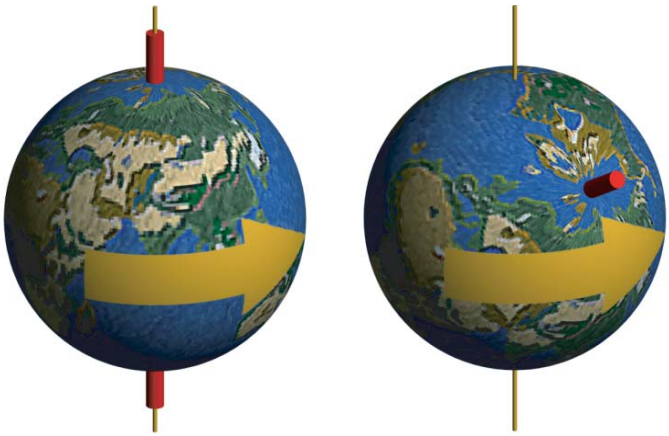


Figure 4: True Polar Wander. Red axis corresponds to the original rotational axis, and the gold axis corresponds to the instantaneous rotational axis.

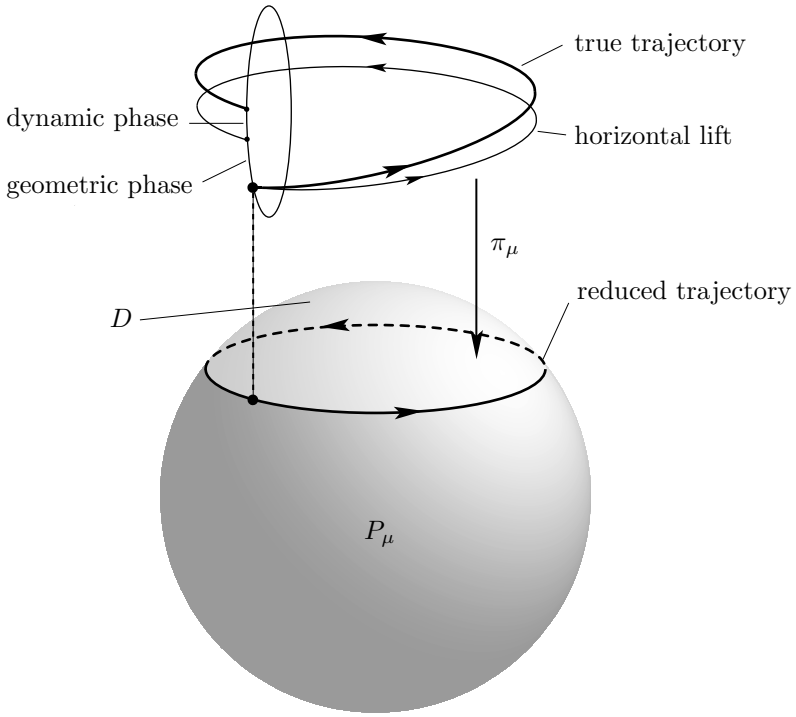


Figure 5: Geometry of rigid-body phase.

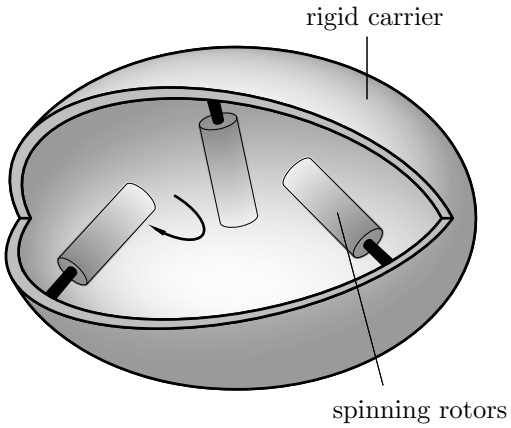


Figure 6: Rigid body with internal rotors.

2. General Theory of Bundles

Before considering the discrete analogue of connections on principal bundles, we will review some basic material on the general theory of bundles, fiber bundles, and principal fiber bundles. A more in-depth discussion of fiber bundles can be found in Steenrod [1951] and Kobayashi and Nomizu [1963].

A bundle $\mathcal{Q}$ consists of a triple (Q, S, π), where Q and S are topological spaces, respectively referred to as the ***bundle space*** and the ***base space***, and $\pi : Q \to S$ is a continuous map called the ***projection***. We may assume, without loss of generality, that π is surjective, by considering the bundle over the image $\pi(Q) \subset S$.

The ***fiber over the point*** $x \in S$, denoted F_x, is given by, $F_x = \pi^{-1}(x)$. In most situations of practical interest, the fiber at every point is homeomorphic to a common space F, in which case, F is the ***fiber*** of the bundle, and the bundle is a ***fiber bundle***. The geometry of a fiber bundle is illustrated in Figure 7.

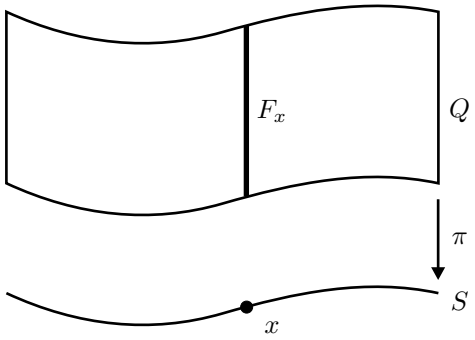


Figure 7: Geometry of a fiber bundle.

A bundle (Q, S, π) is a ***G-bundle*** if G acts on Q by left translation, and it is isomorphic to $(Q, Q/G, \pi_{Q/G})$, where Q/G is the orbit space of the G action on Q, and $\pi_{Q/G}$ is the natural projection.

If G acts freely on Q, then (Q, S, π) is called a ***principal G-bundle***, or ***principal bundle***, and G is its ***structure group***. G acting freely on Q implies that each orbit is homeomorphic to G, and therefore, $\mathcal{Q}$ is a fiber bundle with fiber G.

To make the setting for the rest of this chapter more precise, we will adopt the following definition of a principal bundle,

Definition 2.1. *A **principal bundle** is a manifold Q with a free left action, $\rho : G \times Q \to Q$, of a Lie group G, such that the natural projection, $\pi : Q \to Q/G$, is a submersion. The base space Q/G is often referred to as the **shape space** S, which is a terminology originating from reduction theory.*

We will now consider a few standard techniques for combining bundles together to form new bundles. These methods include the fiber product, Whitney sum, and the associated bundle construction.

Fiber Product. Given two bundles with the same base space, we can construct a new bundle, referred to as the ***fiber product***, which has the same base space, and a fiber which is the direct product of the fibers of the original two bundles. More formally, we have,

Definition 2.2. *Given two bundles $\pi_i : Q_i \to S$, $i = 1, 2$, the **fiber product** is the bundle,*

$$\pi_1 \times_S \pi_2 : Q_1 \times_S Q_2 \to S,$$

where $Q_1 \times_S Q_2$ is the set of all elements $(q_1, q_2) \in Q_1 \times Q_2$ such that $\pi_1(q_1) = \pi_2(q_2)$, and the projection $\pi_1 \times_S \pi_2$ is naturally defined by $\pi_1 \times_S \pi_2(q_1, q_2) = \pi_1(q_1) = \pi_2(q_2)$. The fiber is given by $(\pi_1 \times_Q \pi_2)^{-1}(x) = \pi_1^{-1}(x) \times \pi_2^{-1}(x)$.

Whitney Sum. The ***Whitney sum*** combines two vector bundles using the fiber product construction.

Definition 2.3. *Given two vector bundles $\tau_i : V_i \to Q$, $i = 1, 2$, with the same base, their **Whitney sum** is their fiber product, and it is a vector bundle over Q, and is denoted $V_1 \oplus V_2$. This bundle is obtained by taking the fiberwise direct sum of the fibers of V_1 and V_2.*

Associated Bundle. Given a principal bundle, $\pi : Q \to Q/G$, and a left action, $\rho : G \times M \to M$, of the Lie group G on a manifold M, we can construct the associated bundle.

Definition 2.4. *An **associated bundle** $\tilde{M}$ with standard fiber M is,*

$$\tilde{M} = Q \times_G M = (Q \times M)/G,$$

where the action of G on $Q \times M$ is given by $g(q, m) = (gq, gm)$. The class (or orbit) of (q, m) is denoted $[q, m]_G$ or simply $[q, m]$. The projection $\pi_M : Q \times_G M \to Q/G$ is given by,

$$\pi_M : ([q, m]_G) = \pi(q)\,,$$

and it is easy to check that it is well-defined and is a surjective submersion.

3. Connections and Bundles

Before formally introducing the precise definition of a connection, we will attempt to develop some intuition and motivation for the concept. As alluded to in the introduction to this chapter, a connection describes the curvature of a space. In the classical Riemannian setting used by Einstein in his theory of general relativity, the curvature of the space is constructed out of the connection, in terms of the Christoffel symbols that encode the connection in coordinates.

In the context of principal bundles, the connection provides a means of decomposing the tangent space to the bundle into complementary spaces, as show in Figure 8. Directions in the bundle that project to zero on the base space are called ***vertical directions***, and a ***connection*** specifies a set of directions, called ***horizontal directions***, at each point, which complements the space of vertical directions.

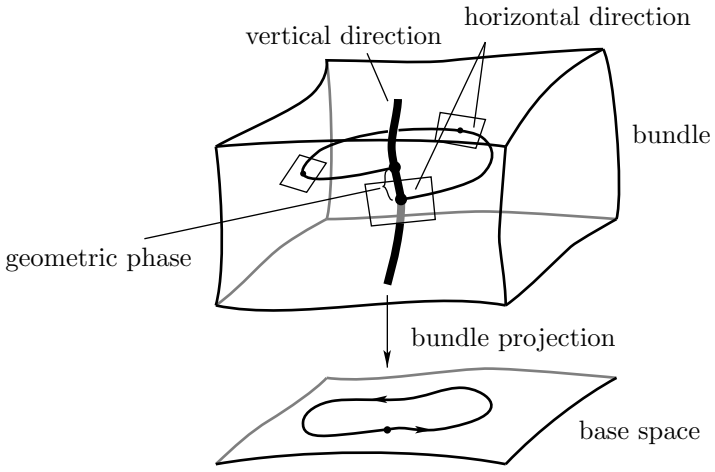


Figure 8: Geometric phase and connections.

In the rest of this section, we will formally define connections on principal bundles, and in the next section, discrete connections will be introduced in a parallel fashion.

Short Exact Sequence. This decomposition of the tangent space TQ into horizontal and vertical subspaces yields the following short exact sequence of vector bundles over Q,

$$0 \longrightarrow VQ \longrightarrow TQ \xrightarrow{\pi_*} \pi^*TS \longrightarrow 0\,,$$

where VQ is the vertical subspace of TQ, and π^*TS is the pull-back of TS by the projection $\pi : Q \to S$.

Atiyah Sequence. When the short exact sequence above is quotiented modulo G, we obtain an exact sequence of vector bundles over S,

$$0 \longrightarrow \tilde{\mathfrak{g}} \xrightarrow{i} TQ/G \xrightarrow{\pi_*} TS \longrightarrow 0\,,$$

which is called the ***Atiyah sequence*** (see, for example Atiyah [1957], Almeida and Molino [1985], Mackenzie [1995]). Here, $\tilde{\mathfrak{g}}$ is the ***adjoint bundle***, which is a special case of an associated bundle (see Definition 2.4). In particular,

$$\tilde{\mathfrak{g}} = Q \times_G \mathfrak{g} = (Q \times \mathfrak{g})/G\,,$$

where the action of G on $Q \times \mathfrak{g}$ is given by $g(q, \xi) = (gq, Ad_g\xi)$, and $\pi_{\mathfrak{g}} : \tilde{\mathfrak{g}} \to S$ is given by $\pi_{\mathfrak{g}}([q, \xi]_G) = \pi(q)$.

The maps in the Atiyah sequence, $i : (Q \times \mathfrak{g})/G \to TQ/G$ and $\pi_* : TQ/G \to TS$, are given by

$$i([q, \xi]_G) = [\xi_Q(q)]_G,$$

and

$$\pi_*([v_q]_G) = T\pi(v_q).$$

Connection 1-form. Given a connection on a principal fiber bundle $\pi : Q \to Q/G$, we can represent this as a Lie algebra-valued ***connection 1-form***, $\mathcal{A} : TQ \to \mathfrak{g}$, constructed as follows (see, for example, Kobayashi and Nomizu [1963]). Given an element of the Lie algebra $\xi \in \mathfrak{g}$, the infinitesimal generator map $\xi \mapsto \xi_Q$ yields a linear isomorphism between $\mathfrak{g}$ and V_qQ for each $q \in Q$. For each $v_q \in T_qQ$, we define $\mathcal{A}(v_q)$ to be the unique $\xi \in \mathfrak{g}$ such that ξ_Q is equal to the vertical component of v_q.

Proposition 3.1. *The connection 1-form, $\mathcal{A} : TQ \to \mathfrak{g}$, of a connection satisfies the following conditions.*

(1) *The* 1*-form is* G*-equivariant, that is,*

$$\mathcal{A} \circ TL_g = \mathrm{Ad}_g \circ \mathcal{A}\,,$$

for every $g \in G$*, where* Ad *denotes the adjoint representation of* G *in* $\mathfrak{g}$*.*

(2) *The* 1*-form induces a splitting of the Atiyah sequence, that is,*

$$\mathcal{A}(\xi_Q) = \xi\,,$$

for every $\xi \in \mathfrak{g}$*.*

Conversely, given a $\mathfrak{g}$*-valued* 1*-form* $\mathcal{A}$ *on* Q *satisfying conditions 1 and 2, there is a unique connection in* Q *whose connection* 1*-form is* $\mathcal{A}$*.*

Proof. See page 64 of Kobayashi and Nomizu [1963]. □

Horizontal Lift. The ***horizontal lift*** of a vector field $X \in \mathfrak{X}(S)$ is the unique vector field $X^h \in \mathfrak{X}(Q)$ which is horizontal and which projects onto X, that is, $T\pi_q(X^h_q) = X_{\pi(q)}$ for all $q \in Q$. The horizontal lift is in one-to-one correspondence with the choice of a connection on Q, as the following proposition states.

Proposition 3.2. *Given a connection in* Q*, and a vector field* $X \in \mathfrak{X}(S)$*, there is a unique horizontal lift* X^h *of* X*. The lift* X^h *is left-invariant under the action of* G*. Conversely, every horizontal vector field* X^h *on* Q *that is left-invariant by* G *is the lift of a vector field* $X \in \mathfrak{X}(S)$*.*

Proof. See page 65 of Kobayashi and Nomizu [1963]. □

Connection as a Splitting of the Atiyah Sequence. Consider the continuous Atiyah sequence,

$$0 \longrightarrow \tilde{\mathfrak{g}} \underset{(\pi_1,\mathcal{A})}{\overset{i}{\leftrightarrows}} TQ/G \underset{X^h}{\overset{\pi_*}{\leftrightarrows}} TS \longrightarrow 0$$

We see that the connection 1-form, $\mathcal{A} : TQ \to \mathfrak{g}$, induces a splitting of the continuous Atiyah sequence, since

$$(\pi_1, \mathcal{A}) \circ i([q,\xi]_G) = (\pi_1, \mathcal{A})([\xi_Q(q)]_g) = [q, \mathcal{A}(\xi_Q(q))]_G = [q,\xi]_G, \qquad \text{for all } q \in Q, \xi \in \mathfrak{g},$$

which is to say that $(\pi_1, \mathcal{A}) \circ i = 1_{\tilde{\mathfrak{g}}}$. Conversely, given a splitting of the continuous Atiyah sequence, we can extend the map, by equivariance, to yield a connection 1-form.

The horizontal lift also induces a splitting on the continuous Atiyah sequence, since, by definition, the horizontal lift of a vector field $X \in \mathfrak{X}(S)$ projects onto X, which is to say that $\pi_* \circ X^h = 1_{TS}$. The horizontal lift and the connection are related by the fact that

$$1_{TQ/G} = i \circ (\pi_1, \mathcal{A}) + X^h \circ \pi_*,$$

which is a simple consequence of the fact that the two splittings are part of the following commutative diagram,

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \tilde{\mathfrak{g}} & \underset{(\pi_1,\mathcal{A})}{\overset{i}{\leftrightarrows}} & TQ/G & \underset{X^h}{\overset{\pi_*}{\leftrightarrows}} & TS & \longrightarrow & 0 \\
 & & \| 1_{\tilde{\mathfrak{g}}} & & \downarrow \alpha_{\mathcal{A}} & & \| 1_{TS} & & \\
0 & \longrightarrow & \tilde{\mathfrak{g}} & \underset{\pi_1}{\overset{i_1}{\leftrightarrows}} & \tilde{\mathfrak{g}} \oplus TS & \underset{i_2}{\overset{\pi_2}{\leftrightarrows}} & TS & \longrightarrow & 0
\end{array}$$

where $\alpha_{\mathcal{A}}$ is an isomorphism. The isomorphism is given in the following lemma.

Lemma 3.3. *The map* $\alpha_{\mathcal{A}} : TQ/G \to \tilde{\mathfrak{g}} \oplus TS$ *defined by*

$$\alpha_{\mathcal{A}}([q,\dot{q}]_G) = [q, \mathcal{A}(q,\dot{q})]_G \oplus T\pi(q,\dot{q}),$$

is a well-defined vector bundle isomorphism. The inverse of $\alpha_{\mathcal{A}}$ *is given by*

$$\alpha^{-1}_{\mathcal{A}}([q,\xi]_G \oplus (x,\dot{x})) = [(x,\dot{x})^h_q + \xi q]_G.$$

Proof. See page 15 of Cendra et al. [2001]. □

This lemma, and its higher-order generalization, that identifies $T^{(2)}Q/G$ with $T^{(2)}S \times_S 2\tilde{\mathfrak{g}}$, is critical in allowing us to construct the Lagrange–Poincaré operator, which is an intrinsic method of expressing the reduced equations arising from Lagrangian reduction.

In the next section, we will develop the theory of discrete connections on principal bundles in a parallel fashion to the way we introduced continuous connections.

4. Discrete Connections

Discrete variational mechanics is based on the idea of approximating the tangent bundle TQ of Lagrangian mechanics with the pair groupoid $Q \times Q$. As such, the purpose of a discrete connection is to decompose the subset of $Q \times Q$ that projects to a neighborhood of the diagonal of $S \times S$ into horizontal and vertical spaces.

The reason why we emphasize that the construction is only valid for the subset of $Q \times Q$ that projects to a neighborhood of the diagonal of $S \times S$ is that there are topological obstructions to globalizing the construction to all of $Q \times Q$ except in the case that Q is a trivial bundle.

One of the challenges of dealing with the discrete space modelled by the pair groupoid $Q \times Q$ is that it is not a linear space, in contrast to TQ. As we shall see, the standard pair groupoid composition is not sufficient to make sense of the notion of an element $(q_0, q_1) \in Q \times Q$ being the composition of a horizontal and a vertical element. We will propose a natural notion of composing an element with a vertical element that makes sense of the horizontal and vertical decomposition.

In the subsequent sections, we will use the discrete connection to extend the pair groupoid composition even further, and explore its applications to the notion of curvature in discrete geometry.

Intrinsic Representation of the Tangent Bundle. The intuition underlying our construction of discrete horizontal and vertical spaces is best developed by considering the intrinsic representation of the tangent bundle. This representation is obtained by identifying a tangent vector at a point on the manifold with the equivalence class of curves on the manifold going through the point, such that the tangent to the curve at the point is given by the tangent vector. This notion is illustrated in Figure 9.

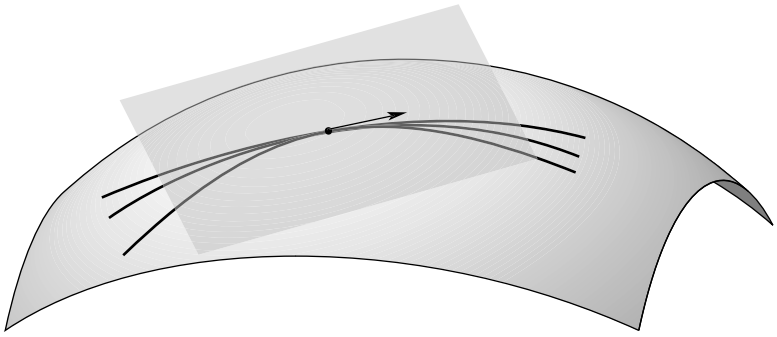


Figure 9: Intrinsic representation of the tangent bundle.

Given a vector $v_q \in TQ$, we identify it with the family of curves $q : \mathbb{R} \to Q$, such that $q(0) = q$, and $\dot{q}(0) = v$. The equivalence class $[\,\cdot\,]$ identifies curves with the same basepoint, and the same velocity at the basepoint.

With this representation, it is natural to consider $(q_0, q_1) \in Q \times Q$ to be an approximation of $[q(\cdot)] = v_q \in TQ$, in the sense that,

$$q_0 = q(0), \qquad q_1 = q(h),$$

for some fixed time step h, and where $q(\cdot)$ is a representative curve corresponding to v_q in the intrinsic representation of the tangent bundle.

4.1. **Horizontal and Vertical Subspaces of $Q \times Q$.** Recall that the vertical subspace at a point q, denoted V_q, is given by

$$V_q = \{v_q \in TQ \mid \pi_*(v_q) = 0\} = \{\xi_Q \mid \xi \in \mathfrak{g}\}.$$

Notice that the vertical space is precisely that subspace of TQ which maps under the lifted projection map to the embedded copy of S in TS. We proceed in an analogous fashion to define a discrete vertical subspace at a point q.

The natural discrete analogue of the lifted projection map π_* is the diagonal action of the projection map on $Q \times Q$, $(\pi, \pi) : Q \times Q \to Q \times Q$, where $(q_0, q_1) \mapsto (\pi q_0, \pi q_1)$. This is because

$$\pi_*(v_q) = \pi_*([q(\cdot)]) = [\pi(q(\cdot))].$$

In the same way that we embed S into TS by the map $x \mapsto [x] = 0_x$, S naturally embeds itself into the diagonal of $S \times S$, $x \mapsto (x, x) = e_{S\times S}$, which we recall is the identity subspace of the pair groupoid.

The alternative description of the vertical space is in terms of the embedding of $Q \times \mathfrak{g}$ into TQ, by $(q, \xi) \mapsto \xi_Q(q)$, using the infinitesimal generator construction,

$$\xi_Q(q) = [\exp(\xi t)q].$$

In an analogous fashion, we construct a ***discrete generator*** map, which is given in the following definition.

Definition 4.1. *The **discrete generator** is the map $i : Q \times G \to Q \times Q$, given by*

$$i(q, g) = (q, gq),$$

which we also denote by $i_q(g) = i(q, g) = (q, gq)$.

Then, we have the following definition of the ***discrete vertical space***.

Definition 4.2. *The **discrete vertical space** is given by*

$$\begin{aligned}\mathrm{Ver}_q &= \{(q, q') \in Q \times Q \mid (\pi, \pi)(q, q') = e_{S\times S}\} \\ &= \{i_q(g) \mid g \in G\}.\end{aligned}$$

This is the discrete analogue of the statement $\mathrm{Ver}_q = \{v_q \in TQ \mid \pi_*(v_q) = 0\} = \{\xi_Q \mid \xi \in \mathfrak{g}\}$.

Since the pair groupoid composition is only defined on the space of composable pairs, we need to extend the composition to make sense of how the discrete horizontal space is complementary to the discrete vertical space. In particular, we define the composition of a vertical element with an arbitrary element of $Q \times Q$ as follows.

Definition 4.3. *The composition of an arbitrary element $(q_0, q_1) \in Q \times Q$ with a vertical element is given by*

$$i_{q_0}(g) \cdot (q_0, q_1) = (e, g)(q_0, q_1) = (q_0, gq_1).$$

An elementary consequence of this definition is that it makes the discrete generator map a homomorphism.

Lemma 4.1. *The discrete generator, i_q, is a homomorphism. This is a discrete analogue of the statement in the continuous theory that $(\xi + \chi)_Q = \xi_Q + \chi_Q$.*

Proof. We compute,

$$\begin{aligned}i_q(g) \cdot i_q(h) &= i_q(g) \cdot (q, hq) \\ &= (e, g)(q, hq) \\ &= (q, ghq) \\ &= i_q(gh).\end{aligned}$$

Therefore, i_q is a homomorphism. □

If we define the G action on $Q \times G$ to be $h(q, g) = (hq, hgh^{-1})$, we find that the composition of a vertical element with an arbitrary element is G-equivariant.

Lemma 4.2. *The composition of a vertical element with an arbitrary element of $Q \times Q$ is G-equivariant,*

$$i_{hq_0}(hgh^{-1}) \cdot (hq_0, hq_1) = h \cdot i_{q_0}(g) \cdot (q_0, q_1)\,.$$

Proof. Consider the following computation,

$$\begin{aligned} i_{hq_0}(hgh^{-1}) \cdot (hq_0, hq_1) &= (hq_0, hgh^{-1}hq_1) \\ &= (hq_0, hgq_1) \\ &= h(q_0, gq_1) \\ &= h \cdot i_{q_0}(g) \cdot (q_0, q_1)\,. \end{aligned}$$ □

Having made sense of how to compose an arbitrary element of $Q \times Q$ with a vertical element, we are in a position to introduce the notion of a discrete connection.

A ***discrete connection*** is a G-equivariant choice of a subset of $Q \times Q$ called the ***discrete horizontal space***, that is complementary to the discrete vertical space. In particular, given $(q_0, q_1) \in Q \times Q$, a discrete connection decomposes this into the ***horizontal component***, $\mathrm{hor}(q_0, q_1)$, and the ***vertical component***, $\mathrm{ver}(q_0, q_1)$, such that

$$\mathrm{ver}(q_0, q_1) \cdot \mathrm{hor}(q_0, q_1) = (q_0, q_1)\,,$$

in the sense of the composition of a vertical element with an arbitrary element we defined previously. Furthermore, the G-equivariance condition states that

$$\mathrm{hor}(gq_0, gq_1) = g \cdot \mathrm{hor}(q_0, q_1)\,,$$

and

$$\mathrm{ver}(gq_0, gq_1) = g \cdot \mathrm{ver}(q_0, q_1)\,.$$

4.2. **Discrete Atiyah Sequence.** Recall that we obtain a short exact sequence corresponding to the decomposition of TQ into horizontal and vertical spaces. Due to the equivariant nature of the decomposition, quotienting this short exact sequence yields the Atiyah sequence. In this subsection, we will introduce the analogous discrete objects.

Short Exact Sequence. The decomposition of the pair groupoid $Q \times Q$, into discrete horizontal and vertical spaces, yields the following short exact sequence of bundles over Q.

$$0 \longrightarrow \mathrm{Ver}\, Q \xrightarrow{\ i\ } Q \times Q \xrightarrow{\ (\pi,\pi)\ } (\pi, \pi)^* S \times S \longrightarrow 0,$$

where $\mathrm{Ver}\, Q$ is the discrete vertical subspace of $Q \times Q$, and $(\pi, \pi)^* S \times S$ is the pull-back of $S \times S$ by the projection $(\pi, \pi) : Q \times Q \to S \times S$.

Discrete Atiyah Sequence. When the short exact sequence above is quotiented modulo G, we obtain an exact sequence of bundles over S,

$$0 \longrightarrow \tilde{G} \xrightarrow{\ i\ } (Q \times Q)/G \xrightarrow{\ (\pi,\pi)\ } S \times S \longrightarrow 0\,,$$

which we call the ***discrete Atiyah sequence***. Here, $\tilde{G}$ is an associated bundle (see Definition 2.4). In particular,

$$\tilde{G} = Q \times_G G = (Q \times G)/G\,,$$

where the action of G on $Q \times G$ is given by $g(q,h) = (gq, ghg^{-1})$, which is the natural discrete analogue of the adjoint action of $\mathfrak{g}$ on $Q \times \mathfrak{g}$. Furthermore, $\pi_G : \tilde{G} \to S$ is given by $\pi_G([q,g]_G) = \pi(q)$.

The maps in the discrete Atiyah sequence $i : \tilde{G} \to (Q \times Q)/G$, and $(\pi, \pi) : (Q \times Q)/G] \to S \times S$, are given by

$$i([q,g]_G) = [q, gq]_G = [i_q(g)]_G \,,$$

and

$$(\pi, \pi)([q_0, q_1]_g) = (\pi q_0, \pi q_1) \,.$$

4.3. **Equivalent Representations of a Discrete Connection.** In addition to the discrete connection which arises from a G-equivariant decomposition of the pair groupoid $Q \times Q$ into a discrete horizontal and vertical space, we have equivalent representations in terms of splittings of the discrete Atiyah sequence, as well as maps on the unreduced short exact sequence.

Maps on the Unreduced Short Exact Sequence. These correspond to discrete analogues of the connection 1-form, and the horizontal lift.

- Discrete connection 1-form, $\mathcal{A}_d : Q \times Q \to G$.
- Discrete horizontal lift, $(\cdot,\cdot)_q^h : S \times Q \to Q \times Q$.

Maps That Yield a Splitting of the Discrete Atiyah Sequence.

- $(\pi_1, \mathcal{A}_d) : (Q \times Q)/G \to \tilde{G}$, which is related to the discrete connection 1-form.
- $(\cdot,\cdot)^h : S \times S \to (Q \times Q)/G$, which is related to the discrete horizontal lift.

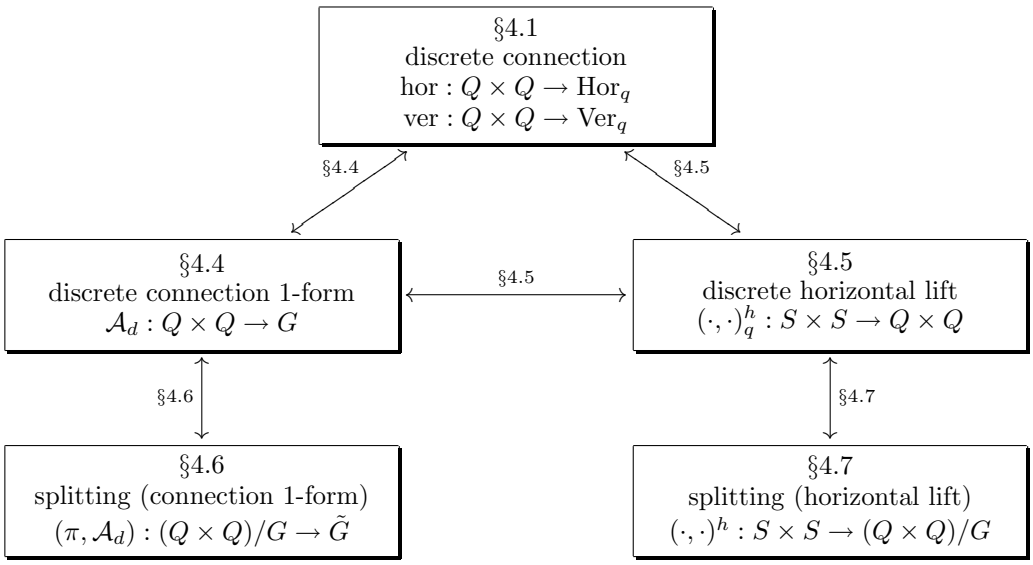
Relating the Two Sets of Representations. These two sets of representations are related in the following way:

- The maps on the unreduced short exact sequence are equivariant, and hence drop to the discrete Atiyah sequence, where they induce splittings of the short exact sequence.
- The maps that yield splittings of the discrete Atiyah sequence can be extended equivariantly to recover the maps on the unreduced short exact sequence.

Furthermore, standard results from homological algebra yield an equivalence between the two splittings of the discrete Atiyah sequence.

In the rest of this section, we will also discuss in detail the method of moving between the various representations of the discrete connection. The organization of the rest of the section, and the subsections

in which we relate the various representations are given in the following diagram.



4.4. **Discrete Connection 1-Form.** Given a discrete connection on a principal fiber bundle $\pi : Q \to Q/G$, we can represent this as a Lie group-valued ***discrete connection* 1*-form***, $\mathcal{A}_d : Q \times Q \to G$, which is a natural generalization of the Lie algebra-valued connection 1-form on tangent bundles, $\mathcal{A} : TQ \to \mathfrak{g}$, to the discrete context.

Discrete Connection 1-Forms from Discrete Connections. The discrete connection 1-form is constructed as follows. Given an element of the Lie group $g \in G$, the discrete generator map $g \mapsto i_q(g)$ yields an isomorphism between G and Ver_q for each $q \in Q$. For each $(q_0, q_1) \in Q \times Q$, we define $\mathcal{A}_d(q_0, q_1)$ to be the unique $g \in G$ such that $i_q(g)$ is equal to the vertical component of (q_0, q_1). In particular, this is equivalent to the condition that the following statement holds,

$$(q_0, q_1) = i_{q_0}(\mathcal{A}_d(q_0, q_1)) \cdot \mathrm{hor}(q_0, q_1).$$

Remark 4.1. *It follows from the above identity that the discrete horizontal space can also be expressed as*

$$\begin{aligned}\mathrm{Hor}_{q_0} &= \{(q_0, q_1) \in Q \times Q \mid \mathrm{hor}(q_0, q_1) = (q_0, q_1)\} \\ &= \{(q_0, q_1) \in Q \times Q \mid \mathcal{A}_d(q_0, q_1) = e\}.\end{aligned}$$

We will now establish a few properties of the discrete connection 1-form.

Proposition 4.3. *The* ***discrete connection* 1*-form***, $\mathcal{A}_d : Q \times Q \to G$, *satisfies the following properties.*

(1) *The 1-form is G-equivariant, that is,*

$$\mathcal{A}_d \circ L_g = I_g \circ \mathcal{A}_d,$$

which is the discrete analogue of the G-equivariance of the continuous connection, $\mathcal{A} \circ TL_g = Ad_g \circ \mathcal{A}$.

(2) *The 1-form induces a splitting of the Discrete Atiyah sequence, that is,*

$$\mathcal{A}_d(i_q(g)) = \mathcal{A}_d(q_0, gq_0) = g,$$

which is the discrete analogue of $\mathcal{A}(\xi_Q) = \xi$.

Proof. The proof relies on the properties of a discrete connection, and the definition of the discrete connection 1-form.

(1) The discrete connection 1-form satisfies the condition

$$(q_0, q_1) = i_{q_0}(\mathcal{A}_d(q_0, q_1)) \cdot \mathrm{hor}(q_0, q_1)\,.$$

If we denote $\mathrm{hor}(q_0, q_1)$ by $(q_0, \bar{q}_1)$, we have that

$$(q_0, q_1) = (q_0, \mathcal{A}_d(q_0, q_1)\bar{q}_1)\,.$$

Similarly, we have,

$$\begin{aligned}(gq_0, gq_1) &= i_{gq_0}(\mathcal{A}_d(gq_0, gq_1)) \cdot \mathrm{hor}(gq_0, gq_1)\\ &= i_{gq_0}(\mathcal{A}_d(gq_0, gq_1)) \cdot g \cdot \mathrm{hor}(q_0, q_1)\\ &= (e, \mathcal{A}_d(gq_0, gq_1))(gq_0, g\bar{q}_1)\\ &= (gq_0, \mathcal{A}_d(gq_0, gq_1)g\bar{q}_1)\,,\end{aligned}$$

where we have used the G-equivariance of the discrete horizontal space. By looking at the expressions for gq_1 and q_1, we conclude that

$$\begin{aligned}\mathcal{A}_d(gq_0, gq_1)g\bar{q}_1 &= gq_1\\ &= g\mathcal{A}_d(q_0, q_1)\bar{q}_1\,,\\ \mathcal{A}_d(gq_0, gq_1)g &= g\mathcal{A}_d(q_0, q_1)\,,\\ \mathcal{A}_d(gq_0, gq_1) &= g\mathcal{A}_d(q_0, q_1)g^{-1}\,,\end{aligned}$$

which is precisely the statement that $\mathcal{A}_d \circ L_g = I_g \circ \mathcal{A}_d$, that is to say that $\mathcal{A}_d$ is G-equivariant.

(2) Recall that $i_q(g)$ is an element of the discrete vertical space. Since the discrete horizontal space is complementary to the discrete vertical space, it follows that $\mathrm{ver}(i_q(g)) = i_q(g)$. Then, by the construction of the discrete connection 1-form, $\mathcal{A}_d(i_q(g))$ is the unique element of G such that

$$i_q(\mathcal{A}_d(i_q(g))) = \mathrm{ver}(i_q(g)) = i_q(g)\,.$$

Since i_q is an isomorphism between G and the discrete vertical space, we conclude that $\mathcal{A}_d(i_q(g)) = g$, as desired. □

The second result is equivalent to the map recovering the discrete Euler–Poincaré connection when restricted to a G-fiber, that is, $\mathcal{A}_d(x, g_0, x, g_1) = g_1 g_0^{-1}$. In particular, it follows that the map is trivial when restricted to the diagonal space, that is, $\mathcal{A}_d(q, q) = e$.

The properties of a discrete connection are discrete analogues of the properties of a continuous connection in the sense that if a discrete connection has a given property, the corresponding continuous connection which is induced in the infinitesimal limit has the analogous continuous property. The precise sense in which a discrete connection induces a continuous connection will be discussed in §5.2.

Discrete Connections from Discrete Connection 1-Forms. Having shown how to obtain a discrete connection 1-form from a discrete connection, let us consider the converse case of obtaining a discrete connection from a discrete connection 1-form with the properties above. We do this by constructing the discrete horizontal and vertical components as follows.

Definition 4.4. *Given a discrete connection 1-form, $\mathcal{A}_d : Q \times Q \to G$ that is G-equivariant and induces a splitting of the discrete Atiyah sequence, we define the* ***horizontal component*** *to be*

$$\mathrm{hor}(q_0, q_1) = i_{q_0}((\mathcal{A}_d(q_0, q_1))^{-1}) \cdot (q_0, q_1)\,.$$

The ***vertical component*** *is given by*

$$\mathrm{ver}(q_0, q_1) = i_{q_0}(\mathcal{A}_d(q_0, q_1))\,.$$

Proposition 4.4. *The discrete connection we obtain from a discrete connection 1-form has the following properties.*

(1) *The discrete connection yields a horizontal and vertical decomposition of $Q \times Q$, in the sense that*

$$(q_0, q_1) = \mathrm{ver}(q_0, q_1) \cdot \mathrm{hor}(q_0, q_1)\,,$$

for all $(q_0, q_1) \in Q \times Q$.

(2) *The discrete connection is G-equivariant, in the sense that*

$$\mathrm{hor}(gq_0, gq_1) = g \cdot \mathrm{hor}(q_0, q_1)\,,$$

and

$$\mathrm{ver}(gq_0, gq_1) = g \cdot \mathrm{ver}(q_0, q_1)\,.$$

Proof. The proof relies on the properties of the discrete connection 1-form, and the definitions of the discrete horizontal and vertical spaces.

(1) Consider the following computation,

$$\begin{aligned}
\mathrm{ver}(q_0, q_1) \cdot \mathrm{hor}(q_0, q_1) &= i_{q_0}(\mathcal{A}_d(q_0, q_1)) \cdot i_{q_0}((\mathcal{A}_d(q_0, q_1))^{-1}) \cdot (q_0, q_1) \\
&= i_{q_0}(\mathcal{A}_d(q_0, q_1)(\mathcal{A}_d(q_0, q_1))^{-1}) \cdot (q_0, q_1) \\
&= i_{q_0}(e) \cdot (q_0, q_1) \\
&= (q_0, q_1)\,,
\end{aligned}$$

where we used that i_q is a homomorphism (see Lemma 4.1).

(2) We compute,

$$\begin{aligned}
\mathrm{hor}(gq_0, gq_1) &= i_{gq_0}((\mathcal{A}_d(gq_0, gq_1))^{-1}) \cdot (gq_0, gq_1) \\
&= i_{gq_0}(g(\mathcal{A}_d(q_0, q_1))^{-1} g^{-1}) \cdot (gq_0, gq_1) \\
&= (e, g(\mathcal{A}_d(q_0, q_1))^{-1} g^{-1})(gq_0, gq_1) \\
&= (gq_0, g(\mathcal{A}_d(q_0, q_1))^{-1} q_1) \\
&= g \cdot i_{q_0}((\mathcal{A}_d(q_0, q_1))^{-1}) \cdot (q_0, q_1) \\
&= g \cdot \mathrm{hor}(q_0, q_1)\,,
\end{aligned}$$

where we have used the fact that the composition of a vertical element with an arbitrary element is G-equivariant (see Lemma 4.2). Similarly, we compute,

$$\begin{aligned}
\mathrm{ver}(gq_0, gq_1) &= i_{gq_0}(\mathcal{A}_d(gq_0, gq_1)) \\
&= i_{gq_0}(g\mathcal{A}_d(q_0, q_1) g^{-1}) \\
&= (gq_0, g\mathcal{A}_d(q_0, q_1) g^{-1} gq_0) \\
&= g \cdot (q_0, \mathcal{A}_d(q_0, q_1) q_0) \\
&= g \cdot i_{q_0}(\mathcal{A}_d(q_0, q_1)) \\
&= g \cdot \mathrm{ver}(q_0, q_1)\,.
\end{aligned}$$

□

Local Representation of the Discrete Connection 1-Form. Since the discrete connection 1-form can be thought of as comparing group fiber quantities at different base points, we obtain the natural identity that

$$\mathcal{A}_d(gq_0, hq_1) = h\mathcal{A}_d(q_0, q_1) g^{-1}.$$

In a local trivialization, this corresponds to

$$\mathcal{A}_d(x_0, g_0, x_1, g_1) = g_1 \mathcal{A}_d(x_0, e, x_1, e) g_0^{-1}.$$

We define

$$A(x_0, x_1) = \mathcal{A}_d(x_0, e, x_1, e),$$

which yields the local representation of the discrete connection 1-form.

Definition 4.5. *Given a discrete connection* 1*-form,* $\mathcal{A}_d : Q \times Q \to G$*, its* ***local representation*** *is given by*

$$\mathcal{A}_d(x_0, g_0, x_1, g_1) = g_1 A(x_0, x_1) g_0^{-1}\,,$$

where

$$A(x_0, x_1) = \mathcal{A}_d(x_0, e, x_1, e)\,.$$

Lemma 4.5. *The local representation of a discrete connection is* G*-equivariant.*

Proof. Consider the following computation,

$$\begin{aligned}\mathcal{A}_d(g(x_0, g_0), g(x_1, g_1)) &= \mathcal{A}_d((x_0, gg_0), (x_1, gg_1))\\ &= gg_1 A(x_0, x_1)(gg_0)^{-1}\\ &= g(g_1 A(x_0, x_1) g_0^{-1}) g^{-1}\\ &= g\mathcal{A}_d((x_0, g_0), (x_1, g_1)) g^{-1}\,,\end{aligned}$$

which shows that the local representation is G-equivariant, as expected. □

Notice also that in the pure group case, where $Q = G$, this recovers the discrete Euler–Poincaré connection, as we would expect, since the shape space is trivial. In particular, $x_0 = x_1 = e$, which implies that $A(x_0, x_1) = \mathcal{A}_d(e, e, e, e) = e$, and $\mathcal{A}_d(g_0, g_1) = g_1 g_0^{-1}$.

Example 4.1. *As an example, we construct the natural discrete analogue of the mechanical connection,* $\mathcal{A} : TQ \to \mathfrak{g}$*, by the following procedure, which yields a discrete connection* 1*-form,* $\mathcal{A}_d : Q \times Q \to G$*.*

1. *Given the point* $(q_0, q_1) \in Q \times Q$*, we construct the geodesic path* $q_{01} : [0, 1] \to Q$ *with respect to the kinetic energy metric, such that* $q_{01}(0) = q_0$*, and* $q_{01}(1) = q_1$*.*
2. *Project the geodesic path to the shape space,* $x_{01}(t) \equiv \pi q_{01}(t)$*, to obtain the curve* x_{01} *on* S*.*
3. *Taking the horizontal lift of* x_{01} *to* Q *using the connection* $\mathcal{A}$ *yields* $\tilde{q}_{01}$*.*
4. *There is a unique* $g \in G$ *such that* $q_{01}(1) = g \cdot \tilde{q}_{01}(1)$*.*
5. *Define* $\mathcal{A}_d(q_0, q_1) = g$*.*

This discrete connection is consistent with the classical notion of a connection in the limit that q_1 *approaches* q_0*, in the usual sense in which discrete mechanics on* $Q \times Q$ *converges to continuous Lagrangian mechanics on* TQ*. As mentioned before, this statement is made more precise in §5.2.*

4.5. **Discrete Horizontal Lift.** The ***discrete horizontal lift*** of an element $(x_0, x_1) \in S \times S$ is the subset of $Q \times Q$ that are horizontal elements, and project to (x_0, x_1). Once we specify the base point $q \in Q$, the discrete horizontal lift is unique, and we introduce the map $(\cdot, \cdot)_q^h : S \times S \to Q \times Q$.

Discrete Horizontal Lifts from Discrete Connections. The discrete horizontal lift can be constructed once the discrete horizontal space is defined by a choice of discrete connection.

Definition 4.6. *The* ***discrete horizontal lift*** *is the unique map* $(\cdot, \cdot)_q^h : S \times S \to Q \times Q$*, such that*

$$(\pi, \pi) \cdot (x_0, x_1)_q^h = (x_0, x_1)\,,$$

and

$$(x_0, x_1)_q^h \in \mathrm{Hor}_q\ .$$

Lemma 4.6. *The discrete horizontal lift is* G*-equivariant, which is to say that*

$$(x_0, x_1)_{gq}^h = g \cdot (x_0, x_1)_q^h\,.$$

Proof. Given $(x_0, x_1) \in S \times S$, denote $(x_0, x_1)^h_{q_0}$ by (q_0, q_1). Then, by the definition of the discrete horizontal lift, we have that

$$(\pi, \pi) \cdot (q_0, q_1) = (x_0, x_1)\,,$$

and it follows that

$$(\pi, \pi) \cdot (gq_0, gq_1) = (x_0, x_1)\,.$$

Also, from the definition of the discrete horizontal lift,

$$(q_0, q_1) \in \mathrm{Hor}_{q_0}\,,$$

and by the G-equivariance of the horizontal space,

$$(gq_0, gq_1) \in g \cdot \mathrm{Hor}_{q_0} = \mathrm{Hor}_{gq_0}\,.$$

This implies that (gq_0, gq_1) satisfies the conditions for being the discrete horizontal lift of (x_0, x_1) with basepoint gq_0. Therefore, $(x_0, x_1)^h_{gq_0} = (gq_0, gq_1) = g \cdot (q_0, q_1) = g \cdot (x_0, x_1)^h_{q_0}$, as desired. □

Discrete Connections from Discrete Horizontal Lifts. Conversely, given a discrete horizontal lift, we can recover a discrete connection.

Definition 4.7. *Given a discrete horizontal lift, we define the* ***horizontal component*** *to be*

$$\mathrm{hor}(q_0, q_1) = (\pi(q_0, q_1))^h_{q_0}\,,$$

and the ***vertical component*** *is given by*

$$\mathrm{ver}(q_0, q_1) = i_{q_0}(g)\,,$$

where g is the unique group element such that

$$(q_0, q_1) = i_{q_0}(g) \cdot \mathrm{hor}(q_0, q_1)\,.$$

The last expression simply states that the discrete horizontal and vertical space are complementary with respect to the composition we defined between a vertical element and an arbitrary element of $Q \times Q$.

Discrete Horizontal Lifts from Discrete Connection 1-Forms. We wish to construct a discrete horizontal lift $(\cdot, \cdot)^h : S \times S \to (Q \times Q)/G$, given a discrete connection $\mathcal{A}_d : Q \times Q \to G$. We state the construction of such a discrete horizontal lift as a proposition.

Proposition 4.7. *Given a discrete connection 1-form, $\mathcal{A}_d : Q \times Q \to G$, the discrete horizontal lift is given by*

$$(x_0, x_1)^h = [\pi^{-1}(x_0, x_1) \cap \mathcal{A}_d^{-1}(e)]_G.$$

Furthermore, the discrete horizontal lift satisfies the following identity,

$$i_{q_0}(\mathcal{A}_d(q_0, q_1)) \cdot (\pi(q_0, q_1))^h_{q_0} = (q_0, q_1),$$

which implies that the discrete connection 1-form and the discrete horizontal lift induces a horizontal and vertical decomposition of $Q \times Q$.

The horizontal lift can be expressed in a local trivialization, where $q_0 = (x_0, g_0)$, using the local expression for the discrete connection,

$$(x_0, x_1)^h_{q_0} = (x_0, g_0, x_1, g_0(A(x_0, x_1))^{-1}).$$

Proof. We will show that this operation is well-defined on the quotient space. Using the local representation of the discrete connection in the local trivialization (see Definition 4.5), we have,

$$\begin{aligned}&\mathcal{A}_d^{-1}(e) \cap \pi^{-1}(x_0, x_1)\\ &\quad = \{(\tilde{x}_0, g, \tilde{x}_1, g \cdot (A(\tilde{x}_0, \tilde{x}_1))^{-1}) \mid \tilde{x}_0, \tilde{x}_1 \in S, g \in G\}\end{aligned}$$

$$\begin{aligned}
&\cap \{(x_0, h_0, x_1, h_1) \mid h_0, h_1 \in G\} \\
&= \{(x_0, g, x_1, g \cdot (A(x_0, x_1))^{-1}) \mid h \in G\} \\
&= G \cdot (x_0, e, x_1, (A(x_0, x_1))^{-1}),
\end{aligned}$$

which is a well-defined element of $(Q \times Q)/G$. Since this is true in a local trivialization, and both the discrete connection and projection operators are globally defined, this inverse coset is globally well-defined as an element of $(Q \times Q)/G$.

In particular, the computation above allows us to obtain a local expression for the discrete horizontal lift in terms of the local representation of the discrete connection. That is,

$$\begin{aligned}
(x_0, x_1)^h_{(x_0,e)} &= (x_0, e, x_1, (A(x_0, x_1))^{-1}), \\
(x_0, x_1)^h &= [(x_0, e, x_1, (A(x_0, x_1))^{-1})]_G.
\end{aligned}$$

By the properties of the discrete horizontal lift, this extends to $\pi^{-1}(x_0, x_1) \subset Q \times Q$,

$$\begin{aligned}
(x_0, x_1)^h_{(x_0,g)} &= (x_0, x_1)^h_{g(x_0,e)} \\
&= g \cdot (x_0, x_1)^h_{(x_0,e)} \\
&= g(x_0, e, x_1, (A(x_0, x_1))^{-1}) \\
&= (x_0, g, x_1, g(A(x_0, x_1))^{-1}).
\end{aligned}$$

To prove the second claim, we have in the local trivialization of $Q \times Q$, $(q_0, q_1) = (x_0, g_0, x_1, g_1)$. Then, by the result above,

$$(\pi(q_0, q_1))^h_{q_0} = (\pi(q_0, q_1))^h_{(x_0,g_0)} = (x_0, g_0, x_1, g_0(A(x_0, x_1))^{-1}).$$

Also, by the local representation of the discrete connection,

$$\mathcal{A}_d(q_0, q_1) = g_1 A(x_0, x_1) g_0^{-1}.$$

Therefore,

$$\begin{aligned}
i_{q_0}(\mathcal{A}_d(q_0, q_1) \cdot (\pi(q_0, q_1))^h_{q_0} &= (e, \mathcal{A}_d(q_0, q_1)) \cdot (\pi(q_0, q_1))^h_{q_0} \\
&= (e, g_1 A(x_0, x_1) g_0^{-1}) \cdot (x_0, g_0, x_1, g_0(A(x_0, x_1))^{-1}) \\
&= (x_0, g_0, x_1, (g_1 A(x_0, x_1) g_0^{-1})(g_0(A(x_0, x_1))^{-1})) \\
&= (x_0, g_0, x_1, g_1) \\
&= (q_0, q_1),
\end{aligned}$$

as claimed. □

Discrete Connection 1-Forms from Discrete Horizontal Lifts. Given a horizontal lift $(\cdot, \cdot)^h_q : S \times S \to Q \times Q$, we wish to construct a discrete connection 1-form, $\mathcal{A}_d : Q \times Q \to G$.

Lemma 4.8. *Given a discrete horizontal lift, $(\cdot, \cdot)^h_q : S \times S \to Q \times Q$, the **discrete connection** 1-**form**, $\mathcal{A}_d : Q \times Q \to G$, is uniquely defined by the following identity,*

$$i_{q_0}(\mathcal{A}_d(q_0, q_1)) \cdot (\pi(q_0, q_1))^h_{q_0} = (q_0, q_1).$$

Proof. To show that this construction is well-defined, we note that $\pi_1(q_0, q_1) = \pi_1(\pi(q_0, q_1))^h_{q_0}$, by the construction of $(\cdot, \cdot)^h_{q_0}$ from $(\cdot, \cdot)^h : S \times S \to (Q \times Q)/G$. Furthermore, $\pi_2(q_0, q_1)$ and $\pi_2(\pi(q_0, q_1))^h_{q_0}$ are in the same fiber of the principal bundle $\pi : Q \to Q/G$ and are therefore related by a unique element $g \in G$. Since this element is unique, $\mathcal{A}_d(q_0, q_1)$ is uniquely defined by the identity. □

4.6. **Splitting of the Discrete Atiyah Sequence (Connection 1-Form).** Consider the discrete Atiyah sequence,

$$0 \longrightarrow \tilde{G} \underset{(\pi_1,\mathcal{A}_d)}{\overset{(q,gq)}{\leftarrow\!\!-\!\!\rightarrow}} (Q\times Q)/G \underset{(\cdot,\cdot)^h}{\overset{(\pi,\pi)}{\leftarrow\!\!-\!\!\rightarrow}} S\times S \longrightarrow 0 \ .$$

Given a short exact sequence

$$0 \longrightarrow A_1 \underset{k}{\overset{f}{\leftarrow\!\!-\!\!\rightarrow}} B \underset{h}{\overset{g}{\leftarrow\!\!-\!\!\rightarrow}} A_2 \longrightarrow 0 \ ,$$

there are three equivalent conditions under which the exact sequence is split. They are as follows,

(1) There is a homomorphism $h : A_2 \to B$ with $g \circ h = 1_{A_2}$;
(2) There is a homomorphism $k : B \to A_1$ with $k \circ f = 1_{A_1}$;
(3) The given sequence is isomorphic (with identity maps on A_1 and A_2) to the direct sum short exact sequence,
$$0 \longrightarrow A_1 \xrightarrow{i_1} A_1 \oplus A_2 \xrightarrow{\pi_2} A_2 \longrightarrow 0 \ ,$$
and in particular, $B \cong A_1 \oplus A_2$.

We will address all three conditions in this and the next two subsections.

Splittings from Discrete Connection 1-Forms. A discrete connection 1-form, $\mathcal{A}_d : Q\times Q \to G$, induces a splitting of the discrete Atiyah sequence, in the sense that

$$(\pi_1, \mathcal{A}_d) \circ i = 1_{\tilde{G}} \, .$$

Lemma 4.9. *Given a discrete connection 1-form, $\mathcal{A}_d : Q\times Q \to G$, we obtain a splitting of the discrete Atiyah sequence, $\varphi : (Q\times Q)/G \to \tilde{G}$, which is given by*

$$\varphi([q_0, q_1]_G) = [q_0, \mathcal{A}_d(q_0, q_1)]_G.$$

We denote this map by $(\pi_1, \mathcal{A}_d)$.

Proof. This expression is well-defined, as the following computation shows,

$$\begin{aligned}\varphi([gq_0, gq_1]_G) &= [gq_0, \mathcal{A}_d(gq_0, gq_1)]_G \\ &= [gq_0, g\mathcal{A}_d(q_0, q_1)g^{-1}]_G \\ &= \varphi([q_0, q_1]_G).\end{aligned}$$

Furthermore, since

$$\begin{aligned}(\pi_1, \mathcal{A}_d) \circ i([q, g]_G) &= (\pi_1, \mathcal{A}_d)([q, gq]_G) \\ &= [\pi_1(q, gq), \mathcal{A}_d(q, gq)]_G \\ &= [q, g]_G,\end{aligned}$$

it follows that we obtain a splitting of the discrete Atiyah sequence. □

Discrete Connection 1-Forms from Splittings. Given a splitting of the discrete Atiyah sequence, we can obtain a discrete connection 1-form using the following construction.

Given $[q_0, q_1]_G \in (Q\times Q)/G$, we obtain from the splitting of the discrete Atiyah sequence an element, $[q, g]_G \in \tilde{G}$. Viewing $[q, g]_G$ as a subset of $Q \times G$, consider the unique $\tilde{g}$ such that $(q_0, \tilde{g}) \in [q, g]_G \subset Q \times G$. Then, we define

$$\mathcal{A}_d(x_0, e, x_1, g_0^{-1}g_1) = \tilde{g}.$$

We extend this definition to the whole of $Q \times Q$ by equivariance,

$$\mathcal{A}_d(x_0, g_0, x_1, g_1) = g_0\tilde{g}g_0^{-1}.$$

Lemma 4.10. *Given a splitting of the discrete Atiyah sequence, the construction above yields a discrete connection* 1*-form with the requisite properties.*

Proof. To show that the $\mathcal{A}_d$ satisfies the properties of a discrete connection 1-form, we first note that equivariance follows from the construction.

Since we have a splitting, it follows that $\varphi([q, gq]_G) = [q, g]_G$, as φ composed with the map from $\tilde{G}$ to $(Q \times Q)/G$ is the identity on $\tilde{G}$. Using a local trivialization, we have,

$$\begin{aligned}[q_0, g]_G &= \varphi([q_0, gq_0]_G) \\ &= \varphi([(x_0, e), (x_0, g_0^{-1} g g_0)]_G) \\ &= [(x_0, e), \tilde{g}]_G \\ &= [(x_0, g_0), g_0 \tilde{g} g_0^{-1}]_G.\end{aligned}$$

Then, by definition,

$$\mathcal{A}_d((x_0, e), (x_0, g_0^{-1} g g_0)) = \tilde{g},$$

and furthermore, $g = g_0 \tilde{g} g_0^{-1}$. From this, we conclude that

$$\begin{aligned}\mathcal{A}_d(q_0, gq_0) &= \mathcal{A}_d((x_0, g_0), (x_0, g g_0)) \\ &= g_0 \mathcal{A}_d((x_0, e), (x_0, g_0^{-1} g g_0)) g_0^{-1} \\ &= g_0 \tilde{g} g_0^{-1} \\ &= g.\end{aligned}$$

Therefore, we have that $\mathcal{A}_d(q_0, gq_0) = g$, which together with equivariance implies that $\mathcal{A}_d$ is a discrete connection 1-form. □

4.7. **Splitting of the Discrete Atiyah Sequence (Horizontal Lift).** As was the case with the discrete connection 1-form, the discrete horizontal lift is in one-to-one correspondence with splittings of the discrete Atiyah sequence, and they are related by taking the quotient, or extending by G-equivariance, as appropriate.
Splittings from Discrete Horizontal Lifts. Given a discrete horizontal lift, we obtain a splitting by taking its quotient.

Lemma 4.11. *Given a discrete horizontal lift,* $(\cdot, \cdot)_q^h : S \times S \to Q \times Q$*, the map* $(\cdot, \cdot)^h : S \times S \to (Q \times Q)/G$*, which is given by*

$$(x_0, x_1)^h = [(x_0, x_1)^h_{(x_0, e)}]_G \,,$$

induces a splitting of the discrete Atiyah sequence.

Proof. We compute,

$$\begin{aligned}(\pi, \pi) \circ (x_0, x_1)^h &= (\pi, \pi)([(x_0, x_1)^h_{(x_0, e)}]_G) \\ &= (x_0, x_1) \,,\end{aligned}$$

where we used the G-equivariance of the discrete horizontal lift, and the property that $(\pi, \pi) \cdot (x_0, x_1)^h_q = (x_0, x_1)$ for any $q \in Q$. This implies that $(\pi, \pi) \circ (\cdot, \cdot)^h = 1_{S \times S}$, as desired. □

Discrete Horizontal Lifts from Splittings. Given a splitting, $(\cdot, \cdot)^h : S \times S \to (Q \times Q)/G$, we obtain a discrete horizontal lift, $(\cdot, \cdot)_q^h : S \times S \to Q \times Q$, using the following construction.

We denote by $(x_0, x_1)^h_{q_0}$ the unique element in $(x_0, x_1)^h$, thought of as a subset of $Q \times Q$, such that the first component is q_0. This is the discrete horizontal lift of the point $(x_0, x_1) \in S \times S$ where the base point is specified.

Lemma 4.12. *Given a splitting of the discrete Atiyah sequence, the construction above yields a discrete horizontal lift with the requisite properties.*

Proof. Since the quotient space $(Q \times Q)/G$ is obtained by the diagonal action of G on $Q \times Q$, it follows that if $(x_0, x_1)^h_{q_0} \in (x_0, x_1)^h \subset Q \times Q$, then $g \cdot (x_0, x_1)^h_{q_0} \in (x_0, x_1)^h \subset Q \times Q$. Since the first component of $G \cdot (x_0, x_1)^h_{q_0}$ is gq_0, and $g \cdot (x_0, x_1)^h_{q_0} \in (x_0, x_1)^h \subset Q \times Q$, we have that

$$(x_0, x_1)^h_{gq_0} = g \cdot (x_0, x_1)^h_{q_0},$$

which is to say that the discrete horizontal lift constructed above is G-equivariant.

Since $(\cdot, \cdot)^h$ is a splitting of the discrete Atiyah sequence, we have that $(\pi, \pi) \circ (\cdot, \cdot)^h = 1_{S \times S}$, and this implies that any element in $(x_0, x_1)^h$, viewed as a subset of $Q \times Q$, projects to (x_0, x_1). Therefore, the discrete horizontal lift we constructed above has the requisite properties. □

4.8. **Isomorphism between $(Q \times Q)/G$ and $(S \times S) \oplus \tilde{G}$.** The notion of a discrete connection is motivated by the desire to construct a global diffeomorphism between $(Q \times Q)/G \to S$ and $(S \times S) \oplus \tilde{G} \to S$. This is the discrete analogue of the identification between $TQ/G \to Q/G$ and $T(Q/G) \oplus \tilde{\mathfrak{g}} \to Q/G$ which is the context for Lagrangian Reduction in Cendra et al. [2001]. Since a choice of discrete connection corresponds to a choice of splitting of the discrete Atiyah sequence, we have the following commutative diagram, where each row is a short exact sequence.

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \tilde{G} & \underset{(\pi_1, \mathcal{A}_d)}{\overset{(q,gq)}{\longleftrightarrow}} & (Q \times Q)/G & \underset{(\cdot,\cdot)^h}{\overset{(\pi,\pi)}{\longleftrightarrow}} & S \times S & \longrightarrow & 0 \\
 & & \| {\scriptstyle 1_{\tilde{G}}} & & \downarrow {\scriptstyle \alpha_{\mathcal{A}_d}} & & \| {\scriptstyle 1_{S\times S}} & & \\
0 & \longrightarrow & \tilde{G} & \underset{\pi_1}{\overset{i_1}{\longleftrightarrow}} & \tilde{G} \oplus (S \times S) & \underset{i_2}{\overset{\pi_2}{\longleftrightarrow}} & S \times S & \longrightarrow & 0
\end{array}$$

Here, we see how the identification between $(Q \times Q)/G$ and $(S \times S) \oplus \tilde{G}$ are naturally related to the discrete connection and the discrete horizontal lift.

Recall that the discrete adjoint bundle $\tilde{G}$ is the associated bundle one obtains when $M = G$, and ρ_g acts by conjugation. Furthermore, the action of G on $Q \times Q$ is by the diagonal action, and the action of $G \times G$ on $Q \times Q$ is component-wise.

Proposition 4.13. *The map $\alpha_{\mathcal{A}_d} : (Q \times Q)/G \to (S \times S) \oplus \tilde{G}$ defined by*

$$\alpha_{\mathcal{A}_d}([q_0, q_1]_G) = (\pi q_0, \pi q_1) \oplus [q_0, \mathcal{A}_d(q_0, q_1)]_G,$$

is a well-defined bundle isomorphism. The inverse of α_{A_d} is given by

$$\alpha^{-1}_{\mathcal{A}_d}((x_0, x_1) \oplus [q, g]_G) = [(e, g) \cdot (x_0, x_1)^h_q]_G,$$

for any $q \in Q$ such that $\pi q = x_0$.

Proof. To show that $\alpha_{\mathcal{A}_d}$ is well-defined, note that for any $g \in G$, we have that

$$(\pi g q_0, \pi g q_1) = (\pi q_0, \pi q_1),$$

and also,

$$\begin{aligned}
[gq_0, \mathcal{A}_d(gq_0, gq_1)]_G &= [gq_0, g\mathcal{A}_d(q_0, q_1)g^{-1}]_G \\
&= [q_0, \mathcal{A}_d(q_0, q_1)]_G.
\end{aligned}$$

Then, we see that

$$\alpha_{\mathcal{A}_d}([gq_0, gq_1]_G) = \alpha_{\mathcal{A}_d}([q_0, q_1]_G).$$

To show that $\alpha^{-1}_{\mathcal{A}_d}$ is well-defined, note that for any $k \in G$,

$$(x_0, x_1)^h_{kq} = k \cdot (x_0, x_1)^h_q,$$

and that

$$\begin{aligned}\alpha^{-1}_{\mathcal{A}_d}((x_0,x_1)\oplus[kq,kgk^{-1}]_G) &= [(e,kgk^{-1})\cdot(x_0,x_1)^h_{kq}]_G\\ &= [(e,kgk^{-1})\cdot k\cdot(x_0,x_1)^h_q]_G\\ &= [(ek,kgk^{-1}k)\cdot(x_0,x_1)^h_q]_G\\ &= [(ke,kg)\cdot(x_0,x_1)^h_q]_G\\ &= [k\cdot(e,g)\cdot(x_0,x_1)^h_q]_G\\ &= [(e,g)\cdot(x_0,x_1)^h_q]_G\\ &= \alpha^{-1}_{\mathcal{A}_d}((x_0,x_1)\oplus[q,g]_G).\end{aligned}$$ □

Example 4.2. *It is illustrative to consider the notion of a discrete connection, and the isomorphism, in the degenerate case when $Q = G$, which is the context of discrete Euler–Poincaré reduction. Here, the isomorphism is between $(G\times G)/G$ and $\tilde{G}$, and the connection $\mathcal{A}_d : G\times G\to G$ is given by*

$$\mathcal{A}_d(g_0,g_1) = g_1\cdot g_0^{-1}.$$

Then, we have that

$$\begin{aligned}\alpha_{\mathcal{A}_d}([g_0,g_1]_G) &= (\pi g_0,\pi g_1)\oplus[g_0,\mathcal{A}_d(q_0,q_1)]_G\\ &= (e,e)\oplus[g_0,g_1g_0^{-1}]_G\,.\end{aligned}$$

Taking the inverse, we have,

$$\begin{aligned}\alpha^{-1}_{\mathcal{A}_d}([g_0,g_1g_0^{-1}]_G) &= [(e,g_1g_0^{-1}]\cdot(e,e)^h_{g_0}]_G\\ &= [(e,g_1g_0^{-1}]\cdot(g_0,g_0)]_G\\ &= [eg_0,g_1g_0^{-1}g_0]_G\\ &= [g_0,g_1]_G\,,\end{aligned}$$

as expected.

4.9. **Discrete Horizontal and Vertical Subspaces Revisited.** Having now fully introduces all the equivalent representations of a discrete connection, we can revisit the notion of discrete horizontal and vertical subspaces in light of the new structures we have introduced.

Consider the following split exact sequence,

$$0\longrightarrow A_1 \underset{k}{\overset{f}{\underset{\longleftarrow}{\longrightarrow}}} B \underset{h}{\overset{g}{\underset{\longleftarrow}{\longrightarrow}}} A_2\longrightarrow 0\ .$$

We can decompose any element in B into a A_1 and A_2 term by considering the following isomorphism,

$$B\cong f\circ k(B)\oplus h\circ g(B).$$

Similarly, in the discrete Atiyah sequence, we can decompose an element of $(Q\times Q)/G$ into a horizontal and vertical piece by performing the analogous construction on the split exact sequence

$$0\longrightarrow \tilde{G} \underset{(\pi_1,\mathcal{A}_d)}{\overset{(q,gq)}{\underset{\longleftarrow}{\longrightarrow}}} (Q\times Q)/G \underset{(\cdot,\cdot)^h}{\overset{(\pi,\pi)}{\underset{\longleftarrow}{\longrightarrow}}} S\times S\longrightarrow 0\ .$$

This allows us to define horizontal and vertical spaces associated with the pair groupoid $Q\times Q$, in terms of all the structures we have introduced.

Definition 4.8. *The **horizontal space** is given by*

$$\mathrm{Hor}_q = \{(q,q')\in Q\times Q \mid \mathcal{A}_d(q,q') = e\}$$

$$= \{(\pi q, x_1)_q^h \in Q \times Q \mid x_1 \in S\}.$$

This is the discrete analogue of the statement $\mathrm{Hor}_q = \{v_q \in TQ \mid \mathcal{A}(v_q) = 0\} = \{(v_{\pi q})_q^h \in TQ \mid v_{\pi q} \in TS\}$.

Definition 4.9. *The* ***vertical space*** *is given by*

$$\begin{aligned}\mathrm{Ver}_q &= \{(q, q') \in Q \times Q \mid (\pi, \pi)(q, q') = e_{S\times S}\} \\ &= \{i_q(g) \mid g \in G\}.\end{aligned}$$

This is the discrete analogue of the statement $\mathrm{Ver}_q = \{v_q \in TQ \mid \pi_*(v_q) = 0\} = \{\xi_Q \mid \xi \in \mathfrak{g}\}$.

In particular, we can decompose an element of $Q \times Q$ into a horizontal and vertical component.

Definition 4.10. *The* ***horizontal component*** *of* $(q_0, q_1) \in Q \times Q$ *is given by*

$$\mathrm{hor}(q_0, q_1) = ((\cdot, \cdot)^h \circ (\pi, \pi))(q_0, q_1) = (\pi q_0, \pi q_1)_{q_0}^h.$$

Definition 4.11. *The* ***vertical component*** *of* $(q_0, q_1) \in Q \times Q$ *is given by*

$$\mathrm{ver}(q_0, q_1) = (i \circ (\pi_1, \mathcal{A}_d))(q_0, q_1) = (q_0, \mathcal{A}_d(q_0, q_1)q_0) = i_{q_0}(\mathcal{A}_d(q_0, q_1)).$$

Lemma 4.14. *The horizontal component can be expressed as*

$$\mathrm{hor}(q_0, q_1) = i_{q_0}((\mathcal{A}_d(q_0, q_1))^{-1}) \cdot (q_0, q_1).$$

Proof.

$$\begin{aligned}i_{q_0}(\mathcal{A}_d(q_0, q_1)^{-1}) \cdot (q_0, q_1) &= (q_0, (\mathcal{A}_d(q_0, q_1))^{-1} q_0) \cdot (q_0, q_1) \\ &= (e, (\mathcal{A}_d(q_0, q_1))^{-1})(q_0, q_1) \\ &= (q_0, (\mathcal{A}_d(q_0, q_1))^{-1} q_1).\end{aligned}$$

Clearly, $(\pi, \pi)(q_0, (\mathcal{A}_d(q_0, q_1))^{-1} q_1) = (\pi, \pi)(q_0, q_1)$. Furthermore,

$$\mathcal{A}_d(q_0, (\mathcal{A}_d(q_0, q_1))^{-1} q_1) = (\mathcal{A}_d(q_0, q_1))^{-1} \mathcal{A}_d(q_0, q_1) = e.$$

Therefore, by definition, $(q_0, (\mathcal{A}_d(q_0, q_1))^{-1} q_1) = \mathrm{hor}(q_0, q_1)$. □

Lemma 4.15. *The horizontal and vertical operators satisfy the following identity,*

$$\mathrm{ver}(q_0, q_1) \cdot \mathrm{hor}(q_0, q_1) = (q_0, q_1).$$

Proof.

$$\begin{aligned}\mathrm{ver}(q_0, q_1) \cdot \mathrm{hor}(q_0, q_1) &= i_{q_0}(\mathcal{A}_d(q_0, q_1)) \cdot (i_{q_0}((\mathcal{A}_d(q_0, q_1))^{-1}) \cdot (q_0, q_1)) \\ &= (e, \mathcal{A}_d(q_0, q_1))(e, (\mathcal{A}_d(q_0, q_1))^{-1})(q_0, q_1) \\ &= (e, \mathcal{A}_d(q_0, q_1))(q_0, (\mathcal{A}_d(q_0, q_1))^{-1} q_1) \\ &= (q_0, \mathcal{A}_d(q_0, q_1)(\mathcal{A}_d(q_0, q_1))^{-1} q_1) \\ &= (q_0, q_1),\end{aligned}$$

as desired. □

5. Geometric Structures Derived from the Discrete Connection

In this section, we will introduce some of the additional geometric structures that can be derived from a choice of discrete connection. These structures include an extension of the pair groupoid composition to take into account the principal bundle structure, continuous connections that are a limit of a discrete connection, and higher-order connection-like structures.

5.1. **Extending the Pair Groupoid Composition.** Recall that the composition of a vertical element (q_0, gq_0) with an element (q_0, q_1) is given by

$$(q_0, gq_0) \cdot (q_0, q_1) = (q_0, gq_1).$$

The choice of a discrete connection allows us to further extend the composition, in a manner that is relevant in describing the curvature of a discrete connection. The decomposition of an element of $Q \times Q$ into a horizontal and vertical piece naturally suggests a generalization of the composition operation on $Q \times Q$ (viewed as a pair groupoid), by using the discrete connection, and the principal bundle structure of Q.

We wish to define a composition on $Q \times Q$ such that the composition of $(q_0, q_1) \cdot (\tilde{q}_0, \tilde{q}_1)$ is defined whenever $\pi q_1 = \pi \tilde{q}_0$. Furthermore, we require that the extended composition be consistent with the vertical composition we introduced previously, as well as the pair groupoid composition, whenever their domains of definition coincide.

The extended composition is obtained by left translating $(\tilde{q}_0, \tilde{q}_1)$ by a group element h, such that $q_1 = h\tilde{q}_0$, and then using the pair groupoid composition on (q_0, q_1) and the left translated term $h(\tilde{q}_0, \tilde{g}_1)$. This yields the following intrinsic definition of the extended composition.

Definition 5.1. *The **extended pair groupoid composition** of $(q_0, q_1), (\tilde{q}_0, \tilde{q}_1) \in Q \times Q$ is defined whenever $\pi q_1 = \pi \tilde{q}_0$, and it is given by*

$$(q_0, q_1) \cdot (\tilde{q}_0, \tilde{q}_1) = (q_0, \mathcal{A}_d(\tilde{q}_0, q_1)\tilde{q}_1).$$

As the following lemmas show, this extended composition is consistent with the vertical composition and the pair groupoid composition.

Lemma 5.1. *The extended pair groupoid composition is consistent with the composition of a vertical element with an arbitrary element.*

Proof. Consider the composition of a vertical element with an arbitrary element,

$$(q_0, gq_0) \cdot (q_0, q_1) = (q_0, gq_1).$$

This is consistent with the result using the extended composition,

$$\begin{aligned}(q_0, gq_0) \cdot (q_0, q_1) &= (q_0, \mathcal{A}_d(q_0, gq_0)q_1) \\ &= (q_0, gq_1)\,,\end{aligned}$$

where we used that the discrete connection yields a splitting of the Atiyah sequence. □

Lemma 5.2. *The extended pair groupoid composition is consistent with the pair groupoid composition.*

Proof. The pair groupoid composition is given by

$$(q_0, q_1) \cdot (q_1, q_2) = (q_0, q_2)\,.$$

This is consistent with the extended composition,

$$\begin{aligned}(q_0, q_1) \cdot (q_1, q_2) &= (q_0, \mathcal{A}_d(q_1, q_1)q_2) \\ &= (q_0, eq_2) \\ &= (q_0, q_2)\,.\end{aligned}$$

□

The extended composition is G-equivariant, and is well-defined on the quotient space, as the following lemma shows.

Lemma 5.3. *The composition $\cdot : (Q \times Q) \times (Q \times Q) \to (Q \times Q)$ is G-equivariant, that is,*

$$(gq_0, gq_1) \cdot (g\tilde{q}_0, g\tilde{q}_1) = g \cdot ((q_0, q_1) \cdot (\tilde{q}_0, \tilde{q}_1)).$$

Furthermore, the composition induces a well-defined quotient composition $\cdot : ((Q \times Q) \times (Q \times Q))/G \to (Q \times Q)/G$.

Proof. Given $g \in G$, we consider,

$$\begin{aligned}(gq_0, gq_1) \cdot (g\tilde{q}_0, g\tilde{q}_1) &= (gq_0, \mathcal{A}_d(g\tilde{q}_0, gq_1)g\tilde{q}_1)\\ &= (gq_0, g\mathcal{A}_d(\tilde{q}_0, q_1)g^{-1}g\tilde{q}_1)\\ &= (gq_0, g\mathcal{A}_d(\tilde{q}_0, q_1)\tilde{q}_1)\\ &= g \cdot (q_0, \mathcal{A}_d(\tilde{q}_0, q_1)\tilde{q}_1)\\ &= g \cdot ((q_0, q_1) \cdot (\tilde{q}_0, \tilde{q}_1))\,,\end{aligned}$$

where we used the equivariance of the discrete connection. It follows that the composition is equivariant. Furthermore,

$$[(g\tilde{q}_0, g\tilde{q}_1) \cdot (gq_0, gq_1)]_G = [(\tilde{q}_0, \tilde{q}_1) \cdot (q_0, q_1)]_G,$$

which means that $\cdot : ((Q \times Q) \times (Q \times Q))/G \to (Q \times Q)/G$ is well-defined. □

Corollary 5.4. *The composition of n-terms is G-equivariant. That is to say,*

$$\begin{aligned}(gq_0^1, gq_1^1) \cdot (gq_0^2, gq_1^2) \cdot \ldots \cdot (gq_0^{n-1}, gq_1^{n-1}) \cdot (gq_0^n, gq_1^n)&\\ = g \cdot ((q_0^1, q_1^1) \cdot (q_0^2, q_1^2) \cdot \ldots \cdot (q_0^{n-1}, q_1^{n-1}) \cdot (q_0^n, q_1^n)).&\end{aligned}$$

Proof. The result follows by induction on the previous lemma. □

We find that the extended composition we have constructed on the pair groupoid is associative. However, as we shall see in §7.3, composing pair groupoid elements about a loop in the shape space will not in general yield the identity element $e_{Q\times Q}$, and the defect represents the holonomy about the loop, which is related to curvature. This may yield the discrete analogue of the expression giving the geometric phase in terms of a loop integral (in shape space) of the curvature of the connection.

Lemma 5.5. *The composition $\cdot : (Q \times Q) \times (Q \times Q) \to (Q \times Q)$ is associative. That is,*

$$((q_0^0, q_1^0) \cdot (q_0^1, q_1^1)) \cdot (q_0^2, q_1^2) = (q_0^0, q_1^0) \cdot ((q_0^1, q_1^1) \cdot (q_0^2, q_1^2))\,.$$

Proof. Evaluating the left-hand side, we obtain

$$\begin{aligned}((q_0^0, q_1^0) \cdot (q_0^1, q_1^1)) \cdot (q_0^2, q_1^2) &= (q_0^0, \mathcal{A}_d(q_0^1, q_1^0)q_1^1) \cdot (q_0^2, q_1^2)\\ &= (q_0^0, \mathcal{A}_d(q_0^2, \mathcal{A}_d(q_0^1, q_1^0)q_1^1)q_1^2)\\ &= (q_0^0, \mathcal{A}_d(q_0^1, q_1^0)\mathcal{A}_d(q_0^2, q_1^1)q_1^2)\,,\end{aligned}$$

and the right-hand side is given by

$$\begin{aligned}(q_0^0, q_1^0) \cdot ((q_0^1, q_1^1) \cdot (q_0^2, q_1^2)) &= (q_0^0, q_1^0) \cdot (q_0^1, \mathcal{A}_d(q_0^2, q_1^1)q_1^2)\\ &= (q_0^0, \mathcal{A}_d(q_0^1, q_1^0)\mathcal{A}_d(q_0^2, q_1^1)q_1^2)\,.\end{aligned}$$

Therefore,

$$((q_0^0, q_1^0) \cdot (q_0^1, q_1^1)) \cdot (q_0^2, q_1^2) = (q_0^0, q_1^0) \cdot ((q_0^1, q_1^1) \cdot (q_0^2, q_1^2))\,,$$

and the extended groupoid composition is associative. □

5.2. **Continuous Connections from Discrete Connections.** Given a discrete G-valued connection 1-form, $\mathcal{A}_d : Q \times Q \to G$, we associate to it a continuous $\mathfrak{g}$-valued connection 1-form, $\mathcal{A} : TQ \to \mathfrak{g}$, by the following construction,

$$\mathcal{A}([q(\cdot)]) = [\mathcal{A}_d(q(0), q(\cdot))],$$

where $[\cdot]$ denotes the equivalence class of curves associated with a tangent vector.

This uses the intrinsic representation of the tangent bundle, which is obtained by identifying a tangent vector at a point on the manifold with the equivalence class of curves on the manifold going through the

point, such that the tangent to the curve at the point is given by the tangent vector, which was illustrated earlier in Figure 9 on page 10.

More explicitly, given $v_q \in TQ$, we consider an associated curve $q : [0,1] \to Q$, and construct the curve $g : [0,1] \to G$, given by

$$g(t) = \mathcal{A}_d(q(0), q(t)).$$

Then,

$$\mathcal{A}(v_q) = \left.\frac{d}{dt}\right|_{t=0} g(t).$$

When computing the equations in discrete reduction theory, it is often necessary to consider horizontal and vertical variations, which we introduce below.

Definition 5.2. *We introduce the* ***vertical variation*** *of a point $(q_0, q_1) \in Q \times Q$. Given a curve $q_1^\epsilon : [0,1] \to Q$, such that $q_1^\epsilon(0) = q_1$, the vertical variation is given by*

$$\operatorname{ver} \delta q = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \operatorname{ver}(q_0, q_1^\epsilon) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} i_{q_0}(\mathcal{A}_d(q_0, q_1^\epsilon)) \,.$$

Definition 5.3. *We introduce the* ***horizontal variation*** *of a point $(q_0, q_1) \in Q \times Q$. Given a curve $q_1^\epsilon : [0,1] \to Q$, such that $q_1^\epsilon(0) = q_1$, the horizontal variation is given by*

$$\operatorname{hor} \delta q = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \operatorname{hor}(q_0, q_1^\epsilon) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (\pi(q_0, q_1^\epsilon))^h_{q_0} \,.$$

5.3. **Connection-Like Structures on Higher-Order Tangent Bundles.** Given a continuous connection, we can construct connection-like structures on higher-order tangent bundles. This construction is described in detail in Lemma 3.2.1 of Cendra et al. [2001]. In particular, given a connection 1-form, $\mathcal{A} : TQ \to \mathfrak{g}$, we obtain a well-defined map, $\mathcal{A}^k : T^{(k)}Q \to k\mathfrak{g}$.

As we will see later, these connection-like structures on higher-order tangent bundles will provide an intrinsic method of characterizing the order of approximation of a continuous connection by a discrete connection.

We will describe the discrete analogue of this construction. To begin, the discrete analogue of the k-th order tangent bundle, $T^{(k)}Q$, is $k+1$ copies of Q, namely Q^{k+1}. Intermediate spaces between $T^{(k)}Q$ and Q^{k+1} arise in the general theory of multi-spaces, which is introduced in Olver [2001].

The discrete analogue of tangent lifts, and their higher-order analogues, are obtained by componentwise application of the map, since a tangent lift of a map is computed by applying the map to a representative curve, and taking its equivalence class. Therefore, given a map $f : M \to N$, we have the naturally induced map,

$$T^{(k)}f : M^{k+1} \to N^{k+1} \quad \text{given by} \quad T^{(k)}f(m_0, \ldots, m_k) = (f(m_0), \ldots, f(m_k)).$$

And in particular, the group action is lifted to the diagonal group action on the product space.

The discrete connection can be extended to Q^{k+1} in the natural way, $\mathcal{A}_d^k : Q^{k+1} \to \oplus_{l=0}^{k-1} G \equiv kG$,

$$\mathcal{A}_d^k(q_0, \ldots, q_k) = \oplus_{l=0}^{k-1} \mathcal{A}_d(q_l, q_{l+1}).$$

Similarly, we can define the map from Q^{k+1} to the Whitney sum of k copies of the conjugate bundle $\tilde{G}$ by

$$Q^{k+1} \to k\tilde{G} \quad \text{by} \quad (q_0, \ldots, q_k) \mapsto \oplus_{l=0}^{k-1} [q_0, \mathcal{A}_d(q_l, q_{l+1})]_G.$$

In a natural way, we have the following proposition.

Proposition 5.6. *The map*

$$\alpha_{\mathcal{A}_d^k} : Q^{k+1} \to (Q/G)^{k+1} \times_{Q/G} k\tilde{G}$$

defined by

$$\alpha_{\mathcal{A}_d^k}(q_0, \ldots, q_k) = (\pi q_0, \ldots, \pi q_k) \times_{Q/G} \oplus_{l=0}^{k-1} [q_0, \mathcal{A}_d(q_l, q_{l+1})]_G,$$

is a well-defined bundle isomorphism. The inverse of $\alpha_{\mathcal{A}_d^k}$ *is given by*

$$\begin{aligned}\alpha^{-1}_{\mathcal{A}_d^k}\left((x_0,\ldots,x_k)\times_{Q/G}\oplus_{l=0}^{k-1}[q_l,g_l]_G\right)\\ =[(e,g_0,g_1g_0,\ldots,g_{k-1}\cdots g_0))\cdot(x_0,\ldots,x_k)^h_{q_0}]_G,\end{aligned}$$

where $(x_0,\ldots,x_k)^h_{q_0}=(\bar{q}_0,\ldots,\bar{q}_k)$ *is defined by the conditions:*

$$\begin{aligned}\bar{q}_0&=q_0,\\ \pi\bar{q}_l&=x_l,\\ \mathcal{A}_d(\bar{q}_l,\bar{q}_{l+1})&=e.\end{aligned}$$

Remark 5.1. *In a local trivialization, where* $q_0=(h_0,x_0)$*, we have,*

$$\bar{q}_{l+1}=\left((A(x_l,x_{l+1})\cdots A(x_0,x_1))^{-1}h_0,x_{l+1}\right).$$

6. Computational Aspects

While we saw in the previous section that a discrete connection induces a continuous connection in the limit, we are often concerned with constructing discrete connections that approximate a continuous connection to a given order of approximation. This section will address the question of what it means for a discrete connection to approximate a continuous connection to a given order, as well as introduce methods for constructing such discrete connections.

6.1. Exact Discrete Connection.

It is interesting from the point of view of computation to construct an ***exact discrete connection*** associated with a prescribed continuous connection, so that we can make sense of the statement that a given discrete connection is a k-th order approximation of a continuous connection.

Additional Structure. To construct the exact discrete connection, we require that the configuration manifold Q be endowed with a bi-invariant Riemannian metric, with the property that the associated exponential map,

$$\exp:TQ\to Q,$$

is consistent with the group action, in the sense that

$$\exp(\xi_Q(q))=\exp(\xi)\cdot q.$$

We extend the exponential to $Q\times Q$ as follows,

$$\begin{aligned}\overline{\exp}:TQ&\to Q\times Q,\\ v_q&\mapsto(q,\exp(v_q)),\end{aligned}$$

and denote the inverse by $\overline{\log}:Q\times Q\to TQ$, which is defined in a neighborhood of the diagonal of $Q\times Q$.

Construction. Having introduced the appropriate structure on the configuration manifold, we define the exact discrete connection as follows.

Definition 6.1. *The* ***Exact Discrete Connection*** $\mathcal{A}_d^E$ *associated with a prescribed continuous connection* $\mathcal{A}:TQ\to\mathfrak{g}$ *and a Riemannian metric is given by*

$$\mathcal{A}_d^E(q_0,q_1)=\exp(\mathcal{A}(\overline{\log}(q_0,q_1))).$$

This construction is more clearly illustrated in the following diagram,

$$Q\times Q\xrightarrow[\overline{\log}]{}TQ\xrightarrow[\mathcal{A}]{}\mathfrak{g}\xrightarrow[\exp]{}G,\qquad Q\times Q\xrightarrow{\mathcal{A}_d^E}G$$

Properties. The exact discrete connection satisfies the properties of the discrete connection 1-form, in that it is G-equivariant, and it induces a splitting of the discrete Atiyah sequence. The equivariance of the exact discrete connection arises from the fact that each of the composed maps is equivariant, and the splitting condition,

$$\mathcal{A}_d^E(i_q(g)) = g,$$

is a consequence of the compatibility condition,

$$\exp(\xi_Q(q)) = \exp(\xi) \cdot q.$$

Since the logarithm map is only defined on a neighborhood of the diagonal of $Q \times Q$, it follows that the exact discrete connection will have the same restriction on its domain of definition.

Example 6.1 (Discrete Mechanical Connection)**.** *The continuous mechanical connection is defined by the following diagram,*

$$\begin{array}{ccc} T^*Q & \xrightarrow{J} & \mathfrak{g}^* \\ {\scriptstyle \mathbb{F}L}\uparrow & & \uparrow{\scriptstyle \mathbb{I}} \\ TQ & \xrightarrow[\mathcal{A}]{} & \mathfrak{g} \end{array}$$

Correspondingly, the discrete mechanical connection is defined by the following diagram,

$$\begin{array}{ccccccc} & & & \overset{J_d}{\frown} & & & \\ Q \times Q & \xrightarrow[\mathbb{F}L_d]{} & T^*Q & \xrightarrow[J]{} & \mathfrak{g}^* & & \\ \| & & & & \uparrow{\scriptstyle \mathbb{I}} & & \\ Q \times Q & & \longrightarrow & & \mathfrak{g} & \xrightarrow{\exp} & G \\ & & & \underset{\mathcal{A}_d}{\smile} & & & \end{array}$$

This is consistent with our notion of an exact discrete mechanical connection as the following diagram illustrates,

$$\begin{array}{ccccccc} & & & \overset{J_d}{\frown} & & & \\ Q \times Q & \xrightarrow[\mathbb{F}L_d]{} & T^*Q & \xrightarrow[J]{} & \mathfrak{g}^* & & \\ \| & & {\scriptstyle \mathbb{F}L}\uparrow & & \uparrow{\scriptstyle \mathbb{I}} & & \\ Q \times Q & \xrightarrow{\overline{\log}} & TQ & \xrightarrow{\mathcal{A}} & \mathfrak{g} & \xrightarrow{\exp} & G \\ & & & \underset{\mathcal{A}_d}{\smile} & & & \end{array}$$

where the portion in the dotted box recovers the continuous mechanical connection. In checking G-equivariance, we use the equivariance of $\exp : \mathfrak{g} \to G$, $J_d : Q \times Q \to \mathfrak{g}^*$, *and the equivariance of* $\mathbb{I}$ *in the sense of a map* $\mathbb{I} : Q \to L(\mathfrak{g}, \mathfrak{g}^*)$, *namely,* $\mathbb{I}(gq) \cdot \mathrm{Ad}_g \, \xi = \mathrm{Ad}^*_{g^{-1}} \, \mathbb{I}(q) \cdot \xi$.

6.2. **Order of Approximation of a Connection.** We have the necessary constructions to consider the order to which a discrete connection approximates a continuous connection. There are two equivalent ways of defining the order of approximation of a continuous connection by a discrete connection, the first is more analytical, and is given by the order of convergence in an appropriate norm on the group.

Definition 6.2 (Order of Connection, Analytic)**.** *A discrete connection $\mathcal{A}_d$ is a k-th order discrete connection if, k is the maximum integer for which*

$$\exists 0 < c < \infty,$$
$$\exists h_0 > 0,$$

such that

$$\sup_{\substack{v_q \in TQ,\\ |v_q|=1}} \|\mathcal{A}_d^E(q, \exp(hv_q))(\mathcal{A}_d(q, \exp(hv_q)))^{-1}\| \leq ch^{k+1}, \quad \forall h < h_0.$$

The second definition is more intrinsic, and is related to considering the infinitesimal limit of a discrete connection to connection-like structures on higher-order tangent bundles, without the need for the introduction of the exact discrete connection.

Recall from §5.2 that we can construct a continuous connection from a discrete connection by the following construction,

$$\mathcal{A}([q(\cdot)]) = [\mathcal{A}_d(q(0), q(\cdot))].$$

Given $\mathcal{A}_d^k : Q^{k+1} \to kG$, we can obtain the continuous limit $\mathcal{A}^k : T^{(k)}Q \to k\mathfrak{g}$ in a similar fashion.

Definition 6.3 (Order of Connection, Intrinsic)**.** *A discrete connection $\mathcal{A}_d$ is a k-th order approximation to $\mathcal{A}$ if, k is the maximum integer for which the diagram holds,*

$$\begin{array}{ccc} \mathcal{A}_d : Q \times Q \to G & & \mathcal{A} : TQ \to \mathfrak{g} \\ \Downarrow & & \Downarrow \\ \mathcal{A}_d^k : Q^{k+1} \to kG & \dashrightarrow & \mathcal{A}^k : T^{(k)}Q \to k\mathfrak{g} \end{array}$$

Here, the double arrows represent the higher-order structures induced by the connections, and the dotted arrow represents convergence in the limit.

6.3. **Discrete Connections from Exponentiated Continuous Connections.** To apply the exponential and logarithm approach to construct a discrete connection from a prescribed connection, in the sense of the diagram,

$$Q \times Q \xrightarrow[\overline{\log}]{} TQ \xrightarrow[\mathcal{A}]{} \mathfrak{g} \xrightarrow[\exp]{} G \ , \qquad Q \times Q \xrightarrow{\mathcal{A}_d} G$$

we can rely on explicit expressions for the exponential and logarithm, or we can rely on approximations to the exponential and logarithm.

The explicit formulas for the special Euclidean group are particularly useful for applying the theory of discrete connections to the geometric control of problems such as robotic manipulators, and clusters of satellites. In dealing with other configuration manifolds, approximants to the exponential and logarithm may be required due to the absence of explicit formulas.

Even when explicit formulas are available, it may be desirable to rely on a more computationally efficient approximation, such as the Cayley transformation, methods based on Padé approximants (see, for example, Cardoso and Silva Leite [2001], Higham [2001]), or Lie group techniques (see, for example, Celledoni and Iserles [2000, 2001], Zanna and Munthe-Kaas [2001/02]). Clearly, these will yield a discrete connection that has an order of approximation equal to the lower of the two orders of approximation of the numerical schemes used for the exponential and the logarithm.

When used in the context of geometric control, high-order approximations to the continuous connection may not be necessary, since the optimal trajectory is often recomputed at each step, and in such instances, a low-order approximation suffices.

6.4. **Discrete Mechanical Connections and Discrete Lagrangians.** We will introduce a discrete mechanical connection that is consistent with the structure of discrete variational mechanics, and will yield a discrete connection that has an order of approximation that is equation to that obtained from the discrete mechanics.

Consider a G-invariant k-th order discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$, which is to say that

$$L_d(gq_0, gq_1) = L_d(q_0, q_1),$$

and

$$L_d = L_d^{\text{exact}} + \mathcal{O}(h^{k+1}),$$

where the exact discrete Lagrangian, $L_d^{\text{exact}} : Q \times Q \to \mathbb{R}$, is given by

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

Here, $q_{01} : [0, h] \to Q$ is the solution of the Euler–Lagrange equations with $q_{01}(0) = q_0$, and $q_{01}(h) = q_1$. The exact discrete Lagrangian is a generator of the symplectic flow, coming from the Jacobi solution of the Hamilton–Jacobi equation.

This k-th order discrete Lagrangian yields a k-th order accurate numerical update scheme, through the discrete Euler–Lagrange equations,

$$D_2L_d(q_0, q_1) + D_1L_d(q_1, q_2) = 0,$$

which implicitly defines a discrete flow $\Phi : (q_0, q_1) \mapsto (q_1, q_2)$. By pushing this numerical scheme forward to T^*Q using the discrete fiber derivative $\mathbb{F}L_d : Q \times Q \to T^*Q$, which maps $(q_0, q_1) \mapsto (q_0, -D_1L_d(q_0, q_1))$, we can obtain a Symplectic Partitioned Runge–Kutta scheme of the same order.

We also introduce the discrete momentum map, $J_d : Q \times Q \to \mathfrak{g}^*$, given by

$$\langle J_d(q_0, q_1), \xi \rangle = -D_1L_d(q_0, q_1) \cdot \xi_Q(q_0).$$

The discrete Lagrangian is G-invariant, which implies that for any $\xi \in \mathfrak{g}$, we have,

$$\begin{aligned} L_d(q_0, q_1) &= L_d(\exp(\xi t) \cdot q_0, \exp(\xi t) \cdot q_1), \\ 0 &= \left.\frac{d}{dt}\right|_{t=0} L_d(\exp(\xi t) \cdot q_0, \exp(\xi t) \cdot q_1) \\ &= D_1L_d(q_0, q_1) \cdot \xi_Q(q_0) + D_2L_d(q_0, q_1) \cdot \xi_Q(q_1). \end{aligned}$$

If we restrict to the flow of the discrete Euler–Lagrange equations, we have that

$$(D_1L_d(q_0, q_1) + D_2L_dq(q_1, q_2)) \cdot \xi_Q(q_1) = 0,$$

which upon substitution into the previous equation, yields

$$\begin{aligned} D_1L_d(q_0, q_1) \cdot \xi_Q(q_0) - D_1L_d(q_1, q_2) \cdot \xi_Q(q_1) &= 0, \\ -D_1L_d(q_1, q_2) \cdot \xi_Q(q_1) &= -D_1L_d(q_0, q_1) \cdot \xi_Q(q_0), \\ \langle J_d(q_1, q_2), \xi_Q(q_1) \rangle &= \langle J_d(q_0, q_1), \xi_Q(q_0) \rangle, \\ \Phi^* J_d &= J_d. \end{aligned}$$

which is the statement of the discrete Noether theorem, that the discrete momentum map is preserved by the discrete Euler–Lagrange flow.

We note that the mechanical connection corresponds to a choice of horizontal space corresponding to the zero momentum surface. That is to say that the horizontal distribution corresponding to the continuous mechanical connection is

$$\text{Hor}_q = \{v_q \in TQ \mid J(v_q) = 0\},$$

and the discrete horizontal subspace corresponding to the discrete mechanical connection is

$$\mathrm{Hor}_q^d = \{(q, q') \in Q \times Q \mid J_d(q, q') = 0\}$$

Remark 6.1. *For the discrete horizontal subspace we defined above to have the correct dimensionality, the discrete Lagrangian needs to satisfy certain non-degeneracy conditions, which dictates the size of the neighborhood of the diagonal that the discrete connection is defined on.*

Since the continuous momentum map is preserved by the continuous Euler–Lagrange flow, and the discrete momentum map is preserved by the discrete Euler–Lagrange flow, it follows that the order of approximation of the continuous mechanical connection by the discrete mechanical connection is equal to the order of approximation of the continuous Euler–Lagrange flow by the discrete Euler–Lagrange flow. To construct a discrete mechanical connection of a prescribed order, we use the following procedure.

(1) Consider a k-th order G-invariant discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$,
$$L_d = L_d^{\text{exact}} + \mathcal{O}(h^{k+1}).$$
(2) Construct the corresponding discrete momentum map, $J_d(q_0, q_1) \to \mathfrak{g}^*$,
$$\langle J_d(q_0, q_1), \xi \rangle = -D_1 L_d(q_0, q_1) \cdot \xi_Q(q_0).$$
(3) Then, the k-th order discrete mechanical connection is given implicitly by considering the condition for the discrete horizontal space,
$$\mathcal{A}_d(q_0, q_1) = e \qquad \text{iff} \qquad J_d(q_0, q_1) = 0,$$
and then extending the construction to the domain of definition by G-equivariance.
(4) More explicitly, given $(q_0, q_1) \in Q \times Q$, we consider a local trivialization, in which $(q_0, q_1) = (x_0, g_0, x_1, g_1)$, and we find the unique $g \in G$ such that
$$J_d(x_0, g_0, x_1, g) = 0.$$
Then, we have that
$$\mathcal{A}_d(x_0, g_0, x_1, g) = e,$$
from which we conclude that
$$\begin{aligned} \mathcal{A}_d(x_0, g_0, x_1, g_1) &= \mathcal{A}_d(x_0, g_0, x_1, g_1 g^{-1} g) \\ &= g_1 g^{-1} \cdot \mathcal{A}_d(x_0, g_0, x_1, g) \\ &= g_1 g^{-1}. \end{aligned}$$

7. Applications

This section will sketch some of the applications of the mathematical machinery of discrete connections and discrete exterior calculus to problems in computational geometric mechanics, geometric control theory, and discrete Riemannian geometry.

7.1. **Discrete Lagrangian Reduction.** Lagrangian reduction, which is the Lagrangian analogue of Poisson reduction on the Hamiltonian side, is associated with the reduction of Hamilton's variational principle for systems with symmetry.

The variation of the action integral associated with a variation in the curve can be expressed in terms of the Euler–Lagrange operator, $\mathcal{EL} : T^{(2)}Q \to T^*Q$. When the Lagrangian is G-invariant, the associated Euler–Lagrange operator is G-equivariant, and this induces a reduced Euler–Lagrange operator, $[\mathcal{EL}]_G : T^{(2)}Q/G \to T^*Q/G$. The choice of a connection allows us to construct intrinsic coordinates on $T^{(2)}/G$ and T^*Q/G, and the representation of the reduced Euler–Lagrange operator in these coordinates correspond to the Lagrange–Poincaré operator, $\mathcal{LP} : T^{(2)}(Q/G) \times_{Q/G} 2\tilde{\mathfrak{g}} \to T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$.

The reduced equations obtained by reduction tend to have non-canonical symplectic structures. As such, naïvely applying standard symplectic algorithms to reduced equations can have undesirable consequences for the longtime behavior of the simulation, since it preserves the canonical symplectic form on the reduced space, as opposed to the reduced (non-canonical) symplectic form that is invariant under the reduced dynamics.

This sends a cautionary message, that it is important to understand the reduction of discrete variational mechanics, since applying standard numerical algorithms to the reduced equations obtained from continuous reduction theory may yield undesirable results, inasmuch as long-term stability is concerned.

Discrete connections on principal bundles provide the appropriate geometric structure to construct a discrete analogue of Lagrangian reduction. We first introduce the discrete Euler–Lagrange operator, which is constructed as follows.

Discrete Euler–Lagrange Operator. The discrete Euler–Lagrange operator, $\mathcal{EL}_d : Q^3 \to T^*Q$ satisfies the following property,

$$\mathbf{d}\mathfrak{S}_d(L_d) \cdot \delta\mathbf{q} = \sum \mathcal{EL}_d(L_d)(q_{k-1}, q_k, q_{k+1}) \cdot \delta q_k.$$

In coordinates, the discrete Euler–Lagrange operator has the form

$$[D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1})]\, dq_k.$$

Discrete Lagrange–Poincaré Operator. The map $\mathcal{EL}_d(L_d) : Q^3 \to T^*Q$, being G-equivariant, induces a quotient map

$$[\mathcal{EL}_d(L_d)]_G : Q^3/G \to T^*Q/G,$$

which depends only on the reduced discrete Lagrangian $l_d : (Q \times Q)/G \to \mathbb{R}$. We can therefore identify $[\mathcal{EL}_d(L_d)]_G$ with an operator $\mathcal{EL}_d(l_d)$ which we call the ***reduced discrete Euler–Lagrange operator***.

If in addition to the principal bundle structure, we have a discrete principal connection as described in the previous section, we can identify

$$Q^3/G \quad \text{with} \quad (Q/G)^3 \times_{Q/G} (\tilde{G} \oplus \tilde{G}).$$

The isomorphism between these two spaces is a consequence of Proposition 5.6, which is higher-order generalization of Proposition 4.13. The discrete mechanical connection which was introduced in §6.4 is a particularly natural choice of discrete connection, since it does not require any *ad hoc* choices, as it is constructed directly from the discrete Lagrangian.

Furthermore, each discrete G-valued connection 1-form, $\mathcal{A}_d : Q \times Q \to G$, induces in the infinitesimal limit a continuous $\mathfrak{g}$-valued connection 1-form, $\mathcal{A} : TQ \to \mathfrak{g}$, as shown in §5.2. This continuous principal connection allows us to identify

$$T^*Q/G \quad \text{with} \quad T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*.$$

The discrete Lagrange–Poincaré operator, $\mathcal{LP}_d(l_d) : (Q/G)^3 \times_{Q/G} (\tilde{G} \oplus \tilde{G}) \to T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$, is obtained from the reduced discrete Euler–Lagrange operator by making the identifications obtained from the discrete connection structure.

The splitting of the range space of $\mathcal{LP}_d(l_d)$ as a direct product (as in §3.3 of Cendra et al. [2001]) naturally induces a decomposition of the discrete Lagrange–Poincaré operator,

$$\mathcal{LP}_d(l_d) = \mathrm{Hor}(\mathcal{LP}_d(l_d)) \oplus \mathrm{Ver}(\mathcal{LP}_d(l_d)),$$

and this allows the discrete reduced equations to be decomposed in horizontal and vertical equations.

7.2. **Geometric Control Theory and Formations.** There are well-established control algorithms for actuating a control system to achieve a desired reference configuration. In many problems of practical interest, the actuation of the mechanical system decomposes into shape and group variables in a natural fashion.

A canonical example of this is a satellite in motion about the Earth, where the orientation of the satellite is controlled by internal rotors through the use of holonomy and geometric phases, and the position is controlled by chemical propulsion.

In this example, the configuration space is SE(3), the group is SO(3), and the shape space is $\mathbb{R}^3$. The group variable corresponds to the orientation, and the shape variable corresponds to the position. When given an initial and desired configuration, it is desirable, while computing the control inputs, to decompose the relative motion into a shape component and a group component, so that they can be individually actuated.

Since the discrete connection is used here to provide an efficient choice of local coordinates for optimal control, the discrete connection is most naturally obtained by exponentiating the continuous connection, in the manner described in §6.3, in conjunction with the explicit formulas for the exponential and logarithm for SE(3). The natural choice of the continuous connection is one in which the horizontal space is given by the momentum surface corresponding to the current value of the momentum.

To illustrate why it may not be desirable from a control-theoretic point of view to decompose the space using a trivial connection, consider a satellite that is in a tidally locked orbit about the Earth, with the initial and desired configuration as illustrated in Figure 10.

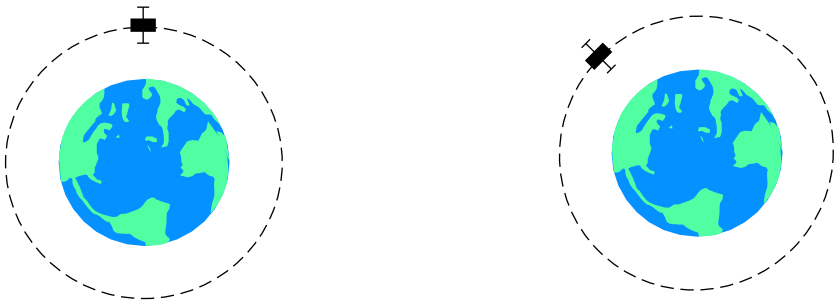


Figure 10: Application of discrete connections to control.

Here, if we choose a trivial connection, then the relative group element would be a rotation by $\pi/4$, but this choice is undesirable, since the motion is tidally locked, and moving the center of mass to the new location would result in a shift in the orientation by precisely the desired amount. In this case, the optimal control input should therefore only actuate in the shape variables, and the relative group element assigned to this pair of configurations should be the identity element.

As such, the extension of mechanically relevant connections to pairs of points in the configuration space with finite separation, through the use of a discrete connection, can be of immense value in geometric control theory.

Similarly, in the case of formations, discrete connections allow for the orientation coordination problem to be handled in a more efficient manner, by taking into account the dynamic coupling of the shape and group motions automatically through the use of the discrete mechanical connection.

7.3. **Discrete Levi-Civita Connection.** Vector bundle connections can be cast in the language of connections on principal bundles by considering the frame bundle consisting of oriented orthonormal frames over the manifold M, which is a principal $\mathrm{SO}(n)$ bundle, as originally proposed by Cartan [1983, 2001]. For related work on discrete connections on triangulated manifolds with applications to algebraic topology and the computation of Chern classes, please see Novikov [2003].

To construct our model of a discrete Riemannian manifold, we first trivialize the frame bundle to yield $\mathrm{SO}(n) \times M$. Then, $Q = \mathrm{SO}(n) \times M$, and $G = \mathrm{SO}(n)$.

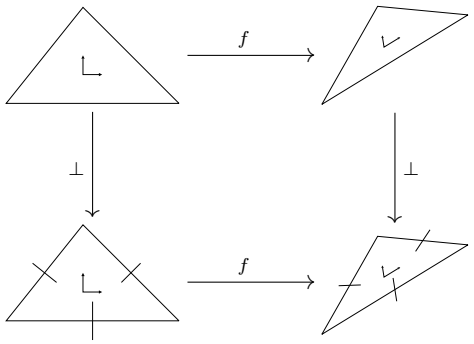
Here, we introduce the notion of a semidiscretized principal bundle, where the shape space, $S = Q/G$, is discretized as a simplicial complex K, and the structure group G remains continuous. In this context, the semidiscretization of the trivialization of the frame bundle is given by $Q = \mathrm{SO}(n) \times K$.

A discrete connection is a map $\mathcal{A}_d : Q \times Q \to G$, and we can construct a candidate for the Levi-Civita connection on a simplicial complex K, using the discrete analogue of the frame bundle described above. However, we will first introduce the notion of a discrete Riemannian manifold.

Definition 7.1. *A **discrete dual Riemannian manifold** is a simplicial complex where each n-simplex σ^n is endowed with a constant Riemannian metric tensor g, such that the restriction of the metric tensor to a common face with an adjacent n-simplex is consistent.*

This is referred to as a discrete dual Riemannian manifold as we can equivalently think of associating a Riemannian metric tensor to each dual vertex, and as we shall see, by adopting Cartan's method of orthogonal frames (see, for example, Cartan [1983, 2001]), the connection is a $\mathrm{SO}(n)$-valued discrete dual 1-form, and the curvature is a $\mathrm{SO}(n)$-valued discrete dual 2-form.

For each n-simplex σ^n, consider an invertible transformation f of $\mathbb{R}^n$ such that $f^*g = I$. In the orthonormal space, we have a normal operator that maps a $(n-1)$-dimensional subspace to a generator of the orthogonal complement, denoted by $\perp$. Then, we obtain a normal operator on the faces of σ^n by making the following diagram commute.



The coordinate axes in the diagram represent the normalized eigenvectors of the metric, scaled by their respective eigenvalues.

The local representation of the discrete connection is given by

$$\mathcal{A}_d((\sigma_0^n, R_0), (\sigma_1^n, R_1)) = R_1 A(\sigma_0^n, \sigma_1^n) R_0^{-1},$$

and so the discrete connection is uniquely defined if we specify $A(\sigma_0^n, \sigma_1^n)$, where σ_0^n and σ_1^n are adjacent n-simplices. Since they are adjacent, they share a $(n-1)$-simplex, denoted σ^{n-1}. In particular, this can then be thought of as a $\mathrm{SO}(n)$-valued discrete dual 1-form, since to each dual 1-cell, $\star\sigma^{n-1}$, we associate an element of $\mathrm{SO}(n)$.

This element of $\mathrm{SO}(n)$ is computed as follows.

(1) In each of the n-simplices, we have a normal direction associated with σ^{n-1}, denoted by
$$\perp (\sigma^{n-1}, \sigma_i^n) \in \mathbb{R}^n.$$
(2) If these two normal directions are parallel, we set
$$\langle A, \star\sigma^{n-1} \rangle = I,$$
otherwise, we continue.

(3) Construct the $(n-2)$-dimensional hyperplane P^{n-2}, given by the orthogonal complement to the span of the two normal directions.

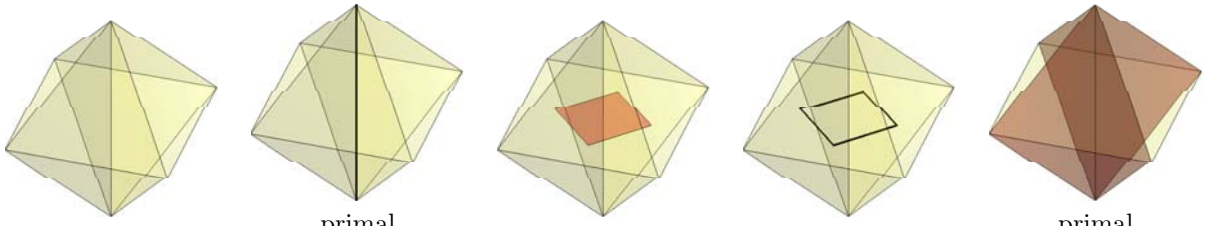
$$P^{n-2} = \perp (\text{span}(\perp (\sigma^{n-1}, \sigma_0^n), \perp (\sigma^{n-1}, \sigma_1^n))).$$

(4) If $\star\sigma^{n-1}$ is oriented from σ_0^n to σ_1^n, we set

$$\langle A, \star\sigma^{n-1}\rangle = \{R \in \text{SO}(n) \mid R|_{P^{n-2}} = I_{P^{n-2}}, R(\perp (\sigma^{n-1}, \sigma_0^n)) = \perp (\sigma^{n-1}, \sigma_1^n)\}.$$

The curvature of this discrete Levi-Civita connection is then a $\text{SO}(n)$-valued discrete dual 2-form. There is however the curious property that the boundary operator for a dual cell complex may not necessarily agree with the standard notion of boundary, since that may not be expressible in terms of a chain in the dual cell complex. This is primarily an issue on the boundary of the simplicial complex, and if we are in the interior, this is not a problem.

Since curvature is a dual 2-form, it is associated with the dual of a codimension-two simplex, given by $\star\sigma^{n-2}$. Consider the example illustrated in Figure 11.



Simplicial Complex, K | primal $(n-2)$-simplex, σ^{n-2} | dual 2-cell, $\star\sigma^{n-2}$ | dual 1-chain, $\partial \star \sigma^{n-2}$ | primal $(n-1)$-chain, $\star\partial \star \sigma^{n-2}$

Figure 11: Discrete curvature as a discrete dual 2-form.

The curvature $\mathcal{B}$ of the discrete Levi-Civita connection is given by

$$\langle \mathcal{B}, \star\sigma^{n-2}\rangle = \langle \mathbf{d}A, \star\sigma^{n-2}\rangle = \langle A, \partial \star \sigma^{n-2}\rangle.$$

As can be seen from the geometric region $\star\partial \star \sigma^{n-2}$, the curvature associated with $\star\sigma^{n-2}$ is given by the ordered product of the connection associated with the dual cells, $\star\sigma^{n-1}$, where $\sigma^{n-1} \succ \sigma^{n-2}$. This also suggests that we can also think of the discrete connection as a primal $(n-1)$-form, and the curvature as a primal $(n-2)$-form, where the curvature is obtained from the connection using the codifferential.

When the group G is nonabelian, we see that the curvature is only defined up to conjugation, since we need to specify a dual vertex on the dual one-chain $\partial \star \sigma^{n-2}$ from which to start composing group elements. To make this well-defined, we can adopt the approach used in defining the simplicial cup product, and assume that there is a partial ordering on dual vertices, which would make the curvature unambiguous.

As a quality measure for simplicial triangulations of a Riemannian manifold, having the curvature defined up to conjugation may be sufficient if we have a norm on $\text{SO}(n)$ which is invariant under conjugation. As an example, taking the logarithm to get an element of the Lie algebra $\mathfrak{so}(n)$, and then using a norm on this vector space yields a conjugation-invariant norm on $\text{SO}(n)$. This allows us to detect regions of the mesh with high curvature, and selectively subdivide the triangulation in such regions.

Similarly, we can define a discrete primal Riemannian manifold, where the Riemannian metric tensor is associated with primal vertices, and the connection is a G-valued primal 1-form, and the curvature is a G-valued primal 2-form.

Abstract Simplicial Complex with a Local Metric. In Desbrun et al. [2003], the notion of an abstract simplicial complex with a local metric, defined on pairs of vertices that are adjacent, was introduced.

In this situation, we can compute the curvature around a loop in the mesh using local embeddings. We start with an initial n-simplex, which we endow with an orthonormal frame. By locally embedding adjacent n-simplices into Euclidean space, and parallel transporting the orthonormal frame, we will eventually transport the frame back to the initial simplex.

The relative orientation between the original frame and the transported frame yields the integral of the curvature of the surface which is bounded by the traversed curve. This results from a simple application of the Generalized Stokes' theorem, and the fact that the curvature is given by the exterior derivative of the connection 1-form.

8. Conclusions and Future Work

We have introduced a complete characterization of discrete connections, in terms of horizontal and vertical spaces, discrete connection 1-forms, horizontal lifts, and splittings of the discrete Atiyah sequence. Geometric structures that can be derived from a given discrete connection have been discussed, including continuous connections, an extended pair groupoid composition, and higher-order analogues of the discrete connection. In addition, we have explored computational issues, such as order of accuracy, and the construction of discrete connections from continuous connections.

Applications to discrete reduction theory, geometric control theory, and discrete geometry, have also been discussed, and it would be desirable to systematically apply the machinery of discrete connections to these problems.

In addition, connections play a crucial role in representing the nonholonomic constraint distribution in nonholonomic mechanics, particularly when considering nonholonomic mechanical systems with symmetry, wherein the nonholonomic connection enters (see, for example, Bloch [2003]). There has been recent progress on constructing nonholonomic integrators in the work of Cortés [2002] and McLachlan and Perlmutter [2003], but an intrinsically discrete notion of a connection remains absent from their work, and they do not consider the role of symmetry reduction in discrete nonholonomic mechanics.

It would be very interesting to apply the general theory of discrete connections on principal bundles to nonholonomic mechanical systems with symmetry, and to cast the notion of a discrete nonholonomic constraint distribution and the nonholonomic connection in the language of discrete connections, and thereby develop a discrete theory of nonholonomic mechanics with symmetry. This would be particularly important for the numerical implementation of geometric control algorithms.

The role of discrete connections in the study of discrete geometric phases would also be an area worth pursuing. A discrete analogue of the rigid-body phase formula, that involves the discrete mechanical connection, that is exact for rigid-body simulations that use discrete variational mechanics, would yield significant insights into the geometric structure-preservation properties of variational integrators. In particular, it would provide much needed insight into how discretization interacts with geometric phases, and yield an understanding how much of the phase drift observed in a numerical simulation is due to the underlying geometry of the mechanical system, and how much is due to the process of discretizing the system.

References

R. Almeida and P. Molino. Suites d'Atiyah et feuilletages transversalement complets. *C. R. Acad. Sci. Paris Sér. I Math.*, 300(1):13–15, 1985.

M. F. Atiyah. Complex analytic connections in fibre bundles. *Trans. Amer. Math. Soc.*, 85:181–207, 1957.

M. Berry. Anticipations of the geometric phase. *Phys. Today*, pages 34–40, December 1990.

A. M. Bloch. *Nonholonomic Mechanics and Control*. Interdisciplinary Applied Mathematics. Springer-Verlag, 2003.

J. R. Cardoso and F. Silva Leite. Theoretical and numerical considerations about Padé approximants for the matrix logarithm. *Linear Algebra Appl.*, 330(1-3):31–42, 2001.

É. Cartan. *Geometry of Riemannian spaces*. Math. Sci. Press, Brookline, CA, 1983. (translated from French).

É. Cartan. *Riemannian Geometry in an Orthogonal Frame*. World Scientific, 2001. (translated from French).

E. Celledoni and A. Iserles. Approximating the exponential from a Lie algebra to a Lie group. *Math. Comp.*, 69(232):1457–1480, 2000.

E. Celledoni and A. Iserles. Methods for the approximation of the matrix exponential in a lie-algebraic setting. *IMA J. Num. Anal.*, 21(2):463–488, 2001.

H. Cendra, J. E. Marsden, and T. S. Ratiu. Lagrangian reduction by stages. *Mem. Amer. Math. Soc.*, 152 (722), 2001.

J. Cortés. *Geometric, Control and Numerical Aspects of Nonholonomic Systems*, volume 1793 of *Lecture Notes in Mathematics*. Springer-Verlag, 2002.

M. Desbrun, A. N. Hirani, M. Leok, and J. E. Marsden. Discrete exterior calculus. (in preparation), 2003.

P. Goldreich and A. Toomre. Some remarks on polar wandering. *J. Geophys. Res.*, 10:2555–2567, 1969.

N. J. Higham. Evaluating Padé approximants of the matrix logarithm. *SIAM J. Matrix Anal. Appl.*, 22(4): 1126–1135 (electronic), 2001.

S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*, volume 1. Wiley, 1963.

M. Leok. A mathematical model of true polar wander. Caltech SURF Report, 1998.

K. C. H. Mackenzie. Lie algebroids and Lie pseudoalgebras. *Bull. London Math. Soc.*, 27(2):97–147, 1995.

J. E. Marsden. Geometric mechanics, stability, and control. In L. Sirovich, editor, *Applied Mathematical Sciences*, volume 100, pages 265–291. Springer-Verlag, 1994.

J. E. Marsden. Geometric foundations of motion and control. In *Motion, Control and Geometry*, pages 3–19. National Academy Press, 1997.

J. E. Marsden, R. Montgomery, and T. S. Ratiu. Reduction, symmetry, and phases in mechanics. *Mem. Amer. Math. Soc.*, 88(436):1–110, 1990.

J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, second edition, 1999.

R. McLachlan and M. Perlmutter. Integrators for nonholonomic mechanical systems. (preprint), 2003.

R. Montgomery. How much does a rigid body rotate? A Berry's phase from the eighteenth century. *Amer. J. Phys.*, 59:394–398, 1991.

S. P. Novikov. Discrete connections on the triangulated manifolds and difference linear equations. arXiv, math-ph/0303035, 2003.

P. J. Olver. Geometric foundations of numerical algorithms and symmetry. *Appl. Algebra Engrg. Comm. Comput.*, 11(5):417–436, 2001. Special issue "Computational geometry for differential equations".

A. Shapere and F. Wilczek. *Geometric Phases in Physics*, volume 5 of *Advanced Series in Mathematical Physics*. World Scientific, 1989.

N. Steenrod. *The Topology of Fibre Bundles*. Princeton University Press, 1951.

V. Vedral. Geometric phases and topological quantum computation. *Int. J. Quantum Inform.*, 1(1):1–23, 2003.

G. Walsh and S. Sastry. On reorienting linked rigid bodies using internal motions. *IEEE Trans. Robotic Autom.*, 11:139–146, 1993.

A. Zanna and H. Z. Munthe-Kaas. Generalized polar decompositions for the approximation of the matrix exponential. *SIAM J. Matrix Anal. Appl.*, 23(3):840–862 (electronic), 2001/02.

107-81, Control and Dynamical Systems, Caltech, Pasadena, CA 91125.
E-mail address: mleok@cds.caltech.edu

107-81, Control and Dynamical Systems, Caltech, Pasadena, CA 91125.
E-mail address: marsden@cds.caltech.edu

Department of Mathematics, University of California, Berkeley, CA 94720.
E-mail address: alanw@math.berkeley.edu

DISCRETE EXTERIOR CALCULUS

MATHIEU DESBRUN, ANIL N. HIRANI, MELVIN LEOK, AND JERROLD E. MARSDEN

Abstract. We present a theory and applications of discrete exterior calculus on simplicial complexes of arbitrary finite dimension. This can be thought of as calculus on a discrete space. Our theory includes not only discrete differential forms but also discrete vector fields and the operators acting on these objects. This allows us to address the various interactions between forms and vector fields (such as Lie derivatives) which are important in applications. Previous attempts at discrete exterior calculus have addressed only differential forms. We also introduce the notion of a circumcentric dual of a simplicial complex. The importance of dual complexes in this field has been well understood, but previous researchers have used barycentric subdivision or barycentric duals. We show that the use of circumcentric duals is crucial in arriving at a theory of discrete exterior calculus that admits both vector fields and forms.

Contents

1. Introduction

This work presents a theory of ***discrete exterior calculus*** (DEC) motivated by potential applications in computational methods for field theories such as elasticity, fluids, and electromagnetism. In addition, it provides much needed mathematical machinery to enable a systematic development of numerical schemes that mirror the approach of geometric mechanics.

This theory has a long history that we shall outline below in §2, but we aim at a comprehensive, systematic, as well as useful, treatment. Many previous works, as we shall review, are incomplete both in terms of the objects that they treat as well as the types of meshes that they allow.

Our vision of this theory is that it should proceed *ab initio* as a discrete theory that parallels the continuous one. General views of the subject area of DEC are common in the literature (see, for instance, Mattiussi [2000]), but they usually stress the process of discretizing a continuous theory and the overall approach is tied to this goal. However, if one takes the point of view that the discrete theory can, and indeed should, stand in its own right, then the range of application areas naturally is enriched and increases.

Convergence and consistency considerations alone are inadequate to discriminate between the various choices of discretization available to the numerical analyst, and only by requiring, when appropriate, that the discretization exhibits discrete analogues of continuous properties of interest can we begin to address the question of what makes a discrete theory a canonical discretization of a continuous one.

Applications to Variational Problems. One of the major application areas we envision is to variational problems, be they in mechanics or optimal control. One of the key ingredients in this direction that we imagine will play a key role in the future is that of AVI's (asynchronous variational integrators) designed for the numerical integration of mechanical systems, as in Lew et al. [2003]. These are integration algorithms that respect some of the key features of the continuous theory, such as their multi-symplectic nature and exact conservation laws. They do so by discretizing the underlying variational principles of mechanics rather than discretizing the equations. It is well-known (see the reference just mentioned for some of the literature) that variational problems come equipped with a rich exterior calculus structure and so on the discrete level, such structures will be enhanced by the availability of a discrete exterior calculus. One of the objectives of this chapter is to fill this gap.

Structured Constraints. There are many constraints in numerical algorithms that naturally involve differential forms, such as the divergence constraint for incompressibility of fluids, as well as the fact that differential forms are naturally the fields in electromagnetism, and some of Maxwell's equations are expressed in terms of the divergence and curl operations on these fields. Preserving, as in the mimetic differencing literature, such features directly on the discrete level is another one of the goals, overlapping with our goals for variational problems.

Lattice Theories. Periodic crystalline lattices are of important practical interest in material science, and the anisotropic nature of the material properties arises from the geometry and connectivity of the intermolecular bonds in the lattice. It is natural to model these lattices as inherently discrete objects, and an understanding of discrete curvature that arises from DEC is particularly relevant, since part of the potential energy arises from stretched bonds that can be associated with discrete curvature in the underlying relaxed configuration of the lattice. In particular, this could yield a more detailed geometric understanding of what happens at grain boundaries. Lattice defects can also be associated with discrete curvature when appropriately interpreted. The introduction of a discrete notion of curvature will lay the foundations for a better understanding of the role of geometry in the material properties of solids.

Some of the Key Theoretical Accomplishments. Our development of discrete exterior calculus includes discrete differential forms, the Hodge star operator, the wedge product, the exterior derivative, as well as contraction and the Lie derivative. For example, this approach leads to the proper definition of

discrete divergence and curl operators and has already resulted in applications like a discrete Hodge type decomposition of 3D vector fields on irregular grids—see Tong et al. [2003].

Context. We present the theory and some applications of DEC in the context of simplicial complexes of arbitrary finite dimension.

Methodology. We believe that the correct way to proceed with this program is to develop, as we have already stressed, *ab initio*, a calculus on discrete manifolds which parallels the calculus on smooth manifolds of arbitrary finite dimension. Chapters 6 and 7 of Abraham et al. [1988] are a good source for the concepts and definitions in the smooth case. However we have tried to make this chapter as self-contained as possible. Indeed, one advantage of developing a calculus on discrete manifolds, as we do here, is pedagogical. By using concrete examples of discrete two- and three-dimensional spaces one can explain most of calculus on manifolds at least formally as we will do using the examples in this chapter. The machinery of Riemannian manifolds and general manifold theory from the smooth case is, strictly speaking, not required in the discrete world. The technical terms that are used in this introduction will be defined in subsequent sections, but they should be already familiar to someone who knows the usual exterior calculus on smooth manifolds.

The Objects in DEC. To develop a discrete theory, one must define discrete differential forms along with vector fields and operators involving these. Once discrete forms and vector fields are defined, a calculus can be developed by defining the discrete exterior derivative ($\mathbf{d}$), codifferential ($\boldsymbol{\delta}$) and Hodge star ($*$) for operating on forms, discrete wedge product ($\wedge$) for combining forms, discrete flat ($\flat$) and sharp ($\sharp$) operators for going between vector fields and 1-forms and discrete contraction operator ($\mathbf{i}_X$) for combining forms and vector fields. Once these are done, one can then define other useful operators. For example, a discrete Lie derivative ($\pounds_X$) can be *defined* by requiring that the Cartan magic (or homotopy) formula hold. A discrete divergence in any dimension can be defined. A discrete Laplace–deRham operator (Δ) can be defined using the usual definition of $\mathbf{d}\boldsymbol{\delta} + \boldsymbol{\delta}\mathbf{d}$. When applied to functions, this is the same as the discrete Laplace–Beltrami operator (∇^2), which is the defined as $\mathrm{div} \circ \mathrm{curl}$. We define all these operators in this chapter.

The discrete manifolds we work with are simplicial complexes. We will recall the standard formal definitions in §3 but familiar examples of simplicial complexes are meshes of triangles embedded in $\mathbb{R}^3$ and meshes made up of tetrahedra occupying a portion of $\mathbb{R}^3$. We will assume that the angles and lengths on such discrete manifolds are computed in the embedding space $\mathbb{R}^N$ using the standard metric of that space. In other words, in this chapter we do not address the issue of how to discretize a given smooth Riemannian manifold, and how to embed it in $\mathbb{R}^N$, since there may be many ways to do this. For example, SO(3) can be embedded in $\mathbb{R}^9$ with a constraint, or as the unit quaternions in $\mathbb{R}^4$. Another potentially important consideration in discretizing the manifold is that the topology of the simplicial complex should be the same as the manifold to be discretized. This can be verified using the methods of computational homology (see, for example, Kaczynski et al. [2004]), or discrete Morse theory (see, for example, Forman [2002], Wood [2003]). For the purposes of discrete exterior calculus, only local metric information is required, and we will comment towards the end of §3 how to address the issue of embedding in a local fashion, as well as the criterion for a good global embedding.

Our development in this chapter is for the most part formal in that we choose appropriate geometric definitions of the various objects and quantities involved. For the most part, we do not prove that these definitions converge to the smooth counterparts. The definitions are chosen so as to make some important theorems like the generalized Stokes' theorem true by definition. Moreover, in the cases where previous results are available, we have checked that the operators we obtain match the ones obtained by other means, such as variational derivations.

2. History and Previous Work

The use of simplicial chains and cochains as the basic building blocks for a discrete exterior calculus has appeared in several papers. See, for instance, Sen et al. [2000], Adams [1996], Bossavit [2002c], and references therein. These authors view forms as linearly interpolated versions of smooth differential forms, a viewpoint originating from Whitney [1957], who introduced the Whitney and deRham maps that establish an isomorphism between simplicial cochains and Lipschitz differential forms.

We will, however, view discrete forms as real-valued linear functions on the space of chains. These are inherently discrete objects that can be paired with chains of oriented simplices, or their geometric duals, by the bilinear pairing of evaluation. In the next chapter, where we consider applications involving the curvature of a discrete space, we will relax the condition that discrete forms are real-valued, and consider group-valued forms.

Intuitively, this natural pairing of evaluation can be thought of as integration of the discrete form over the chain. This difference from the work of Sen et al. [2000] and Adams [1996] is apparent in the definitions of operations like the wedge product as well.

There is also much interest in a discrete exterior calculus in the computational electromagnetism community, as represented by Bossavit [2001, 2002a,b,c], Gross and Kotiuga [2001], Hiptmair [1999, 2001a,b, 2002], Mattiussi [1997, 2000], Nicolaides and Wang [1998], Teixeira [2001], and Tonti [2002].

Many of the authors cited above, for example, Bossavit [2002c], Sen et al. [2000], and Hiptmair [2002], also introduce the notions of dual complexes in order to construct the Hodge star operator. With the exception of Hiptmair, they use barycentric duals. This works if one develops a theory of discrete forms and does not introduce discrete vector fields. We show later that to introduce discrete vector fields into the theory the notion of circumcentric duals seems to be important.

Other authors, such as Moritz [2000], Moritz and Schwalm [2001], Schwalm et al. [1999], have incorporated vector fields into the cochain based approach to exterior calculus by identifying vector fields with cochains, and having them supported on the same mesh. This is ultimately an unsatisfactory approach, since dual meshes are essential as a means of encoding physically relevant phenomena such as fluxes across boundaries.

The use of primal and dual meshes arises most often as staggered meshes in finite volume and finite difference methods. In fluid computations, for example, the density is often a cell-centered quantity, which can either be represented as a primal object by being associated with the 3-cell, or as a dual object associated with the 0-cell at the center of the 3-cell. Similarly, the flux across boundaries can be associated with the 2-cells that make up the boundary, or the 1-cell which is normal to the boundary.

Another approach to a discrete exterior calculus is presented in Dezin [1995]. He defines a one-dimensional discretization of the real line in much the same way we would. However, to generalize to higher dimensions he introduces a tensor product of this space. This results in logically rectangular meshes. Our calculus, however, is defined over simplicial meshes. A further difference is that like other authors in this field, Dezin [1995] does not introduce vector fields into his theory.

A related effort for three-dimensional domains with logically rectangular meshes is that of Mansfield and Hydon [2001], who established a variational complex for difference equations by constructing a discrete homotopy operator. We construct an analogous homotopy operator for simplicial meshes in proving the discrete Poincaré lemma.

3. Primal Simplicial Complex and Dual Cell Complex

In constructing the discretization of a continuous problem in the context of our formulation of discrete exterior calculus, we first discretize the manifold of interest as a simplicial complex. While this is typically in the form of a simplicial complex that is embedded into Euclidean space, it is only necessary to have an abstract simplicial complex, along with a local metric defined on adjacent vertices. This abstract setting will be addressed further toward the end of this section.

We will now recall some basic definitions of simplices and simplicial complexes, which are standard from simplicial algebraic topology. A more extensive treatment can be found in Munkres [1984].

Definition 3.1. *A k-**simplex** is the convex span of $k+1$ geometrically independent points,*

$$\sigma^k = [v_0, v_1, \ldots, v_k] = \left\{ \sum_{i=0}^{k} \alpha^i v_i \;\middle|\; \alpha^i \geq 0, \sum_{i=0}^{n} \alpha^i = 1 \right\}.$$

*The points $v_0, \ldots, v_k$ are called the **vertices** of the simplex, and the number k is called the **dimension** of the simplex. Any simplex spanned by a (proper) subset of $\{v_0, \ldots, v_k\}$ is called a **(proper) face** of σ^k. If σ^l is a proper face of σ^k, we denote this by $\sigma^l \prec \sigma^k$.*

Example 3.1. *Consider 3 non-collinear points v_0, v_1 and v_2 in $\mathbb{R}^3$. Then, these three points individually are examples of 0-simplices, to which an orientation is assigned through the choice of a sign. Examples of 1-simplices are the oriented line segments $[v_0, v_1]$, $[v_1, v_2]$ and $[v_0, v_2]$. By writing the vertices in that order we have given orientations to these 1-simplices, i.e., $[v_0, v_1]$ is oriented from v_0 to v_1. The triangle $[v_0, v_1, v_2]$ is a 2-simplex oriented in counterclockwise direction. Note that the orientation of $[v_0, v_2]$ does not agree with that of the triangle.*

Definition 3.2. *A **simplicial complex** K in $\mathbb{R}^N$ is a collection of simplices in $\mathbb{R}^N$, such that,*

1. *Every face of a simplex of K is in K.*
2. *The intersection of any two simplices of K is a face of each of them.*

Definition 3.3. *A **simplicial triangulation** of a polytope $|K|$ is a simplicial complex K such that the union of the simplices of K recovers the polytope $|K|$.*

Definition 3.4. *If L is a subcollection of K that contains all faces of its elements, then L is a simplicial complex in its own right, and it is called a **subcomplex** of K. One subcomplex of K is the collection of all simplices of K of dimension at most k, which is called the k-**skeleton** of K, and is denoted $K^{(k)}$.*

Circumcentric Subdivision. We will also use the notion of a circumcentric dual or Voronoi mesh of the given primal mesh. We will point to the importance of this choice later on in §7 and 9. We call the Voronoi dual a circumcentric dual since the dual of a simplex is its circumcenter (equidistant from all vertices of the simplex).

Definition 3.5. *The **circumcenter** of a k-simplex σ^k is given by the center of the k-circumsphere, where the k-circumsphere is the unique k-sphere that has all $k+1$ vertices of σ^k on its surface. Equivalently, the circumcenter is the unique point in the k-dimensional affine space that contains the k-simplex that is equidistant from all the $k+1$ nodes of the simplex. We will denote the circumcenter of a simplex σ^k by $c(\sigma^k)$.*

The circumcenter of a simplex σ^k can be obtained by taking the intersection of the normals to the $(k-1)$-dimensional faces of the simplex, where the normals are emanating from the circumcenter of the face. This allows us to recursively compute the circumcenter.

If we are given the nodes which describe the primal mesh, we can construct a simplicial triangulation by using the Delaunay triangulation, since this ensures that the circumcenter of a simplex is always a point within the simplex. Otherwise we assume that a nice mesh has been given to us, i.e., it is such that the circumcenters lie within the simplices. While this is not be essential for our theory it makes some proofs simpler. For some computations the Delaunay triangulation is desirable in that it reduces the maximum aspect ratio of the mesh, which is a factor in determining the rate at which the corresponding numerical scheme converges. But in practice there are many problems for which Delaunay triangulations are a bad idea. See, for example, Schewchuck [2002]. We will address such computational issues in a separate work.

Definition 3.6. *The* ***circumcentric subdivision*** *of a simplicial complex is given by the collection of all simplices of the form*

$$[c(\sigma_0), \ldots, c(\sigma_k)],$$

where $\sigma_0 \prec \sigma_1 \prec \ldots \prec \sigma_k$*, or equivalently, that* σ_i *is a proper face of* σ_j *for all* $i < j$*.*

Circumcentric Dual. We construct a circumcentric dual to a k-simplex using the circumcentric duality operator, which is introduced below.

Definition 3.7. *The* ***circumcentric duality operator*** *is given by*

$$\star\left(\sigma^k\right) = \sum_{\sigma^k \prec \sigma^{k+1} \prec \ldots \prec \sigma^n} \epsilon_{\sigma^k,\ldots,\sigma^n} \left[c(\sigma^k), c(\sigma^{k+1}), \ldots, c(\sigma^n)\right],$$

where the $\epsilon_{\sigma^k,\ldots,\sigma^n}$ *coefficient ensures that the orientation of* $\left[c(\sigma^k), c(\sigma^{k+1}), \ldots, c(\sigma^n)\right]$ *is consistent with the orientation of the primal simplex, and the ambient volume-form.*

Orienting σ^k *is equivalent to choosing a ordered basis, which we shall denote by* $dx^1 \wedge \ldots \wedge dx^k$*. Similarly,* $\left[c(\sigma^k), c(\sigma^{k+1}), \ldots, c(\sigma^n)\right]$ *has an orientation denoted by* $dx^{k+1} \wedge \ldots \wedge dx^n$*. If the orientation corresponding to* $dx^1 \wedge \ldots \wedge dx^n$ *is consistent with the volume-form on the manifold, then* $\epsilon_{\sigma^k,\ldots,\sigma^n} = 1$*, otherwise it takes the value* -1*.*

We immediately see from the construction of the circumcentric duality operator that the dual elements can be realized as a submesh of the first circumcentric subdivision, since it consists of elements of the form $[c(\sigma_0), \ldots, c(\sigma_k)]$, which are, by definition, part of the first circumcentric subdivision.

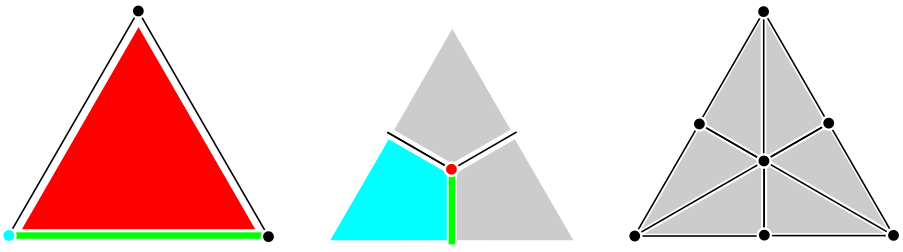
Example 3.2. *The circumcentric duality operator maps a* 0*-simplex into the convex hull generated by the circumcenters of* n*-simplices that contain the* 0*-simplex,*

$$\star(\sigma^0) = \left\{ \sum \alpha_{\sigma^n} c\left(\sigma^n\right) \;\middle|\; \alpha_{\sigma^n} \geq 0, \sum \alpha_{\sigma^n} = 1, \sigma^0 \prec \sigma^n \right\},$$

and the circumcentric duality operator maps a $n-$*simplex into the circumcenter of the* $n-$*simplex,*

$$\star(\sigma^n) = c(\sigma^n).$$

This is more clearly illustrated in Figure 1, where the primal and dual elements are color coded to represent the dual relationship between the elements in the primal and dual mesh.



(a) Primal (b) Dual (c) First subdivision

Figure 1: Primal, and dual meshes, as chains in the first circumcentric subdivision.

The choice of a circumcentric dual is significant, since it allows us to recover geometrically important objects such as normals to $(n-1)$-dimensional faces, which are obtained by taking their circumcentric dual,

whereas, if we were to use a barycentric dual, the dual to a $(n-1)$-dimensional face would not be normal to it.

Orientation of the Dual Cell. Notice that given an oriented simplex σ^k, which is represented by $[v_0, \ldots, v_k]$, the orientation is equivalently represented by $(v_1 - v_0) \wedge (v_2 - v_1) \wedge \ldots \wedge (v_k - v_{k-1})$, which we denote by,

$$[v_0, \ldots, v_k] \sim (v_1 - v_0) \wedge (v_2 - v_1) \wedge \ldots \wedge (v_k - v_{k-1}),$$

which is an equivalence at the level of orientation. It would be nice to express our criterion for determining the orientation of the dual cell in terms of the $(k+1)$-vertex representation.

To determine the orientation of the $(n-k)$-simplex given by $[c(\sigma^k), c(\sigma^{k+1}), \ldots, c(\sigma^n)]$, or equivalently, $dx^{k+1} \wedge \ldots \wedge dx^n$, we consider the n-simplex given by $[c(\sigma^0), \ldots, c(\sigma^n)]$, where $\sigma^0 \prec \ldots \prec \sigma^k$. This is related to the expression $dx^1 \wedge \ldots \wedge dx^n$, up to a sign determined by the relative orientation of $[c(\sigma^0), \ldots, c(\sigma^k)]$ and σ^k. Thus, we have that

$$dx^1 \wedge \ldots \wedge dx^n \sim \operatorname{sgn}([c(\sigma^0), \ldots, c(\sigma^k)], \sigma^k)[c(\sigma^0), \ldots, c(\sigma^n)]\,.$$

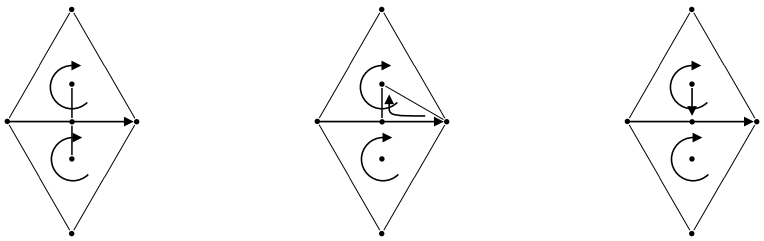
Then, we need to check that $dx^1 \wedge \ldots \wedge dx^n$ is consistent with the volume-form on the manifold, which is represented by the orientation of σ^n. Thus, we have that the correct orientation for the $[c(\sigma^k), c(\sigma^{k+1}), \ldots, c(\sigma^n)]$ term is given by,

$$\operatorname{sgn}([c(\sigma^0), \ldots, c(\sigma^k)], \sigma^k) \cdot \operatorname{sgn}([c(\sigma^0), \ldots, c(\sigma^n)], \sigma^n).$$

These two representations of the choice of orientation for the dual cells are equivalent, but the combinatorial definition above might be preferable for the purposes of implementation.

Example 3.3. *We would like to compute the orientation of the dual of a* 1*-simplex, in two dimensions, given the orientation of the two neighboring* 2*-simplices.*

Given a simplicial complex, as shown in Figure 2(a), we consider a 2*-simplex of the form* $[c(\sigma^0), c(\sigma^1), c(\sigma^2)]$*, which is illustrated in Figure 2(b).*



(a) Simplicial complex (b) 2-simplex (c) $\star\sigma^1$

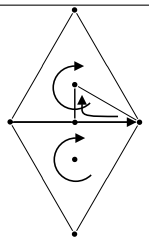
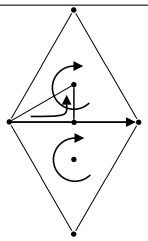
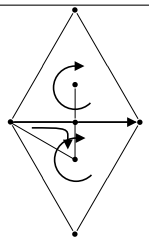
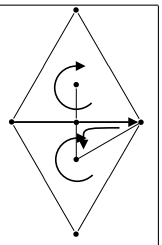
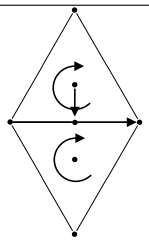
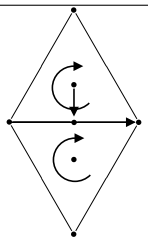
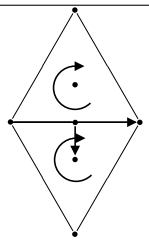
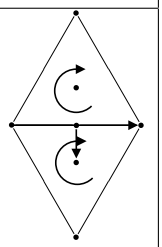
Figure 2: Orienting the dual of a cell.

Notice that the orientation is consistent with the given orientation of the 2*-simplex, but it is not consistent with the orientation of the primal* 1*-simplex, so the orientation should be reversed, to give the dual cell illustrated in Figure 2(c).*

We summarize the results for the induced orientation of dual cells for the other 2*-simplices of the form* $[c(\sigma^0), c(\sigma^1), c(\sigma^2)]$*, in Table 1.*

Orientation of the Dual of a Dual Cell. While the circumcentric duality operator is a map from the primal simplicial complex to the dual cell complex, we can formally extend the circumcentric duality operator

TABLE 1. Determining the induced orientation of a dual cell.

$[c(\sigma^0), c(\sigma^1), c(\sigma^2)]$				
$\operatorname{sgn}([c(\sigma^0), c(\sigma^1)], \sigma^1)$	$-$	$+$	$+$	$-$
$\operatorname{sgn}([c(\sigma^0), c(\sigma^1), c(\sigma^2)], \sigma^2)$	$+$	$-$	$+$	$-$
$\operatorname{sgn}([c(\sigma^0), c(\sigma^1)], \sigma^1)$ $\cdot \operatorname{sgn}([c(\sigma^0), c(\sigma^1), c(\sigma^2)], \sigma^2)$ $\cdot [c(\sigma^1), c(\sigma^2)]$				

to a map from the dual cell complex to the primal simplicial complex. However, we need to be slightly careful about the orientation of primal simplex we recover from applying the circumcentric duality operator twice.

We have that, $\star\star\,(\sigma^k) = \pm\sigma^k$, where the sign is chosen to ensure the appropriate choice of orientation. If, as before, σ^k has an orientation represented by $dx^1 \wedge \ldots \wedge dx^k$, and $\star\sigma^k$ has an orientation represented by $dx^{k+1} \wedge \ldots \wedge dx^n$, then the orientation of $\star\star\,(\sigma^k)$ is chosen so that $dx^{k+1} \wedge \ldots \wedge dx^n \wedge dx^1 \wedge \ldots \wedge dx^k$ is consistent with the ambient volume-form. Since, by construction, $\star(\sigma^k)$, $dx^1 \wedge \ldots \wedge dx^n$ has an orientation consistent with the ambient volume-form, we need only compare $dx^{k+1} \wedge \ldots \wedge dx^n \wedge dx^1 \wedge \ldots \wedge dx^k$ with $dx^1 \wedge \ldots \wedge dx^n$. Notice that it takes $n-k$ transpositions to get the dx^1 term in front of the $dx^{k+1} \wedge \ldots \wedge dx^n$ terms, and we need to do this k times for each term of $dx^1 \wedge \ldots \wedge dx^k$, so it follows that the sign is simply given by $(-1)^{k(n-k)}$, or equivalently,

$$\star\star\,(\sigma^k) = (-1)^{k(n-k)}\sigma^k. \tag{3.1}$$

A similar relationship holds if we use a dual cell instead of the primal simplex σ^k.

Support Volume of a Primal Simplex and Its Dual Cell. We can think of a cochain as being constructed out of a basis consisting of cosimplices or cocells with value 1 on a single simplex or cell, and 0 otherwise. The way to visualize this cosimplex is that it is associated with a differential form that has support on what we will refer to as the ***support volume*** associated with a given simplex or cell.

Definition 3.8. *The **support volume** of a simplex σ^k is a n-volume given by the convex hull of the geometric union of the simplex and its circumcentric dual. This is given by*

$$V_{\sigma^k} = \operatorname{convexhull}(\sigma^k, \star\sigma^k) \cap |K|.$$

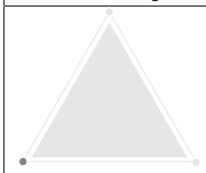
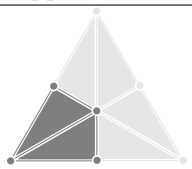
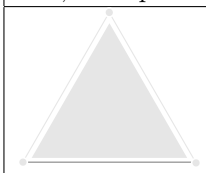
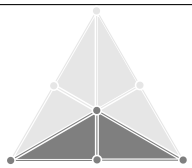
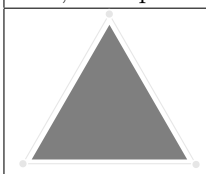
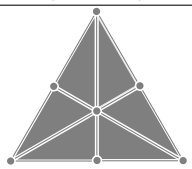
The intersection with $|K|$ is necessary to ensure that the support volume does not extend beyond the polytope $|K|$ which would otherwise occur if $|K|$ is nonconvex.

We extend the notion of a support volume to a dual cell $\star\sigma^k$ by similarly defining

$$V_{\star\sigma^k} = \operatorname{convexhull}(\star\sigma^k, \star\star\,\sigma^k) \cap |K| = V_{\sigma^k}.$$

To clarify this definition, we will consider some examples of simplices, their dual cells, and their corresponding support volumes. For two-dimensional simplicial complexes, this is illustrated in Table 2.

TABLE 2. Primal simplices, dual cells, and support volumes in two dimensions.

Primal Simplex	Dual Cell	Support Volume
		
σ^0, 0-simplex	$\star\sigma^0$, 2-cell	$V_{\sigma^0} = V_{\star\sigma^0}$
		
σ^1, 1-simplex	$\star\sigma^1$, 1-cell	$V_{\sigma^1} = V_{\star\sigma^1}$
		
σ^2, 2-simplex	$\star\sigma^2$, 0-cell	$V_{\sigma^2} = V_{\star\sigma^2}$

The support volume has the nice property that at each dimension, it partitions the polytope $|K|$ into distinct non-intersecting regions associated with each individual k-simplex. For any two distinct k-simplices, the intersection of their corresponding support volumes have measure zero, and the union of the support volumes of all k-simplices recovers the original polytope $|K|$.

Notice, from our construction, that the support volume of a simplex and its dual cell are the same, which suggests that there is an identification between cochains on k-simplices and cochains on $(n-k)$-cells. This is indeed the case, and is a concept associated with the Hodge star for differential forms.

Examples of simplices, their dual cells, and the corresponding support volumes in three dimensions are given in Table 3.

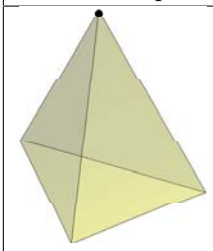
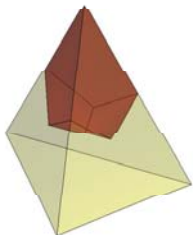
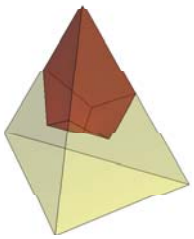
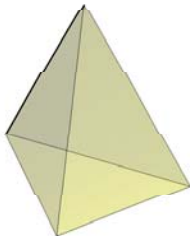
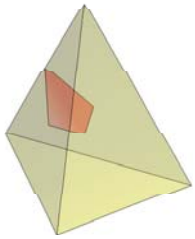
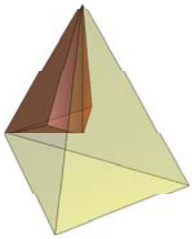
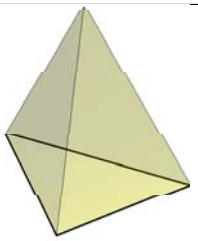
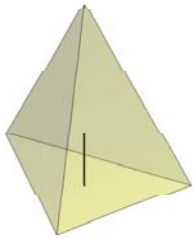
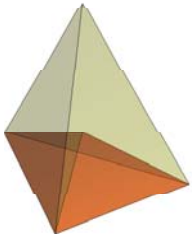
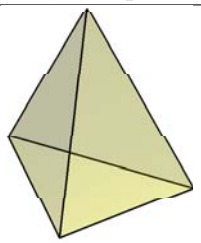
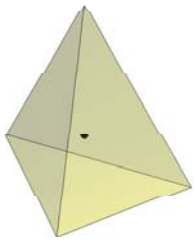
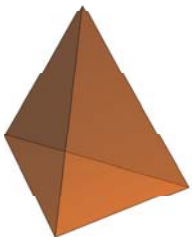
In our subsequent discussion, we will assume that we are given a simplicial complex K of dimension n in $\mathbb{R}^N$. Thus, the highest-dimensional simplex in the complex is of dimension n and each 0-simplex (vertex) is in $\mathbb{R}^N$. One can obtain this, for example, by starting from 0-simplices, i.e., vertices, and then constructing a Delaunay triangulation, using the vertices as sites. Often, our examples will be for two-dimensional discrete surfaces in $\mathbb{R}^3$ made up of triangles (here $n = 2$ and $N = 3$) or three-dimensional manifolds made of tetrahedra, possibly embedded in a higher-dimensional space.

Cell Complexes. The circumcentric dual of a primal simplicial complex is an example of a cell complex. The definition of a cell complex follows.

Definition 3.9. *A **cell complex** $\star K$ in $\mathbb{R}^N$ is a collection of cells in $\mathbb{R}^N$ such that,*

1. *There is a partial ordering of cells in $\star K$, $\hat{\sigma}^k \prec \hat{\sigma}^l$, which is read as $\hat{\sigma}^k$ is a face of $\hat{\sigma}^l$.*
2. *The intersection of any two cells in $\star K$, is either a face of each of them, or it is empty.*
3. *The boundary of a cell is expressible as a sum of its proper faces.*

TABLE 3. Primal simplices, dual cells, and support volumes in three dimensions.

Primal Simplex	Dual Cell	Support Volume
 σ^0, 0-simplex	 $\star\sigma^0$, 3-cell	 $V_{\sigma^0} = V_{\star\sigma^0}$
 σ^1, 1-simplex	 $\star\sigma^1$, 2-cell	 $V_{\sigma^1} = V_{\star\sigma^1}$
 σ^2, 2-simplex	 $\star\sigma^2$, 1-cell	 $V_{\sigma^2} = V_{\star\sigma^2}$
 σ^3, 3-simplex	 $\star\sigma^3$, 0-cell	 $V_{\sigma^3} = V_{\star\sigma^3}$

We will see in the next section that the notion of boundary in the circumcentric dual has to be modified slightly from the geometric notion of a boundary in order for the circumcentric dual to be made into a cell complex.

4. Local and Global Embeddings

While it is computationally more convenient to have a global embedding of the simplicial complex into a higher-dimensional ambient space to account for non-flat manifolds it suffices to have an abstract simplicial complex along with a local metric on vertices. The metric is local in the sense that distances between two

vertices are only defined if they are part of a common n-simplex in the abstract simplicial complex. Then, the local metric is a map $d : \{(v_0, v_1) \mid v_0, v_1 \in K^{(0)}, [v_0, v_1] \prec \sigma^n \in K\} \to \mathbb{R}$.

The axioms for a ***local metric*** are as follows,

Positive.: $d(v_0, v_1) \geq 0$, and $d(v_0, v_0) = 0$, $\forall [v_0, v_1] \prec \sigma^n \in K$.
Strictly Positive.: If $d(v_0, v_1) = 0$, then $v_0 = v_1$, $\forall [v_0, v_1] \prec \sigma^n \in K$.
Symmetry.: $d(v_0, v_1) = d(v_1, v_0)$, $\forall [v_0, v_1] \prec \sigma^n \in K$.
Triangle Inequality.: $d(v_0, v_2) \leq d(v_0, v_1) + d(v_1, v_2)$, $\forall [v_0, v_1, v_2] \prec \sigma^n \in K$.

This allows us to embed each n-simplex locally into $\mathbb{R}^n$, and thereby compute all the necessary metric dependent quantities in our formulation. For example, the volume of a k-dual cell will be computed as the sum of the k-volumes of the dual cell restricted to each n-simplex in its local embedding into $\mathbb{R}^n$.

This notion of local metrics and local embeddings is consistent with the point of view that exterior calculus is a local theory with operators that operate on objects in the tangent and cotangent space of a fixed point. The issue of comparing objects in different tangent spaces is addressed in the discrete theory of connections on principal bundles in Leok et al. [2003].

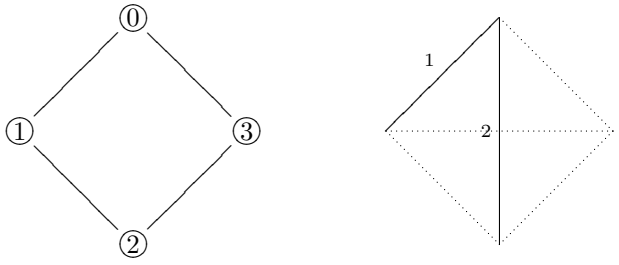
This also provides us with a criterion for evaluating a global embedding. The embedding should be such that the metric of the ambient space $\mathbb{R}^N$ restricted to the vertices of the complex, thought of as points in $\mathbb{R}^N$, agrees with the local metric imposed on the abstract simplicial complex. A global embedding that satisfies this condition will produce the same numerical results in discrete exterior calculus as that obtained using the local embedding method.

It is essential that the metric condition we impose is local, since the notion of distances between points in a manifold which are far away is not a well-defined concept, nor is it particularly useful for embeddings. As the simple example below illustrates, there may not exist any global embeddings into Euclidean space that satisfies a metric constraint imposed for all possible pairs of vertices.

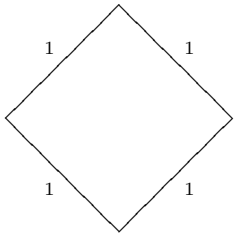
Example 4.1. *Consider a circle, with the distance between two points given by the minimal arc length. Consider a discretization given by 4 equidistant points on the circle, labelled $v_0, \ldots, v_3$, with the metric distances as follows,*

$$d(v_i, v_{i+1}) = 1, d(v_i, v_{i+2}) = 2,$$

where the indices are evaluated modulo 4, and this distance function is extended to a metric on all pairs of vertices by symmetry. It is easy to verify that this distance function is indeed a metric on vertices.



If we only use the local metric constraint, then we only require that adjacent vertices are separated by 1*, and the following is an embedding of the simplicial complex into $\mathbb{R}^2$,*



If, however, we use the metric defined on all possible pairs of vertices, by considering v_0, v_1, v_2, we have that $d(v_0, v_1) + d(v_1, v_2) = d(v_0, v_2)$. Since we are embedding these points into a Euclidean space, it follows that v_0, v_1, v_2 are collinear.

Similarly, by considering v_0, v_2, v_3, we conclude that they are collinear as well, and that v_1, v_3 are coincident, which contradicts $d(v_1, v_3) = 2$. Thus, we find that there does not exist a global embedding of the circle into Euclidean space if we require that the embedding is consistent with the metric on vertices defined for all possible pairs of vertices.

5. Differential Forms and Exterior Derivative

We will now define discrete differential forms. We will use some terms (which we will define) from algebraic topology, but it will become clear by looking at the examples that one can gain a clear and working notion of what a discrete form is without any algebraic topology. We start with a few definitions for which more details can be found on page 26 and 27 of Munkres [1984].

Definition 5.1. *Let K be a simplicial complex. We denote the free abelian group generated by a basis consisting of oriented k-simplices by $C_k(K;\mathbb{Z})$. This is the space of finite formal sums of the k-simplices, with coefficients in $\mathbb{Z}$. Elements of $C_k(K;\mathbb{Z})$ are called k-**chains**.*

Example 5.1. *Figure 3 shows examples of 1-chains and 2-chains.*

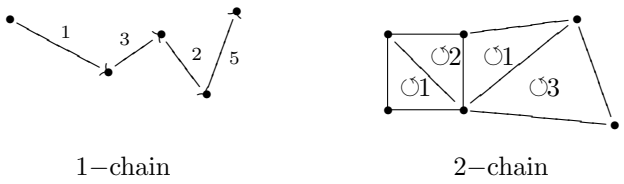


Figure 3: Examples of chains.

We view discrete k-forms as maps from the space of k-chains to $\mathbb{R}$. Recalling that the space of k-chains is a group, we require that these maps be homomorphisms into the additive group $\mathbb{R}$. Thus, discrete forms are what are called cochains in algebraic topology. We will define cochains below in the definition of forms but for more context and more details readers can refer to any algebraic topology text, for example, page 251 of Munkres [1984].

This point of view of forms as cochains is not new. The idea of defining forms as cochains appears, for example, in the works of Adams [1996], Dezin [1995], Hiptmair [1999], and Sen et al. [2000]. Our point of departure is that the other authors go on to develop a theory of discrete exterior calculus of forms only by introducing interpolation of forms, which we will be able to avoid. The formal definition of discrete forms follows.

Definition 5.2. *A **primal discrete** k-**form** α is a homomorphism from the chain group $C_k(K;\mathbb{Z})$ to the additive group $\mathbb{R}$. Thus, a discrete k-form is an element of $\mathrm{Hom}(C_k(K),\mathbb{R})$, the space of **cochains**. This space becomes an abelian group if we add two homomorphisms by adding their values in $\mathbb{R}$. The standard notation for $\mathrm{Hom}(C_k(K),\mathbb{R})$ in algebraic topology is $C^k(K;\mathbb{R})$. But we will often use the notation $\Omega_d^k(K)$ for this space as a reminder that this is the space of discrete (hence the d subscript) k-forms on the simplicial complex K. Thus,*

$$\Omega_d^k(K) := C^k(K;\mathbb{R}) = \mathrm{Hom}(C_k(K),\mathbb{R}).$$

Note that, by the above definition, given a k-chain $\sum_i a_i c_i^k$ (where $a_i \in \mathbb{Z}$) and a discrete k-form α, we have that

$$\alpha\left(\sum_i a_i c_i^k\right) = \sum_i a_i \alpha(c_i^k)\,,$$

and for two discrete k-forms $\alpha, \beta \in \Omega_d^k(K)$ and a k-chain $c \in C_k(K;\mathbb{Z})$,

$$(\alpha+\beta)(c) = \alpha(c) + \beta(c)\,.$$

In the usual exterior calculus on smooth manifolds integration of k-forms on a k-dimensional manifold is defined in terms of the familiar integration in $\mathbb{R}^k$. This is done roughly speaking by doing the integration in local coordinates, and showing that the value is independent of the choice of coordinates, due to the change of variables theorem in $\mathbb{R}^k$. For details on this, see the first few pages of Chapter 7 of Abraham et al. [1988]. We will not try to introduce the notion of integration of discrete forms on a simplicial complex. Instead the fundamental quantity that we will work with is the natural bilinear pairing of cochains and chains, defined by evaluation. More formally, we have the following definition.

Definition 5.3. *The **natural pairing** of a k-form α and a k-chain c is defined as the bilinear pairing*

$$\langle \alpha, c\rangle = \alpha(c).$$

As mentioned above, in discrete exterior calculus, this natural pairing plays the role that integration of forms on chains plays in the usual exterior calculus on smooth manifolds. The two are related by a procedure done at the time of discretization. Indeed, consider a simplicial triangulation K of a polyhedron in $\mathbb{R}^n$, i.e., consider a "flat" discrete manifold. If we are discretizing a continuous problem, we will have some smooth forms defined in the space $|K| \subset \mathbb{R}^n$. Consider such a smooth k-form α^k. In order to define the discrete form α_d^k corresponding to α^k, one would integrate α^k on all the k-simplices in K. Then, the evaluation of α_d^k on a k-simplex σ^k is defined by $\alpha_d^k(\sigma^k) := \int_{\sigma^k} \alpha^k$. Thus, discretization is the only place where integration plays a role in our discrete exterior calculus.

In the case of a non-flat manifold, the situation is somewhat complicated by the fact that the smooth manifold, and the simplicial complex, as geometric sets embedded in the ambient space do not coincide. A smooth differential form on the manifold can be discretized into the cochain representation by identifying the vertices of the simplicial complex with points on the manifold, and then using a local chart to identify k-simplices with k-volumes on the manifold.

There is the possibility of k-volumes overlapping even when their corresponding k-simplices do not intersect, and this introduces a discretization error that scales like the mesh size. One can alternatively construct geodesic boundary surfaces in an inductive fashion, which yields a partition of the manifold, but this can be computationally prohibitive to compute.

Now we can define the discrete exterior derivative which we will call $\mathbf{d}$, as in the usual exterior calculus. The discrete exterior derivative will be defined as the dual, with respect to the natural pairing defined above, of the boundary operator, which is defined below.

Definition 5.4. *The **boundary** operator $\partial_k : C_k(K;\mathbb{Z}) \to C_{k-1}(K;\mathbb{Z})$ is a homomorphism defined by its action on a simplex $\sigma^k = [v_0,\ldots,v_k]$,*

$$\partial_k \sigma^k = \partial_k([v_0,\ldots,v_k]) = \sum_{i=0}^{k} (-1)^i [v_0,\ldots,\hat{v}_i,\ldots,v_k]\,,$$

where $[v_0,\ldots,\hat{v}_i,\ldots,v_k]$ is the $(k-1)$-simplex obtained by omitting the vertex v_i. Note that $\partial_k \circ \partial_{k+1} = 0$.

Example 5.2. *Given an oriented triangle $[v_0,v_1,v_2]$ the boundary, by the above definition, is the chain $[v_1,v_2]-[v_0,v_2]+[v_0,v_1]$, which are the three boundary edges of the triangle.*

Definition 5.5. *On a simplicial complex of dimension n, a **chain complex** is a collection of chain groups and homomorphisms ∂_k, such that,*

$$0 \longrightarrow C_n(K) \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_{k+1}} C_k(K) \xrightarrow{\partial_k} \cdots \xrightarrow{\partial_1} C_0(K) \longrightarrow 0\,,$$

and $\partial_k \circ \partial_{k+1} = 0$.

Definition 5.6. *The **coboundary operator**, $\delta^k : C^k(K) \to C^{K+1}(K)$, is defined by duality to the boundary operator, with respect to the natural bilinear pairing between discrete forms and chains. Specifically, for a discrete form $\alpha^k \in \Omega_d^k(K)$, and a chain $c_{k+1} \in C_{k+1}(K;\mathbb{Z})$, we define δ^k by*

$$\left\langle \delta^k \alpha^k, c_{k+1} \right\rangle = \left\langle \alpha^k, \partial_{k+1} c_{k+1} \right\rangle . \tag{5.1}$$

That is to say

$$\delta^k(\alpha^k) = \alpha^k \circ \partial_{k+1}\,.$$

*This definition of the coboundary operator induces the **cochain complex**,*

$$0 \longleftarrow C^n(K) \xleftarrow{\delta^{n-1}} \cdots \xleftarrow{\delta^k} C^k(K) \xleftarrow{\delta^{k-1}} \cdots \xleftarrow{\delta^0} C^0(K) \longleftarrow 0\,,$$

where it is easy to see that $\delta^{k+1} \circ \delta^k = 0$.

Definition 5.7. *The **discrete exterior derivative** denoted by $\mathbf{d} : \Omega_d^k(K) \to \Omega_d^{k+1}(K)$ is defined to be the coboundary operator δ^k.*

Remark 5.1. *With the above definition of the exterior derivative, $\mathbf{d} : \Omega_d^k(K) \to \Omega_d^{k+1}(K)$, and the relationship between the natural pairing and integration, one can regard equation 5.1 as a discrete **generalized Stokes' theorem**. Thus, given a k-chain c, and a discrete k-form α, the discrete Stokes' theorem, which is true by definition, states that*

$$\langle \mathbf{d}\alpha, c \rangle = \langle \alpha, \partial c \rangle\ .$$

Furthermore, it also follows immediately that $\mathbf{d}^{k+1}\mathbf{d}^k = 0$.

Dual Discrete Forms. Everything we have said above in terms of simplices and the simplicial complex K can be said in terms of the cells that are duals of simplices and elements of the dual complex $\star K$. One just has to be a little more careful in the definition of the boundary operator, and the definition we construct below is well-defined on the dual cell complex. This gives us the notion of cochains of cells in the dual complex and these are the ***dual discrete forms***.

Definition 5.8. *The **dual boundary** operator, $\partial_k : C_k(\star K;\mathbb{Z}) \to C_{k-1}(\star K;\mathbb{Z})$, is a homomorphism defined by its action on a dual cell $\hat{\sigma}^k = \star\sigma^{n-k} = \star[v_0,\ldots,v_{n-k}]$,*

$$\begin{aligned}\partial\hat{\sigma}^k &= \partial \star [v_0, ..., v_{n-k}] \\ &= \sum_{\sigma^{n-k+1} \succ \sigma^{n-k}} \star\sigma^{n-k+1}\,,\end{aligned}$$

where σ^{n-k+1} is oriented so that it is consistent with the induced orientation on σ^{n-k}.

6. Hodge Star and Codifferential

In the exterior calculus for smooth manifolds, the Hodge star, denoted $*$, is an isomorphism between the space of k-forms and $(n-k)$-forms. The Hodge star is useful in defining the adjoint of the exterior derivative and this is adjoint is called the codifferential. The Hodge star, $* : \Omega^k(M) \to \Omega^{n-k}(M)$, is in the smooth case uniquely defined by the identity,

$$\langle\!\langle \alpha^k, \beta^k \rangle\!\rangle \mathbf{v} = \alpha^k \wedge *\beta^k\,,$$

where $\langle\!\langle\ ,\ \rangle\!\rangle$ is a metric on differential forms, and $\mathbf{v}$ is the volume-form. For a more in-depth discussion, see, for example, page 411 of Abraham et al. [1988].

The appearance of k and $(n-k)$ in the definition of Hodge star may be taken to be a hint that primal and dual meshes will play some role in the definition of a discrete Hodge star, since the dual of a k-simplex is an $(n-k)$-cell. Indeed, this is the case.

Definition 6.1. *The **discrete Hodge Star** is a map $* : \Omega_d^k(K) \to \Omega_d^{n-k}(\star K)$, defined by its action on simplices. For a k-simplex σ^k, and a discrete k-form α^k,*

$$\frac{1}{|\sigma^k|}\langle \alpha^k, \sigma^k \rangle = \frac{1}{|\star\, \sigma^k|}\langle *\alpha^k, \star\sigma^k \rangle.$$

The idea that the discrete Hodge star maps primal discrete forms to dual forms, and vice versa, is well-known. See, for example, Sen et al. [2000]. However, notice we now make use of the volume of these primal and dual meshes. But the definition we have given above does appear in the work of Hiptmair [2002].

The definition implies that the primal and dual *averages* must be equal. This idea has already been introduced, not in the context of exterior calculus, but in an attempt at defining discrete differential geometry operators, see Meyer et al. [2002].

Remark 6.1. *Although we have defined the discrete Hodge star above, we will show in Remark 12.1 of §12 that if an appropriate discrete wedge product and metric on discrete k-forms is defined, then the expression for the discrete Hodge star operator follows from the smooth definition.*

Lemma 6.1. *For a k-form α^k,*

$$** \alpha^k = (-1)^{k(n-k)}\alpha^k\,.$$

Proof. The proof is a simple calculation using the property that for a simplex or a cell σ^k, $\star\star(\sigma^k) = (-1)^{k(n-k)}\sigma^k$ (Equation 3.1). □

Definition 6.2. *Given a simplicial or a dual cell complex K the **discrete codifferential operator**, $\boldsymbol{\delta} : \Omega_d^{k+1}(K) \to \Omega_d^k(K)$, is defined by $\boldsymbol{\delta}(\Omega_d^0(K)) = 0$ and on discrete $(k+1)$-forms to be*

$$\boldsymbol{\delta}\beta = (-1)^{nk+1} * \mathbf{d} * \beta\,.$$

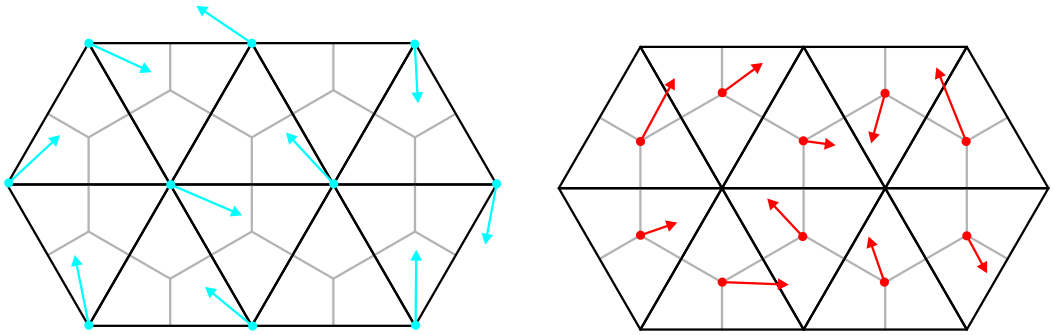
With the discrete forms, Hodge star, $\mathbf{d}$ and $\boldsymbol{\delta}$ defined so far, we already have enough to do an interesting calculation involving the Laplace–Beltrami operator. But, we will show this calculation in §9 after we have introduced discrete divergence operator.

7. Maps between 1-Forms and Vector Fields

Just as discrete forms come in two flavors, primal and dual (being linear functionals on primal chains or chains made up of dual cells), discrete vector fields also come in two flavors. Before formally defining primal and dual discrete vector fields, consider the examples illustrated in Figure 4. The distinction lies in the choice of basepoints, be they primal or dual vertices, to which we assign vectors.

Definition 7.1. *Let K be a flat simplicial complex, that is, the dimension of K is the same as that of the embedding space. A **primal discrete vector field** X on a flat simplicial complex K is a map from the zero-dimensional primal subcomplex $K^{(0)}$ (i.e., the primal vertices) to $\mathbb{R}^N$. We will denote the space of such vector fields by $\mathfrak{X}_d(K)$. The value of such a vector field is piecewise constant on the dual n-cells of $\star K$. Thus, we could just as well have called such vector fields dual and defined them as functions on the n-cells of $\star K$.*

Definition 7.2. *A **dual discrete vector field** X on a simplicial complex K is a map from the zero-dimensional dual subcomplex $(\star K)^{(0)}$ (i.e, the circumcenters of the primal n simplices) to $\mathbb{R}^N$ such that its value on each dual vertex is tangential to the corresponding primal n-simplex. We will denote the space of*



(a) Primal vector field (b) Dual vector field

Figure 4: Discrete vector fields.

such vector fields by $\mathfrak{X}_d(\star K)$. The value of such a vector field is piecewise constant on the n-simplices of K. Thus, we could just as well have called such vector fields primal and defined them as functions on the n-simplices of K.

Remark 7.1. *In this paper we have defined the primal vector fields only for flat meshes. We will address the issue of non-flat meshes in separate work.*

As in the smooth exterior calculus, we want to define the flat ($\flat$) and sharp ($\sharp$) operators that relate forms to vector fields. This allows one to write various vector calculus identities in terms of exterior calculus.

Definition 7.3. *Given a simplicial complex K of dimension n, the **discrete flat operator on a dual vector field**, $\flat : \mathfrak{X}_d(\star K) \to \Omega^d(K)$, is defined by its evaluation on a primal 1 simplex σ^1,*

$$\langle X^\flat, \sigma^1 \rangle = \sum_{\sigma^n \succ \sigma^1} \frac{|\star \sigma^1 \cap \sigma^n|}{|\star \sigma^1|} X \cdot \vec{\sigma}^1 \,,$$

where $X \cdot \vec{\sigma}^1$ is the usual dot product of vectors in $\mathbb{R}^N$, and $\vec{\sigma}^1$ stands for the vector corresponding to σ^1, and with the same orientation. The sum is over all σ^n containing the edge σ^1. The volume factors are in dimension n.

Definition 7.4. *Given a simplicial complex K of dimension n, the **discrete sharp operator on a primal 1-form**, $\sharp : \Omega^d(K) \to \mathfrak{X}_d(\star K)$, is defined by its evaluation on a given vertex v,*

$$\alpha^\sharp(v) = \sum_{[v,\sigma^0]} \langle \alpha, [v, \sigma^0] \rangle \sum_{\sigma^n \succ [v,\sigma^0]} \frac{|\star v \cap \sigma^n|}{|\sigma^n|} \hat{n}_{[v,\sigma^0]} \,,$$

where the outer sum is over all 1-simplices containing the vertex v, and the inner sum is over all n-simplices containing the 1-simplex $[v, \sigma^0]$. The volume factors are in dimension n, and the vector $\hat{n}_{[v,\sigma^0]}$ is the normal vector to the simplex $[v, \sigma^0]$, pointing into the n-simplex σ^n.

For a discussion of the proliferation of discrete sharp and flat operators that arise from considering the interpolation of differential forms and vector fields, please see Hirani [2003].

8. Wedge Product

As in the smooth case, the wedge product we will construct is a way to build higher degree forms from lower degree ones. For information about the smooth case, see the first few pages of Chapter 6 of Abraham et al. [1988].

Definition 8.1. *Given a primal discrete k-form $\alpha^k \in \Omega_d^k(K)$, and a primal discrete l-form $\beta^l \in \Omega_d^l(K)$, the* ***discrete primal-primal wedge product****, $\wedge : \Omega_d^k(K) \times \Omega_d^l(K) \to \Omega_d^{k+l}(K)$, defined by the evaluation on a $(k+l)$-simplex $\sigma^{k+l} = [v_0, \dots, v_{k+l}]$ is given by*

$$\langle \alpha^k \wedge \beta^l, \sigma^{k+l} \rangle = \frac{1}{(k+l)!} \sum_{\tau \in S_{k+l+1}} \operatorname{sign}(\tau) \frac{|\sigma^{k+l} \cap \star v_{\tau(k)}|}{|\sigma^{k+l}|} \alpha \smile \beta(\tau(\sigma^{k+l}))\,,$$

where S_{k+l+1} is the permutation group, and its elements are thought of as permutations of the numbers $0, \dots, k+l+1$. The notation $\tau(\sigma^{k+l})$ stands for the simplex $[v_{\tau(0)}, \dots, v_{\tau(k+l)}]$. Finally, the notation $\alpha \smile \beta(\tau(\sigma^{k+l}))$ is borrowed from algebraic topology (see, for example, page 206 of Hatcher [2001]) and is defined as

$$\alpha \smile \beta(\tau(\sigma^{k+l})) := \langle \alpha, [v_{\tau(0)}, \dots, v_{\tau(k)}] \rangle \langle \beta, [v_{\tau(k)}, \dots, v_{\tau(k+l)}] \rangle\,.$$

Example 8.1. *When we take the wedge product of two discrete 1-forms, we obtain terms in the sum that are graphically represented in Figure 5.*

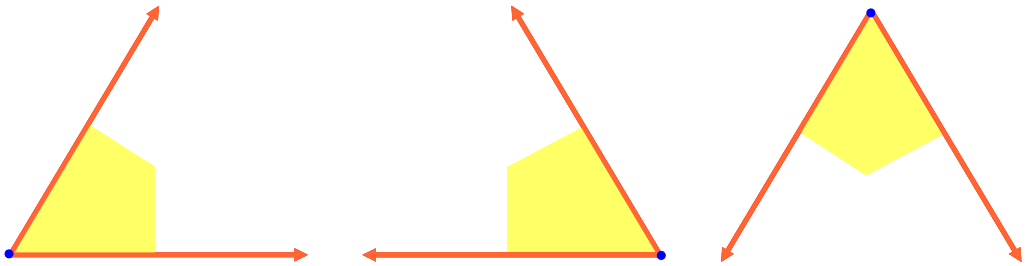


Figure 5: Terms in the wedge product of two discrete 1-forms.

Definition 8.2. *Given a dual discrete k-form $\hat{\alpha}^k \in \Omega_d^k(\star K)$, and a primal discrete l-form $\hat{\beta}^l \in \Omega_d^l(\star K)$, the* ***discrete dual-dual wedge product****, $\wedge : \Omega_d^k(\star K) \times \Omega_d^l(\star K) \to \Omega_d^{k+l}(\star K)$, defined by the evaluation on a $(k+l)$-cell $\hat{\sigma}^{k+l} = \star \sigma^{n-k-l}$, is given by*

$$\begin{aligned}
\langle \hat{\alpha}^k \wedge \hat{\beta}^l, \hat{\sigma}^{k+l} \rangle =& \langle \hat{\alpha}^k \wedge \hat{\beta}^l, \star \sigma^{n-k-l} \rangle \\
=& \sum_{\sigma^n \succ \sigma^{n-k-l}} \operatorname{sign}(\sigma^{n-k-l}, [v_{k+l}, \dots, v_n]) \sum_{\tau \in S_{k+l}} \operatorname{sign}(\tau) \\
& \cdot \langle \hat{\alpha}^k, \star [v_{\tau(0)}, \dots, v_{\tau(l-1)}, v_{k+l}, \dots, v_n] \rangle \langle \hat{\beta}^l, \star [v_{\tau(l)}, \dots, v_{\tau(k+l-1)}, v_{k+l}, \dots, v_n] \rangle
\end{aligned}$$

where $\sigma^n = [v_0, \dots, v_n]$, and, without loss of generality, assumed that $\sigma^{n-k-l} = \pm [v_{k+l}, \dots, v_n]$.

Anti-Commutativity of the Wedge Product.

Lemma 8.1. *The discrete wedge product, $\wedge : C^k(K) \times C^l(K) \to C^{k+l}(K)$, is anti-commutative, i.e.,*

$$\alpha^k \wedge \beta^k = (-1)^{kl} \beta^l \wedge \alpha^k\,.$$

Proof. We first rewrite the expression for the discrete wedge product using the following computation,

$$\sum_{\bar{\tau} \in S_{k+l+1}} \operatorname{sign}(\bar{\tau}) |\sigma^{k+l} \cap \star v_{\bar{\tau}(k)}| \langle \alpha^k, \bar{\tau}[v_0, \dots, v_k] \rangle \beta^l, \bar{\tau}[v_k, \dots, v_{k+l}] \rangle$$

$$
\begin{aligned}
&= \sum_{\bar{\tau}\in S_{k+l+1}} (-1)^{k-1}\operatorname{sign}(\bar{\tau})|\sigma^{k+l}\cap\star v_{\bar{\tau}(k)}|\ \langle\alpha^k,\bar{\tau}[v_1,\ldots,v_0,v_k]\rangle\langle\beta^l,\bar{\tau}[v_k,\ldots,v_{k+l}]\rangle\\
&= \sum_{\bar{\tau}\in S_{k+l+1}} (-1)^{k-1}\operatorname{sign}(\bar{\tau})|\sigma^{k+l}\cap\star v_{\bar{\tau}\rho(0)}|\langle\alpha^k,\bar{\tau}\rho[v_1,\ldots,v_k,v_0]\rangle\langle\beta^l,\bar{\tau}\rho[v_0,v_{k+1},\ldots,v_{k+l}]\rangle\\
&= \sum_{\bar{\tau}\in S_{k+l+1}} (-1)^{k-1}(-1)^{k}\operatorname{sign}(\bar{\tau})|\sigma^{k+l}\cap\star v_{\bar{\tau}\rho(0)}|\langle\alpha^k,\bar{\tau}\rho[v_0,\ldots,v_k]\rangle\langle\beta^l,\bar{\tau}\rho[v_0,v_{k+1},\ldots,v_{k+l}]\rangle\\
&= \sum_{\bar{\tau}\rho\in S_{k+l+1}\rho} (-1)^{k-1}(-1)^{k}(-1)\operatorname{sign}(\bar{\tau}\rho)|\sigma^{k+l}\cap\star v_{\bar{\tau}\rho(0)}|\\
&\qquad\qquad \cdot\langle\alpha^k,\bar{\tau}\rho[v_0,\ldots,v_k]\rangle\langle\beta^l,\bar{\tau}\rho[v_0,v_{k+1},\ldots,v_{k+l}]\rangle\\
&= \sum_{\tau\in S_{k+l+1}} \operatorname{sign}(\tau)|\sigma^{k+l}\cap\star v_{\tau(0)}|\langle\alpha^k,\tau[v_0,\ldots,v_k]\rangle\langle\beta^l,\tau[v_0,v_{k+1},\ldots,v_{k+l}]\rangle\,.
\end{aligned}
$$

Here, we used the elementary fact, from permutation group theory, that a $k+1$ cycle can be written as the product of k transpositions, which accounts for the $(-1)^k$ factors. Also, ρ is a transposition of 0 and k. Then, the discrete wedge product can be rewritten as

$$
\begin{aligned}
\langle\alpha^k\wedge\beta^l,\sigma^{k+l}\rangle = \frac{1}{(k+l)!}\sum_{\tau\in S_{k+l+1}}\operatorname{sign}(\tau)\frac{|\sigma^{k+l}\cap\star v_{\tau(0)}|}{|\sigma^{k+l}|}\\
\cdot\langle\alpha^k,[v_{\tau(0)},\ldots,v_{\tau(k)}]\rangle\langle\beta^l,[v_{\tau(0),\tau(k+1)},\ldots,v_{\tau(k+l)}]\rangle.
\end{aligned}
$$

For ease of notation, we denote $[v_0,\ldots,v_k]$ by σ^k, and $[v_0,v_{k+1},\ldots,v_{k+l}]$ by σ^l. Then, we have

$$
\langle\alpha^k\wedge\beta^l,\sigma^{k+l}\rangle = \frac{1}{(k+l)!}\sum_{\tau\in S_{k+l+1}}\operatorname{sign}(\tau)\frac{|\sigma^{k+l}\cap\star v_{\tau(0)}|}{|\sigma^{k+l}|}\langle\alpha^k,\tau(\sigma^k)\rangle\langle\beta^l,\tau(\sigma^l)\rangle.
$$

Furthermore, we denote $[v_0,v_{l+1}\ldots,v_{k+l}]$ by $\bar{\sigma}^k$, and $[v_0,v_1,\ldots,v_l]$ by $\bar{\sigma}^l$. Then,

$$
\langle\beta^l\wedge\alpha^k,\sigma^{k+l}\rangle = \frac{1}{(k+l)!}\sum_{\bar{\tau}\in S_{k+l+1}}\operatorname{sign}(\bar{\tau})\frac{|\sigma^{k+l}\cap\star v_{\bar{\tau}(0)}|}{|\sigma^{k+l}|}\langle\alpha^k,\bar{\tau}(\bar{\sigma}^k)\rangle\langle\beta^l,\bar{\tau}(\bar{\sigma}^l)\rangle.
$$

Consider the permutation $\theta\in S_{k+l+1}$, given by

$$
\theta = \begin{pmatrix} 0 & 1 & \ldots & k & k+1 & \ldots & k+l\\ 0 & l+1 & \ldots & k+l & 1 & \ldots & l\end{pmatrix},
$$

which has the property that

$$
\begin{aligned}
\bar{\sigma}^k &= \theta(\sigma^k),\\
\bar{\sigma}^l &= \theta(\sigma^l).
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
\langle\beta^l\wedge\alpha^k,\sigma^{k+l}\rangle &= \frac{1}{(k+l)!}\sum_{\bar{\tau}\in S_{k+l+1}}\operatorname{sign}(\bar{\tau})\frac{|\sigma^{k+l}\cap\star v_{\bar{\tau}(0)}|}{|\sigma^{k+l}|}\langle\alpha^k,\bar{\tau}(\bar{\sigma}^k)\rangle\langle\beta^l,\bar{\tau}(\bar{\sigma}^l)\rangle\\
&= \frac{1}{(k+l)!}\sum_{\bar{\tau}\in S_{k+l+1}}\operatorname{sign}(\bar{\tau})\frac{|\sigma^{k+l}\cap\star v_{\bar{\tau}\theta(0)}|}{|\sigma^{k+l}|}\langle\alpha^k,\bar{\tau}\theta(\sigma^k)\rangle\langle\beta^l\bar{\tau}\theta(\sigma^l)\rangle\\
&= \frac{1}{(k+l)!}\sum_{\bar{\tau}\theta\in S_{k+l+1}\theta}\operatorname{sign}(\bar{\tau}\theta)\operatorname{sign}(\theta)\frac{|\sigma^{k+l}\cap\star v_{\bar{\tau}\theta(0)}|}{|\sigma^{k+l}|}\langle\alpha^k,\bar{\tau}\theta(\sigma^k)\rangle\langle\beta^l,\bar{\tau}\theta(\sigma^l)\rangle.
\end{aligned}
$$

By making the substitution, $\tau = \bar{\tau}\theta$, and noting that $S_{k+l+1}\theta = S_{k+l+1}$, we obtain

$$\begin{aligned}\langle \beta^l \wedge \alpha^k, \sigma^{k+l}\rangle &= \operatorname{sign}(\theta)\frac{1}{(k+l)!}\sum_{\tau\in S_{k+l+1}} \operatorname{sign}(\tau)\frac{|\sigma^{k+l}\cap \star v_{\tau(0)}|}{|\sigma^{k+l}|}\langle \alpha^k, \tau(\sigma^k)\rangle\langle \beta^l, \tau(\sigma^l)\rangle \\ &= \operatorname{sign}(\theta)\langle \alpha^k \wedge \beta^l, \sigma^{k+l}\rangle\,.\end{aligned}$$

To obtain the desired result, we simply need to compute the sign of θ, which is given by

$$\operatorname{sign}(\theta) = (-1)^{kl}.$$

This follows from the observation that in order to move each of the last l vertices of σ^{k+l} forward, we require k transpositions with $v_1, \ldots, v_k$. Therefore, we obtain

$$\langle \beta^l \wedge \alpha^k, \sigma^{k+l}\rangle = \operatorname{sign}(\theta)\langle \alpha^k \wedge \beta^l, \sigma^{k+l}\rangle = (-1)^{kl}\langle \alpha^k \wedge \beta^l, \sigma^{k+l}\rangle,$$

and

$$\alpha^k \wedge \beta^l = (-1)^{kl}\beta^l \wedge \alpha^k. \qquad \square$$

Leibniz Rule for the Wedge Product.

Lemma 8.2. *The discrete wedge product satisfies the Leibniz rule,*

$$\mathbf{d}(\alpha^k \wedge \beta^l) = (\mathbf{d}\alpha^k) \wedge \beta^l + (-1)^k \alpha^k \wedge (\mathbf{d}\beta^l).$$

Proof. The proof of the Leibniz rule for discrete wedge products is directly analogous to the proof of the coboundary formula for the simplicial cup product on cochains, which can be found on page 206 of Hatcher [2001]. This is because the discrete exterior derivative is precisely the coboundary operator, and the wedge product is constructed out of weighted sums of cup products.

The cup product satisfies the Leibniz rule for an given partial ordering of the vertices, and the permutations in the signed sum in the discrete wedge product correspond to different choices of partial ordering. We then obtain the Leibniz rule for the discrete wedge product by applying it term-wise for each choice of permutation.

Consider

$$\begin{aligned}\langle (\mathbf{d}\alpha^k) \wedge \beta^l, \sigma^{k+l+1}\rangle = \sum_{i=0}^{k+1}(-1)^i \frac{1}{(k+l)!}\sum_{\tau\in S_{k+l+1}} \operatorname{sign}(\tau)\frac{|\sigma^{k+l}\cap \star v_{\tau(0)}|}{|\sigma^{k+l}|} \\ \cdot \langle \alpha^k, [v_{\tau(0)}, \ldots, \hat{v}_i, \ldots, v_{\tau(k+1)}]\rangle\langle \beta^l, [v_{\tau(k+1)}, \ldots, v_{\tau(k+l+1)}]\rangle,\end{aligned}$$

and

$$\begin{aligned}(-1)^k\langle \alpha^k \wedge (\mathbf{d}\beta^l), \sigma^{k+l+1}\rangle = (-1)^k \sum_{i=k}^{k+l+1}(-1)^{i-k} \frac{1}{(k+l)!}\sum_{\tau\in S_{k+l+1}} \operatorname{sign}(\tau)\frac{|\sigma^{k+l}\cap \star v_{\tau(0)}|}{|\sigma^{k+l}|} \\ \cdot \langle \alpha^k, [v_{\tau(0)}, \ldots, v_{\tau(k)}]\rangle\langle \beta^l, [v_{\tau(k)}, \ldots, \hat{v}_i, \ldots, v_{\tau(k+l+1)}]\rangle.\end{aligned}$$

The last set of terms, $i = k+1$, of the first expression cancels the first set of terms, $i = k$, of the second expression, and what remains is simply $\langle \alpha^k \wedge \beta^l, \partial\sigma^{k+l+1}\rangle$. Therefore, we can conclude that

$$\langle (\mathbf{d}\alpha^k) \wedge \beta^l, \sigma^{k+l+1}\rangle + (-1)^k\langle \alpha^k \wedge (\mathbf{d}\beta^l), \sigma^{k+l+1}\rangle = \langle \alpha^k \wedge \beta^l, \partial\sigma^{k+l+1}\rangle = \langle \mathbf{d}(\alpha^k \wedge \beta^l), \sigma^{k+l+1}\rangle,$$

or simply that the Leibniz rule for discrete differential forms holds,

$$\mathbf{d}(\alpha^k \wedge \beta^l) = (\mathbf{d}\alpha^k) \wedge \beta^l + (-1)^k \alpha^k \wedge (\mathbf{d}\beta^l). \qquad \square$$

Associativity for the Wedge Product. The discrete wedge product which we have introduced is not associative in general. This is a consequence of the fact that the stencil for the two possible triple wedge products are not the same. In the expression for $\langle \alpha^k \wedge (\beta^l \wedge \gamma^m), \sigma^{k+l+m}\rangle$, each term in the double summation consists of a geometric factor multiplied by $\langle \alpha^k, \sigma^k\rangle\langle \beta^l, \sigma^l\rangle\langle \gamma^m, \sigma^m\rangle$ for some k, l, m simplices $\sigma^k, \sigma^l, \sigma^m$.

Since β^l and γ^m are wedged together first, σ^l and σ^m will always share a common vertex, but σ^k could have a vertex in common with only σ^l, or only σ^m, or both. We can represent this in a graph, where the nodes denote the three simplices, which are connected by an edge if, and only if, they share a common vertex. The graphical representation of the terms which arise in the two possible triple wedge products are given in Figure 6.

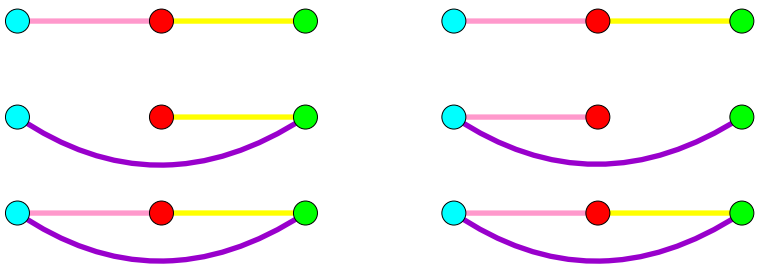


Figure 6: Stencils arising in the double summation for the two triple wedge products.

For the wedge product to be associative for all forms, the two stencils must agree. Since the stencils for the two possible triple wedge products differ, the wedge product is not associative in general. However, in the case of closed forms, we can rewrite the terms in the sum so that all the discrete forms are evaluated on triples of simplices that share a common vertex. This is illustrated graphically in Figure 7.

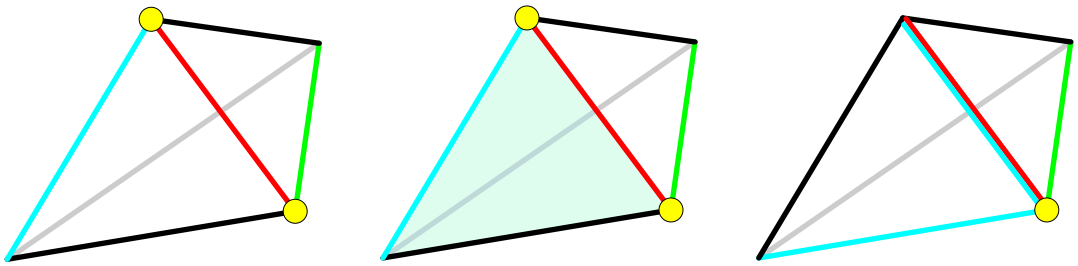


Figure 7: Associativity for closed forms.

This result is proved rigorously in the follow lemma.

Lemma 8.3. *The discrete wedge product is associative for closed forms. That is to say, for $\alpha^k \in C^k(K)$, $\beta^l \in C^l(K)$, $\gamma^m \in C^m(K)$, such that $\mathbf{d}\alpha^k = 0$, $\mathbf{d}\beta^l = 0$, $\mathbf{d}\gamma^m = 0$, we have that*

$$(\alpha^k \wedge \beta^l) \wedge \gamma^m = \alpha^k \wedge (\beta^l \wedge \gamma^m).$$

Proof.

$$\begin{aligned}
&\langle (\alpha^k \wedge \beta^l) \wedge \gamma^m, \sigma^{k+l+m} \rangle \\
&\quad = \sum_{\tau \in S_{k+l+m+1}} \operatorname{sign}(\tau) \langle \alpha^k \wedge \beta^l, \tau[v_0, \ldots, v_{k+l}] \rangle \langle \gamma^m, \tau[v_{k+l}, \ldots, v_{k+l+m}] \rangle \\
&\quad = \sum_{\tau \in S_{k+l+m+1}} \sum_{\rho \in S_{k+l+1}} \operatorname{sign}(\tau) \operatorname{sign}(\rho) \langle \alpha^k, \rho\tau[v_0, \ldots, v_k] \rangle \\
&\qquad\qquad \cdot \langle \beta^l, \rho\tau[v_k, \ldots, v_{k+l}] \rangle \langle \gamma^m, \tau[v_{k+l}, \ldots, v_{k+l+m}] \rangle
\end{aligned}$$

Here, either $\rho\tau(k) = \tau(k+l)$, in which case all three permuted simplices share $v_{\tau(k+l)}$ as a common vertex, or we need to rewrite either $\langle \alpha^k, \rho\tau[v_0, \ldots, v_k]\rangle$ or $\langle \beta^l, \rho\tau[v_k, \ldots, v_{k+l}]\rangle$, using the fact that α^k and β^l are closed forms.

If $v_{\tau(k+l)} \notin \rho\tau[v_0, \ldots, v_k]$, then we need to rewrite $\langle \alpha^k, \rho\tau[v_0, \ldots, v_k]\rangle$ by considering the simplex obtained by adding the vertex $v_{\tau(k+l)}$ to $\rho\tau[v_0, \ldots, v_k]$, which is $[v_{\tau(k+l)}, v_{\rho\tau(0)}, \ldots, v_{\rho\tau(k)}]$. Then, since α^k is closed, we have that

$$\begin{aligned} 0 &= \langle \mathbf{d}\alpha^k, [v_{\tau(k+l)}, v_{\rho\tau(0)}, \ldots, v_{\rho\tau(k)}]\rangle \\ &= \langle \alpha^k, \partial[v_{\tau(k+l)}, v_{\rho\tau(0)}, \ldots, v_{\rho\tau(k)}]\rangle \\ &= \langle \alpha^k, [v_{\rho\tau(0)}, \ldots, v_{\rho\tau(k)}]\rangle - \sum_{i=0}^{k} (-1)^i \langle \alpha^k, [v_{\tau(k+l)}, v_{\rho\tau(0)}, \ldots, \hat{v}_{\rho\tau(i)}, \ldots, v_{\rho\tau(k)}]\rangle \end{aligned}$$

or equivalently,

$$\langle \alpha^k, [v_{\rho\tau(0)}, \ldots, v_{\rho\tau(k)}]\rangle = \sum_{i=0}^{k} (-1)^i \langle \alpha^k, [v_{\tau(k+l)}, v_{\rho\tau(0)}, \ldots, \hat{v}_{\rho\tau(i)}, \ldots, v_{\rho\tau(k)}]\rangle.$$

Notice that all the simplices in the sum, with the exception of the last one, will share two vertices, $v_{\tau(k+l)}$ and $v_{\rho\tau(k)}$ with $\rho\tau[v_k, \ldots, v_{k+l}]$, and so their contribution in the triple wedge product will vanish due to the anti-symmetrized sum.

Similarly, if $v_{\tau(k+l)} \notin \rho\tau[v_k, \ldots, v_{k+l}]$, using the fact that β^l is closed yields

$$\langle \alpha^k, [v_{\rho\tau(k)}, \ldots, v_{\rho\tau(k+l)}]\rangle = \sum_{i=k}^{k+l} (-1)^{(i-k)} \langle \alpha^k, [v_{\tau(k+l)}, v_{\rho\tau(k)}, \ldots, \hat{v}_{\rho\tau(i)}, \ldots, v_{\rho\tau(k+l)}]\rangle.$$

As before, all the simplices in the sum, with the exception of the last one, will share two vertices, $v_{\tau(k+l)}$ and $v_{\rho\tau(k)}$ with $\rho\tau[v_0, \ldots, v_k]$, and so their contribution in the triple wedge product will vanish due to the anti-symmetrized sum.

This allows us to rewrite the triple wedge product in the case of closed forms as

$$\begin{aligned} \langle (\alpha^k \wedge \beta^l) \wedge \gamma^m, \sigma^{k+l+m}\rangle = \sum_{i=0}^{k+l+m} \sum_{\tau \in S_{k+l+m}} &\operatorname{sign}(\rho_i\tau) \langle \alpha^k, \rho_i\tau[v_0, \ldots, v_k]\rangle \langle \beta^l, \rho_i\tau[v_0, v_{k+1}, \ldots, v_{k+l}]\rangle \\ &\cdot \langle \gamma^m, \rho_i\tau[v_0, v_{k+l+1}, \ldots, v_{k+l+m}]\rangle, \end{aligned}$$

where $\tau \in S_{k+l+m}$ is thought of as acting on the set $\{1, \ldots, k+l+m\}$, and ρ_i is a transposition of 0 and i. A similar argument allows us to write $\alpha^k \wedge (\beta^l \wedge \gamma^m)$ in the same form, and therefore, the wedge product is associative for closed forms. □

Remark 8.1. *This lemma is significant, since if we think of a constant smooth differential form, and discretize it to obtain a discrete differential form, this discrete form will be closed. As such, this lemma states that in the infinitesimal limit, the discrete wedge product we have defined will be associative.*

In practice, if we have a mesh with characteristic length Δx, then we will have that

$$\frac{1}{|\sigma^{k+l+m}|} \langle \alpha^k \wedge (\beta^l \wedge \gamma^m) - (\alpha^k \wedge \beta^l) \wedge \gamma^m, \sigma^{k+l+m}\rangle = \mathcal{O}(\Delta x),$$

which is to say that the average of the associativity defect is of the order of the mesh size, and therefore vanishes in the infinitesimal limit.

9. Divergence and Laplace–Beltrami

In this section, we will illustrate the application of some of the DEC operations we have previously defined to the construction of new discrete operators such as the divergence and Laplace–Beltrami operators.

Divergence. The divergence of a vector field is given in terms of the Lie derivative of the volume-form, by the expression, $(\operatorname{div}(\mathbf{X})\mu = \pounds_{\mathbf{X}}\mu$. Physically, this corresponds to the net flow per unit volume of an infinitesimal volume about a point.

We will define the discrete divergence by using the formulas defining them in the smooth exterior calculus. The divergence definition will be valid for arbitrary dimensions. The resulting expressions involve operators that we have already defined and so we can actually perform some calculations to express these quantities in terms of geometric quantities. We will show that the resulting expression in terms of geometric quantities is the same as that derived by variational means in Tong et al. [2003].

Definition 9.1. *For a discrete dual vector field X the divergence* $\operatorname{div}(X)$ *is defined to be*

$$\operatorname{div}(\mathbf{X}) = -\boldsymbol{\delta} X^\flat \,.$$

Remark 9.1. *The above definition is a theorem in smooth exterior calculus. See, for example, page 458 of Abraham et al. [1988].*

As an example, we will now compute the divergence of a discrete dual vector field on a two-dimensional simplicial complex K, as illustrated in Figure 8.

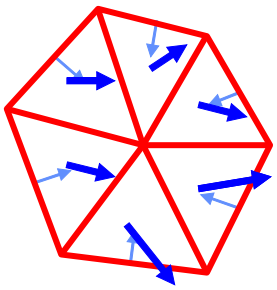


Figure 8: Divergence of a discrete dual vector field.

A similar derivation works in higher dimensions, where one needs to be mindful of the sign that arises from applying the Hodge star twice, $**\alpha^k = (-1)^{k(n-k)}\alpha^k$. Since $\operatorname{div}(X) = -\boldsymbol{\delta} X^\flat$, it follows that $\operatorname{div}(X) = *\mathbf{d} * X^\flat$. Since this is a primal 0-form it can be evaluated on a 0-simplex σ^0, and we have that

$$\langle \operatorname{div}(x), \sigma^0 \rangle = \langle *\mathbf{d} * X^\flat, \sigma^0 \rangle \,.$$

Using the definition of discrete Hodge star, and the discrete generalized Stokes' theorem, we get

$$\begin{aligned}
\frac{1}{|\sigma^0|}\langle \operatorname{div}(X), \sigma^0 \rangle &= \frac{1}{|\star\,\sigma^0|}\langle * * \mathbf{d} * X^\flat, \star\sigma^0 \rangle \\
&= \frac{1}{|\star\,\sigma^0|}\langle \mathbf{d} * X^\flat, \star\sigma^0 \rangle \\
&= \frac{1}{|\star\,\sigma^0|}\langle * X^\flat, \partial(\star\sigma^0) \rangle \,.
\end{aligned}$$

The second equality is obtained by applying the definition of the Hodge star, and the last equality is obtained by applying the discrete generalized Stokes' theorem. But,

$$\partial(\star\sigma^0) = \sum_{\sigma^1 \succ \sigma^0} \star\sigma^1 \,,$$

as given by the expression for the boundary of a dual cell in Equation 5.8. Thus,

$$\begin{aligned}
\frac{1}{|\sigma^0|}\langle \mathrm{div}(X), \sigma^0\rangle &= \frac{1}{|\star\,\sigma^0|}\langle *X^\flat, \sum_{\sigma^1 \succ \sigma^0} \star\sigma^1\rangle \\
&= \frac{1}{|\star\,\sigma^0|} \sum_{\sigma^1 \succ \sigma^0} \langle *X^\flat, \star\sigma^1\rangle \\
&= \frac{1}{|\star\,\sigma^0|} \sum_{\sigma^1 \succ \sigma^0} \frac{|\star\,\sigma^1|}{|\sigma^1|}\langle X^\flat, \sigma^1\rangle \\
&= \frac{1}{|\star\,\sigma^0|} \sum_{\sigma^1 \succ \sigma^0} \frac{|\star\,\sigma^1|}{|\sigma^1|} \sum_{\sigma^2 \succ \sigma^1} \frac{|\star\,\sigma^1 \cap \sigma^2|}{|\star\,\sigma^1|} X \cdot \vec{\sigma}^1 \\
&= \frac{1}{|\star\,\sigma^0|} \sum_{\sigma^1 \succ \sigma^0} \sum_{\sigma^2 \succ \sigma^1} \frac{|\star\,\sigma^1 \cap \sigma^2|}{|\sigma^1|} X \cdot \vec{\sigma}^1 \\
&= \frac{1}{|\star\,\sigma^0|} \sum_{\sigma^1 \succ \sigma^0} |\star\,\sigma^1 \cap \sigma^2| \, (X \cdot \frac{\vec{\sigma}^1}{|\sigma^1|}) \,.
\end{aligned}$$

This expression has the nice property that the divergence theorem holds on any dual n-chain, which, as a set, is a simply connected subset of $|K|$. Furthermore, the coefficients we computed for the discrete divergence operator are the unique ones for which a discrete divergence theorem holds.

Laplace–Beltrami. The Laplace–Beltrami operator is the generalization of the Laplacian to curved spaces. In the smooth case the Laplace–Beltrami operator on smooth functions is defined to be $\nabla^2 = \mathrm{div} \circ \mathrm{curl} = \boldsymbol{\delta} d$. See, for example, page 459 of Abraham et al. [1988]. Thus, in the smooth case, the Laplace–Beltrami on functions is a special case of the more general Laplace–deRham operator, $\Delta : \Omega^k(M) \to \Omega^k(M)$, defined by $\Delta = \mathbf{d}\boldsymbol{\delta} + \boldsymbol{\delta}\mathbf{d}$.

As an example, we compute Δf on a primal vertex σ^0, where $f \in \Omega^0_d(K)$, and K is a (not necessarily flat) triangle mesh in $\mathbb{R}^3$, as illustrated in Figure 9.

This calculation is done below.

$$\begin{aligned}
\frac{1}{|\sigma^0|}\langle \Delta f, \sigma^0\rangle &= \langle \boldsymbol{\delta}\mathbf{d} f, \sigma^0\rangle \\
&= -\langle *\mathbf{d} * \mathbf{d} f, \sigma^0\rangle \\
&= -\frac{1}{|\star\,\sigma^0|}\langle \mathbf{d} * \mathbf{d} f, \star\sigma^0\rangle \\
&= -\frac{1}{|\star\,\sigma^0|}\langle *\mathbf{d} f, \partial(\star\sigma^0)\rangle \\
&= -\frac{1}{|\star\,\sigma^0|}\langle *\mathbf{d} f, \sum_{\sigma^1 \succ \sigma^0} \star\sigma^1\rangle \\
&= -\frac{1}{|\star\,\sigma^0|} \sum_{\sigma^1 \succ \sigma^0} \langle *\mathbf{d} f, \star\sigma^1\rangle
\end{aligned}$$

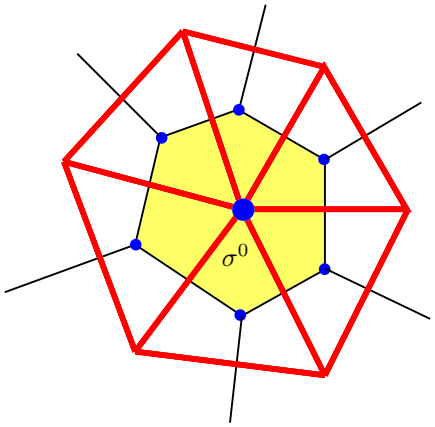


Figure 9: Laplace–Beltrami of a discrete function.

$$= -\frac{1}{|\star\sigma^0|}\sum_{\sigma^1\succ\sigma^0}\frac{|\star\sigma^1|}{|\sigma^1|}\langle \mathbf{d}f,\sigma^1\rangle$$
$$= -\frac{1}{|\star\sigma^0|}\sum_{\sigma^1\succ\sigma^0}\frac{|\star\sigma^1|}{|\sigma^1|}(f(v)-f(\sigma^0))\,,$$

where $\partial\sigma^1 = v - \sigma^0$. But, the above is the same as the formula involving cotangents found by Meyer et al. [2002] without using discrete exterior calculus.

Another interesting aspect, which will be discussed in §12, is that the characterization of harmonic functions as those functions which vanish when the Laplace–Beltrami operator is applied is equivalent to that obtained from a discrete variational principle using DEC as the means of discretizing the Lagrangian.

10. Contraction and Lie Derivative

In this section we will discuss some more operators that involve vector fields, namely contraction, and Lie derivatives.

For contraction, we will first define the usual smooth contraction algebraically, by relating it to Hodge star and wedge products. This yields one potential approach to defining discrete contraction. However, since in the discrete theory we are only concerned with integrals of forms, we can use the interesting notion of extrusion of a manifold by the flow of a vector field to define the integral of a contracted discrete differential form.

We learned about this definition of contraction via extrusion from Bossavit [2002b], who goes on to define discrete extrusion in his paper. Thus, he is able to obtain a definition of discrete contraction. Extrusion turns out to be a very nice way to define integrals of operators involving vector fields, and we will show how to define integrals of Lie derivatives via extrusion, which will yield discrete Lie derivatives.

Definition 10.1. *Given a manifold M, and S, a k-dimensional submanifold of M, and a vector field $X \in \mathfrak{X}(M)$, we call the manifold obtained by sweeping S along the flow of X for time t as the* ***extrusion*** *of S by X for time t, and denote it by $E_X^t(S)$. The manifold S carried by the* ***flow*** *for time t will be denoted $\varphi_X^t(S)$.*

Example 10.1. *Figure 10 illustrates the 2-simplex that arises from the extrusion of a 1-simplex by a discrete vector field that is interpolated using a linear shape function.*

Contraction (Extrusion). We first establish an integral property of the contraction operator.

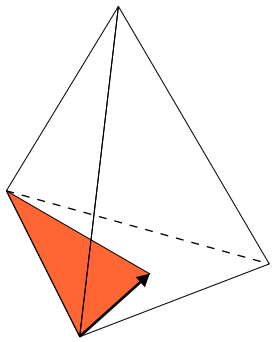


Figure 10: Extrusion of 1*-simplex by a discrete vector field.*

Lemma 10.1.

$$\int_S \mathbf{i}_X \beta = \left.\frac{d}{dt}\right|_{t=0} \int_{E_X^t(S)} \beta$$

Proof. Prove instead that

$$\int_0^t \left[\int_{S_\tau} \mathbf{i}_X \beta \right] d\tau = \int_{E_X^t(S)} \beta \,.$$

Then, by first fundamental theorem of calculus, the desired result will follow. To prove the above, simply take coordinates on S and carry them along with the flow and define the transversal coordinate to be the flow of X. This proof is sketched in Bossavit [2002b]. □

This lemma allows us to interpret contraction as being the dual, under the integration pairing between k-forms and k-volumes, to the geometric operation of extrusion. The discrete contraction operator is then given by

$$\langle \mathbf{i}_X \alpha^{k+1}, \sigma^k \rangle = \left.\frac{d}{dt}\right|_{t=0} \langle \alpha^{k+1}, E_X^t(\sigma^k) \rangle,$$

where the evaluation of the RHS will typically require that the discrete differential form and the discrete vector field are appropriately interpolated.

Remark 10.1. *Since the dynamic definition of the contraction operator only depends on the derivative of pairing of the differential form with the extruded region, it will only depend on the vector field in the region S, and not on its extension into the rest of the domain.*

In addition, if the interpolation for the discrete vector field satisfies a superposition principle, then the discrete contraction operator will satisfy a corresponding superposition principle.

Contraction (Algebraic). Contraction is an operator that allows one to combine vector fields and forms. For a smooth manifold M, the contraction of a vector field $X \in \mathfrak{X}(M)$ with a $(k+1)$-form $\alpha \in \Omega^{k+1}(M)$ is written as $\mathbf{i}_X \alpha$, and for vector fields $X_1, \ldots, X_k \in \mathfrak{X}(M)$, the contraction in smooth exterior calculus is defined by

$$\mathbf{i}_X \alpha(X_1, \ldots, X_k) = \alpha(X, X_1, \ldots, X_k) \,.$$

We define contraction by using an identity that is true in smooth exterior calculus. This identity originally appeared in Hirani [2003], and we state it here with proof.

Lemma 10.2 (Hirani [2003])**.** *Given a smooth manifold M of dimension n, a vector field $X \in \mathfrak{X}(M)$, and a k-form $\alpha \in \Omega^k(M)$, we have that*

$$\mathbf{i}_X \alpha = (-1)^{k(n-k)} * (* \alpha \wedge X^\flat) \,.$$

Proof. Recall that for a smooth function $f \in \Omega^0(M)$, we have that $\mathbf{i}_X \alpha = f \mathbf{i}_X \alpha$. This, and the multilinearity of α, implies that it is enough to show the result in terms of basis elements. In particular, let $\tau \in S_n$ be a permutation of the numbers $1, \dots n$, such that $\tau(1) < \ldots < \tau(k)$, and $\tau(k+1) < \ldots < \tau(n)$. Let $X = e_{\tau(j)}$, for some $j \in 1, \dots, n$. Then, we have to show that

$$\mathbf{i}_{e_{\tau(j)}} e^{\tau(1)} \wedge \ldots \wedge e^{\tau(k)} = (-1)^{k(n-k)} * \left(*(e^{\tau(1)} \wedge \ldots \wedge e^{\tau(k)}) \wedge e^{\tau(j)} \right).$$

It is easy to see that the LHS is 0 if $j > k$, and it is

$$(-1)^{j-1} (e^{\tau(1)} \wedge \ldots \wedge \widehat{e^{\tau(j)}} \ldots \wedge e^{\sigma(k)}),$$

otherwise, where $\widehat{e^{\tau(j)}}$ means that $e^{\tau(j)}$ is omitted from the wedge product. Now, on the RHS of Equation 10.2, we have that

$$*(e^{\tau(1)} \wedge \ldots \wedge e^{\tau(k)}) = \operatorname{sign}(\tau)(e^{\tau(k+1)} \wedge \ldots \wedge e^{\tau(n)}).$$

Thus, the RHS is equal to

$$(-1)^{k(n-k)} \operatorname{sign}(\tau) * (e^{\tau(k+1)} \wedge \ldots \wedge e^{\tau(n)} \wedge e^{\tau(j)}),$$

which is 0 as required if $j > k$. So, assume that $1 \leq j \leq k$. We need to compute

$$*(e^{\tau(k+1)} \wedge \ldots \wedge e^{\tau(n)} \wedge e^{\tau(j)}),$$

which is given by

$$s\, e^{\tau(1)} \wedge \ldots \widehat{e^{\tau(j)}} \wedge \ldots \wedge e^{\tau(k)},$$

where the sign $s = \pm 1$, such that the equation,

$$s\, e^{\tau(k+1)} \wedge \ldots \wedge e^{\tau(n)} \wedge e^{\tau(j)} \wedge e^{\tau(1)} \wedge \ldots \wedge \widehat{e^{\tau(j)}} \wedge \ldots \wedge e^{\tau(k)} = \mu,$$

holds for the standard volume-form, $\mu = e^1 \wedge \ldots \wedge e^n$. This implies that

$$s = (-1)^{j-1} (-1)^{k(n-k)} \operatorname{sign}(\tau).$$

Then, RHS = LHS as required. □

Since we have expressions for the discrete Hodge star ($*$), wedge product ($\wedge$), and flat ($\flat$), we have the necessary ingredients to use the algebraic expression proved in the above lemma to construct a discrete contraction operator.

One has to note, however, that the wedge product is only associative for closed forms, and as a consequence, the Leibniz rule for the resulting contraction operator will only hold for closed forms as well. This is, however, sufficient to establish that the Leibniz rule for the discrete contraction will hold in the limit as the mesh is refined.

Lie Derivative (Extrusion). As was the case with contraction, we will establish a integral identity that allows the Lie derivative to be interpreted as the dual of a geometric operation on a volume. This involves the flow of a volume by a vector field, and it is illustrated in the following example.

Example 10.2. *Figure 11 illustrates the flow of a* 1*-simplex by a discrete vector field interpolated using a linear shape function.*

Lemma 10.3.

$$\int_S \pounds_X \beta = \left. \frac{d}{dt} \right|_{t=0} \int_{\varphi_X^t(S)} \beta.$$

Proof.

$$F_t^* (\pounds_X \beta) = \frac{d}{dt} F_t^* \beta$$

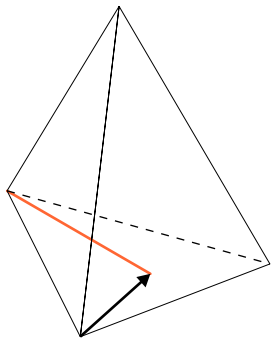


Figure 11: Flow of a 1-simplex by a discrete vector field.

$$\int_0^t F_\tau^*(\pounds_X \beta) d\tau = F_t^*\beta - \beta$$
$$\int_S \int_0^t F_\tau^*(\pounds_X \beta) d\tau = \int_S F_t^*\beta - \int_S \beta$$
$$\int_0^t \int_{\varphi_X^\tau(S)} \pounds_X \beta d\tau = \int_{\varphi_X^t(S)} \beta - \int_S \beta\,. \qquad \square$$

This lemma allows us to define a discrete Lie derivative as follows,

$$\langle \pounds_X \beta^k, \sigma^k \rangle = \left.\frac{d}{dt}\right|_{t=0} \langle \beta^k, \varphi_X^t(\sigma^k) \rangle\,,$$

where, as before, evaluating the RHS will require the discrete differential form and discrete vector field to be appropriately interpolated.

Lie Derivative (Algebraic). Alternatively, as we have expressions for the discrete contraction operator ($\mathbf{i}_X$), and exterior derivative ($\mathbf{d}$), we can construct a discrete Lie derivative using the Cartan magic formula,

$$\pounds_X \omega = \mathbf{i}_X \mathbf{d}\omega + \mathbf{d}\mathbf{i}_X \omega.$$

As is the case with the algebraic definition of the discrete contraction, the discrete Lie derivative will only satisfy a Leibniz rule for closed forms. As before, this is sufficient to establish that the Leibniz rule will hold in the limit as the mesh is refined.

11. Discrete Poincaré Lemma

In this section, we will prove the discrete Poincaré lemma by constructing a homotopy operator though a generalized cocone construction. This section is based on the work in Desbrun et al. [2003].

The standard cocone construction fails at the discrete level, since the cone of a simplex is not, in general, expressible as a chain in the simplicial complex. As such, the standard cocone does not necessarily map k-cochains to $(k-1)$-cochains.

An example of how the standard cone construction fails to map chains to chains is illustrated in Figure 12. Given the simplicial complex on the left, consisting of triangles, edges and nodes, we wish, in the center figure, to consider the cone of the bold edge with respect to the top most node. Clearly, the resulting cone in the right figure, which is shaded grey, cannot be expressed as a combination of the triangles in the original complex.

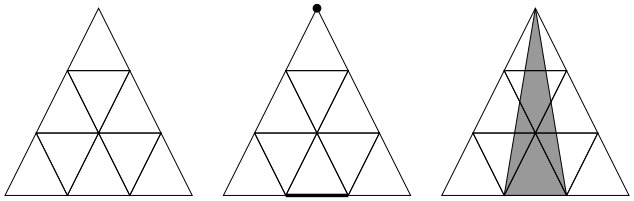


Figure 12: The cone of a simplex is, in general, not expressible as a chain.

In this subsection, a generalized cone operator that is valid for chains is developed which has the essential homotopy properties to yield a discrete analogue of the Poincaré lemma.

We will first consider the case of trivially star-shaped complexes, followed by logically star-shaped complexes, before generalizing the result to contractible complexes.

Definition 11.1. *Given a k-simplex $\sigma^k = [v_0, \ldots, v_k]$ we construct the **cone** with vertex w and base σ^k, as follows,*

$$w \diamond \sigma^k = [w, v_0, \ldots, v_k].$$

Lemma 11.1. *The geometric cone operator satisfies the following property,*

$$\partial(w \diamond \sigma^k) + w \diamond (\partial \sigma^k) = \sigma^k.$$

Proof. This is a standard result from simplicial algebraic topology. □

Trivially Star-Shaped Complexes.

Definition 11.2. *A complex K is called **trivially star-shaped** if there exists a vertex $w \in K^{(0)}$, such that for all $\sigma^k \in K$, the cone with vertex w and base σ^k is expressible as a chain in K. That is to say,*

$$\exists w \in K^{(0)} \mid \forall \sigma^k \in K, w \diamond \sigma^k \in C_{k+1}(K).$$

We can then denote the cone operation with respect to w as $p : C_k(K) \to C_{k+1}(K)$.

Lemma 11.2. *In trivially star-shaped complexes, the cone operator, $p : C_k(K) \to C_{k+1}(K)$, satisfies the following identity,*

$$p\partial + \partial p = I,$$

at the level of chains.

Proof. Follows immediately from the identity for cones, and noting that the cone is well-defined at the level of chains on trivially star-shaped complexes. □

Definition 11.3. *The **cocone** operator, $H : C^k(K) \to C^{k-1}(K)$, is defined by*

$$\langle H\alpha^k, \sigma^{k-1} \rangle = \langle \alpha^k, p(\sigma^{k-1}) \rangle.$$

This operator is well-defined on trivially star-shaped simplicial complexes.

Lemma 11.3. *The cocone operator, $H : C^k(K) \to C^{k-1}(K)$, satisfies the following identity,*

$$H\mathbf{d} + \mathbf{d}H = I,$$

at the level of cochains.

Proof. A simple duality argument applied to the cone identity,

$$p\partial + \partial p = I,$$

yields the following,

$$\langle \alpha^k, \sigma^k \rangle = \langle \alpha^k, (p\partial + \partial p)\sigma^k \rangle$$

$$
\begin{aligned}
&= \langle \alpha^k, p\partial\sigma^k\rangle + \langle \alpha^k, \partial p\sigma^k\rangle \\
&= \langle H\alpha^k, \partial\sigma^k\rangle + \langle \mathbf{d}\alpha^k, p\sigma^k\rangle \\
&= \langle (\mathbf{d}H\alpha^k, \sigma^k\rangle + \langle H\mathbf{d}\alpha^k, \sigma^k\rangle \\
&= \langle (\mathbf{d}H + H\mathbf{d})\alpha^k, \sigma^k\rangle.
\end{aligned}
$$

Therefore,

$$H\mathbf{d} + \mathbf{d}H = I,$$

at the level of cochains. □

Corollary 11.4 (Discrete Poincaré Lemma for Trivially Star-shaped Complexes)**.** *Given a closed cochain α^k, that is to say, $\mathbf{d}\alpha^k = 0$, there exists a cochain β^{k-1}, such that, $\mathbf{d}\beta^{k-1} = \alpha^k$.*

Proof. Applying the identity for cochains,

$$H\mathbf{d} + \mathbf{d}H = I,$$

we have,

$$\langle \alpha^k, \sigma^k\rangle = \langle (H\mathbf{d} + \mathbf{d}H)\alpha^k, \sigma^k\rangle,$$

but, $\mathbf{d}\alpha^k = 0$, so,

$$\langle \alpha^k, \sigma^k\rangle = \langle \mathbf{d}(H\alpha^k), \sigma^k\rangle.$$

Therefore, $\beta^{k-1} = H\alpha^k$ is such that $\mathbf{d}\beta^{k-1} = \alpha^k$ at the level of cochains. □

Example 11.1. *We demonstrate the construction of the tetrahedralization of the cone of a $(n-1)$-simplex over the origin.*

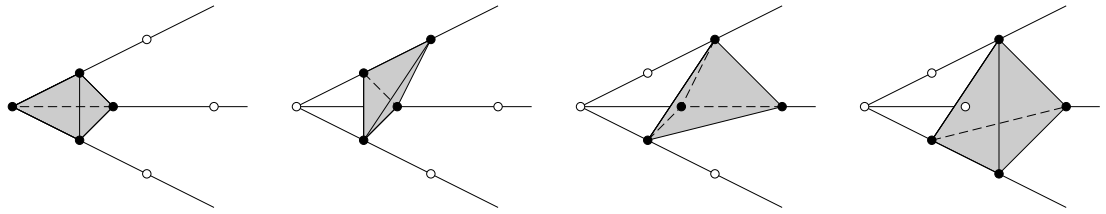
If we denote by v_i^k, the projection of the v_i vertex to the k-th concentric sphere, where the 0-th concentric sphere is simply the central point, then we fill up the cone $[c, v_1, ...v_n]$ with simplices as follows,

$$[v_1^0, v_1^1, \ldots, v_n^1], [v_1^2, v_1^1, \ldots, v_n^1], [v_1^2, v_2^2, v_2^1, \ldots, v_n^1], \ldots, [v_1^2, \ldots, v_n^2, v_n^1].$$

Since S^{n-1} is orientable, we can use a consistent triangulation of S^{n-1} and these n-cones to consistently triangulate B^n such that the resulting triangulation is star-shaped.

This fills up the region to the 1st concentric sphere, and we repeat the process by leapfrogging at the last vertex to add $[v_1^2, ..., v_n^2, v_n^3]$, and continuing the construction, to fill up the annulus between the 1st and 2nd concentric sphere. Thus, we can keep adding concentric shells to create an arbitrarily dense triangulation of a n-ball about the origin.

In three dimensions, these simplices are given by



$[c, v_1^1, v_2^1, v_3^1]$, $[v_1^2, v_1^1, v_2^1, v_3^1]$, $[v_1^2, v_2^2, v_2^1, v_3^1]$, $[v_1^2, v_2^2, v_3^2, v_3^1]$.

Putting them together, we obtain Figure 13.

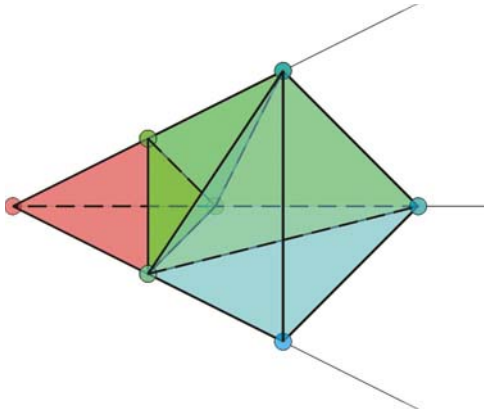


Figure 13: Triangulation of a three-dimensional cone.

This example is significant, since we have demonstrated that for any n-dimensional ball about a point, we can construct a trivially star-shaped triangulation of the ball, with arbitrarily high resolution. This allows us to recover the smooth Poincaré lemma in the limit of an infinitely fine mesh, using the discrete Poincaré lemma for trivially star-shaped complexes.

Logically Star-Shaped Complexes.

Definition 11.4. *A simplicial complex L is **logically star-shaped** if it is isomorphic, at the level of an abstract simplicial complex, to a trivially star-shaped complex K.*

Example 11.2. *We see two simplicial complexes, in Figure 14, which are clearly isomorphic as abstract simplicial complexes.*

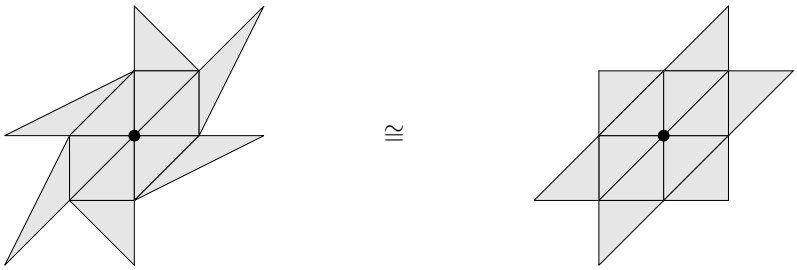


Figure 14: Trivially star-shaped complex (left); Logically star-shaped complex (right).

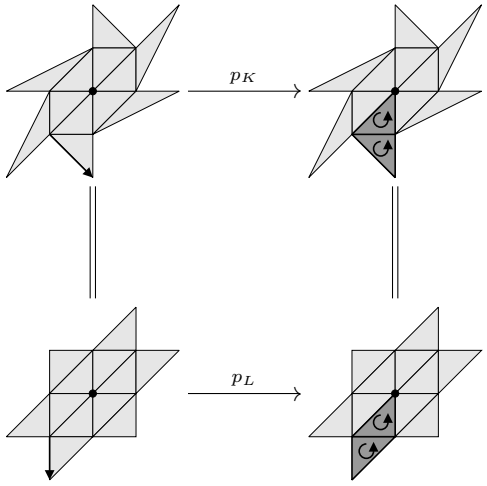
Definition 11.5. *The **logical cone** operator $p : C^k(L) \to C^{k+1}(L)$ is defined by making the following diagram commute,*

$$\begin{array}{ccc} C^k(K) & \xrightarrow{p_K} & C^{k+1}(K) \\ \| & & \| \\ C^k(L) & \xrightarrow{p_L} & C^{k+1}(L) \end{array}$$

Which is to say that, given the isomorphism $\varphi : K \to L$, we define

$$p_L = \varphi \circ p_K \circ \varphi^{-1}.$$

Example 11.3. *We show an example of the construction of the logical cone operator.*



This definition of the logical cone operator results in identities for the cone and cocone operator that follow from the trivially star-shaped case, and we record the results as follows.

Lemma 11.5. *In logically star-shaped complexes, the logical cone operator satisfies the following identity,*

$$p\partial + \partial p = I,$$

at the level of chains.

Proof. Follows immediately by pushing forward the result for trivially star-shaped complexes using the isomorphism. □

Lemma 11.6. *In logically star-shaped complexes, the logical cocone operator satisfies the following identity,*

$$H\mathbf{d} + \mathbf{d}H = I,$$

at the level of cochains.

Proof. Follows immediately by pushing forward the result for trivially star-shaped complexes using the isomorphism. □

Similarly, we have a Discrete Poincaré Lemma for logically star-shaped complexes.

Corollary 11.7 (Discrete Poincaré Lemma for Logically Star-shaped Complexes)**.** *Given a closed cochain α^k, that is to say, $\mathbf{d}\alpha^k = 0$, there exists a cochain β^{k-1}, such that, $\mathbf{d}\beta^{k-1} = \alpha^k$.*

Proof. Follows from the above lemma using the proof for the trivially star-shaped case. □

Contractible Complexes. For arbitrary contractible complexes, we construct a generalized cone operator such that it satisfies the identity,

$$p\partial + \partial p = I,$$

which is the crucial property of the cone operator, from the point of view of proving the discrete Poincaré lemma.

The trivial cone construction gives a clue as to how to proceed in the construction of a ***generalized cone operator***. Notice that if a σ^{k+1} is a term in $p(\sigma^k)$, then $p(\sigma^{k+1}) = \varnothing$. This suggests how we can use the cone identity to inductively construct the generalized cone operator.

To define $p(\sigma^k)$, we consider $\sigma^{k+1} \succ \sigma^k$, such that, σ^{k+1} and σ^k are consistently oriented. We apply $p\partial + \partial p$ to σ^{k+1}. Then, we have

$$\sigma^{k+1} = p(\sigma^k) + p(\partial\sigma^{k+1} - \sigma^k) + \partial p(\sigma^{k+1}).$$

If we set $p(\sigma^{k+1}) = \varnothing$,

$$\begin{aligned}\sigma^{k+1} &= p(\sigma^k) + p(\partial\sigma^{k+1} - \sigma^k) + \partial(\varnothing)\\ &= p(\sigma^k) + p(\partial\sigma^{k+1} - \sigma^k).\end{aligned}$$

Rearranging, we have

$$p(\sigma^k) = \sigma^{k+1} - p(\partial\sigma^{k+1} - \sigma^k),$$

and

$$p(\sigma^{k+1}) = \varnothing.$$

We are done, so long as the simplices in the chain $\partial\sigma^{k+1} - \sigma^k$ already have p defined on it. This then reduces to enumerating the simplices in such a way that in the right hand side of the equation, we never evoke terms that are undefined.

We now introduce a method of augmenting a complex so that the enumeration condition is always satisfied.

Definition 11.6. *Given a n-complex K, consider a $(n-1)$-chain c_{n-1} that is contained on the boundary of K, and is included in the one-ring of some vertex on ∂K. Then, the **one-ring cone augmentation** of K is the complex obtained by adding the n-cone $w \diamond c_{n-1}$, and all its faces to the complex.*

Definition 11.7. *A complex is **generalized star-shaped** if it can be constructed by repeatedly applying the one-ring augmentation procedure.*

We will explicitly show in Examples 11.4, and 11.7, how to enumerate the vertices in two and three dimensions. And in Examples 11.6, and 11.8, we will introduce regular triangulations of $\mathbb{R}^2$ and $\mathbb{R}^3$ that can be constructed by inductive one-ring cone augmentation.

Remark 11.1. *Notice that a non-contractible complex cannot be constructed by inductive one-ring cone augmentation, since it will involve adding a cone to a vertex that has two disjoint base chains. This prevents us from enumerating the simplices in such as way that all the terms in $\partial\sigma^{k+1} - \sigma^k$ have had p defined on them, and we see in Example 11.9 how this causes the cone identity, and hence the discrete Poincaré lemma to break.*

Example 11.4. *In two dimensions, the one-ring condition implies that the base of the cone consists of either one or two 1-simplices. To aid in visualization, consider Figure 15.*

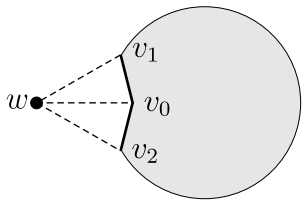


Figure 15: One-ring cone augmentation of a complex in two dimensions.

In the case of one 1-simplex, $[v_0, v_1]$, when we augment using the cone construction with the new vertex w, we define,

$$\begin{aligned} p([w]) &= [v_0, w] + p([v_0]), & p([v_0, w]) &= \varnothing,\\ p([v_1, w]) &= [v_0, v_1, w] - p([v_0, v_1]), & p([v_0, v_1, w]) &= \varnothing.\end{aligned}$$

In the case of two 1*-simplices,* $[v_0, v_1]$*,* $[v_0, v_2]$*, we have,*

$$\begin{aligned} p([w]) &= [v_0, w] + p([v_0]), & p([v_0, w]) &= \varnothing, \\ p([v_1, w]) &= [v_0, v_1, w] - p([v_0, v_1]), & p([v_0, v_1, w]) &= \varnothing, \\ p([v_2, w]) &= [v_0, v_2, w] - p([v_0, v_2]), & p([v_0, v_2, w]) &= \varnothing. \end{aligned}$$

Example 11.5. *We will now explicitly utilize the one-ring cone augmentation procedure to compute the generalized cone operator for part of a regular two-dimensional triangulation that is not logically star-shaped.*

As a preliminary, we shall consider a logically star-shaped complex, and augment with a new vertex, as seen in Figure 16.

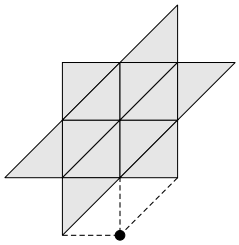
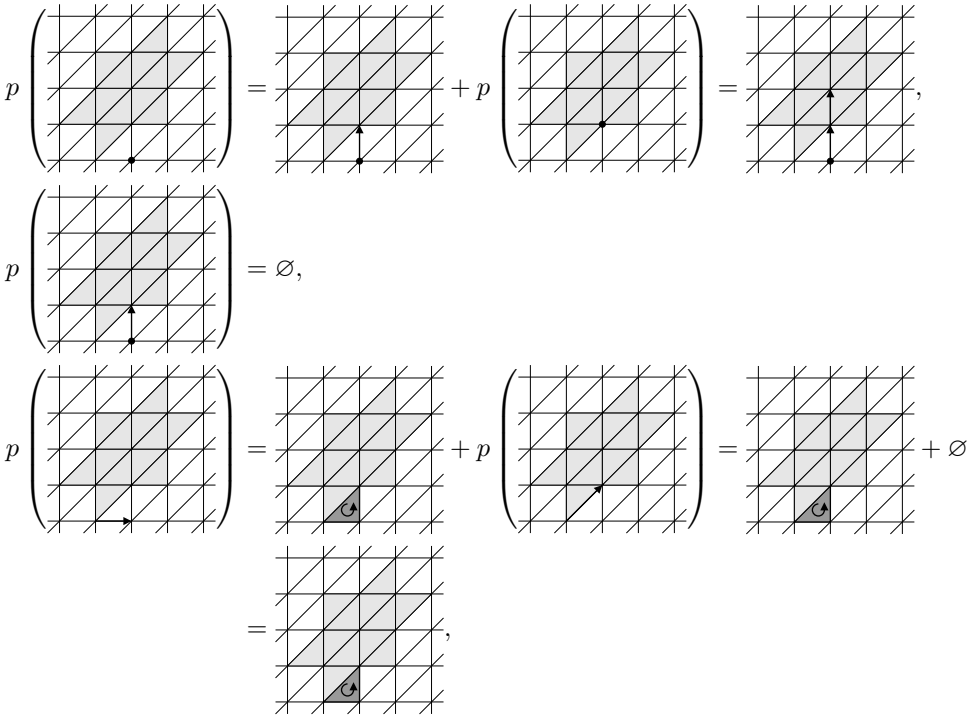
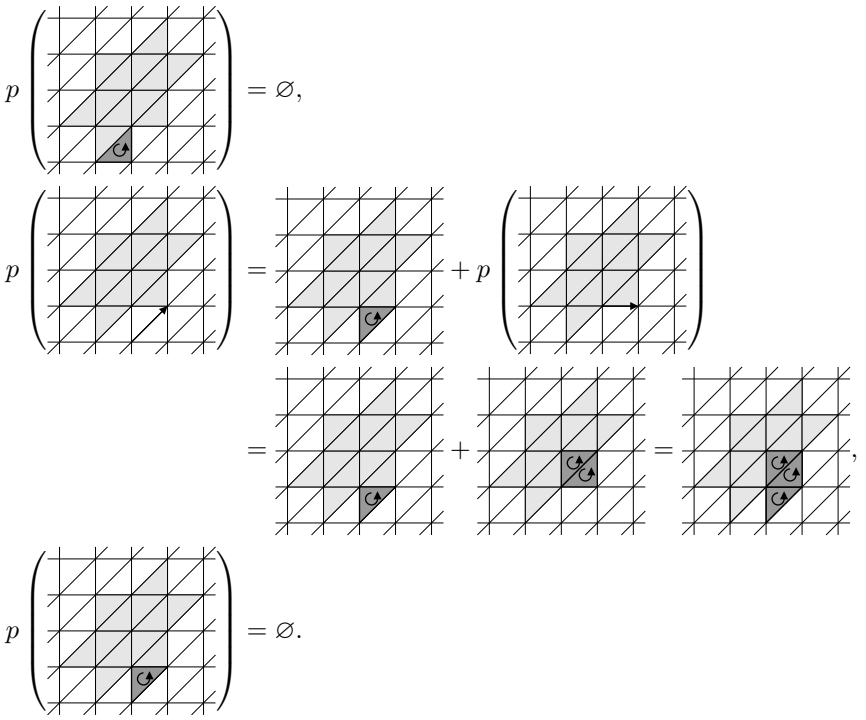


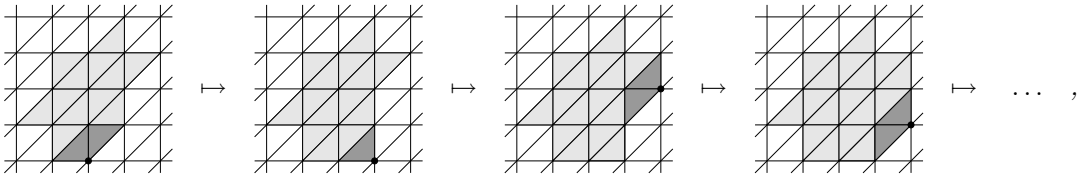
Figure 16: Logically star-shaped complex augmented by cone.

We use the logical cone operator for the subcomplex that is logically star-shaped, and the augmentation rules in the example above for the newly introduced simplices. This yields,





Example 11.6. *Clearly, the regular two-dimensional triangulation can be obtained by the successive application of the one-ring cone augmentation procedure, as the following sequence illustrates,*



which means that the discrete Poincaré lemma can be extended to the entire regular triangulation of the plane.

Example 11.7. *We consider the case of augmentation in three dimensions. Denote by v_0, the center of the one-ring on the two-surface, to which we are augmenting the new vertex w. The other vertices of the one-ring are enumerated in order, $v_1, \ldots, v_m$. To aid in visualization, consider Figure 17.*

If the one-ring does not go completely around v_0, we shall denote the missing term by $[v_0, v_1, v_m]$. The generalized cone operators are given as follows.

k=0,

$$p([w]) = [v_0, w] + p([v_0]), \qquad p([v_0, w]) = \varnothing,$$

k=1,

$$p([v_1, w]) = [v_0, v_1, w] - p([v_0, v_1]), \qquad p([v_0, v_1, w]) = \varnothing,$$
$$p([v_m, w]) = [v_0, v_m, w] - p([v_0, v_m]), \qquad p([v_0, v_m, w]) = \varnothing,$$

k=2,

$$p([v_1, v_2, w]) = [v_0, v_1, v_2, w] + p([v_0, v_1, v_2]), \qquad p([v_1, v_2, w]) = \varnothing,$$

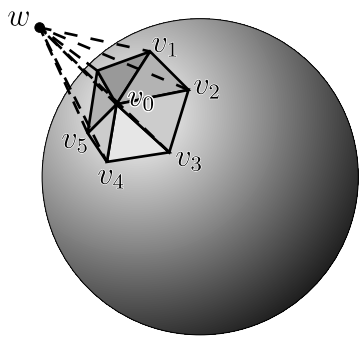


Figure 17: One-ring cone augmentation of a complex in three dimensions.

$$p([v_{m-1}, v_m, w]) = [v_0, v_{m-1}, v_m, w] + p([v_0, v_{m-1}, v_m]), \qquad p([v_{m-1}, v_m, w]) = \varnothing.$$

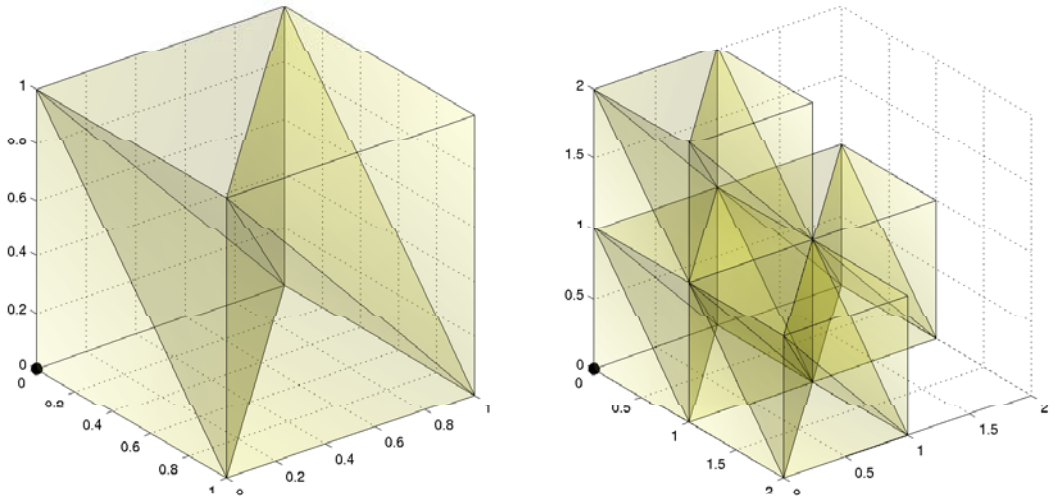
If it does go around completely,

$$p([v_m, v_1, w]) = [v_0, v_m, v_1, w] + p([v_0, v_m, v_1]), \qquad p([v_0, v_m, v_1, w]) = \varnothing.$$

Example 11.8. *We provide a tetrahedralization of the unit cube that can be tiled to yield a regular tetrahedralization of* $\mathbb{R}^3$*. The* 3*-simplices are as follows,*

$$[v_{000}, v_{001}, v_{010}, v_{10}], [v_{001}, v_{010}, v_{100}, v_{101}], [v_{001}, v_{010}, v_{011}, v_{101}],$$
$$[v_{010}, v_{100}, v_{101}, v_{110}], [v_{010}, v_{011}, v_{101}, v_{110}], [v_{011}, v_{101}, v_{110}, v_{111}].$$

The tetrahedralization of the unit cube can be seen in Figure 18.



(a) Tileable tetrahedralization of the unit cube

(b) Partial tiling of $\mathbb{R}^3$

Figure 18: Regular tiling of $\mathbb{R}^3$ *that admits a generalized cone operator.*

Since this regular tetrahedralization can be constructed by the successive application of the one-ring cone augmentation procedure, the Discrete Poincaré lemma can be extended to the entire regular tetrahedralization of $\mathbb{R}^3$.

In higher dimensions, we can extend the construction of the generalized cone operator inductively using the one-ring cone augmentation by choosing an appropriate enumeration of the base chain. Topologically, the base chain will be the cone of S^{n-2} (with possibly an open $(n-2)$-ball removed) with respect to the central point.

By spiraling around S^{n-2}, starting from around the boundary of the $n-2$ ball, and covering the rest of S^{n-2}, as in Figure 19, we obtain the higher-dimensional generalization of the procedure we have taken in Examples 11.4, and 11.7.

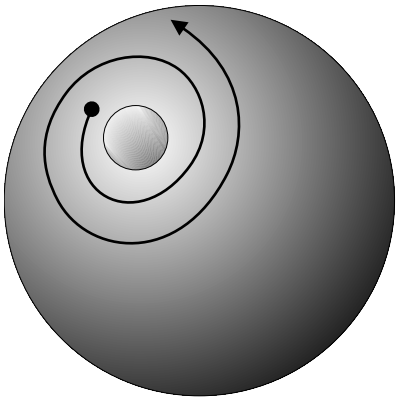


Figure 19: Spiral enumeration of S^{n-2}, $n = 4$.

Notice that $n = 2$ is distinguished, since $S^{2-2} = S^0$ is disjoint, which is why in the two-dimensional case, we were not able to use the spiraling technique to enumerate the simplices.

Since we have constructed the generalized cone operator such that the cone identity holds, we have,

Lemma 11.8. *In generalized star-shaped complexes, the generalized cone operator satisfies the following identity,*

$$p\partial + \partial p = I,$$

at the level of chains.

Proof. By construction. □

Lemma 11.9. *In generalized star-shaped complexes, the generalized cocone operator satisfies the following identity,*

$$H\mathbf{d} + \mathbf{d}H = I,$$

at the level of cochains.

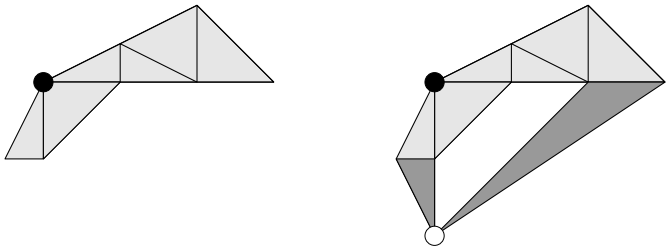
Proof. Follows immediately from applying the proof in the trivially star-shaped case, and using the identity in the previous lemma. □

Similarly, we have a discrete Poincaré lemma for generalized star-shaped complexes.

Corollary 11.10 (Discrete Poincaré Lemma for Generalized Star-shaped Complexes)**.** *Given a closed cochain* α^k*, that is to say,* $\mathbf{d}\alpha^k = 0$*, there exists a cochain* β^{k-1}*, such that,* $\mathbf{d}\beta^{k-1} = \alpha^k$.

Proof. Follows from the above lemma using the proof for the trivially star-shaped case. □

Example 11.9. *We will consider an example of how the Poincaré lemma fails in the case when the complex is not contractible. Consider the following trivially star-shaped complex, and augment by one vertex so as to make the region non-contractible, as show in Figure 20.*



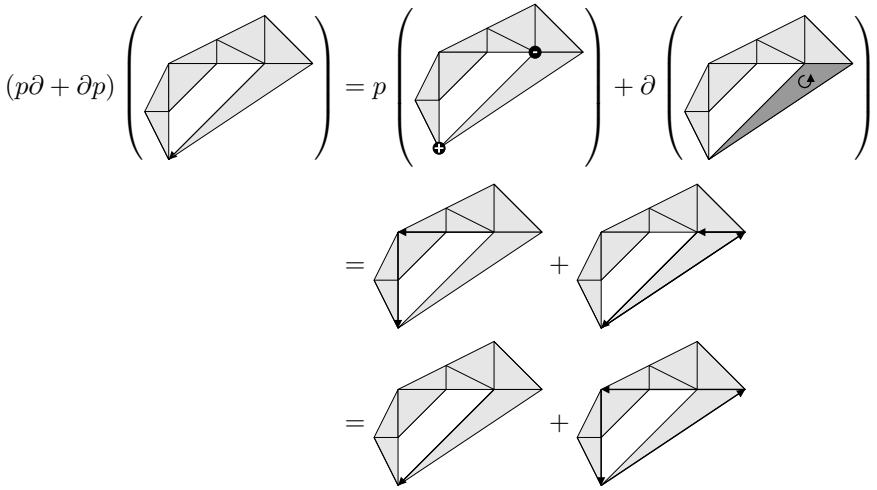
(a) Trivially star-shaped complex (b) Non-contractible complex

Figure 20: Counter-example for the discrete Poincaré lemma for a non-contractible complex.

Now we attempt to verify the identity,

$$p\partial + \partial p = I,$$

and we will see how this is only true up to a chain that is homotopic to the inner boundary.

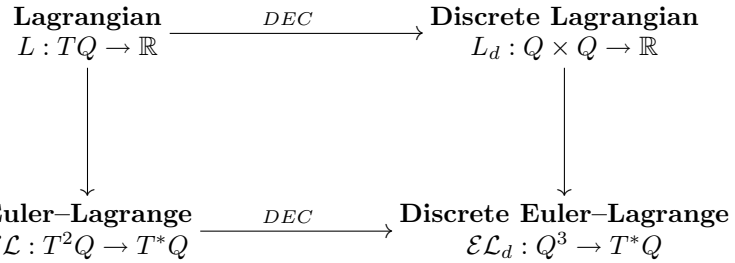


Since the second term cannot be expressed as the boundary of a 2*-chain, it will contribute a non-trivial effect, even on closed discrete forms, and therefore the discrete Poincaré lemma does not hold for non-contractible complexes, as expected.*

12. Discrete Variational Mechanics and DEC

We recall that discrete variational mechanics is based on a discrete analogue of Hamilton's principle, and they yield the discrete Euler–Lagrange equations. A particularly interesting property of DEC arises when it is used to construct the discrete Lagrangian for harmonic functions, and Maxwell's equations.

In particular, for these examples, the following diagram commutes,



Which is to say that directly discretizing the differential equations for harmonic functions, and Maxwell's equations using DEC results in the same expressions as the discrete Euler–Lagrange equations associated with a discrete Lagrangian which is discretized from the corresponding continuous Lagrangian by using DEC as the discretization scheme.

This is significant, since it implies that when DEC is used to discretize these equations, the corresponding numerical scheme which is obtained is variational, and consequently exhibits excellent structure-preserving properties.

In the variational principles for both harmonic functions and Maxwell's equations, we require the L^2 norm obtained from the L^2 inner product on $\Omega^k(M)$, which is given by

$$\langle \alpha^k, \beta^k \rangle = \int_M \alpha \wedge *\beta \,.$$

The discrete analogue of this requires a primal-dual wedge product, which is given below for forms of complementary dimension.

Definition 12.1. *Given a primal discrete k-form $\alpha^k \in \Omega_d^k(K)$, and a dual discrete $(n-k)$-form $\hat{\beta}^{n-k} \in \Omega_d^{n-k}(\star K)$, the **discrete primal-dual wedge product** is defined as follows,*

$$\begin{aligned}\langle \alpha^k \wedge \hat{\beta}^{n-k}, V_{\sigma^k} \rangle &= \frac{|V_{\sigma^k}|}{|\sigma^k||\star\,\sigma^k|} \langle \alpha^k, \sigma^k \rangle \langle \hat{\beta}^{n-k}, \star\sigma^k \rangle \\ &= \frac{1}{n} \langle \alpha^k, \sigma^k \rangle \langle \hat{\beta}^{n-k}, \star\sigma^k \rangle \,,\end{aligned}$$

where V_{σ^k} is the n-dimensional support volume obtained by taking the convex hull of the simplex σ^k and its dual cell $\star\sigma^k$.

The corresponding L^2 inner product is as follows.

Definition 12.2. *Given two primal discrete k-forms, $\alpha^k, \beta^k \in \Omega_d^k(K)$, their **discrete L^2 inner product**, $\langle \alpha^k, \beta^k \rangle_d$, is given by*

$$\begin{aligned}\langle \alpha^k, \beta^k \rangle_d &= \sum_{\sigma^k \in K} \frac{|V_{\sigma^k}|}{|\sigma^k||\star\,\sigma^k|} \langle \alpha^k, \sigma^k \rangle \langle *\beta, \star\sigma^k \rangle \\ &= \frac{1}{n} \sum_{\sigma^k \in K} \langle \alpha^k, \sigma^k \rangle \langle *\beta, \star\sigma^k \rangle \,.\end{aligned}$$

Remark 12.1. *Notice that it would have been quite natural from the smooth theory to propose the following metric tensor $\langle\!\langle \, , \, \rangle\!\rangle$ for differential forms,*

$$\langle \, \langle\!\langle \alpha^k, \beta^k \rangle\!\rangle \mathbf{v}, V_{\sigma^k} \rangle = |V_{\sigma^k}| \frac{\langle \alpha^k, \sigma^k \rangle}{|\sigma^k|} \frac{\langle \beta^k, \sigma^k \rangle}{|\sigma^k|} \,,$$

where the $|V_{\sigma^k}|$ is the factor arising from integrating the volume-form over V_{σ^k}, and

$$\frac{\langle \alpha^k, \sigma^k \rangle}{|\sigma^k|} \frac{\langle \beta^k, \sigma^k \rangle}{|\sigma^k|}$$

is what we would expect for $\langle\!\langle \alpha^k, \beta^k \rangle\!\rangle$, if the forms α^k and β^k were constant on σ^k, which is the product of the average values of α^k, and β^k.

If we adopt this as our definition of the metric tensor for forms, we can recover the definition we obtained in §6 for the Hodge star operator. Starting from the definition from the smooth theory,

$$\int \langle\!\langle \alpha^k, \beta^k \rangle\!\rangle \mathbf{v} = \int \alpha^k \wedge * \beta^k \,,$$

and expanding this in terms of the metric tensor for discrete forms, and the primal-dual wedge operator, we obtain

$$\langle \langle\!\langle \alpha^k, \beta^k \rangle\!\rangle \mathbf{v}, V_{\sigma^k} \rangle = \langle \alpha^k \wedge * \beta^k, V_{\sigma^k} \rangle \,,$$
$$|V_{\sigma^k}| \frac{\langle \alpha^k, \sigma^k \rangle}{|\sigma^k|} \frac{\langle \beta^k, \sigma^k \rangle}{|\sigma^k|} = \frac{|V_{\sigma^k}|}{|\sigma^k||\star \sigma^k|} \langle \alpha^k, \sigma^k \rangle \langle * \beta^k, \star \sigma^k \rangle \,.$$

When we eliminate common factors from both sides, we obtain the expression,

$$\frac{1}{|\sigma^k|} \langle \beta^k, \sigma^k \rangle = \frac{1}{|\star \sigma^k|} \langle * \beta^k, \star \sigma^k \rangle \,,$$

which is the expression we previously obtained in Definition 6.1 of §6.

The L^2 norm for discrete differential forms is given below.

Definition 12.3. *Given a primal discrete k-form $\alpha^k \in \Omega_d^k(K)$, its **discrete** L^2 **norm** is given by*

$$\begin{aligned} \|\alpha^k\|_d^2 &= \langle \alpha^k, \alpha^k \rangle_d \\ &= \frac{1}{n} \sum_{\sigma^k \in K} \langle \alpha^k, \sigma^k \rangle \langle * \alpha^k, \star \sigma^k \rangle \\ &= \frac{1}{n} \sum_{\sigma^k \in K} \frac{|\star \sigma^k|}{|\sigma^k|} \langle \alpha^k, \sigma^k \rangle^2 \,. \end{aligned}$$

Given these definitions, we can now reproduce some computations that were originally shown in Castrillón-López [2003].

Harmonic Functions. Harmonic functions $\phi : M \to \mathbb{R}$ can be characterized in a variational fashion as extremals of the following action functional,

$$\mathcal{S}(\phi) = \frac{1}{2} \int_M \|\mathbf{d}\phi\|^2 \mathbf{v},$$

where $\mathbf{v}$ is a Riemannian volume-form in M. The corresponding Euler–Lagrange equation is given by

$$* \mathbf{d} * \mathbf{d}\phi = -\Delta \phi = 0,$$

which is the familiar characterization of harmonic functions in terms of the Laplace–Beltrami operator.

The discrete action functional can be expressed in terms of the L^2 norm we introduced above for discrete forms,

$$\begin{aligned} \mathcal{S}_d(\phi) &= \frac{1}{2} \|\mathbf{d}\phi\|_d^2 \\ &= \frac{1}{2n} \sum_{\sigma^1 \in K} \frac{|\star \sigma^1|}{|\sigma^1|} \langle \mathbf{d}\phi, \sigma^1 \rangle^2. \end{aligned}$$

The basic variations needed for the determination of the discrete Euler–Lagrange operator are obtained from variations that vary the value of the function ϕ at a given vertex v_0, leaving the other values fixed. These variations have the form,

$$\phi_\varepsilon = \phi + \varepsilon\tilde{\eta},$$

where $\tilde{\eta} \in \Omega^0(M;\mathbb{R})$ is such that $\langle\tilde{\eta}, v_0\rangle = 1$, and $\langle\tilde{\eta}, v\rangle = 0$, for any $v \in K^{(0)} - \{v_0\}$. This family of variations is enough to establish the variational principle. That is, we have

$$\begin{aligned} 0 &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{S}_d(\phi_\varepsilon) \\ &= \frac{1}{n}\sum_{\sigma^1\in K} \frac{\left|\star\sigma^1\right|}{|\sigma^1|}\langle \mathbf{d}\phi, \sigma^1\rangle\langle \mathbf{d}\tilde{\eta}, \sigma^1\rangle \\ &= \frac{1}{n}\sum_{v_0\prec\sigma^1} \frac{\left|\star\sigma^1\right|}{|\sigma^1|}\langle \mathbf{d}\phi, \sigma^1\rangle \operatorname{sgn}(\sigma^1; v_0), \end{aligned} \tag{12.1}$$

where $\operatorname{sgn}(\sigma^1; v)$ stands for the sign of σ^1 with respect to v. Which is to say, $\operatorname{sgn}(\sigma^1; v) = 1$ if $\sigma^1 = [v', v]$, and $\operatorname{sgn}(\sigma^1; v) = -1$ if $\sigma^1 = [v, v']$. On the other hand,

$$\begin{aligned} \langle *\mathbf{d} * \mathbf{d}\phi, v_0\rangle &= \frac{1}{|\star v_0|}\langle \mathbf{d} * \mathbf{d}\phi, \star v_0\rangle \\ &= \frac{1}{|\star v_0|}\langle *\mathbf{d}\phi, \partial \star v_0\rangle \\ &= \frac{1}{|\star v_0|}\sum_{v_0\prec\sigma^1} \langle *\mathbf{d}\phi, \star\sigma^1\rangle \operatorname{sgn}(\sigma^1; v_0) \\ &= \frac{1}{|\star v_0|}\sum_{v_0\prec\sigma^1} \frac{|\star\sigma^1|}{|\sigma^1|}\langle \mathbf{d}\phi, \sigma^1\rangle \operatorname{sgn}(\sigma^1; v_0), \end{aligned}$$

where in the second to last equality, one has to note that the border of the dual cell of a vertex v_0 consists, up to orientation, in the dual of all the 1-simplices starting from v_0. This is illustrated in Figure 21, and follows from a general expression for the boundary of a dual cell that was given in Definition 5.8.

The sign factor comes from the relation between the orientation of the dual of the 1-simplices and that of $\partial * v_0$. From this, we conclude that the variational discrete equation, given in Equation 12.1, is equivalent to the vanishing of the discrete Laplace–Beltrami operator,

$$*\mathbf{d} * \mathbf{d}\phi = -\Delta = 0.$$

Maxwell Equations. We can formulate the Maxwell equations of electromagnetism in a covariant fashion by considering the 1-form A (the potential) as our fundamental variable in a Lorentzian manifold X. The action functional for a Lagrangian formulation of electromagnetism is given by,

$$\mathcal{S}(A) = \frac{1}{2}\int_X \|\mathbf{d}A\|^2 \mathbf{v}\,,$$

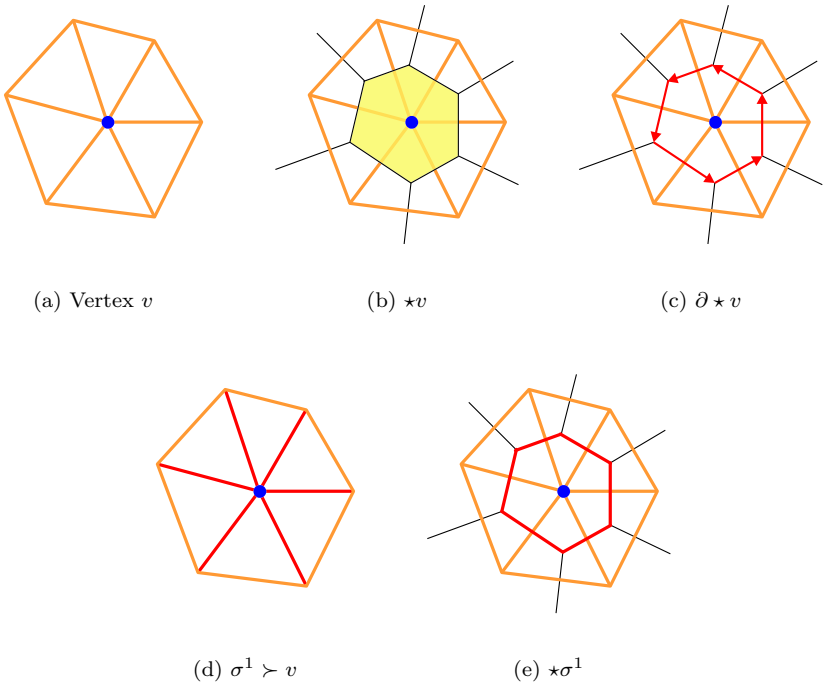
where $\|\cdot\|$ is the norm on forms induced by the Lorentzian metric on X, and $\mathbf{v}$ is the pseudo-Riemannian volume-form. The 1-form A is related to the 4-vector potential encountered in the relativistic formulation of electromagnetism (see, for example Jackson [1998]).

The Euler–Lagrange equation corresponding to this action functional is given by

$$*\mathbf{d} * \mathbf{d}A = 0\,.$$

In terms of the field strength, $F = \mathbf{d}A$, the last equation is usually rewritten as

$$\mathbf{d}F = 0\,, \qquad *\mathbf{d} * F = 0\,,$$



(a) Vertex v (b) $\star v$ (c) $\partial \star v$

(d) $\sigma^1 \succ v$ (e) $\star\sigma^1$

Figure 21: Boundary of a dual cell.

which is the geometric formulation of the Maxwell equations.

For the purposes of simplicity of exposition, we consider the special case where the Lorentzian manifold decomposes into $X = M \times \mathbb{R}$, where (M, g) is a compact Riemannian 3-manifold. In formulating the discrete version of this variational problem, we need to generalize the notion of a discrete Hodge dual to take into account the pseudo-Riemannian metric structure. This can be subtle in practice, and to overcome this, we consider a special family of complexes instead.

Let K' be a simplicial complex modelling M. For the sake of simplicity we consider $M = \mathbb{R}^3$ although this is not strictly necessary. We now consider a discretization $\{t_n\}_{n\in\mathbb{Z}}$ of $\mathbb{R}$. We define the complex K, modelling $X = \mathbb{R}^4$, the cells of which are the sets $\sigma = \sigma' \times \{t_n\} \subset \mathbb{R}^3 \times \mathbb{R}$, and $\sigma = \sigma' \times (t_n, t_{n+1}) \subset \mathbb{R}^3 \times \mathbb{R}$ for any $\sigma' \in K'$ and $n \in \mathbb{Z}$. Of course, this is not a simplicial complex but rather a "prismal" complex, as shown in Figure 22.

The advantage of these cell complexes is the existence of the Voronoi dual. More precisely, given any prismal cells $\sigma' \times \{t_n\} \in K$ and $\sigma' \times (t_n, t_{n+1}) \in K$, the Lorentz orthonormal to any of its edges coincide with the Euclidean one in $\mathbb{R}^4$ and the existence of the circumcenter is thus guaranteed. In other words, the Lorentz circumcentric dual $\star K$ to K is the same as the Euclidean one in $\mathbb{R}^4$.

Remark 12.2. *Much of the construction above can be carried out more generally by considering arbitrary cell complexes in* $\mathbb{R}^4$ *that are not necessarily prismal, as long as none of its* 1*-cells are lightlike. This causality condition is necessary to ensure that the circumcentric dual complex is well-behaved. However, it is sufficient for computational purposes that the complex is well-centered, in the sense that the Lorentzian circumcenter of each cell is contained inside the cell. These issues will be addressed in future work.*

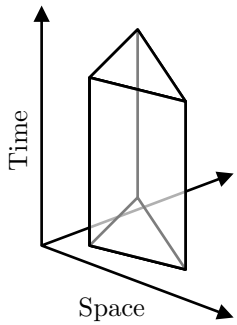


Figure 22: Prismal cell complex decomposition of space-time.

Recall that the Hodge star $*$ is uniquely defined by satisfying the following expression,

$$\alpha \wedge *\beta = \langle\!\langle \alpha, \beta \rangle\!\rangle \mathbf{v}\,,$$

for all $\alpha, \beta \in \Omega^k(X)$. The upshot of this is that the Hodge star operator depends on the metric, and since we have a pseudo-Riemannian metric, there is a sign that is introduced in our expression for the discrete Hodge star (Definition 6.1) that depends on whether the cell it is applied to is either spacelike or timelike. The discrete Hodge star for prismal complexes in Lorentzian space is given below.

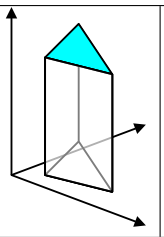
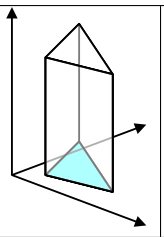
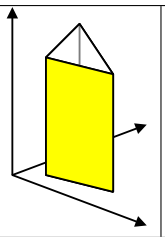
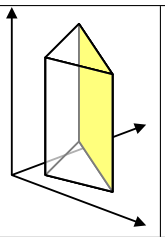
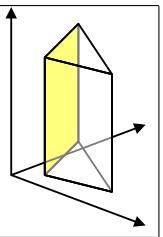
Definition 12.4. *The **discrete Hodge star for prismal complexes in Lorentzian space** is a map $* : \Omega_d^k(K) \to \Omega_d^k(*K)$ defined by giving its action on cells in a prismal complex as follows,*

$$\frac{1}{|\star\, \sigma^k|} \langle *\alpha^k, \star\sigma^k \rangle = \kappa(\sigma^k) \frac{1}{|\sigma^k|} \langle \alpha^k, \sigma^k \rangle,$$

*where $|\cdot|$ stands for the volume and the **causality sign** $\kappa(\sigma^k)$ is defined to be $+1$ if all the edges of σ^k are spacelike, and -1 otherwise.*

The causality sign of 2-cells in a $(2+1)$-space-time is summarized in Table 4. We should note that the causality sign for a 0-simplex, $\kappa(\sigma^0)$, is always 1. This is because a 0-simplex has no edges, and as such the statement that *all* of its edges are spacelike is trivially true.

TABLE 4. Causality sign of 2-cells in a $(2+1)$-space-time.

σ^2					
$\kappa(\sigma^2)$	$+1$	$+1$	-1	-1	-1

This causality term in the discrete Hodge star has consequences for the expression for the discrete norm (Definition 12.3), which is now given.

Definition 12.5. *Given a primal discrete k-form $\alpha^k \in \Omega^k_d(K)$, its* ***discrete*** L^2 ***Lorentzian norm*** *is given by,*

$$\begin{aligned}\|\alpha^k\|^2_{\mathrm{Lor},d} &= \frac{1}{n}\sum_{\sigma^k\in K}\langle\alpha^k,\sigma^k\rangle\langle{*}\alpha^k,{\star}\sigma^k\rangle \\ &= \frac{1}{n}\sum_{\sigma^k\in K}\kappa(\sigma^k)\frac{|\star\sigma^k|}{|\sigma^k|}\langle\alpha^k,\sigma^k\rangle^2\,.\end{aligned}$$

Having defined the discrete Lorentzian norm, we can express the discrete action as

$$\begin{aligned}\mathcal{S}_d(A) &= \frac{1}{2}\|\mathbf{d}A\|^2_{\mathrm{Lor},d} \\ &= \frac{1}{8}\sum_{\sigma^2\in K}\langle\mathbf{d}A,\sigma^2\rangle\langle{*}\mathbf{d}A,{\star}\sigma^2\rangle \\ &= \frac{1}{8}\sum_{\sigma^2\in K}\kappa(\sigma^2)\frac{|\star\sigma^2|}{|\sigma^2|}\langle\mathbf{d}A,\sigma^2\rangle^2\,.\end{aligned}$$

The basic variations needed to determine the discrete Euler–Lagrange operator are obtained from variations that vary the value of the 1-form A at a given 1-simplex σ^1_0, leaving the other values fixed. These variations have the form,

$$A_\varepsilon = A_\varepsilon + \varepsilon\tilde{\eta},$$

where $\tilde{\eta}\in\Omega^1_d(K)$ is given by $\langle\tilde{\eta},\sigma^1_0\rangle = 1$ for a fixed interior $\sigma^1_0\in K$ and $\langle\tilde{\eta},\sigma^1\rangle = 0$ for $\sigma^1\neq\sigma^1_0$. The derivation of the variational principle gives

$$\begin{aligned}\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\mathcal{S}_d(A_\varepsilon) &= \frac{1}{4}\sum_{\sigma^2\in K}\frac{|\star\sigma^2|}{|\sigma^2|}\kappa(\sigma^2)\langle\mathbf{d}A,\sigma^2\rangle\langle\mathbf{d}\tilde{\eta},\sigma^2\rangle \\ &= \frac{1}{4}\sum_{\sigma^1_0\prec\sigma^2}\frac{|\star\sigma^2|}{|\sigma^2|}\kappa(\sigma^2)\langle\mathbf{d}A,\sigma^2\rangle\langle\mathbf{d}\tilde{\eta},\sigma^2\rangle \\ &= \frac{1}{4}\sum_{\sigma^1_0\prec\sigma^2}\frac{|\star\sigma^2|}{|\sigma^2|}\kappa(\sigma^2)\langle\mathbf{d}A,\sigma^2\rangle\,\mathrm{sgn}(\sigma^2,\sigma^1_0),\end{aligned}$$

which vanishes for all the basic variations above. On the other hand, we now expand the discrete 1-form ${*}\mathbf{d}{*}\mathbf{d}A$. For any $\sigma^1_0\in K$, we have that

$$\begin{aligned}\langle{*}\mathbf{d}{*}\mathbf{d}A,\sigma^1_0\rangle &= \frac{|\sigma^1_0|}{|\star\sigma^1_0|}\kappa(\sigma^1_0)\langle\mathbf{d}{*}\mathbf{d}A,{\star}\sigma^1_0\rangle \\ &= \frac{|\sigma^1_0|}{|\star\sigma^1_0|}\kappa(\sigma^1_0)\langle{*}\mathbf{d}A,\partial{\star}\sigma^1_0\rangle \\ &= \frac{|\sigma^1_0|}{|\star\sigma^1_0|}\kappa(\sigma^1_0)\sum_{\sigma^1_0\prec\sigma^2}\mathrm{sgn}(\sigma^2,\sigma^1_0)\langle{*}\mathbf{d}A,{\star}\sigma^2\rangle \\ &= \frac{|\sigma^1_0|}{|\star\sigma^1_0|}\kappa(\sigma^1_0)\sum_{\sigma^1_0\prec\sigma^2}\frac{|\star\sigma^2|}{|\sigma^2|}\kappa(\sigma^2)\langle\mathbf{d}A,\sigma^2\rangle\,\mathrm{sgn}(\sigma^2,\sigma^1_0),\end{aligned}$$

where the sign $\mathrm{sgn}(\sigma^2,\sigma^1)$ stands for the relative orientation between σ^2 and σ^1. Which is to say, $\mathrm{sgn}(\sigma^2,\sigma^1) = 1$ if the orientation induced by σ^2 on σ^1 coincides with the orientation of σ^1, and $\mathrm{sgn}(\sigma^2,\sigma^1) = -1$ otherwise. For the second to last equality, one has to note that the border of the dual cell of an edge σ^1_0 consists, conveniently oriented with the sgn operator, of the union of the duals of all the 2-simplices containing σ^1_0.

This statement is the content of Definition 5.8, which gives the expression for the boundary of a dual cell, and was illustrated in Figure 21 for the case of n-dimensional dual cells.

By comparing the two computations, we find that for an arbitrary choice of $\sigma_0^1 \in K$, $\langle *\mathbf{d} * \mathbf{d}A, \sigma_0^1 \rangle$ is equal (up to a non-zero constant) to $\delta \mathcal{S}_d(A)$, which always vanishes. It follows that the variational discrete equations obtained above is equivalent to the discrete Maxwell equations,

$$*\mathbf{d} * \mathbf{d}A = 0.$$

13. Extensions to Dynamic Problems

It is desirable to leverage the exactness properties of the operators of discrete exterior calculus to construct numerical algorithms with discrete conservation properties. For these purposes, it is appropriate to extend the scope of DEC to incorporate dynamical behavior, by addressing the issue of discrete diffeomorphisms and flows.

As discussed in the previous section, DEC and discrete mechanics have interesting synergistic properties, and in this section we will explore a groupoid interpretation of discrete mechanics that is particularly appropriate to formulating the notion of pull-back and push-forward of discrete differential forms.

13.1. **Groupoid Interpretation of Discrete Variational Mechanics.** The groupoid formulation of discrete mechanics is particularly fruitful and natural, and it serves as a unifying tool for understanding the variational formulation of discrete Lagrangian mechanics, and discrete Euler–Poincaré reduction, as discussed in the work of Weinstein [1996] and Marsden et al. [1999, 2000].

The groupoid interpretation of discrete mechanics is most clearly illustrated if we consider the discretization of trajectories on TQ in two stages. Given a curve $\gamma : \mathbb{R}^+ \to TQ$, we consider a discrete sampling given by

$$g_i = \gamma(ih) \in TQ.$$

We then approximate TQ by $Q \times Q$, and associate to g_i two elements in Q. We denote this by

$$g_i \mapsto (q_i^0, q_i^1).$$

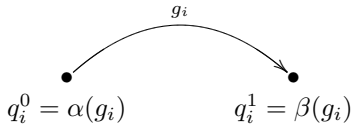
Or equivalently, in the language of groupoids, see Cannas da Silva and Weinstein [1999], Weinstein [2001], we have

$$\begin{array}{ccc} & G & \\ {\scriptstyle \alpha}\swarrow & & \searrow{\scriptstyle \beta} \\ Q & & \end{array}$$

where α is the ***source*** map, and β is the ***target*** map. Then,

$$g_i \mapsto (\alpha(g_i), \beta(g_i)) = (q_i^0, q_i^1).$$

This can be visualized as



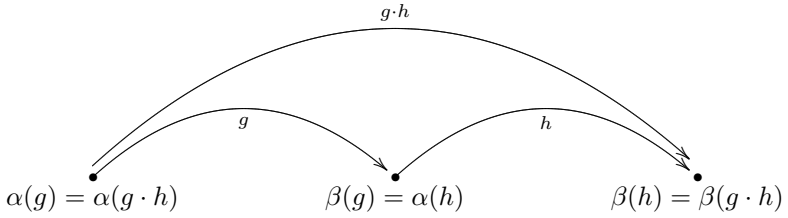
A product $\cdot : G^{(2)} \to G$ is defined on the set of composable pairs,

$$G^{(2)} := \{(g, h) \in G \times G \mid \beta(g) = \alpha(h)\}.$$

The ***groupoid composition*** $g \cdot h$ is defined by

$$\alpha(g \cdot h) = \alpha(g),$$
$$\beta(g \cdot h) = \beta(h).$$

This can be represented graphically as follows,



The set of composable pairs is the discrete analogue of the set of second-order curves on TQ. A curve $\gamma : \mathbb{R}^+ \to TQ$ is said to be second-order if there exists a curve $q : \mathbb{R}^+ \to Q$, such that,

$$\gamma(t) = (q(t), \dot{q}(t)).$$

The corresponding condition for discrete curves is that given a sequence of points in $Q \times Q$, $(q_1^0, q_1^1), \ldots, (q_p^0, q_p^1)$, we require that

$$q_i^1 = q_{i+1}^0.$$

This implies that the discrete curve on $Q \times Q$ is derived from a $(p+1)$-pointed curve $(q_0, \ldots, q_p)$ on Q, where

$$q_i = \begin{cases} q_{i+1}^0, & \text{if } 0 \leq i < p; \\ q_i^1, & \text{if } i = p. \end{cases}$$

This condition has a direct equivalent in groupoids,

$$\beta(g_i) = q_i^1 = q_{i+1}^0 = \alpha(g_{i+1}).$$

Which is to say that the sequence of points in $Q \times Q$ are composable. In general, this hierarchy of sets is denoted by

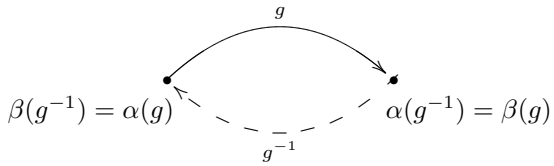
$$G^{(p)} := \{(g_1, \ldots, g_p) \in G^p \mid \beta(g_i) = \alpha(g_{i+1})\},$$

where $G^{(0)} \simeq Q$.

In addition, the ***groupoid inverse*** is defined by the following,

$$\alpha(g^{-1}) = \beta(g),$$
$$\beta(g^{-1}) = \alpha(g).$$

This is represented as follows,



Visualizing Groupoids. In summary, composition of groupoid elements, and the inverse of groupoid elements can be illustrated by Figure 23. As we will see in the next subsection, representing discrete diffeomorphisms as pair groupoids is the natural method of ensuring that the mesh remains nondegenerate.

13.2. **Discrete Diffeomorphisms and Discrete Flows.** We will adopt the point of view of representing a discrete diffeomorphism as a groupoid, which was first introduced in Pekarsky and West [2003], and appropriately modify it to reflect the simplicial nature of our mesh. In addition, we will address the induced action of a discrete diffeomorphism on the dual mesh.

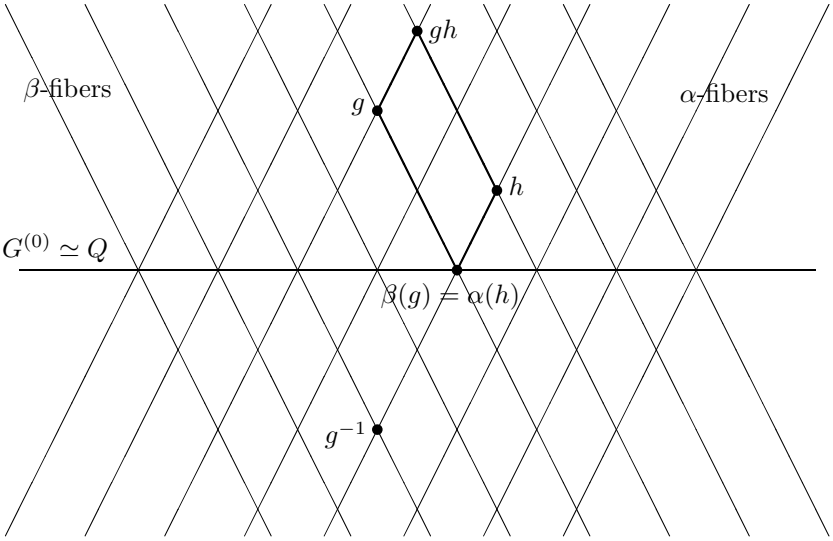


Figure 23: Groupoid composition and inverses.

Definition 13.1. *Given a complex K embedded in V, and its corresponding abstract simplicial complex M, a **discrete diffeomorphism**, $\varphi \in \mathrm{Diff}_d(M)$, is a pair of simplicial complexes K_1, K_2, which are realizations of M in the ambient space V. This is denoted by $\varphi(M) = (K_1, K_2)$.*

Definition 13.2. *A **one-parameter family of discrete diffeomorphisms** is a map $\varphi : I \to \mathrm{Diff}_d(M)$, such that,*

$$\pi_1(\varphi(t)) = \pi_1(\varphi(s)), \qquad \forall s, t \in I.$$

Since we are concerned with evolving equations represented by these discrete diffeomorphisms, and mesh degeneracy causes the numerics to fail, we introduce the notion of non-degenerate discrete diffeomorphisms,

Definition 13.3. *A **non-degenerate discrete diffeomorphism** $\varphi = (K_1, K_2)$ is such that K_1 and K_2 are non-degenerate realizations of the abstract simplicial complex M in the ambient space V.*

Notice that it is sufficient to define the discrete diffeomorphism on the vertices of the abstract complex M, since we can extend it to the entire complex by the relation

$$\varphi([v_0, ..., v_k]) = ([\pi_1\varphi(v_0), ..., \pi_1\varphi(v_k)], [\pi_2\varphi(v_0), ..., \pi_2\varphi(v_k)]).$$

If $X \in K^{(0)}$ is a material vertex of the manifold, corresponding to the abstract vertex w, that is to say, $\pi_1\varphi_t(w) = X, \forall t \in I$, the corresponding trajectory followed by X in space is $x = \pi_2\varphi_t(w)$. Then, the ***material velocity*** $V(X, t)$ is given by

$$V(\pi_1(w), t) = \left.\frac{\partial \pi_2\varphi_s(w)}{\partial s}\right|_{s=t},$$

and the ***spatial velocity*** $v(x, t)$ is given by

$$v(\pi_2(w), t) = V(\pi_1(w), t) = \left.\frac{\partial \varphi_s(\varphi_t^{-1}(x))}{\partial s}\right|_{s=t}.$$

The distinction between the spatial and material representation is illustrated in Figure 24.

The material velocity field can be thought of as a discrete vector field with the vectors based at the vertices of K, which is to say that $T\varphi_t \in \mathfrak{X}_d(K)$, is a discrete primal vector field. Notice that φ_t on K induces a

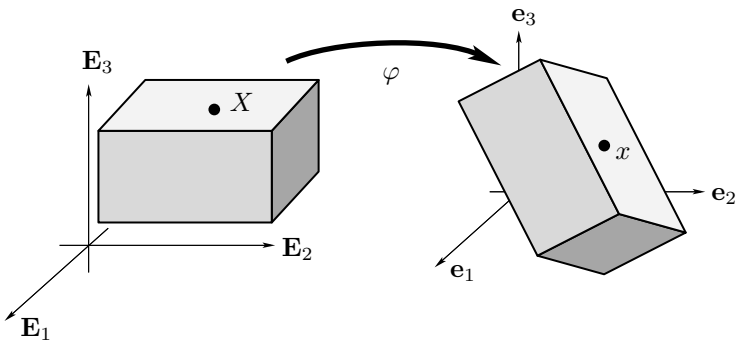


Figure 24: Spatial and material representations.

map $\star\varphi_t$ on the vertices of the dual $\star K$, by the following,

$$\star\varphi_t(c[v_0,\ldots,v_n]) = (c[\pi_1\varphi_t(v_0),\ldots,\pi_1\varphi_t(v_n)], c[\pi_2\varphi_t(v_0),\ldots,\pi_2\varphi_t(v_n)]).$$

Similarly then, $T\star\varphi_t \in \mathfrak{X}_d(\star K)$ is a discrete dual vector field.

Comparison with Interpolatory Methods. At first glance, the groupoid formulation seems like a cumbersome way to define a one-parameter family of discrete diffeomorphisms, and one may be tempted to think of extending φ_t to the ambient space. We would then be thinking of $\varphi_t : V \to V$. This is undesirable since given φ_t and ψ_s which are non-degenerate flows, their composition $\varphi_t \circ \psi_s$, which is defined, may result in a degenerate mesh when applied to K. Thus, non-degenerate flows are not closed under this notion of composition.

If we adopt groupoid composition instead at the level of vertices, we can always be sure that if we compose two nondegenerate discrete diffeomorphisms, they will remain a nondegenerate discrete diffeomorphism.

Discrete Diffeomorphisms as Pair Groupoids. The space of discrete diffeomorphisms naturally has the structure of a pair groupoid. The discrete analogue of $T\operatorname{Diff}(M)$ from the point of view of temporal discretization is the pair groupoid $\operatorname{Diff}(M) \times \operatorname{Diff}(M)$. In addition, we discretize $\operatorname{Diff}(M)$ using $\operatorname{Diff}_d(M)$, which is in turn a pair groupoid involving realizations of an abstract simplicial complex in an ambient space.

13.3. **Push-Forward and Pull-Back of Discrete Vector Fields and Discrete Forms.** For us to construct a discrete theory of exterior calculus that admits dynamic problems, it is critical that we introduce the notion of push-forward and pull-back of discrete vector fields and discrete forms under a discrete flow.

Push-Forward and Pull-Back of Discrete Vector Fields.

The push-forward of a discrete vector field satisfies the following commutative diagram,

$$\begin{array}{ccccc} K & \xrightarrow{\star} & \star K & \xrightarrow{X} & \mathbb{R}^N \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \star f} & & \downarrow{\scriptstyle Tf} \\ L & \xrightarrow{\star} & \star L & \xrightarrow{f_*X} & \mathbb{R}^N \end{array}$$

and the pull-back satisfies the following commutative diagram,

$$\begin{array}{ccccc} K & \xrightarrow{\star} & \star K & \xrightarrow{f^*X} & \mathbb{R}^N \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \star f} & & \downarrow{\scriptstyle Tf} \\ L & \xrightarrow{\star} & \star L & \xrightarrow{X} & \mathbb{R}^N \end{array}$$

By appropriately following the diagram around its boundary, we obtain the following expressions for the push-forward and pull-back of a discrete vector field.

Definition 13.4. *The **push-forward of a dual discrete vector field** $X \in \mathfrak{X}_d(\star K)$, under the map $f : K \to L$, is given by its evaluation on a dual vertex $\hat{\sigma}_0 = \star\sigma^n \in (\star L)^{(0)}$,*

$$f_* X(\star\sigma^n) = Tf \cdot X(\star(f^{-1}(\sigma^n))).$$

Definition 13.5. *The **pull-back of a dual discrete vector field** $X \in \mathfrak{X}_d(\star L)$, under the map $f : K \to L$, is given by its evaluation on a dual vertex $\hat{\sigma}_0 = \star\sigma^n \in (\star K)^{(0)}$,*

$$f^* X(\star\sigma^n) = (f^{-1})_* X(\star\sigma^n) = T(f^{-1}) \cdot X(\star(f(\sigma^n))).$$

Pull-Back and Push-Forward of Discrete Forms. A natural operation involving exterior calculus in the context of dynamic problems is the pull-back of a differential form by a flow. We define the pull-back of a discrete form as follows.

Definition 13.6. *The **pull-back of a discrete form** $\alpha^k \in \Omega_d^k(L)$, under the map $f : K \to L$, is defined so that the **change of variables formula** holds,*

$$\langle f^* \alpha^k, \sigma^k \rangle = \langle \alpha^k, f(\sigma^k) \rangle,$$

where $\sigma^k \in K$.

We can define the push-forward of a discrete form as its pull-back under the inverse map as follows.

Definition 13.7. *The **push-forward of a discrete form** $\alpha^k \in \Omega_d^k(K)$, under the map $f : K \to L$ is defined by its action on $\sigma^k \in L$,*

$$\langle f_* \alpha^k, \sigma^k \rangle = \langle (f^{-1})^* \alpha^k, \sigma^k \rangle = \langle \alpha^k, f^{-1}(\sigma^k) \rangle.$$

Naturality under Pull-Back of Wedge Product. We find that the discrete wedge product we introduced in §8 is not natural under pull-back, which is to say that the relation

$$f^*(\alpha \wedge \beta) = f^* \alpha \wedge f^* \beta \,,$$

does not hold in general. However, a metric independent definition that is natural under pull-back was proposed in Castrillón-López [2003].

Definition 13.8 (Castrillón-López [2003])**.** *Given a primal discrete k-form $\alpha^k \in \Omega_d^k(K)$, and a primal discrete l-form $\beta^l \in \Omega_d^l(K)$, the **natural discrete primal-primal wedge product**, $\wedge : \Omega_d^k(K) \times \Omega_d^l(K) \to \Omega_d^{k+l}(K)$, is defined by its evaluation on a $(k+l)$-simplex $\sigma^{k+l} = [v_0, \ldots, v_{k+l}]$,*

$$\langle \alpha^k \wedge \beta^l, \sigma^{k+l} \rangle = \frac{1}{(k+l+1)!} \sum_{\tau \in S_{k+l+1}} \operatorname{sign}(\tau) \alpha \smile \beta(\tau(\sigma^{k+l})) \,.$$

In contrasting this definition to that given by Definition 8.1, we see that the geometric factor

$$\frac{|\sigma^{k+l} \cap \star v_{\tau(k)}|}{|\sigma^{k+l}|} \,,$$

has been replaced by

$$\frac{1}{k+l+1}$$

in this alternative definition. By replacing the geometric factor which is metric dependent with a constant factor, Definition 13.8 becomes natural under pull-back.

The proofs in §8 that the discrete wedge product is anti-commutative, and satisfies a Leibniz rule, remain valid for this alternative discrete wedge product, with only trivial modifications. As for the proof of the

associativity of the wedge product for closed forms, we note the following identity,

$$\sum_{\tau \in S_{k+l+1}} \frac{|\sigma^{k+l} \cap \star v_{\tau(k)}|}{|\sigma^{k+l}|} = \sum_{\tau \in S_{k+l+1}} \frac{1}{k+l+1} = (k+l)!\,,$$

which is a crucial observation for the original proof to apply to the alternative wedge product.

14. Remeshing Cochains and Multigrid Extensions

It is sometimes desirable, particularly in the context of multigrid, multiscale, and multiresolution computations, to be able to represent a discrete differential form which is given as a cochain on a prescribed mesh, as one which is supported on a new mesh. Given a differential form $\omega^k \in \Omega^k(K)$, and a new mesh M such that $|K| = |M|$, we can define it at the level of cosimplices,

$$\forall \tau^k \in M^{(k)}, \qquad \langle \omega^k, \tau^k \rangle = \sum_{\sigma^k \in K^{(k)}} \operatorname{sgn}(\tau^k, \sigma^k) \frac{|V_{\tau^k} \cap V_{\sigma^k}|}{|V_{\sigma^k}|} \langle \omega^k, \sigma^k \rangle,$$

and extend this by linearity to cochains. Here, $\operatorname{sgn}(\tau^k, \sigma^k)$ is $+1$ if the orientation of τ^k and σ^k are consistent, and -1 otherwise. Since k-skeletons of meshes that are not related by subdivision may not have nontrivial intersections, intersections of support volumes are used in the remeshing formula, as opposed to intersections of the k-simplices.

We denote this transformation at the level of cochains as, $T_{K,M} : C^k(K) \to C^k(M)$. This has the natural property that if we have a k-volume U^k that can be represented as a chain in either the complex K or the complex M, that is to say, $U^k = \sigma_1^k + \ldots + \sigma_l^k = \tau_1^k + \ldots + \tau_l^k$, then we have

$$\begin{aligned}
\langle \omega^k, \tau_1^k + \ldots + \tau_m^k \rangle &= \sum_{i=1}^{m} \langle \omega, \tau_i^k \rangle = \sum_{i=1}^{m} \sum_{\sigma^k \in K^{(k)}} \operatorname{sgn}(\tau_i^k, \sigma^k) \frac{|V_{\tau_i^k} \cap V_{\sigma^k}|}{|V_{\sigma^k}|} \langle \omega^k, \sigma^k \rangle \\
&= \sum_{i=1}^{m} \sum_{j=1}^{l} \operatorname{sgn}(\tau_i^k, \sigma_j^k) \frac{|V_{\tau_i^k} \cap V_{\sigma_j^k}|}{|V_{\sigma_j^k}|} \langle \omega^k, \sigma_j^k \rangle \\
&= \sum_{j=1}^{l} \sum_{i=1}^{m} \operatorname{sgn}(\tau_i^k, \sigma_j^k) \frac{|V_{\tau_i^k} \cap V_{\sigma_j^k}|}{|V_{\sigma_j^k}|} \langle \omega^k, \sigma_j^k \rangle \\
&= \sum_{j=1}^{l} \langle \omega^k, \sigma_j^k \rangle = \langle \omega^k, \sigma_1^k + \ldots + \sigma_l^k \rangle.
\end{aligned}$$

Which is to say that the integral of the differential form over U^k is well-defined, and independent of the representation of the differential form.

Note that, in particular, if we choose to coarsen the mesh, the value the form takes on a cell in the coarser mesh is simply the sum of the values the form takes on the old cells of the fine mesh which make up the new cell in the coarser mesh.

Non-Flat Manifolds. The case of non-flat manifolds presents a challenge in remeshing akin to that encountered in the discretization of differential forms. In particular, if the two meshes represent different discretizations of a non-flat manifold, they will in general correspond to different polyhedral regions in the embedding space, and not have the same support region.

We assume that our discretization of the manifold is sufficiently fine that for every simplex, all its vertices are contained in some chart. Then, by using these local charts, we can identify support volumes in the computational domain with n-volumes in the manifold, and thereby make sense of the remeshing formula.

15. Conclusions and Future Work

We have presented a framework for discrete exterior calculus using the cochain representation of discrete differential forms, and introduced combinatorial representations of discrete analogues of differential operators on discrete forms and discrete vector fields. The role of primal and dual cell complexes in the theory are developed in detail. In addition, extensions to dynamic problems and multi-resolution computations are discussed.

In the next few paragraphs, we will describe some of the future directions that emanate from the current work on discrete exterior calculus.

Relation to Computational Algebraic Topology Since we have introduced a discrete Laplace-deRham operator, one can hope to develop a discrete Hodge-deRham theory, and relate the deRham cohomology of a simplicial complex to its simplicial cohomology.

Extensions to Non-Flat Manifolds. The intrinsic notion of what constitutes the discrete tangent space to a node on a non-flat mesh remains an open question. It is possible that this notion is related to a choice of discrete connection on the mesh, and it is an issue that deserves further exploration.

Generalization to Arbitrary Tensors. The discretization of differential forms as cochains is particularly natural, due to the pairing between forms and volumes by integration. When attempting to discretize an arbitrary tensor, the natural discrete analogue is unclear. In particular, while it is possible to expand an arbitrary tensor using the tensor product of covariant and contravariant one-tensors, this would be cumbersome to represent on a mesh. In Leok et al. [2003], which is on discrete connections, we will see Lie group-valued discrete 1-forms, and one possible method of discretizing a (p, q)-tensor that is alternating in the contravariant indices, is to consider it as a $(0, q)$-tensor-valued discrete p-form.

It would be particularly interesting to explore this in the context of the elasticity complex (see, for example, Arnold [2002]),

$$\mathfrak{se}(3) \hookrightarrow C^\infty(\Omega, \mathbb{R}^3) \xrightarrow{\epsilon} C^\infty(\Omega, \mathbb{S}) \xrightarrow{J} C^\infty(\Omega, \mathbb{S}) \xrightarrow{\text{div}} C^\infty(\Omega, \mathbb{R}^3) \longrightarrow 0\,,$$

where $\mathbb{S}$ is the space of 3×3 symmetric matrices. One approach to discretize this was suggested in Arnold [2002], which cites the use of the Bernstein–Gelfand–Gelfand resolution in Eastwood [2000] to derive the elasticity complex from the deRham complex. Alternatively, it might be appropriate in the context of the elasticity complex to consider Lie algebra-valued discrete differential forms.

Convergence and Higher-Order Theories. The natural question from the point of view of numerical analysis would be to carefully analyze the convergence properties of these discrete differential geometric operators. In addition, higher-order analogues of the discrete theory of exterior calculus are desirable from the point of view of computational efficiency, but the cochain representation is attractive due to its conceptual simplicity and the elegance of representing discrete operators as combinatorial operations on the mesh.

It would therefore be desirable to reconcile the two, by ensuring that high-order interpolation and combinatorial operations are consistent. As a low-order example, Whitney forms, which are used to interpolate differential forms on a simplicial mesh, have the nice property that taking the Whitney form associated with the coboundary of a simplicial cochain is equal to taking the exterior derivative of the Whitney form associated with the simplicial cochain. As such, the coboundary operation, which is a combinatorial operation akin to finite differences, is an exact discretization of the exterior derivative, when applied to the degrees of freedom associated to the finite-dimensional function space of Whitney forms.

It would be interesting to apply subdivision surface techniques to construct interpolatory spaces that are compatible with differential geometric operations that are combinatorial operations on the degrees of freedom. This will result in a massively simplified approach to higher-order theories of discrete exterior

calculus, by avoiding the use of symbolic computation, which would otherwise be necessary to compute the action of continuous exterior differential operators on the polynomial expansions for differential forms.

References

R. Abraham, J. E. Marsden, and T. S. Ratiu. *Manifolds, Tensor Analysis and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, second edition, 1988.

D. H. Adams. R-torsion and linking numbers from simplicial abelian gauge theories. arXiv, hep-th/9612009, 1996.

D. N. Arnold. Differential complexes and numerical stability. In *Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002)*, pages 137–157, Beijing, 2002. Higher Ed. Press.

A. Bossavit. Generalized finite differences in computational electromagnetics. *Progress in Electromagnetics Research, PIER*, 32:45–64, 2001.

A. Bossavit. Applied differential geometry (a compendium). URL `http://www.icm.edu.pl/edukacja/mat/Compendium.php`. (preprint), 2002a.

A. Bossavit. Extrusion, contraction: Their discretization via Whitney forms. (preprint), 2002b.

A. Bossavit. On "generalized finite differences": Discretization of electromagnetic problems. (preprint), 2002c.

A. Cannas da Silva and A. Weinstein. *Geometric Models for Noncommutative Algebras*, volume 10 of *Berkeley Mathematics Lecture Notes*. American Mathematical Society, 1999.

M. Castrillón-López. Discrete variational problems on forms. (in preparation), 2003.

M. Desbrun, A. N. Hirani, M. Leok, and J. E. Marsden. Discrete Poincaré lemma. *Appl. Numer. Math.*, 2003. (submitted).

A. A. Dezin. *Multidimensional Analysis and Discrete Models*. CRC Press, 1995.

M. Eastwood. A complex from linear elasticity. In *The Proceedings of the 19th Winter School "Geometry and Physics" (Srní, 1999)*, number 63 in Rend. Circ. Mat. Palermo (2) Suppl., pages 23–29, 2000.

R. Forman. Discrete Morse theory and the cohomology ring. *Trans. Amer. Math. Soc.*, 354(12):5063–5085 (electronic), 2002.

P. W. Gross and P. R. Kotiuga. Data structures for geometric and topological aspects of finite element algorithms. *Progress in Electromagnetics Research, PIER*, 32:151–169, 2001.

A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2001.

R. Hiptmair. Canonical construction of finite elements. *Math. Comp.*, 68(228):1325–1346, 1999.

R. Hiptmair. Discrete Hodge-operators: An algebraic perspective. *Progress in Electromagnetics Research, PIER*, 32:247–269, 2001a.

R. Hiptmair. Higher order Whitney forms. *Progress in Electromagnetics Research, PIER*, 32:271–299, 2001b.

R. Hiptmair. Finite elements in computational electromagnetism. In *Acta Numerica*, volume 11, pages 237–339. Cambridge University Press, 2002.

A. N. Hirani. *Discrete Exterior Calculus*. PhD thesis, California Institute of Technology, 2003.

J. D. Jackson. *Classical Electrodynamics*. Wiley, third edition, 1998.

T. Kaczynski, K. Mischaikow, and M. Mrozek. *Computational homology*, volume 157 of *Applied Mathematical Sciences*. Springer-Verlag, 2004.

M. Leok, J. E. Marsden, and A. Weinstein. A discrete theory of connections on principal bundles. (in preparation), 2003.

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

E. L. Mansfield and P. E. Hydon. On a variational complex for difference equations. In *The Geometrical Study of Differential Equations (Washington, DC, 2000)*, volume 285 of *Contemporary Mathematics*, pages 121–129. American Mathematical Society, 2001.

J. E. Marsden, S. Pekarsky, and S. Shkoller. Discrete Euler–Poincaré and Lie–Poisson equations. *Nonlinearity*, 12(6):1647–1662, 1999.

J. E. Marsden, S. Pekarsky, and S. Shkoller. Symmetry reduction of discrete Lagrangian mechanics on Lie groups. *J. Geom. Phys.*, 36(1-2):140–151, 2000.

C. Mattiussi. An analysis of finite volume, finite element, and finite difference methods using some concepts from algebraic topology. *J. Comput. Phys.*, 133(2):289–309, 1997.

C. Mattiussi. The finite volume, finite difference, and finite element methods as numerical methods for physical field problems. *Adv. Imag. Elect. Phys.*, 113:1–146, 2000.

M. Meyer, M. Desbrun, P. Schröder, and A. H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds. *VisMath*, 2002.

B. Moritz. *Vector Difference Calculus*. PhD thesis, University of North Dakota, 2000.

B. Moritz and W. A. Schwalm. Triangle lattice green functions for vector fields. *J. Phys. A.*, 34(3):589–602, 2001.

J. R. Munkres. *Elements of Algebraic Topology*. Addison-Wesley, 1984.

R. A. Nicolaides and D. -Q. Wang. Convergence analysis of a covolume scheme for Maxwell's equations in three dimensions. *Math. Comp.*, 67(223):947–963, 1998.

S. Pekarsky and M. West. Discrete diffeomorphism groupoids and circulation conserving fluid integrators. (in preparation), 2003.

J. R. Schewchuck. What is a good linear finite element? Interpolation, conditioning, anisotropy and quality measures. URL `http://www.cs.berkeley.edu/~jrs/papers/elemj.ps`. (preprint), 2002.

W. A. Schwalm, B. Moritz, M. Giona, and M. K. Schwalm. Vector difference calculus for physical lattice models. *Phys. Rev. E.*, 59(1, part B):1217–1233, 1999.

S. Sen, S. Sen, J. C. Sexton, and D. H. Adams. Geometric discretization scheme applied to the abelian Chern–Simons theory. *Phys. Rev. E*, 61(3):3174–3185, 2000.

F. L. Teixeira. Geometric aspects of the simplicial discretization of Maxwell's equations. *Progress in Electromagnetics Research, PIER*, 32:171–188, 2001.

Y. Y. Tong, S. Lombeyda, A. N. Hirani, and M. Desbrun. Discrete multiscale vector field decomposition. *ACM Transactions on Graphics (SIGGRAPH)*, July 2003.

E. Tonti. Finite formulation of electromagnetic field. *IEEE Trans. Mag.*, 38:333–336, 2002.

A. Weinstein. Lagrangian mechanics and groupoids. In *Mechanics Day (Waterloo, ON, 1992)*, volume 7 of *Fields Institute Communications*, pages 207–231. American Mathematical Society, 1996.

A. Weinstein. Groupoids: unifying internal and external symmetry. A tour through some examples. In *Groupoids in Analysis, Geometry, and Physics (Boulder, CO, 1999)*, volume 282 of *Contemporary Mathematics*, pages 1–19. American Mathematical Society, 2001.

H. Whitney. *Geometric Integration Theory*. Princeton University Press, 1957.

Z. J. Wood. *Computational Topology Algorithms for Discrete 2-Manifolds*. PhD thesis, California Institute of Technology, 2003.

Computer Science, USC, 941 W. 37th Place Los Angeles, CA 90089.
E-mail address: `desbrun@usc.edu`

256-80, Computer Science, Caltech, Pasadena, CA 91125.
E-mail address: `hirani@caltech.edu`

107-81, Control and Dynamical Systems, Caltech, Pasadena, CA 91125.
E-mail address: `mleok@cds.caltech.edu`

107-81, Control and Dynamical Systems, Caltech, Pasadena, CA 91125.
E-mail address: `marsden@cds.caltech.edu`