# COMPUTATIONAL GEOMETRIC MECHANICS AND CONTROL OF RIGID BODIES

by

Taeyoung Lee

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Aerospace Engineering)
in The University of Michigan
2008

Doctoral Committee:

Professor N. Harris McClamroch, Co-Chair
Assistant Professor Melvin Leok, Purdue University, Co-Chair
Professor Anthony M. Bloch
Professor Jessy W. Grizzle
Professor Daniel J. Scheeres, University of Colorado

*To rachel*

# ACKNOWLEDGEMENTS

It has been a blessing to me that I could do *what I really like to do* under encouragements and supports from my family and friends. As everybody says, a graduation implies a new challenge, which makes me excited and thrilled. But, before setting out on my new journey of academia, I would like to express my gratitude to who have influenced on me.

From the very beginning, my parents have inspired me with courage: when I made any decision, they always supported me and deduced several reasons that made the decision more appropriate. Now, I understand a little bit about raising a child with absolute trust, instead of having expectations or concerning about him. It is easy to tell an instructive story to a child, but it requires a tremendous effort to teach a child by showing a real model of what he ought to be. My parents have been doing it for me ever since I was born, and I am really proud of my parents.

I have had another exclusive privilege of being advised by Professor N. Harris McClamroch and Professor Melvin Leok. It has been an honor to have an academic guidance, instruction, encouragement, and insight from both of them, throughout my doctoral study. I am aware that, in some cases, the relationship between an advisor and a graduate student is similar to that of an employer and an employee. My relationship with them can be described as the exact opposite: they have treated me with respect, as an old colleague rather than a student, and they always let me make an important academic decision by myself. In addition to the technical materials required for this dissertation, more importantly, I have learned what I can do after my graduation and how I should live my life as a researcher while enjoying it. Their support and friendship, which also have been from Ms. Margaret McClamroch, are gratefully acknowledged.

I would like to thank Professor Youdan Kim, whose enthusiasm for control systems engineering has motivated me from my undergraduate studies. I decided to study aerospace engineering by myself, but I chose control systems engineering due to his academic excellence and goodness. I also thank my dissertation committee members, Professor Scheeres, Professor Bloch, and Professor Grizzle for providing valuable suggestions to improve the quality of my dissertation.

I was also lucky to have been interacting with Eugene Fahnestock, whose sincerity and continuous effort have stimulated my research. It has been a great pleasure to be with the ksag.ae members and families. In particular, I would like to thank Dr. Chandeok Park, Dr. Kyungjin Lee and Haewon Lee for their valuable suggestions and for all of the memorable time that we spent together.

Without the constant support of my wife, *Heylim*, I could not have written a single page of this dissertation. Being a wife of an international student implies a sacrifice that any other people cannot fully understand. I am also deeply indebted to her parents. Her contribution to this dissertation cannot be compared with any others, and this dissertation is gratefully dedicated to her. I would like to thank Heylim for being with me, and I would like to say to her, *"I love you, and I will always."*

# TABLE OF CONTENTS

# LIST OF FIGURES

# ABSTRACT

## COMPUTATIONAL GEOMETRIC MECHANICS AND CONTROL OF RIGID BODIES

by

Taeyoung Lee

This dissertation studies the dynamics and optimal control of rigid bodies from two complementary perspectives, by providing theoretical analyses that respect the fundamental geometric characteristics of rigid body dynamics and by developing computational algorithms that preserve those geometric features. This dissertation is focused on developing analytical theory and computational algorithms that are intrinsic and applicable to a wide class of multibody systems.

A geometric numerical integrator, referred to as a Lie group variational integrator, is developed for rigid body dynamics. Discrete-time Lagrangian and Hamiltonian mechanics and Lie group methods are unified to obtain a systematic method for constructing numerical integrators that preserve the geometric properties of the dynamics as well as the structure of a Lie group. It is shown that Lie group variational integrators have substantial computational advantages over integrators that preserve either one of none of these properties. This approach is also extended to mechanical systems evolving on the product of two-spheres.

A computational geometric approach is developed for optimal control of rigid bodies on a Lie group. An optimal control problem is discretized at the problem formulation stage by using a Lie group variational integrator, and discrete-time necessary conditions for optimality are derived using the calculus of variations. The discrete-time necessary conditions inherit the desirable computational properties of the Lie group variational integrator, as they are derived from a symplectic discrete flow. They do not exhibit the numerical dissipation introduced by conventional numerical integration schemes, and consequently, we can efficiently obtain optimal controls that respect the geometric features of the optimality conditions.

The approach that combines computational geometric mechanics and optimal control is illustrated by various examples of rigid body dynamics, which include a rigid body pendulum on a cart, pure bending of an elastic rod, and two rigid bodies connected by a ball joint. Since all of the analytical and computational results developed in this dissertation are coordinate-free, they are independent of a specific choice of local coordinates, and they completely avoid any singularity, ambiguity, and complexity associated with local coordinates. This provides insight into the global dynamics of rigid bodies.

CHAPTER 1

# INTRODUCTION

## 1.1 Motivation and Goal

This dissertation studies dynamics and optimal control problems for rigid bodies from two complementary perspectives, by providing theoretical analyses that respect the fundamental geometric characteristics of rigid body dynamics and by developing computational algorithms that preserve those geometric characteristics.

In control systems engineering, the underlying geometric features of a dynamic system are often not considered carefully. For example, many control systems are developed for the standard form of ordinary differential equations, namely $\dot{x} = f(x, u)$, where the state and the control input are denoted by $x$ and $u$, respectively (see, for example, Khalil 2002; Nijmeijer and van der Schaft 1990). It is assumed that the state and the control input lie in Euclidean spaces, and the system equations are defined in terms of smooth functions between Euclidean spaces. However, for many interesting mechanical systems, the configuration space cannot be expressed globally as a Euclidean space. In addition, general purpose numerical algorithms may not accurately respect fundamental geometric properties (see Hairer et al. 2000; Leimkuhler and Reich 2004).

In this dissertation, dynamics and optimal control problems for rigid bodies are studied, incorporating careful consideration of their geometric features. We explicitly consider the following research questions: what are the geometric properties of dynamics of rigid bodies, how should the configuration of rigid bodies be described, how are the geometric properties utilized in control system analysis and design, and how can the geometric characteristics be preserved in numerical computations. The goal of this dissertation is to develop both analytical tools and computational algorithms for rigid body dynamics and control that respect the fundamental geometric features.

### 1.1.1 Fundamental Geometric Properties of Rigid Body Dynamics

*Lie Group Configuration Manifold.* The configuration of a rigid body can be described by the location of its mass center and the orientation of the rigid body in a three-dimensional space. The location of the rigid body can be expressed in Euclidean space, but the attitude evolves in a nonlinear space that has a certain geometry.

More precisely, the attitude of a rigid body is defined as the direction of a body-fixed frame with respect to a reference frame, considered as a linear transformation on the vector space $\mathbb{R}^3$; the attitude can be represented mathematically by a $3 \times 3$ orthonormal matrix. We require that its

determinant is positive in order to preserve the ordering of the orthonormal axes according to the right-hand rule. The set of $3 \times 3$ orthonormal matrices with positive determinant is a manifold as it is locally diffeomorphic to a Euclidean space, and it also has a group structure with the group action of matrix multiplication. A smooth manifold with a group structure is referred to as a Lie group; the Lie group of $3 \times 3$ orthonormal matrices with positive determinant is referred to as the special orthogonal group, $\mathsf{SO}(3)$ (see, for example, Murray et al. 1993; Varadarajan 1984). The configuration manifold for the combined translational and rotational motion of a rigid body is the special Euclidean group $\mathsf{SE}(3)$, which is a semidirect product $\mathsf{SE}(3) = \mathsf{SO}(3) \circledS \mathbb{R}^3$. A direct product of the Lie groups $\mathsf{SE}(3)$, $\mathsf{SO}(3)$, and $\mathbb{R}^n$ can represent the configuration of multiple rigid bodies, and it is also a Lie group since a product of Lie groups is also a Lie group. Therefore, the configuration manifold of an interconnection of rigid bodies is also a Lie group.

*Lagrangian/Hamiltonian System.* Mechanics studies the dynamics of physical bodies acting under forces and potential fields (see Arnold 1989; Goldstein et al. 2001; Meirovitch 2004). In Lagrangian mechanics, the trajectories are obtained by finding the paths that minimize the integral of a Lagrangian over time, called the action integral. In classical problems, the Lagrangian is chosen as the difference between kinetic energy and potential energy. The Legendre transformation provides an alternative description, referred to as Hamiltonian mechanics.

Rigid body dynamics are characterized by Lagrangian/Hamiltonian dynamics. The dynamics of a Lagrangian or Hamiltonian system has 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. By quotienting out the symmetry, a reduced Lagrangian/Hamiltonian system can be developed (see Marsden 1992).

### 1.1.2 Computational Geometric Mechanics and Control

Geometric mechanics is a modern description of classical mechanics from the perspective of differential geometry (see, for example, Abraham and Marsden 1978; Bloch 2003a; Bullo and Lewis 2005; Jurdjevic 1997; Marsden and Ratiu 1999). It explores the geometric structure of a Lagrangian or Hamiltonian system through the concept of vector fields, symplectic geometry, and symmetry techniques. Geometric mechanics provides fundamental insights into mechanics and yields useful tools for dynamics and control theory. For example, geometric mechanics led to the energy-momentum method in Simo et al. (1990), reduction/reconstruction in Marsden et al. (1990, 2000), and the controlled Lagrangian method in Bloch et al. (2000, 2001).

The goal of computational geometric mechanics is to construct computational algorithms that preserve the geometric properties (see Leok 2004). It applies the fundamental principles of geometric mechanics to discrete-time mechanical system to construct geometric structure-preserving numerical schemes. Since the computational algorithms are developed from discrete-time analogues of physical principles, the geometric properties of the dynamics are preserved naturally. This is in contrast with the perspective that considers a numerical method as an approximation to a

continuous-time equation.

In summary, this dissertation is focused on computational geometric mechanics and control of rigid bodies. We develop computational methods for rigid bodies, that preserve the underlying Lagrangian/Hamiltonian system structure of rigid body dynamics as well as the Lie group structure of the configurations. These methods are applied to numerical integration and optimal control problems. Prior work related to computational geometric mechanics and control of rigid bodies is summarized below, followed by an outline and the contributions of this dissertation.

## 1.2 Literature Review

### 1.2.1 Geometric Numerical Integration

Geometric numerical integration deals with numerical integration methods that preserve geometric properties of the flow of a differential equation, such as invariants, symplecticity, and the configuration manifold (see Hairer et al. 2000; Leimkuhler and Reich 2004; McLachlan and Quispel 2001).

Numerical methods that conserve energy and momentum for mechanical systems are referred to as energy-momentum integrators (see LaBudde and Greenspan 1976; Simo et al. 1992). In these methods, a free parameter is selected to maintain constant angular momentum; energy conservation is typically enforced by a momentum-preserving projection onto the manifold of constant energy.

Numerical integration methods that preserve the symplecticity of a Hamiltonian system have been studied in Lasagni (1988); Sanz-Serna (1992, 1988). Qualitative properties of symplectic integrators are given in Gonzalez and Simo (1996); Gonzalez et al. (1990), and long-time behavior of symplectic methods is addressed in Benettin and Giorgilli (1994); Hairer (1994); Hairer and Lubich (2000). Coefficients of certain Runge-Kutta methods can be chosen to satisfy a symplecticity criterion and order conditions to obtain symplectic Runge-Kutta methods. 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 equations (see Marsden and West 2001). This approach provides a systematic method to develop geometric numerical integrators for Lagrangian / Hamiltonian systems. The resulting integrators have the desirable property that they are symplectic and momentum preserving, and they exhibit good energy behavior for exponentially long times.

The idea of developing a discrete-time mechanical system that conserves the constants of motion appears in the work by Greenspan (1981, 1972); LaBudde and Greenspan (1974), and a discrete-time mechanical system has been developed according to Hamilton's principle by Moser and Veselov (1991); Veselov (1988). The variational view of discrete-time mechanics is further developed by Kane et al. (1999, 2000); Wendlandt and Marsden (1997), and an intrinsic form of discrete-time variational principles is established by Marsden and West (2001).

Geometric integrators that preserve the manifold or Lie group structure have been studied (see,

for example, Budd and Iserles 1999; Hairer and Wanner 1996; Iserles et al. 2000). A natural approach to the numerical solution of differential equations on a manifold is by projection. In the work by Dieci et al. (1994), a solution is updated by a one-step integration method and it is projected to the manifold on which the system evolves at each time step. This projection may destroy desirable long-time behavior of one-step methods, since the projection typically corrupts the numerical trajectory. Numerical methods based on local coordinates of the manifold often result in unnecessary singularities (see Potra and Rheinbold 1991). Differential algebraic approaches have been proposed to solve nonlinear constrained equations at each time step in Hairer and Wanner (1996).

For differential equations that evolve on a Lie group, a group element can be updated by the corresponding group action so that the group structure is preserved naturally. This is referred to as a Lie group method (see Iserles et al. 2000). Among the Lie group methods, the Crouch-Grossman method updates the group elements by multiple evaluations using the exponential map (see Crouch and Grossman 1993), and the Munthe-Kaas method is based on a differential equation on the Lie algebra and uses a single evaluation of the exponential map (see Munthe-Kaas 1995). A homogeneous manifold is a manifold on which a Lie group acts continuously in a transitive way. Lie group methods are extended to homogeneous manifolds in Munthe-Kaas and Zanna (1997).

For mechanical systems evolving on a Lie group, a discrete-time Euler-Poincaré equation has been introduced for a left-invariant Lagrangian by Marsden et al. (1999), with application to the free attitude dynamics of a rigid body. A similar development is presented for the attitude dynamics of an axially symmetric rigid body acting under a gravitational potential in Bobenko and Suris (1999). The idea of using the Lie group structure and the exponential map to numerically compute rigid body dynamics arises in Krysl (2005); Simo et al. (1992). Symplectic integrators with explicit constraints on the Lie group structure are applied to rigid body dynamics in Leimkuhler and Reich (2004).

### 1.2.2 Geometric Optimal Control

Optimal control problems deal with finding trajectories, such that a certain optimality condition is satisfied under prescribed constraints (see, for example, Bryson and Ho 1975; Kirk 1970; Sussmann and Willems 1997). This is typically based on Pontryagin's minimum principle or the calculus of variations. Geometric optimal control forms a theoretical foundation for extensions of the minimum principle to optimal control problems defined on arbitrary differentiable manifolds (see Jurdjevic 1997).

A geometric, intrinsic formulation of the minimum principle is presented in a coordinate-free fashion in Sussmann (1998a,b). A general formulation of optimal control theory for nonholonomic systems on a Riemannian manifold is presented in Bloch and Crouch (1993, 1998, 1995). This approach is applied to both kinematic sub-Riemannian optimal control problems and optimal control problems for mechanical systems by Bloch (2003a,b). A dynamic interpolation problem on a Riemannian manifold is formulated as an optimal control problem in Hussein and Bloch (2004b), and this approach is extended to an optimal control problem on a Riemannian manifold with a potential

in Hussein and Bloch (2004a). An optimal control problem for nonholonomic and under-actuated mechanical systems is considered in Hussein and Bloch (2006).

Controllability, observability and optimal control problems on a Lie group have been studied by Brockett (1973, 1972). A simple closed-form analytic solution for an optimal control problem of right-invariant systems evolving on a matrix Lie group is presented in Baillieul (1978). An optimal control problem for a generalized rigid body on $\mathsf{SO}(n)$ is considered in Bloch and Crouch (1996). A general theory of optimal control problems is developed in Jurdjevic (1998a,b, 1997) together with reachability and controllability conditions; these approaches are based on kinematics equations, and assume that group elements are directly controlled by elements in the Lie algebra. Optimal control problems for the dynamics of a rigid body with application to dynamic coverage problem are studied by Hussein (2005); Hussein and Bloch (2005a,b).

Computational geometric optimal control approaches apply optimal control theory to discrete-time mechanical systems obtained using geometric numerical integrators. A discrete version of the generalized rigid body equations and their formulation as an optimal control problem are presented in Bloch et al. (1998, 2002). Discrete-time optimal control problems for the attitude dynamics of a rigid body on $\mathsf{SO}(3)$ are considered in Bloch et al. (2007); Hussein et al. (2006) based on the variational integrator. A direct optimal control approach is applied to discrete-time mechanical systems in Junge et al. (2005), referred to as Discrete Mechanics and Optimal Control.

## 1.3 Outline of Dissertation

In this dissertation, geometric mechanics and optimal control for rigid bodies are studied, emphasizing computational geometric methods. The outline of the dissertation is summarized by Figure 1.1. Results on geometric mechanics for rigid bodies on a Lie group are presented in Chapter 2, and results on geometric optimal control problems are presented in Chapter 4. Chapter 3 and Chapter 5 present computational geometric algorithms for mechanics and optimal control problems; they can be considered as discrete-time analogues of Chapter 2 and Chapter 4, respectively. In each chapter, a general theory is developed first for dynamic systems on an arbitrary Lie group; this general theory is illustrated by several rigid body systems. The content of each chapter is summarized as follows.

*Geometric Mechanics of Rigid Bodies on a Lie Group.* Euler-Lagrange equations for mechanical systems evolving on an abstract Lie group are developed according to Hamilton's principle. The equivalent Hamilton's equations are presented in Section 2.1. The essential idea is to express variations of a curve on a Lie group in terms of Lie algebra elements using the exponential map. Properties of the Euler-Lagrange equations are discussed, and they are applied to several rigid body systems evolving on a Lie group in Section 2.3. These results are extended to mechanical systems on a product of two-spheres in Section 2.2; corresponding examples are given in Section 2.4.

*Computational Geometric Mechanics of Rigid Bodies on a Lie Group.* This chapter is a discrete-time version of Chapter 2. Discrete-time Euler-Lagrange equations and discrete-time Hamilton's

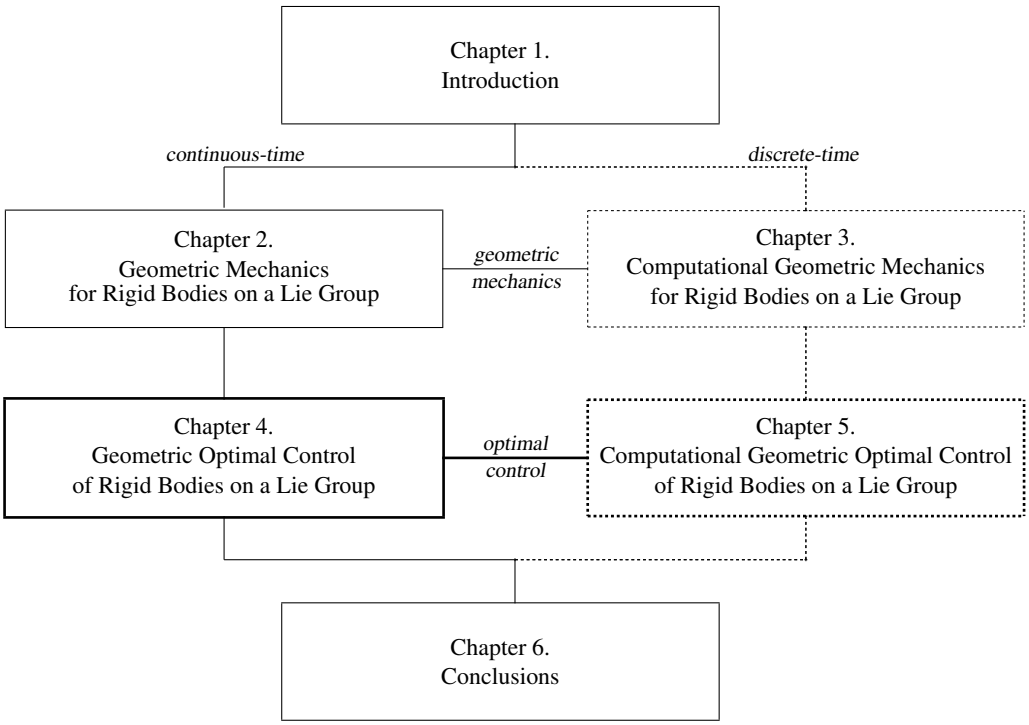



Figure 1.1: Outline of dissertation (solid/dotted: continuous/discrete-time, thin/thick: mechanics/control)

equations, referred to as Lie group variational integrators, are developed in Section 3.1 according to a discrete-time analogue of Hamilton's principle. They are applied to mechanical systems presented in Section 2.3; computational results are summarized in Section 3.3. These results are extended to mechanical systems on a product of two-spheres, to obtain Lie homogeneous variational integrators in Section 3.2; computational results are given in Section 3.4.

*Geometric Optimal Control for Rigid Bodies on a Lie Group.* Based on geometric mechanics on a Lie group developed in Chapter 2, geometric optimal control problems are considered. In Section 4.1, Euler-Lagrange equations are extended to include the effect of control inputs and optimal control problems are formulated. The corresponding necessary conditions for optimality are developed, and they are applied to several optimal control problems for rigid bodies in Section 4.2.

*Computational Geometric Optimal Control for Rigid Bodies on a Lie Group.* This chapter is a discrete-time version of Chapter 4. In Section 5.1, discrete-time forced Euler-Lagrange equations are developed, and a discrete-time optimal control problem is formulated. According to a discrete-time analogue of the calculus of variations, discrete-time necessary conditions for optimality are developed, and they are applied to several discrete-time optimal control problems for rigid bodies in Section 5.2.

## 1.4 Contributions

### 1.4.1 Summary of Contributions

*Coordinate-free approach.* One of the common features of the developments in this dissertation is that the analytical theory and computational methods are developed in terms of a Lie group representation of the configuration of a rigid body system. Therefore, all of the results presented in this dissertation are coordinate-free. Representing geometric objects in terms of coordinates can frequently lead to confusion and complexity, and the corresponding derivations rely on specific choice of coordinates. This dissertation completely avoids local coordinates, thereby expressing the results globally in a compact and elegant manner.

For example, consider the attitude dynamics of a single rigid body. The configuration manifold is $\mathsf{SO}(3)$, but there are numerous attitude parameterizations available (see, for example, Shuster 1993; Stuelpnagel 1964). One of the most popular attitude parameterizations is Euler angles. In addition to the associated singularities, the use of Euler angles can cause confusions since there are 24 types of Euler angles. The use of Euler angles also leads to complicated trigonometric expressions. Other minimal attitude representations have similar difficulties.

Non-minimal representations such as quaternions have no coordinate singularities, but they also introduce certain complications. The group of unit quaternions $\mathsf{SU}(2) \simeq \mathsf{S}^3$ double covers $\mathsf{SO}(3)$, so there is an ambiguity in representing the attitude. While the ambiguity of quaternions is the choice of the sign, there is no consistent way to choose the sign continuously, that is globally valid for $\mathsf{SO}(3)$ (see Marsden and Ratiu 1999). More importantly, the Hamiltonian structure of the attitude dynamics is complicated when it is expressed in terms of quaternions. For instance, it is difficult to express the kinetic energy of a rigid body in terms of a quaternion and its time derivative. It is stated by Leimkuhler and Reich (2004) that although symplectic integration methods based on quaternions can be formulated, approaches based on the rotation matrix are more efficient and conceptually easier to implement. For optimal rigid body control problems, the multiplier equations in necessary conditions for optimality become more complicated if they are written in terms of quaternions (see, for example, Modgalya and Bhat 2006). Therefore, quaternions result in inherent complications when applied to dynamics and control problems for rigid bodies. In many engineering applications, quaternions appear to be simple since they are incorrectly considered to evolve on a flat space, namely $\mathbb{R}^4$, with little attention paid to the unit-length constraint.

In this dissertation, the attitude of a rigid body is represented by a rotation matrix. Geometric numerical integration algorithms and optimal control approaches are directly developed on $\mathsf{SO}(3)$. 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 that allows analysis to be carried out in the Lie algebra that is isomorphic to $\mathbb{R}^3$. For example, necessary conditions for optimality are expressed as compact vector equations on $\mathbb{R}^3$; these equations are more compact than the necessary conditions expressed in terms of quaternions. They also have the advantage of not having singularity or ambiguity.

In summary, this dissertation develops intrinsic, coordinate-free algorithms for computational geometric mechanics and optimal control problems for rigid bodies. All of the analytical and computational results are independent of a specific choice of local coordinates, and they completely avoid any singularity, ambiguity, complexity, and confusion associated with local coordinates.

*Geometric Numerical Integrators on a Lie Group.* The Lie group variational integrators presented in Chapter 3 are geometric numerical integrators for dynamic systems that evolve on a Lie group, such as rigid body dynamics. The variational integrators given in Marsden and West (2001) are geometric integrators that preserve geometric properties of dynamics, but they do not necessarily conserve the nonlinear structure of the configuration manifold. The Lie group method presented in Iserles et al. (2000) is for kinematics equations on a Lie group, so it does not guarantee that the geometric properties of the dynamics are preserved.

This dissertation unifies discrete-time Lagrangian or Hamiltonian mechanics and the Lie group method. A variational integrator is developed in the context of the Lie group method, so that the resulting Lie group variational integrator preserves the geometric properties of the dynamics as well as the structure of the Lie group. For a mechanical system with a Lie group configuration manifold, such as rigid body dynamics, it is shown that the Lie group variational integrator has important computational advantages compared to other geometric integrators that preserve either none or one of these properties (see the numerical example in Section 3.3.6). Due to these superior computational properties, 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. 2006).

Compared with other geometric integrators for a rigid body, as in the work of Hulbert (1992); Krysl (2005); Lewis and Simo (1994); Simo and Wong (1991), the Lie group variational integrator provides a systematic method to obtain a class of numerical integrators that preserve all of the geometric features, rather than developing a specific numerical integrator that preserves only a few geometric characteristics. Compared with discrete-time mechanics on a Lie group developed by Bobenko and Suris (1999); Marsden et al. (1999); Moser and Veselov (1991), the Lie group variational integrator can be applied to a wide class of rigid body dynamics acting under a potential field.

*Optimal Control for Rigid Bodies on a Lie Group.* In Chapter 4, optimal control problems for mechanical systems on a Lie group are formulated, and an intrinsic form of necessary conditions for optimality are developed. Most existing optimal control theory on a Lie group is established based on kinematics equations. For example, an optimal attitude control problem of a rigid body is considered in Jurdjevic (1997) by viewing the angular velocity as a control input. This dissertation deals with optimal control problems of dynamic systems with a Lie group configuration manifold. More precisely, it may be considered as an optimal control problem on a tangent bundle of a Lie group. Compared with the work by Hussein (2005), where optimal control problems on $\mathsf{SO}(3)$ and

$\mathsf{SE}(3)$ are considered, the necessary conditions presented in Section 4.1.3 are applied to a more general class of dynamic systems on an abstract Lie group.

A direct optimal control approach has been applied to discrete-time mechanical systems obtained by variational integrators in Junge et al. (2005, 2006), where control input parameters are optimized using a general constrained parameter optimization scheme such as sequential quadratic programming. The computational geometric optimal control approach presented in Chapter 5 uses a discrete-time analogue of the calculus of variations to derive an intrinsic form of discrete-time optimality conditions, and a computational approach to solve the optimality conditions is presented. Compared with the geometric structure-preserving optimal control approach on $\mathsf{SO}(3)$ by Bloch et al. (2007); Hussein et al. (2006), the discrete-time optimality conditions presented in Section 5.1 can be applied to general optimal control problems on an arbitrary Lie group; they are applied to nontrivial rigid body optimal control problems in Section 5.2.

*Examples of Nontrivial Rigid Body Systems.* In this dissertation, the abstract theory for computational geometric mechanics and optimal control is applied to several nontrivial rigid body systems. For example, in Section 3.3, Lie group variational integrators are developed for a rigid body pendulum, a pendulum with an internal proof mass, a pendulum on a moving cart, rigid bodies acting under mutual potential, and connected rigid bodies. Computational results are also presented for each system. In Section 3.4, several mechanical systems from various scientific fields, such as an elastic rod, magnetic dipoles, and molecular dynamics, are considered. Computational geometric optimal control is applied to minimum time, and minimum fuel optimal control problems of a rigid body, and extended to optimal control problems with symmetry and a combinatorial optimal formation reconfiguration problem in Section 5.2.

New theoretical results in geometric mechanics and control are developed in an abstract form. This dissertation also studies numerous nontrivial rigid body systems that have engineering importance.

### 1.4.2 Publications

The contributions in this dissertation have been published in the following journals, conference proceedings, and book chapter. These contributions that are currently under review/revision are indicated.

**Computational Geometric Mechanics**

- 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*, pages 962–967, 2005.
- E. Fahnestock, T. Lee, M. Leok, N. H. McClamroch, and D. Scheeres. Polyhedral potential and variational integrator computation of the full two body problem. In *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 2006.

- 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.
- 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, 2007.
- T. Lee, M. Leok, and N. H. McClamroch. Lagrangian mechanics and variational integrators on two-spheres. 2007, under revision. URL `http://arxiv.org/abs/0707.0022`.

**Computational Geometric Optimal Control**

- 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.
- T. Lee, M. Leok, and N. H. McClamroch. Optimal control of a rigid body using geometrically exact computations on $\mathsf{SE}(3)$. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2170–2175, 2006.
- T. Lee, M. Leok, and N. H. McClamroch. Discrete control systems. *Invited article for the Encyclopedia of Complexity and System Science*. Springer, 2007.
- T. Lee, M. Leok, and N. H. McClamroch. Computational geometric optimal control of rigid bodies. *Communications in Information and Systems, special issue dedicated to R. W. Brockett*, 2007. submitted.
- T. Lee, M. Leok, and N. H. McClamroch. Optimal attitude control of a rigid body using geometrically exact computations on $\mathsf{SO}(3)$. *Journal of Dynamical and Control Systems*, 2007. accepted.
- T. Lee, M. Leok, and N. H. McClamroch. Optimal attitude control for a rigid body with symmetry. In *Proceedings of the American Control Conference*, pages 1073–1078, 2007.
- T. Lee, M. Leok, and N. H. McClamroch. A combinatorial optimal control problem for spacecraft formation reconfiguration. In *Proceedings of the IEEE Conference on Decision and Control*, pages 5370–5375, 2007.
- T. Lee, M. Leok, and N. H. McClamroch. Time optimal attitude control for a rigid body. In *Proceedings of the American Control Conference*, 2008. accepted.

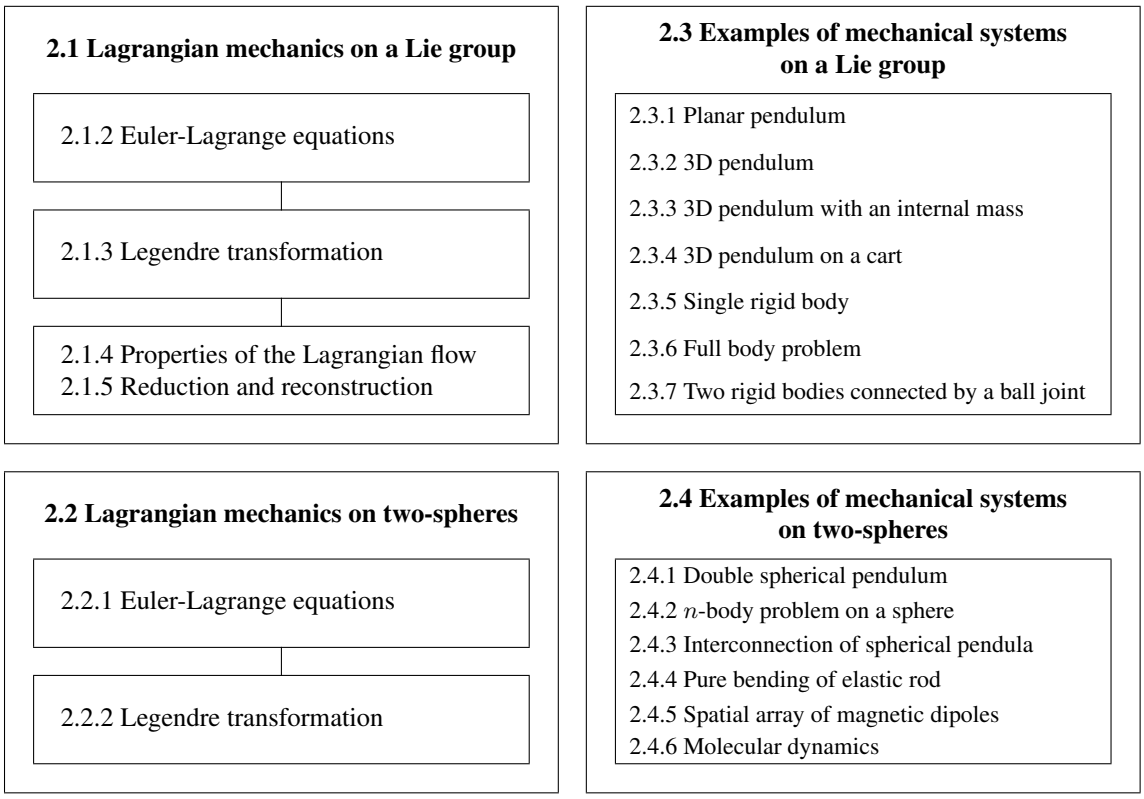
This research has been recognized by the following awards:

- *Rackham International Students Fellowship*, University of Michigan 2006
- *Rackham Predoctoral Fellowship*, University of Michigan 2006-2007
- *SIAM Computational Science and Engineering, BGCE Student Paper Prize, finalist* 2007
- *Distinguished Achievement Award*, College of Engineering, University of Michigan 2008
- *Ivor K. McIvor Award*, College of Engineering, University of Michigan 2008

## CHAPTER 2

# GEOMETRIC MECHANICS FOR RIGID BODIES ON A LIE GROUP

This chapter deals with geometric mechanics for rigid bodies that evolve on a Lie group. The goal is to develop an intrinsic form of Euler-Lagrange equations on an arbitrary Lie group, and to show several properties of Lagrangian flows. These are applied to several rigid body systems evolving on a Lie group, and extended to mechanical systems on a product of two-spheres.




This chapter is organized as follows. In Section 2.1, we develop geometric mechanics on a Lie group; Euler-Lagrange equations are developed and several properties of Lagrangian flow are discussed. These results are applied to several rigid body dynamics in Section 2.3. The remaining part of this chapter develops geometric mechanics on a product of two-spheres. Since the two-sphere is a homogeneous manifold on which a Lie group acts transitively, the Lagrangian mechanics on a Lie group developed in Section 2.1 can be easily extended to the two sphere. Euler-Lagrange equations on a product of two-spheres are developed in Section 2.2, and they are applied to several

mechanical systems in Section 2.4.

## 2.1 Lagrangian Mechanics on a Lie Group

Geometric mechanics is a modern description of classical mechanics from the perspective of differential geometry (see, for example, Abraham and Marsden 1978; Bloch 2003a; Bullo and Lewis 2005; Jurdjevic 1997; Marsden and Ratiu 1999). It explores the geometric structure of a Lagrangian or Hamiltonian system through the concept of vector fields, symplectic geometry, and symmetry techniques. This section develops Lagrange mechanics on a Lie group; Euler-Lagrange equations for a mechanical system evolving on an abstract Lie group are derived, and the symplectic property and symmetry of Lagrangian flow are discussed.

The dynamics of rigid bodies evolve on a Lie group. For example, the configuration manifold for the attitude dynamics of a rigid body is the special orthogonal group $\mathsf{SO}(3)$, and the configuration manifold for combined translational and rotational motion of a rigid body is the special Euclidean group $\mathsf{SE}(3)$. A direct product of Lie groups $\mathsf{SE}(3)$, $\mathsf{SO}(3)$, and $\mathbb{R}^n$ can represent a configuration manifold of multiple rigid bodies, which is also a Lie group.

However, much of the literature on dynamics of rigid bodies relies on local coordinates of a Lie group. For example, a time optimal attitude maneuver of a rigid body is studied in terms of Euler angles by Bilimoria and Wie (1993), and constrained equations of motion in multibody dynamics are developed in terms of local coordinates on a manifold by Yen (1993). As discussed in Section 1.4.1, representing geometric objects in terms of local coordinates frequently leads to confusion and complexity.

The analytical results of this section are coordinate-free; they are independent of a specific choice of local coordinates, and they completely avoid any singularity, ambiguity, and confusion associated with local coordinates. The resulting intrinsic form of the Euler-Lagrange equations are more compact that equations expressed in terms of local coordinates. Since these are developed for mechanical systems that evolves on an arbitrary Lie group, they provides a general framework that can be uniformly applied to dynamics of multiple rigid bodies.

This section is organized as follows. Section 2.1.1 provides preliminaries on a Lie group. Euler-Lagrange equations and Hamilton's equations on an arbitrary Lie group are developed in Section 2.1.2 and in Section 2.1.3, respectively. Properties of Lagrangian flow and Lagrange-Routh reduction are described in Section 2.1.4 and Section 2.1.5.

### 2.1.1 Preliminaries on a Lie Group

We first summarize basic definitions and properties of a Lie group (see, for example, Bloch 2003a; Bullo and Lewis 2005; Marsden and Ratiu 1999; Varadarajan 1984). A *Lie group* is a differentiable manifold that has a group structure such that the group operation is a smooth map. A *Lie algebra* is the tangent space of the Lie group $\mathsf{G}$ at the identity element $e \in \mathsf{G}$, with a *Lie bracket* $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ that is bilinear, skew symmetric, and satisfies the Jacobi identity.

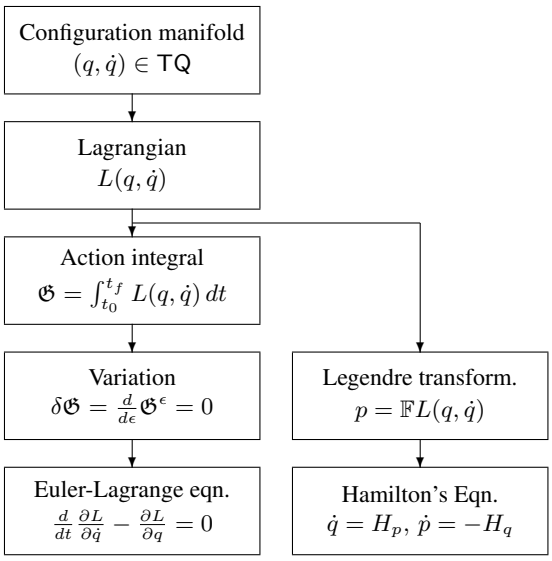



Figure 2.1: Procedures to derive Euler-Lagrange equations

For $g, h \in \mathsf{G}$, the *left translation map* $\mathsf{L}_h : \mathsf{G} \to \mathsf{G}$ is defined as $\mathsf{L}_h g = hg$. Similarly, the *right translation map* $\mathsf{R}_h : \mathsf{G} \to \mathsf{G}$ is defined as $\mathsf{R}_h g = gh$. Given $\xi \in \mathfrak{g}$, define a vector field $X_\xi : \mathsf{G} \to \mathsf{TG}$ such that $X_\xi(g) = \mathsf{T}_e\mathsf{L}_g \cdot \xi$, and let the corresponding unique integral curve passing through the identity $e$ at $t = 0$ be denoted by $\gamma_\xi(t)$. The *exponential map* $\exp : \mathfrak{g} \to \mathsf{G}$ is defined by $\exp \xi = \gamma_\xi(1)$. The $\exp$ is a local diffeomorphism from a neighborhood of zero in $\mathfrak{g}$ onto a neighborhood of $e$ in $\mathsf{G}$.

Define the *inner automorphism* $\mathsf{I}_g : \mathsf{G} \to \mathsf{G}$ as $\mathsf{I}_g(h) = ghg^{-1}$. The *adjoint operator* $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is the differential of $\mathsf{I}_g(h)$ with respect to $h$ at $h = e$ along the direction $\eta \in \mathfrak{g}$, i.e. $\mathrm{Ad}_g\eta = \mathsf{T}_e\mathsf{I}_g \cdot \eta$. The *ad operator* $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ is obtained by differentiating $\mathrm{Ad}_g\eta$ with respect to $g$ at $e$ in the direction $\xi$, i.e. $\mathrm{ad}_\xi\eta = \mathsf{T}_e(\mathrm{Ad}_g\eta) \cdot \xi$. This corresponds to the Lie bracket, i.e. $\mathrm{ad}_\xi\eta = [\xi, \eta]$.

Let $\langle \cdot, \cdot \rangle$ be a pairing between a tangent vector and a cotangent vector. The *coadjoint operator* $\mathrm{Ad}_g^* : G \times \mathfrak{g}^* \to \mathfrak{g}^*$ is defined by $\left\langle \mathrm{Ad}_g^*\alpha,\ \xi \right\rangle = \langle \alpha,\ \mathrm{Ad}_g\xi \rangle$ for $\alpha \in \mathfrak{g}^*$. The *co-ad operator* $\mathrm{ad}^* : \mathfrak{g} \times \mathfrak{g}^* \to \mathfrak{g}^*$ is defined by $\left\langle \mathrm{ad}_\eta^*\alpha,\ \eta \right\rangle = \langle \alpha,\ \mathrm{ad}_\eta\xi \rangle$ for $\alpha \in \mathfrak{g}^*$.

### 2.1.2 Euler-Lagrange Equations

Consider a mechanical system evolving on a Lie group $\mathsf{G}$. We develop the corresponding Euler-Lagrange equations. The procedures to derive Euler-Lagrange equations of a mechanical system are summarized by Figure 2.1: the trajectory of the object is derived by finding the path that minimizes the integral of a Lagrangian over time, called the action integral. The Legendre transformation provides an alternative description of mechanical systems, referred to as Hamiltonian mechanics. The essential idea in applying these procedures on a Lie group $\mathsf{G}$ is expressing the variation of group elements in terms of the Lie algebra $\mathfrak{g}$ using the exponential map.

**Configuration Manifold and Lagrangian**

The configuration manifold is a Lie group $\mathsf{G}$. We identify the tangent bundle $\mathsf{TG}$ with $\mathsf{G} \times \mathfrak{g}$ by left trivialization. For example, a tangent vector $(g, \dot{g}) \in \mathsf{T}_g\mathsf{G}$ is expressed as

$$\dot{g} = \mathsf{T}_e\mathsf{L}_g \cdot \xi = g\xi \tag{2.1}$$

for $\xi \in \mathfrak{g}$. We assume the Lagrangian of the mechanical system is given by $L(g, \xi) : \mathsf{G} \times \mathfrak{g} \to \mathbb{R}$.

**Action Integral**

Define the action integral as

$$\mathfrak{G} = \int_{t_0}^{t_f} L(g, \xi)\, dt.$$

Hamilton's principle states that the variation of the action integral is equal to zero.

$$\delta\mathfrak{G} = \delta \int_{t_0}^{t_f} L(g, \xi)\, dt. \tag{2.2}$$

**Variations**

Let $g(t)$ be a differential curve in $\mathsf{G}$ defined for $t \in [t_0, t_f]$. The variation is a differentiable mapping $g^\epsilon(t) : (-c, c) \times [t_0, t_f] \to \mathsf{G}$ for $c > 0$ such that $g^0(t) = g(t)$ for any $t \in [t_0, t_f]$, and $g^\epsilon(t_0) = g(t_0)$, $g^\epsilon(t_f) = g(t_f)$ for any $\epsilon \in (-c, c)$. We express the variation using the exponential map as

$$g^\epsilon(t) = g \exp \epsilon\eta(t), \tag{2.3}$$

for a curve $\eta(t)$ in $\mathfrak{g}$. It is easy to show that (2.3) is well defined for some constant $c$ as the exponential map is a local diffeomorphism between $\mathfrak{g}$ and $\mathsf{G}$, and it satisfies the properties of the variation provided $\eta(t_0) = \eta(t_f) = 0$. Since this is obtained by a group operation, it is also guaranteed that the variation lies on $\mathsf{G}$ for any $\eta(t)$.

The corresponding infinitesimal variation of $g$ is given by

$$\begin{aligned}\delta g(t) &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} g^\epsilon(t) = \mathsf{T}_e\mathsf{L}_{g(t)} \cdot \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \exp \epsilon\eta(t) \\ &= g(t)\eta(t).\end{aligned} \tag{2.4}$$

For each $t \in [t_0, t_f]$, the infinitesimal variation $\delta g(t)$ lies in the tangent space $\mathsf{T}_{g(t)}\mathsf{G}$. Using this expression and (2.1), the infinitesimal variation of $\xi(t)$ is obtained as follows (see Bloch et al. 1996; Marsden and Ratiu 1999, and Appendix A.4).

$$\delta\xi(t) = \dot{\eta}(t) + \mathrm{ad}_{\xi(t)}\eta(t). \tag{2.5}$$

Equations (2.4) and (2.5) are infinitesimal variations of $(g(t), \xi(t)) : [t_0, t_f] \to G \times \mathfrak{g}$, respectively.

**Euler-Lagrange Equations**

The variation of the Lagrangian can be written as

$$\delta L(g,\xi) = \mathbf{D}_g L(g,\xi)\cdot\delta g + \mathbf{D}_\xi L(g,\xi)\cdot\delta\xi,$$

where $\mathbf{D}_g L \in \mathsf{T}^*\mathsf{G}$ denotes the derivative of the Lagrangian $L$ with respect to $g$, given by

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} L(g^\epsilon,\xi) = \mathbf{D}_g L(g,\xi)\cdot\delta g,$$

and $\mathbf{D}_\xi L(g,\xi) \in \mathfrak{g}^*$ is defined similarly. Since $\mathsf{T}(\mathsf{L}_g \circ \mathsf{L}_{g^{-1}}) = \mathsf{T}\mathsf{L}_g \circ \mathsf{T}\mathsf{L}_{g^{-1}}$ is equal to the identity map on $\mathsf{TG}$, this can be written as

$$\begin{aligned}\delta L(g,\xi) &= \langle \mathbf{D}_g L(g,\xi),\ \delta g\rangle + \langle \mathbf{D}_\xi L(g,\xi),\ \delta\xi\rangle \\ &= \left\langle \mathbf{D}_g L(g,\xi),\ (\mathsf{T}_e\mathsf{L}_g \circ \mathsf{T}_g\mathsf{L}_{g^{-1}})\cdot\delta g\right\rangle + \langle \mathbf{D}_\xi L(g,\xi),\ \delta\xi\rangle\,.\end{aligned}$$

Substituting (2.4) and (2.5), we obtain

$$\begin{aligned}\delta L(g,\xi) &= \langle \mathbf{D}_g L(g,\xi),\ \mathsf{T}_e\mathsf{L}_g\cdot\eta\rangle + \langle \mathbf{D}_\xi L(g,\xi),\ \dot\eta + \mathrm{ad}_\xi\eta\rangle \\ &= \left\langle \mathsf{T}_e^*\mathsf{L}_g\cdot\mathbf{D}_g L(g,\xi) + \mathrm{ad}_\xi^*\cdot\mathbf{D}_\xi L(g,\xi),\ \eta\right\rangle + \langle \mathbf{D}_\xi L(g,\xi),\ \dot\eta\rangle\,. \qquad (2.6)\end{aligned}$$

Since variation and integration commute, the variation of the action integral is given by

$$\delta\mathfrak{G} = \int_{t_0}^{t_f} \delta L(g,\xi)\,dt.$$

Substituting (2.6) and using integration by parts, the variation of the action integral is given by

$$\begin{aligned}\delta\mathfrak{G} &= \int_{t_0}^{t_f} \left\langle \mathsf{T}_e^*\mathsf{L}_g\cdot\mathbf{D}_g L(g,\xi) + \mathrm{ad}_\xi^*\cdot\mathbf{D}_\xi L(g,\xi),\ \eta\right\rangle + \langle \mathbf{D}_\xi L(g,\xi),\ \dot\eta\rangle\ dt \\ &= \langle \mathbf{D}_\xi L(g,\xi),\ \eta\rangle\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \left\langle \mathsf{T}_e^*\mathsf{L}_g\cdot\mathbf{D}_g L(g,\xi) + \mathrm{ad}_\xi^*\cdot\mathbf{D}_\xi L,\ \eta\right\rangle - \left\langle \frac{d}{dt}\mathbf{D}_\xi L(g,\xi),\ \eta\right\rangle dt. \qquad (2.7)\end{aligned}$$

Since $\eta(t) = 0$ at $t = t_0$ and $t = t_f$, the first term of the above equation vanishes. Thus, we obtain

$$\delta\mathfrak{G} = \int_{t_0}^{t_f} \left\langle \mathsf{T}_e^*\mathsf{L}_g\cdot\mathbf{D}_g L(g,\xi) + \mathrm{ad}_\xi^*\cdot\mathbf{D}_\xi L(g,\xi),\ \eta\right\rangle - \left\langle \frac{d}{dt}\mathbf{D}_\xi L(g,\xi),\ \eta\right\rangle dt. \qquad (2.8)$$

From Hamilton's principle, $\delta\mathfrak{G} = 0$ for all $\eta \in \mathfrak{g}$, which yields the Euler-Lagrange equations on $\mathsf{G}$.

**Proposition 2.1** *Consider a mechanical system evolving on a Lie group* $\mathsf{G}$*. We identify the tangent bundle* $\mathsf{TG}$ *with* $\mathsf{G}\times\mathfrak{g}$ *by left trivialization. Suppose that the Lagrangian is defined as* $L(g,\xi)$ :

$\mathsf{G} \times \mathfrak{g} \to \mathbb{R}$. *The corresponding Euler-Lagrange equations are given by*

$$\frac{d}{dt}\mathbf{D}_\xi L(g,\xi) - \mathrm{ad}^*_\xi \cdot \mathbf{D}_\xi L(g,\xi) - \mathsf{T}^*_e\mathsf{L}_g \cdot \mathbf{D}_g L(g,\xi) = 0, \tag{2.9}$$

$$\dot{g} = g\xi. \tag{2.10}$$

**Remark 2.1** The essential idea of this development is expressing the variation of a curve in $\mathsf{G}$ using the exponential map, as given by (2.3). The expression for the variation is carefully chosen such that the varied curve lies on the configuration manifold $\mathsf{G}$. The use of the exponential map $\exp : \mathfrak{g} \to \mathsf{G}$ is desirable in two aspects: (i) since the variation is obtained by a group operation, it is guaranteed to lie on $\mathsf{G}$, and (ii) the variation is parameterized by a curve in a linear vector space $\mathfrak{g}$.

**Remark 2.2** If the Lagrangian is not dependent on $\mathsf{G}$, the third term of (2.9) vanishes. The resulting equation is equivalent to the Euler-Poincaré equation, and (2.10) is a reconstruction equation (see Marsden and Ratiu 1999). Therefore, (2.9) can be considered as a generalization of the Euler-Poincaré equation.

**Remark 2.3** These equations are obtained using the left trivialization. Therefore, the velocity $\xi$ may be considered as a quantity expressed in the body fixed frame. We can develop similar equations using the right trivialization to obtain the equations of motion expressed in the reference frame. This is summarized by the following corollary.

**Corollary 2.1** *Consider a mechanical system evolving on a Lie group* $\mathsf{G}$. *We identify the tangent bundle* $\mathsf{TG}$ *with* $\mathsf{G} \times \mathfrak{g}$ *by right trivialization. Suppose that the Lagrangian is defined as* $L(g,\varsigma) : \mathsf{G} \times \mathfrak{g} \to \mathbb{R}$. *The corresponding Euler-Lagrange equations are given by*

$$\frac{d}{dt}\mathbf{D}_\varsigma L(g,\varsigma) + \mathrm{ad}^*_\varsigma \cdot \mathbf{D}_\varsigma L(g,\varsigma) - \mathsf{T}^*_e\mathsf{R}_g \cdot \mathbf{D}_g L(g,\varsigma) = 0, \tag{2.11}$$

$$\dot{g} = \varsigma g. \tag{2.12}$$

### 2.1.3 Legendre Transformation

We identify the tangent bundle $\mathsf{TG}$ with $\mathsf{G} \times \mathfrak{g}$ using the left trivialization. Using this, the cotangent bundle $\mathsf{T}^*\mathsf{G}$ can be identified with $\mathsf{G} \times \mathfrak{g}^*$. For the given Lagrangian, the Legendre transformation $\mathbb{F}L : \mathsf{G} \times \mathfrak{g} \to \mathsf{G} \times \mathfrak{g}^*$ is defined as

$$\mathbb{F}L(g,\xi) = (g,\mu), \tag{2.13}$$

where $\mu \in \mathfrak{g}^*$ is given by

$$\mu = \mathbf{D}_\xi L(g,\xi). \tag{2.14}$$

If the Legendre transformation is a diffeomorphism, the corresponding Lagrangian is called a hyperregular Lagrangian, which induces a Hamiltonian system on $\mathsf{G} \times \mathfrak{g}^*$.

**Corollary 2.2** *Consider a mechanical system evolving on a Lie group* $\mathsf{G}$*. We identify the tangent bundle* $\mathsf{TG}$ *with* $\mathsf{G} \times \mathfrak{g}$ *by left trivialization. Suppose that the Lagrangian given by* $L(g, \xi) : \mathsf{G} \times \mathfrak{g} \to \mathbb{R}$ *is hyperregular. Then, the Legendre transformation yields Hamilton's equations that are equivalent to the Euler-Lagrange equations presented in Proposition 2.1.*

$$\frac{d}{dt}\mu - \mathrm{ad}^*_\xi \mu - \mathsf{T}^*_e \mathsf{L}_g \cdot \mathbf{D}_g L(g, \xi) = 0, \tag{2.15}$$

$$\dot{g} = g\xi, \tag{2.16}$$

*where* $\mu = \mathbf{D}_\xi L(g, \xi) \in \mathfrak{g}^*$.

### 2.1.4 Properties of the Lagrangian Flow

Here we show two properties of the Lagrangian flow, namely symplecticity and momentum preservation. The subsequent development can be considered as a special form of the general properties of Lagrangian flows, applied to a Lie group configuration manifold (see Marsden and West 2001).

#### Symplecticity

Let $\Theta_L$ be the Lagrangian one-form on $\mathsf{G} \times \mathfrak{g}$ given by

$$\Theta_L(g, \xi) \cdot (\delta g, \delta \xi) = \left\langle \mathbf{D}_\xi L(g, \xi),\ g^{-1}\delta g \right\rangle \tag{2.17}$$

The Lagrangian symplectic two-form $\Omega_L$ is the exterior derivative of the Lagrangian one-forme, i.e. $\Omega_L = \mathbf{d}\Theta_L$. We define the Lagrangian flow map $\mathcal{F}_L : (\mathsf{G} \times \mathfrak{g}) \times [0, t_f - t_0] \to (\mathsf{G} \times \mathfrak{g})$ as the flow of (2.9) and (2.10).

**Proposition 2.2** *The Lagrangian flow preserves the Lagrangian symplectic two-form as follows*

$$(\mathcal{F}_L^T)^* \Omega_L = \Omega_L \tag{2.18}$$

*for* $T = t_f - t_0$.

**Proof.** Define the solution space $\mathcal{C}_L$ to be the set of solutions $g(t) : [t_0, t_f] \to G$ of (2.9) and (2.10). Since an element of $\mathcal{C}_L$ is uniquely determined by the initial condition $(g(0), \xi(0)) \in G \times \mathfrak{g}$, we can identify $\mathcal{C}_L$ with the space of initial conditions $\mathsf{G} \times \mathfrak{g}$. Define the restricted action map $\hat{\mathfrak{G}} : \mathsf{G} \times \mathfrak{g} \to \mathbb{R}$ by

$$\hat{\mathfrak{G}}(g_0, \xi_0) = \mathfrak{G}(g(t)),$$

where $g(t) \in \mathcal{C}_L$ with $(g(0), g^{-1}(0)\dot{g}(0)) = (g_0, \xi_0)$. Since the curve $g(t)$ satisfies (2.9) and (2.10), (2.7) reduces to

$$\mathbf{d}\hat{\mathfrak{G}} \cdot w = ((\mathcal{F}_L^T)^* \Theta_L - \Theta_L) \cdot w \tag{2.19}$$

for any $w = (\delta g_0, \delta \xi_0) \in \mathsf{T}(G \times \mathfrak{g})$. We take the exterior derivative of (2.19). Since exterior derivatives and pull back commute, we obtain

$$\mathbf{d^2}\hat{\mathfrak{G}} = ((\mathcal{F}_L^T)^* \mathbf{d}\Theta_L - \mathbf{d}\Theta_L).$$

Since $\mathbf{d}^2\hat{\mathfrak{G}} = 0$ for any zero-form $\mathfrak{G}$, we obtain (2.18). ■

**Noether's Theorem**

Suppose that a Lie group $\mathsf{H}$, with the Lie algebra $\mathfrak{h}$, acts on $\mathsf{G}$. A left action of $\mathsf{H}$ on $\mathsf{G}$ is a smooth mapping $\Phi : \mathsf{H} \times \mathsf{G} \to \mathsf{G}$ such that $\Phi(e, g) = g$, and $\Phi(h, \Phi(h', g)) = \Phi(hh', g)$ for any $g \in \mathsf{G}$ and $h, h' \in \mathsf{H}$. Let $\Phi_h : \mathsf{G} \to \mathsf{G}$ be defined such that $\Phi_h(g) = \Phi(h, g)$.

Let $\phi_L : \mathsf{TG} \to G \times \mathfrak{g}$ be the left trivialization given by $\phi_L(g, \dot{g}) = (g, g^{-1}\dot{g})$. For $\zeta \in \mathfrak{h}$, the infinitesimal generators $\zeta_{\mathsf{G}} : \mathsf{G} \to \mathsf{G} \times \mathfrak{g}$, and $\zeta_{\mathsf{G}\times\mathfrak{g}} : \mathsf{G} \times \mathfrak{g} \to \mathsf{T}(\mathsf{G} \times \mathfrak{g})$ for the action are defined by

$$\zeta_{\mathsf{G}}(g) = \phi_L \circ \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \Phi_{\exp_{\mathsf{H}} \epsilon\zeta}(g), \tag{2.20}$$

$$\zeta_{\mathsf{G}\times\mathfrak{g}}(g, \xi) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \phi_L \circ \mathsf{T}_g\Phi_{\exp_{\mathsf{H}} \epsilon\zeta}(g) \cdot (\phi_L^{-1}(g, \xi)). \tag{2.21}$$

We define the Lagrangian momentum map $J_L : \mathsf{G} \times \mathfrak{g} \to \mathfrak{h}^*$ to be

$$J_L(g, \xi) \cdot \zeta = \Theta_L \cdot \zeta_{\mathsf{G}\times\mathfrak{g}}(g, \xi). \tag{2.22}$$

**Proposition 2.3** *Suppose that the Lagrangian is infinitesimally invariant under the lifted action, i.e.* $\mathbf{d}L(g, \xi) \cdot \zeta_{\mathsf{G}\times\mathfrak{g}} = 0$ *for any* $\zeta \in \mathfrak{h}$. *Then, the Lagrangian flow preserves the momentum map.*

$$J_L(\mathcal{F}_L^T(g, \xi)) = J_L(g, \xi), \tag{2.23}$$

*This is referred to as Noether's theorem.*

**Proof.** Since the action is the integral of the Lagrangian, $\mathbf{d}L(g, \xi) \cdot \zeta_{G\times\mathfrak{g}} = 0$ implies that $\mathbf{d}\mathfrak{G} \cdot \zeta_{G\times\mathfrak{g}} = 0$, where we consider that the group action $\Phi_h$ is applied to each point of a curve. The invariance of the action integral implies that the action maps a solution curve to another solution curve. Thus, we can restrict $\mathbf{d}\mathfrak{G} \cdot \zeta_{G\times\mathfrak{g}} = 0$ to the solution space to obtain

$$\mathbf{d}\hat{\mathfrak{G}} \cdot \zeta_{G\times\mathfrak{g}} = 0.$$

But, from (2.19), we have

$$\mathbf{d}\hat{\mathfrak{G}} \cdot \zeta_{G\times\mathfrak{g}} = ((\mathcal{F}_L^T)^*\Theta_L - \Theta_L) \cdot \zeta_{G\times\mathfrak{g}} = 0 \tag{2.24}$$

for any $\zeta \in \mathfrak{h}$. Substituting the definition of the momentum map given by (2.22) into this, we obtain (2.23). ■

### 2.1.5 Reduction and Reconstruction

We have shown that if the Lagrangian is infinitesimally invariant under the lifted action of a Lie group $\mathsf{H}$ on $\mathsf{G}$, the corresponding momentum map is preserved along the Lagrangian flow. Suppose that the Lie group $\mathsf{H}$ acts freely and properly on $\mathsf{G}$, and the Lagrangian is invariant under the action $\mathsf{H}$. This is referred to as the symmetry of the Lagrangian. Then, the configuration manifold can be reduced to a quotient space $\mathsf{G}/\mathsf{H}$, referred to as the *shape space*. Because the action is free and proper, it is guaranteed that the shape space is a smooth manifold.

More precisely, for a given curve $g(t)$ in the solution space $\mathcal{C}_L$ and the corresponding value of the momentum map, there exists a unique curve in the shape space $\mathsf{G}/\mathsf{H}$ satisfying reduced Euler-Lagrange equations. If the initial condition $g(t_0)$ is known, the curve $g(t)$ in $\mathsf{G}$ can be reconstructed from the solution of the reduced Euler-Lagrange equations in the shape space $\mathsf{G}/\mathsf{H}$. These are referred to as reduction and reconstruction for mechanical systems.

The procedure for the Lagrangian-Routh reduction and reconstruction is as follows (see Marsden et al. 2000). We define a mechanical connection $\mathcal{A} : \mathsf{G} \times \mathfrak{g} \to \mathfrak{h}$ from the momentum map. This yields a one-form $\mathcal{A}_\nu$ on $\mathsf{G} \times \mathfrak{g}$ paired with the value of the momentum map $\nu \in \mathfrak{h}^*$. The Routhian $R^\nu : \mathsf{G} \times \mathfrak{g} \to \mathbb{R}$ is defined by subtracting the one-form $\mathcal{A}_\nu$ from the Lagrangian. The Routhian satisfies the Lagrange-d'Alembert principle with the magnetic two-form obtained from the exterior derivative of the one-form $\mathcal{A}_\nu$. This form of the variational principle and the Routhian reduce onto the $\nu$-level set of the momentum map, which provides the reduced Euler-Lagrange equations on the shape space.

For a given solution of the reduced Euler-Lagrange equations, we find the horizontal lift of the curve on $\mathsf{G}$. Applying the mechanical connection to the time derivatives of the lifted curve provides a reconstruction equation. A particular example for the Lagrange-Routh reduction and reconstruction is presented in Appendix A.3.

## 2.2 Lagrangian Mechanics on Two-Spheres

In the previous section, we have developed Euler-Lagrange equations for mechanical systems evolving on a Lie group, and the symplectic property and symmetry of the Lagrangian flow are discussed. The essential idea in developing the Euler-Lagrange equations on a Lie group is to express the variation of a curve on a Lie group in terms of a curve on the corresponding Lie algebra using the exponential map.

In this section, we develop Euler-Lagrange equations for mechanical systems evolving on a product of two-spheres

$$\mathsf{S}^2 = \{q \in \mathbb{R}^3 \,|\, q \cdot q = 1\}.$$

The two-sphere $\mathsf{S}^2$ is not a Lie group, but the special orthogonal group $\mathsf{SO}(3) = \{R \in \mathbb{R}^{3\times 3} \,|\, R^T R = I, \det R = 1\}$ acts on the two-sphere transitively, i.e. for any $q_1, q_2 \in \mathsf{S}^2$, there exists a $R \in \mathsf{SO}(3)$ such that $q_2 = Rq_1$. Therefore, we can express the variation of a curve on $\mathsf{S}^2$ in terms of a curve on $\mathfrak{so}(3) \simeq \mathbb{R}^3$ using the exponential map of $\mathsf{SO}(3)$, from which Euler-Lagrange equations can be developed.

In most of the literature that treats dynamic systems on $(\mathsf{S}^2)^n$, either $2n$ angles or $n$ explicit equality constraints enforcing unit length are used to describe the configuration of the system (see, for example, Bendersky and Sandler 2006; Marsden et al. 1993). These descriptions involve complicated trigonometric expressions and introduce additional complexity in analysis and computations.

In this section, we develop Euler-Lagrange equations on $(\mathsf{S}^2)^n$ without need of local parameterizations, constraints, or reprojections. This yields a remarkably compact form of the equations of motion, and also provides insight into the global dynamics on $(\mathsf{S}^2)^n$. A manifold on which a Lie group acts in a transitively way is referred to as a homogeneous manifold. The key idea of this development can be generalized to an abstract homogeneous manifold.

### 2.2.1 Euler-Lagrange equations

The procedures to derive Euler-Lagrange equations are summarized by Figure 2.1: the trajectory of the object is derived by finding the path that minimizes the integral of a Lagrangian over time, called the action integral. The Legendre transformation provides an alternative description of mechanical systems, referred to as Hamiltonian mechanics. The essential idea is to express the variation of a curve on $\mathsf{S}^2$ in terms of the Lie algebra $\mathfrak{so}(3)$ using the exponential map.

#### Configuration Manifold and Lagrangian

The two-sphere is the set of points that have unit length from the origin of $\mathbb{R}^3$, i.e. $\mathsf{S}^2 = \{q \in \mathbb{R}^3 \,|\, q \cdot q = 1\}$. The tangent space $\mathsf{T}_q\mathsf{S}^2$ for $q \in \mathsf{S}^2$ is a plane tangent to the two-sphere at the point $q$. Thus, a curve $q : \mathbb{R} \to \mathsf{S}^2$ and its time derivative satisfy $q \cdot \dot{q} = 0$. The time-derivative of a curve can be written as

$$\dot{q} = \omega \times q, \tag{2.25}$$

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. $q \cdot \dot{\omega} = 0$.

We consider a mechanical system evolving on an $n$ product of two-spheres, $\mathsf{S}^2 \times \cdots \times \mathsf{S}^2 = (\mathsf{S}^2)^n$. We assume that the Lagrangian $L : \mathsf{T}(\mathsf{S}^2)^n \to \mathbb{R}$ is given by the difference between a quadratic kinetic energy and a configuration-dependent potential energy as follows.

$$L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) = \frac{1}{2} \sum_{i,j=1}^{n} M_{ij} \dot{q}_i \cdot \dot{q}_j - U(q_1, \ldots, q_n), \tag{2.26}$$

where $(q_i, \dot{q}_i) \in \mathsf{TS}^2$ 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 $U : (\mathsf{S}^2)^n \to \mathbb{R}$. Here, we assume the inertia matrix is constant, but it can be readily generalized to mechanical systems with a configuration dependent inertia.

### Action Integral

Define the action integral as

$$\mathfrak{G} = \int_{t_0}^{t_f} L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n)\, dt.$$

Hamilton's principle states that the variation of the action integral is equal to zero.

### Variations

Let $q_i(t)$ be a differentiable curve in $\mathsf{S}^2$ defined for $t \in [t_0, t_f]$. The variation is a differentiable mapping $q_i^\epsilon(t) : (-c, c) \times [t_0, t_f] \to \mathsf{S}^2$ for $c > 0$ such that $q_i^0(t) = q_i(t)$ for any $t \in [t_0, t_f]$ and $q_i^\epsilon(t_0) = q_i(t_0)$, $q_i^\epsilon(t_f) = q(t_f)$ for any $\epsilon \in (-c, c)$. Since the special orthogonal group $\mathsf{SO}(3)$ acts on $\mathsf{S}^2$ in a transitive way, we can express the variation of $q_i(t)$ using the exponential map on $\mathsf{SO}(3)$ as follows.

$$q_i^\epsilon(t) = \exp \epsilon\hat{\eta}_i(t)\, q_i(t) \tag{2.27}$$

for a curve $\eta_i(t)$ in $\mathbb{R}^3$. It is easy to show that (2.27) is well defined since the exponential map is a local diffeomorphism between $\mathfrak{so}(3)$ and $\mathsf{SO}(3)$, and $\mathsf{SO}(3)$ acts on $\mathsf{S}^2$. It satisfies other properties of the variation provided that $\eta_i(t_0) = \eta_i(t_f) = 0$. We assume $\eta_i(t) \cdot q_i(t) = 0$ for $t \in [t_0, t_f]$.

The corresponding infinitesimal variation is given by

$$\delta q_i(t) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} q^\epsilon(t) = \hat{\eta}_i(t) q(i) = \eta_i(t) \times q_i(t). \tag{2.28}$$

Since the variation and the differentiation commute, the expression for the infinitesimal variation of $\dot{q}_i(t)$ is given by

$$\delta\dot{q}_i(t) = \dot{\eta}_i(t) \times q_i(t) + \eta_i(t) \times \dot{q}_i(t). \tag{2.29}$$

These expressions are key elements to derive the Euler-Lagrange equations on $(\mathsf{S}^2)^n$.

### Euler-Lagrange Equations

The variation of the Lagrangian can be written as

$$\delta L = \sum_{i,j=1}^{n} \delta\dot{q}_i \cdot M_{ij}\dot{q}_j - \sum_{i=1}^{n} \delta q_i \cdot \frac{\partial U}{\partial q_i},$$

where the symmetric property $M_{ij} = M_{ji}$ is used. Substituting (2.28) and (2.29) 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

$$\delta L = \sum_{i,j=1}^{n} \dot{\eta}_i \cdot (q_i \times M_{ij}\dot{q}_j) + \eta_i \cdot (\dot{q}_i \times M_{ij}\dot{q}_j) - \sum_{i=1}^{n} \eta_i \cdot \left( q_i \times \frac{\partial U}{\partial q_i} \right).$$

Using the above equation and integrating by parts, the variation of the action integral is given by

$$\delta \mathfrak{G} = \sum_{i,j=1}^{n} \eta_i \cdot (q_i \times M_{ij}\dot{q}_j)\Big|_{t_0}^{t_f} - \sum_{i=1}^{n} \int_{t_0}^{t_f} \eta_i \cdot \left[ (q_i \times \sum_{j=1}^{n} M_{ij}\ddot{q}_j) + q_i \times \frac{\partial U}{\partial q_i} \right] dt.$$

From Hamilton's principle $\delta \mathfrak{G} = 0$ for any $\eta_i$ vanishing at $t_0$, and $t_f$. Since $\eta_i$ is orthogonal to $q_i$, the continuous equations of motion are given by

$$(q_i \times \sum_{j=1}^{n} M_{ij}\ddot{q}_j) + q_i \times \frac{\partial U}{\partial q_i} = c_i q_i \tag{2.30}$$

for a curve $c_i(t)$ in $\mathbb{R}$ for $i \in \{1, \ldots, n\}$. Taking the cross product of (2.30) and $q_i$ yields

$$q_i \times (q_i \times \sum_{j=1}^{n} M_{ij}\ddot{q}_j) + q_i \times \left( q_i \times \frac{\partial U}{\partial q_i} \right) = 0. \tag{2.31}$$

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 (2.31), we obtain an expression for $\ddot{q}_i$, which is summarized as follows.

**Proposition 2.4** *Consider a mechanical system on $(\mathsf{S}^2)^n$ whose Lagrangian is expressed as (2.26). The Euler-Lagrange equations are given by*

$$M_{ii}\ddot{q}_i = q_i \times (q_i \times \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{ij}\ddot{q}_j) - (\dot{q}_i \cdot \dot{q}_i)M_{ii}q_i + q_i \times \left( q_i \times \frac{\partial U}{\partial q_i} \right) \tag{2.32}$$

*for $i \in \{1, \ldots, n\}$. Equivalently, this can be written in a matrix form as*

$$\begin{bmatrix} M_{11}I_{3\times 3} & -M_{12}\hat{q}_1\hat{q}_1 & \cdots & -M_{1n}\hat{q}_1\hat{q}_1 \\ -M_{21}\hat{q}_2\hat{q}_2 & M_{22}I_{3\times 3} & \cdots & -M_{2n}\hat{q}_2\hat{q}_2 \\ \vdots & \vdots & & \vdots \\ -M_{n1}\hat{q}_n\hat{q}_n & -M_{n2}\hat{q}_n\hat{q}_n & \cdots & M_{nn}I_{3\times 3} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \vdots \\ \ddot{q}_n \end{bmatrix} = \begin{bmatrix} -(\dot{q}_1 \cdot \dot{q}_1)M_{11}q_1 + \hat{q}_1^2 \frac{\partial U}{\partial q_1} \\ -(\dot{q}_2 \cdot \dot{q}_2)M_{22}q_2 + \hat{q}_2^2 \frac{\partial U}{\partial q_2} \\ \vdots \\ -(\dot{q}_n \cdot \dot{q}_n)M_{nn}q_n + \hat{q}_n^2 \frac{\partial U}{\partial q_n} \end{bmatrix}. \tag{2.33}$$

Since $\dot{q}_i = \omega_i \times q_i$ for the angular velocity $\omega_i$ satisfying $q_i \cdot \omega_i = 0$, we have

$$\ddot{q}_i = \dot{\omega}_i \times q_i + \omega_i \times (\omega_i \times q_i) = \dot{\omega}_i \times q_i - (\omega_i \cdot \omega_i) q_i.$$

Substituting this into (2.30) and using the fact that $q_i \cdot \dot{\omega}_i = 0$, we obtain the Euler-Lagrange equations in terms of the angular velocity.

**Corollary 2.3** *The Euler-Lagrange equations on $(\mathsf{S}^2)^n$ given by (2.32) can be written in terms of the angular velocity as*

$$M_{ii}\dot{\omega}_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} (M_{ij} q_i \times (q_j \times \dot{\omega}_j) + M_{ij}(\omega_j \cdot \omega_j) q_i \times q_j) - q_i \times \frac{\partial U}{\partial q_i}, \tag{2.34}$$

$$\dot{q}_i = \omega_i \times q_i \tag{2.35}$$

*for $i \in \{1, \ldots, n\}$. Equivalently, (2.34) can be written in a matrix form as*

$$\begin{bmatrix} M_{11} I_{3\times 3} & -M_{12}\hat{q}_1\hat{q}_2 & \cdots & -M_{1n}\hat{q}_1\hat{q}_n \\ -M_{21}\hat{q}_2\hat{q}_1 & M_{22} I_{3\times 3} & \cdots & -M_{2n}\hat{q}_2\hat{q}_n \\ \vdots & \vdots & & \vdots \\ -M_{n1}\hat{q}_n\hat{q}_1 & -M_{n2}\hat{q}_n\hat{q}_2 & \cdots & M_{nn} I_{3\times 3} \end{bmatrix} \begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \\ \vdots \\ \dot{\omega}_n \end{bmatrix} = \begin{bmatrix} \sum_{j=2}^{n} M_{1j}(\omega_j \cdot \omega_j)\hat{q}_1 q_j - \hat{q}_1 \frac{\partial U}{\partial q_1} \\ \sum_{j=1, j\neq 2}^{n} M_{2j}(\omega_j \cdot \omega_j)\hat{q}_2 q_j - \hat{q}_2 \frac{\partial U}{\partial q_2} \\ \vdots \\ \sum_{j=1}^{n-1} M_{nj}(\omega_j \cdot \omega_j)\hat{q}_n q_j - \hat{q}_n \frac{\partial U}{\partial q_n} \end{bmatrix}. \tag{2.36}$$

Equations (2.32)–(2.36) are global continuous equations of motion for a mechanical system on $(\mathsf{S}^2)^n$. They avoid singularities completely, and they preserve the structure of $\mathsf{T}(\mathsf{S}^2)^n$ automatically, if an initial condition is chosen properly. These equations are useful to understand global characteristics of the dynamics. In addition, these expressions are remarkably more compact than the equations of motion written in terms of any local parametrization.

We need to check that the $3n \times 3n$ matrices given by the first terms of (2.33) and (2.36) are nonsingular. This is a property of the mechanical system itself, rather than a consequence of these particular form of the 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 nonsingular. One may show a similar property for $n > 2$. Throughout this dissertation, it is assumed that the $3n \times 3n$ matrices given by the first terms of (2.33) and (2.36) are nonsingular.

### 2.2.2 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 chapter for the discrete equations of motion. For $q_i \in \mathsf{S}^2$, the corresponding conjugate momentum $p_i$ lies in the dual space $\mathsf{T}^*_{q_i}\mathsf{S}^2$. We identify the tangent space $\mathsf{T}_{q_i}\mathsf{S}^2$ and its dual space $\mathsf{T}^*_{q_i}\mathsf{S}^2$ 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 $\delta q_i$ 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 $q_i$,

$$\begin{aligned} p_i &= \sum_{j=1}^{n} (M_{ij}\dot{q}_j - (q_i \cdot M_{ij}\dot{q}_j) q_i) = \sum_{j=1}^{n} ((q_i \cdot q_i) M_{ij}\dot{q}_j - (q_i \cdot M_{ij}\dot{q}_j) q_i) \\ &= M_{ii}\dot{q}_i - q_i \times (q_i \times \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{ij}\dot{q}_j). \end{aligned} \tag{2.37}$$

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 (2.32), and using the vector identity $a \times (b \times c) = (a \cdot c)b - (a \cdot b)c$, we obtain the Hamilton's equations.

**Corollary 2.4** *Consider a mechanical system on $(\mathsf{S}^2)^n$ whose Lagrangian is expressed as (2.26). The Hamilton's equations are given by*

$$p_i = M_{ii}\dot{q}_i - q_i \times (q_i \times \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{ij}\dot{q}_j), \tag{2.38}$$

$$\dot{p}_i = -\sum_{j=1}^{n} (\dot{q}_i \cdot M_{ij}\dot{q}_j) q_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} (q_i \cdot M_{ij}\dot{q}_j)\dot{q}_i + q_i \times \left( q_i \times \frac{\partial U}{\partial q_i} \right) \tag{2.39}$$

*for $i \in \{1, \ldots, n\}$. Equivalently, (2.38) can be written in a matrix form as*

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \vdots \\ \dot{q}_n \end{bmatrix} = \begin{bmatrix} M_{11} I_{3\times 3} & -M_{12}\hat{q}_1\hat{q}_1 & \cdots & -M_{1n}\hat{q}_1\hat{q}_1 \\ -M_{21}\hat{q}_2\hat{q}_2 & M_{22} I_{3\times 3} & \cdots & -M_{2n}\hat{q}_2\hat{q}_2 \\ \vdots & \vdots & & \vdots \\ -M_{n1}\hat{q}_n\hat{q}_n & -M_{n2}\hat{q}_n\hat{q}_n & \cdots & M_{nn} I_{3\times 3} \end{bmatrix}^{-1} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}. \tag{2.40}$$

## 2.3 Examples of Mechanical Systems on a Lie Group

In Section 2.1, we have developed Lagrangian mechanics on an abstract Lie group. Since the configuration manifold of dynamics of rigid bodies is a Lie group, the result provides a unified framework that can be applied to various rigid body dynamics.

In this section, we apply the general theory developed in Section 2.1 to the following rigid body dynamics. For each example, a rigid body model is defined, and the corresponding expression for the Lagrangian is derived. Euler-Lagrange equations and Legendre transformations are obtained from Proposition 2.1 and Corollary 2.2.

| Section | Mechanical System | G |
|---|---|---|
| 2.3.1 | Planar pendulum | $\mathsf{SO}(2)$ |
| 2.3.2 | 3D pendulum | $\mathsf{SO}(3)$ |
| 2.3.3 | 3D pendulum with an internal degree of freedom | $\mathsf{SO}(3) \times \mathbb{R}$ |
| 2.3.4 | 3D pendulum on a cart | $\mathsf{SO}(3) \times \mathbb{R}^2$ |
| 2.3.5 | Single rigid body | $\mathsf{SE}(3)$ |
| 2.3.6 | Full body problem | $(\mathsf{SE}(3))^n$ |
| 2.3.7 | Two rigid bodies connected by a ball joint | $\mathsf{SO}(3) \times \mathsf{SO}(3) \times \mathbb{R}^3$ |

### 2.3.1 Planar Pendulum

We consider a planar pendulum model; a mass particle with mass $m$, connected to a frictionless pivot point by a rigid massless link with length $l$ under a uniform gravitational potential.

*Configuration manifold.* Consider a reference frame, and a body fixed frame that is attached to the pendulum. We assume that the origin of these frames is located at the pivot point, and the second axis of the body fixed frame is along the rigid link. The configuration manifold is the one-sphere $\mathsf{S}^1 = \{q \in \mathbb{R}^2 \,|\, q^T q = 1\}$, and we identify it with $\mathsf{SO}(2) = \{R \in \mathbb{R}^{2\times 2} \,|\, R^T R = I_{2\times 2}, \det[R] = 1\}$. A rotation matrix $R \in \mathsf{SO}(3)$ represents the linear transformation from the body fixed frame to the reference frame. The Lie algebra $\mathfrak{so}(2)$ is identified with $\mathbb{R}$ by an isomorphism $\hat{\cdot} : \mathbb{R} \to \mathfrak{so}(2)$ given by

$$\hat{\Omega} = \begin{bmatrix} 0 & -\Omega \\ \Omega & 0 \end{bmatrix},$$

for $\Omega \in \mathbb{R}$. We define an inner product on $\mathfrak{so}(2)$ using the standard inner product on $\mathbb{R}$ as $\left\langle \hat{\Omega}_1, \hat{\Omega}_2 \right\rangle = \frac{1}{2}\mathrm{tr}\left[\hat{\Omega}_1^T \hat{\Omega}_2\right] = \Omega_1 \cdot \Omega_2$. This defines an inner product on $\mathsf{TSO}(2)$ by the left-trivialization as $\langle X,\, Y\rangle = \langle \mathsf{T}_R\mathsf{L}_{R^{-1}} \cdot X, \mathsf{T}_R\mathsf{L}_{R^{-1}} \cdot Y\rangle$ for $X, Y \in \mathsf{T}_R\mathsf{SO}(2)$. The dual space $\mathsf{T}^*\mathrm{SO}(2)$ is identified with $\mathsf{TSO}(2)$ using this inner product. Let $\mathbf{J} : \mathfrak{so}(2) \to \mathfrak{so}(2)^*$ be defined as $\mathbf{J}(\hat{\Omega}) = ml^2\hat{\Omega}$. This induces a metric on $\mathfrak{so}(2)$ as $\langle\langle \Omega_1, \Omega_2 \rangle\rangle = \left\langle \mathbf{J}(\hat{\Omega}_1), \hat{\Omega}_2 \right\rangle$. This defines a metric on $\mathsf{TSO}(2)$ by the left-trivialization. The $\mathrm{ad}$ operation for $\mathsf{SO}(2)$ is zero, i.e. $\mathrm{ad}_\Omega = 0$.

*Lagrangian.* The Lagrangian $L : \mathsf{SO}(2) \times \mathfrak{so}(2) \to \mathbb{R}$ of the planar pendulum is given by

$$L(R, \hat{\Omega}) = \frac{1}{2}\langle\langle \hat{\Omega}, \hat{\Omega} \rangle\rangle + mge_2^T R\rho, \tag{2.41}$$

where $\rho = le_2 \in \mathbb{R}^2$ is a vector from the pivot point to the mass in the body fixed frame, and $e_2 \in \mathbb{R}^2$ is a unit vector along the gravity direction in the inertial frame.

*Euler-Lagrange equations.* The derivatives of the Lagrangian are given by

$$\mathbf{D}_\Omega L(R, \hat{\Omega}) \cdot \delta\hat{\Omega} = \left\langle \mathbf{J}(\hat{\Omega}),\ \delta\hat{\Omega} \right\rangle = ml^2\Omega \cdot \delta\Omega,$$
$$\mathrm{ad}^*_\Omega \cdot \mathbf{D}_\Omega L(R, \hat{\Omega}) = 0,$$
$$\left\langle \mathsf{T}_e\mathsf{L}^*_R \cdot \mathbf{D}_R L(R, \hat{\Omega}),\ \hat{\eta} \right\rangle = mge_2^T R\hat{\eta}le_2 = -mgle_2^T Re_1 \cdot \eta.$$

Substituting these into (2.9) and (2.10), we obtain

$$ml^2\dot{\Omega} + mgle_2^T Re_1 = 0, \tag{2.42}$$
$$\dot{R} = R\hat{\Omega}. \tag{2.43}$$

*Legendre Transformation.* From (2.14), the Legendre transformation is given by $\Pi = \mathbf{D}_\Omega L = ml^2\Omega \in \mathfrak{so}(2)^* \simeq \mathbb{R}^*$, which represents the angular momentum of the pendulum. The Hamilton's equations are given by

$$\dot{\Pi} + mgle_2^T Re_1 = 0, \tag{2.44}$$
$$\dot{R} = R\widehat{\frac{1}{ml^2}\Pi}. \tag{2.45}$$

If we parameterize the rotation matrix as $R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ for $\theta \in \mathsf{S}^1$, these equations are equivalent to

$$ml^2\ddot{\theta} + mgl\sin\theta = 0. \tag{2.46}$$

### 2.3.2 3D Pendulum

A 3D pendulum is a rigid body supported by a frictionless pivot acting under gravitational potential (see Shen et al. 2004). This is a generalization of a planar pendulum and a spherical pendulum, as it has three rotational degrees of freedom. It has been shown that the 3D pendulum may exhibit irregular, possibly chaotic, attitude dynamics (see Chaturvedi et al. 2007).

*Configuration Manifold.* Consider a reference frame, and a body fixed frame that is attached to the 3D pendulum body. We assume that the origin of the body fixed frame is located at the pivot point. The attitude of the pendulum is the orientation of the body fixed frame with respect to the

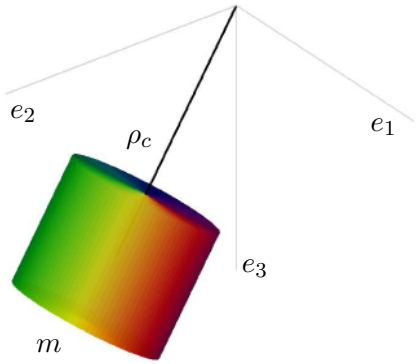



Figure 2.2: 3D Pendulum

reference frame, and it is described by a rotation matrix that represents the linear transformation from the body fixed frame to the reference frame. The configuration manifold of the 3D pendulum is the special orthogonal group

$$\mathsf{SO}(3) = \{R \in \mathbb{R}^{3\times 3} \,|\, R^T R = I, \quad \det R = 1\}. \tag{2.47}$$

The group operation for $\mathsf{SO}(3)$ corresponds to matrix multiplication.

The attitude kinematics equation is given by

$$\dot{R} = R\hat{\Omega}, \tag{2.48}$$

where the angular velocity represented in the body fixed frame is denoted by $\Omega \in \mathbb{R}^3$, and the hat map $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ is an isomorphism between $\mathbb{R}^3$ and the set of $3 \times 3$ skew symmetric matrices, the Lie algebra $\mathfrak{so}(3)$, defined by

$$\hat{\Omega} = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix} \tag{2.49}$$

for $\Omega = [\Omega_1; \Omega_2; \Omega_3] \in \mathbb{R}^3$. The Lie bracket on $\mathfrak{so}(3)$ corresponds to the cross product on $\mathbb{R}^3$, i.e. $[\hat{\Omega}, \hat{\Omega}'] = \Omega \times \Omega'$ for $\Omega, \Omega' \in \mathbb{R}^3$. Several properties of the hat map are summarized in Appendix A.1. Using the kinematics equation, the tangent bundle $\mathsf{TSO}(3)$ is identified with $\mathsf{SO}(3) \times \mathfrak{so}(3)$, and it is further identified with $\mathsf{SO}(3) \times \mathbb{R}^3$ using the hat map. This defines an inner product on $\mathfrak{so}(3)$ using the standard inner product on $\mathbb{R}^3$ as $\left\langle \hat{\Omega}_1, \hat{\Omega}_2 \right\rangle = \frac{1}{2}\mathrm{tr}\left[\hat{\Omega}_1^T \hat{\Omega}_2\right] = \Omega_1 \cdot \Omega_2$ for $\Omega_1, \Omega_2 \in \mathbb{R}^3$. The inner product of $\mathsf{TSO}(3)$ is defined by the left trivialization as $\langle X, Y \rangle = \langle \mathsf{T}_R \mathsf{L}_{R^{-1}} \cdot X, \mathsf{T}_R \mathsf{L}_{R^{-1}} \cdot Y \rangle$ for $X, Y \in \mathsf{T}_R\mathsf{SO}(3)$. The cotangent bundle $\mathsf{T}^*\mathsf{SO}(3)$ is identified with $\mathsf{TSO}(3)$ using this inner product. The $\mathrm{ad}$ operation on $\mathsf{SO}(3)$ is given by $\mathrm{ad}_\Omega \Omega' = \hat{\Omega}\Omega'$, $\mathrm{ad}^*_\Omega \Omega' = -\hat{\Omega}\Omega'$.

*Lagrangian.* The Lagrangian $L : \mathsf{SO(3)} \times \mathfrak{so}(3) \to \mathbb{R}$ is the difference between the kinetic energy $T : \mathsf{SO(3)} \times \mathfrak{so}(3) \to \mathbb{R}$ and the gravitational potential energy $U : \mathsf{SO(3)} \to \mathbb{R}$.

$$L(R, \Omega) = T(R, \Omega) - U(R). \tag{2.50}$$

Let $\rho \in \mathbb{R}^3$ be the vector from the pivot to a mass element represented in the body fixed frame. The mass element has a velocity $\Omega \times \rho$. Thus, the kinetic energy is given by

$$T(\Omega) = \frac{1}{2} \int_{\mathcal{B}} \|\hat{\Omega}\rho\|^2 \, dm(\rho), \tag{2.51}$$

where the body region is denoted by $\mathcal{B}$. Since $\hat{\Omega}\rho = -\hat{\rho}\Omega$, this can be written as

$$T(\Omega) = \frac{1}{2} \int_{\mathcal{B}} \|\hat{\rho}\Omega\|^2 \, dm(\rho) = \frac{1}{2} \int_{\mathcal{B}} \Omega^T \hat{\rho}^T \hat{\rho} \, \Omega \, dm(\rho) = \frac{1}{2} \Omega^T J \Omega, \tag{2.52}$$

where the inertia matrix $J \in \mathbb{R}^3$ is defined as $J = \int_{\mathcal{B}} \hat{\rho}^T \hat{\rho} \, dm$.

Alternatively, using the property $\|x\|^2 = x^T x = \mathrm{tr}\left[xx^T\right]$ for any $x \in \mathbb{R}^3$, equation (2.51) can be written as

$$\begin{aligned} T(\Omega) &= \frac{1}{2} \int_{\mathcal{B}} \mathrm{tr}\left[\hat{\Omega}\rho\rho^T\hat{\Omega}^T\right] \, dm(\rho) \\ &= \frac{1}{2} \mathrm{tr}\left[\hat{\Omega} J_d \hat{\Omega}^T\right], \end{aligned} \tag{2.53}$$

where a nonstandard inertia matrix is defined as $J_d = \int_{\mathcal{B}} \rho\rho^T \, dm$. Therefore, the kinetic energy can be written in the standard form (2.52) or in a non-standard form (2.53). In (2.52) the kinetic energy is expressed as a function of the angular velocity vector with the standard inertia matrix, and in (2.53) it is expressed as a function of the Lie algebra with the non-standard inertia matrix. The relationship between the standard inertia matrix and the nonstandard inertia matrix is summarized in Appendix A.2. It can be shown that

$$\widehat{J\Omega} = \hat{\Omega} J_d + J_d \hat{\Omega}$$

for any $\Omega \in \mathbb{R}^3$. Here, we use the nonstandard inertia matrix, since the corresponding development has a similar structure with the discrete-time Euler-Lagrange equations presented in Section 3.3.2.

The gravitational potential energy is given by

$$U(R) = -mge_3^T R \rho_c, \tag{2.54}$$

where the constants $m, g$ are the mass of the pendulum and the gravitational constant, respectively, and the vector from the pivot to the mass center represented in the body fixed frame is denoted by $\rho_c \in \mathbb{R}^3$.

In summary, the Lagrangian of the attitude dynamics of the 3D pendulum is given by

$$L(R, \Omega) = \frac{1}{2} \mathrm{tr}\left[\hat{\Omega} J_d \hat{\Omega}^T\right] - U(R). \tag{2.55}$$

Let $\mathbf{J} : \mathfrak{so}(3) \to \mathfrak{so}(3)^*$ be the inertia operator defined as

$$\mathbf{J}(\hat{\Omega}) = \hat{\Omega} J_d + J_d \hat{\Omega} = \widehat{J\Omega}. \tag{2.56}$$

The Lagrangian can alternatively written as

$$L(R, \Omega) = \frac{1}{2} \left\langle \mathbf{J}(\hat{\Omega}), \hat{\Omega} \right\rangle - U(R). \tag{2.57}$$

*Euler-Lagrange Equations.* We find expressions for the derivatives of the Lagrangian. We have

$$\mathbf{D}_\Omega L \cdot \delta\hat{\Omega} = -\frac{1}{2} \mathrm{tr}\left[ \delta\hat{\Omega} J_d \hat{\Omega} + \hat{\Omega} J_d \delta\hat{\Omega} \right].$$

Since $\mathrm{tr}[AB] = \mathrm{tr}[BA]$ for any $A, B \in \mathbb{R}^{3\times 3}$, this can be written as

$$\mathbf{D}_\Omega L \cdot \delta\hat{\Omega} = -\frac{1}{2} \mathrm{tr}\left[ \delta\hat{\Omega} (J_d \hat{\Omega} + \hat{\Omega} J_d) \right] = -\frac{1}{2} \mathrm{tr}\left[ \delta\hat{\Omega} \widehat{J\Omega} \right] = \left\langle \widehat{J\Omega},\, \delta\hat{\Omega} \right\rangle. \tag{2.58}$$

This can be derived directly from (2.57) using the symmetry of the inertia operator.

The derivative of the potential is given by

$$\begin{aligned} \mathbf{D}_R U(R) \cdot \delta R &= \frac{d}{d\epsilon} U(R \exp \epsilon\hat{\eta}) \Big|_{\epsilon=0} = \sum_{i,j=1}^{3} \frac{\partial U}{\partial [R]_{ij}} \frac{\partial \left[ R \exp \epsilon\hat{\eta} \right]_{ij}}{\partial \epsilon} \Big|_{\epsilon=0} \\ &= \sum_{i,j=1}^{3} \frac{\partial U}{\partial [R]_{ij}} \left[ R\hat{\eta} \right]_{ij} = -\mathrm{tr}\left[ \hat{\eta} R^T \frac{\partial U}{\partial R} \right], \end{aligned}$$

where $[A]_{ij}$ denotes the $i,j$th element of a matrix $A$, and $\frac{\partial U}{\partial R} \in \mathbb{R}^{3\times 3}$ is defined such that $\left[\frac{\partial U}{\partial R}\right]_{ij} = \frac{\partial U(R)}{\partial [R]_{ij}}$. We use the following identity: since $\mathrm{tr}[\hat{x}B] = -\mathrm{tr}\left[B^T \hat{x}\right] = -\mathrm{tr}\left[\hat{x} B^T\right]$ for any $x \in \mathbb{R}^3, B \in \mathbb{R}^{3\times 3}$,

$$\mathrm{tr}[\hat{x}B] = \frac{1}{2} \mathrm{tr}\left[ \hat{x}(B - B^T) \right] = -\left\langle B - B^T,\, \hat{x} \right\rangle. \tag{2.59}$$

Therefore, the derivative of the potential is given by

$$\mathbf{D}_R U(R) \cdot \delta R = (\mathsf{T}_e^* \mathsf{L}_R \cdot \mathbf{D}_R U(R)) \cdot \hat{\eta} = -\left\langle \hat{M}, \hat{\eta} \right\rangle, \tag{2.60}$$

where the moment due to the potential $M \in \mathbb{R}^3$ is determined by $\hat{M} = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}$. More explicitly, let $r_i$ and $u_i \in \mathbb{R}^{1\times 3}$ be the $i$th row vectors of $R$ and $\frac{\partial U}{\partial R}$, respectively. We have

$$\begin{aligned} \hat{M} &= \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R} \\ &= \begin{bmatrix} & & \\ u_1^T & u_2^T & u_3^T \\ & & \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} - \begin{bmatrix} & & \\ r_1^T & r_2^T & r_3^T \\ & & \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \\ &= \left( u_1^T r_1 - r_1^T u_1 \right) + \left( u_2^T r_2 - r_2^T u_2 \right) + \left( u_3^T r_3 - r_3^T u_3 \right), \end{aligned}$$

Since $(u^T r - r^T u)^\wedge = \widehat{r \times u}$, we obtain

$$\hat{M} = (r_1 \times u_1 + r_2 \times u_2 + r_3 \times u_3)^\wedge. \tag{2.61}$$

Thus, the moment due to the attitude-dependent potential is explicitly given by

$$M = r_1 \times u_1 + r_2 \times u_2 + r_3 \times u_3. \tag{2.62}$$

For the gravitational potential given by (2.54), $\frac{\partial U}{\partial R} = -mge_3\rho_c^T$. Thus, the moment due to the potential is $M = mg\rho_c \times R^T e_3$.

We substitute (2.58) and (2.60) into (2.9) and (2.10) to obtain the Euler-Lagrange equations of the 3D pendulum as

$$J\dot{\Omega} + \Omega \times J\Omega = mg\rho_c \times R^T e_3, \tag{2.63}$$

$$\dot{R} = R\hat{\Omega}. \tag{2.64}$$

*Legendre Transformation.* The Legendre transformation $\mathbb{F}L : (\mathsf{SO(3)} \times \mathfrak{so}(3)) \to (\mathsf{SO(3)} \times \mathfrak{so}(3))$ is given by

$$\begin{aligned}
\mathbb{F}L(R, \hat{\Omega}) \cdot \hat{\eta} &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L(R, \hat{\Omega} + \epsilon\hat{\eta}) \\
&= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \frac{1}{2}\mathrm{tr}\Big[(\hat{\Omega} + \epsilon\hat{\eta})^T J_d(\hat{\Omega} + \epsilon\hat{\eta})\Big] \\
&= \frac{1}{2}\mathrm{tr}\Big[\hat{\eta}^T J_d\hat{\Omega} + \hat{\Omega}^T J_d\hat{\eta}\Big] = \frac{1}{2}\mathrm{tr}\Big[-(J_d\hat{\Omega} + \hat{\Omega}J_d)\hat{\eta}\Big] \\
&= \frac{1}{2}\mathrm{tr}\Big[\widehat{J\Omega}^T\hat{\eta}\Big] = \widehat{J\Omega} \cdot \hat{\eta}.
\end{aligned}$$

This gives the expression for the momentum $\hat{\Pi} = \mathbb{F}L(R, \hat{\Omega}) = \widehat{J\Omega}$, which is the angular momentum expressed in the body fixed frame. Substituting this into (2.63) and (2.64), we obtain Hamilton's equations for the 3D pendulum as

$$\dot{\Pi} + J^{-1}\Pi \times \Pi = mg\rho_c \times R^T e_3, \tag{2.65}$$

$$\dot{R} = R\widehat{J^{-1}\Pi}. \tag{2.66}$$

*Symmetry.* The Lagrangian of the 3D pendulum has a symmetry. It is invariant under an action of $\mathsf{H} = \mathsf{SO}(2) \simeq \mathsf{S}^1$ given by $\Phi : \mathsf{S}^1 \times \mathsf{SO(3)} \to \mathsf{SO(3)}$

$$\Phi(\theta, R) = \exp_{\mathsf{SO(3)}}(\theta\hat{e}_3)R, \tag{2.67}$$

which represents the rotation of the 3D pendulum about the gravity direction $e_3$. The invariance follows from the fact that the kinetic energy is invariant under any left action, and that the gravitational potential is invariant under a rotation about the gravity direction. As a result, the momentum map is preserved, and the configuration manifold can be reduced to $\mathsf{G}/\mathsf{H} = \mathsf{SO(3)}/\mathsf{SO}(2) \simeq \mathsf{S}^2$. Here

we derive the expression for the momentum map of the 3D pendulum using the general expression presented in Section 2.1.4: we find the infinitesimal generator and the Lagrangian one-form from (2.21) and (2.17), respectively, and we combine them to obtain the momentum map from (2.22).

We first find the expression for the infinitesimal generator. We identify $\mathfrak{h} = \mathfrak{so}(2)$ with $\mathbb{R}$ using the hat map introduced in Section 2.3.1, and we identify $\mathfrak{h}^*$ with $\mathbb{R}$. Using (2.21), for $\zeta \in \mathbb{R} \simeq \mathfrak{so}(2)$, the infinitesimal generator $\zeta_{\mathsf{SO}(3)\times\mathfrak{so}(3)} : \mathsf{SO}(3) \times \mathfrak{so}(3) \to \mathsf{T}(\mathsf{SO}(3) \times \mathfrak{so}(3))$ is given by

$$\zeta_{\mathsf{SO}(3)\times\mathfrak{so}(3)}(R, \hat{\Omega}) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \phi_L \circ \mathsf{T}_R\Phi_{\exp_{\mathsf{SO}(2)}\epsilon\zeta}(R) \cdot (R, R\hat{\Omega}). \tag{2.68}$$

The exponential map in $\mathsf{SO}(2)$ is the identity in $\mathbb{R}$. Thus, $\Phi_{\exp_{\mathsf{SO}(2)}\epsilon\zeta}(R) = \Phi_{\epsilon\zeta}(R)$. From (2.67), this is equal to $\Phi_{\epsilon\zeta}(R) = \exp_{\mathsf{SO}(3)}(\epsilon\zeta\hat{e}_3)R$. Its tangent map is given by

$$\begin{aligned}\mathsf{T}_R\Phi_{\exp_{\mathsf{SO}(2)}\epsilon\zeta}(R) \cdot (R, R\hat{\Omega}) &= \left(\exp_{\mathsf{SO}(3)}(\epsilon\zeta\hat{e}_3)R,\ \frac{d}{ds}\bigg|_{s=0} \exp_{\mathsf{SO}(3)}(\epsilon\zeta\hat{e}_3)R\exp_{\mathsf{SO}(3)}(s\hat{\Omega})\right)\\ &= \left(\exp_{\mathsf{SO}(3)}(\epsilon\zeta\hat{e}_3)R,\ \exp_{\mathsf{SO}(3)}(\epsilon\zeta\hat{e}_3)R\hat{\Omega}\right).\end{aligned}$$

The left trivialization of this is given by

$$\phi_L \circ \mathsf{T}_R\Phi_{\exp_{\mathsf{SO}(2)}\epsilon\zeta}(R) \cdot (R, R\hat{\Omega}) = \left(\exp_{\mathsf{SO}(3)}(\epsilon\zeta\hat{e}_3)R,\ \hat{\Omega}\right).$$

Substituting this into (2.68), we obtain the infinitesimal generator $\zeta_{\mathsf{SO}(3)\times\mathfrak{so}(3)}$ as

$$\zeta_{\mathsf{SO}(3)\times\mathfrak{so}(3)}(R, \hat{\Omega}) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \left(\exp_{\mathsf{SO}(3)}(\epsilon\zeta\hat{e}_3)R,\ \hat{\Omega}\right) = \left(\zeta\hat{e}_3R,\ \hat{\Omega}\right). \tag{2.69}$$

Substituting (2.58) into (2.17), we obtain the Lagrangian one-form $\Theta_L$ on $\mathsf{SO}(3) \times \mathfrak{so}(3)$ as

$$\Theta_L(R, \hat{\Omega}) \cdot (\delta R, \delta\hat{\Omega}) = \left\langle \widehat{J\Omega},\ R^T\delta R \right\rangle. \tag{2.70}$$

Now, the expressions for the infinitesimal generator and the Lagrangian one-form are given by (2.69) and (2.70). Substituting these into (2.22), we obtain

$$\begin{aligned}J_L(R, \Omega) \cdot \zeta &= \Theta_L(R, \hat{\Omega}) \cdot \zeta_{\mathsf{SO}(3)\times\mathfrak{so}(3)} = \Theta_L(R, \hat{\Omega}) \cdot \left(\zeta\hat{e}_3R,\ \hat{\Omega}\right)\\ &= \left\langle \widehat{J\Omega},\ \zeta R^T\hat{e}_3R \right\rangle = \left\langle \widehat{J\Omega},\ \zeta\widehat{R^Te_3} \right\rangle = \zeta\, e_3^TRJ\Omega.\end{aligned}$$

Since this is satisfied for any $\zeta \in \mathbb{R}$, the momentum map of the 3D pendulum $J_L : \mathsf{SO}(3)\times\mathfrak{so}(3) \to \mathbb{R}^*$ is given by

$$J_L(R, \Omega) = e_3^TRJ\Omega, \tag{2.71}$$

which represents the angular momentum about the gravity direction. According to Noether's theorem, this is preserved along the solution of (2.63) and (2.64).

*Reduced Euler-Lagrange Equations.* Due to the symmetry of the Lagrangian, the configuration manifold can be reduced to a shape space $\mathsf{SO}(3)/\mathsf{SO}(2) \simeq \mathsf{S}^2$ as discussed in Section 2.1.5. Here, we present the reduced Euler-Lagrange equations on $\mathsf{S}^2$, and the detailed procedure to derive the reduced equations is summarized in Appendix A.3 as well as the reconstruction equation (see also Marsden et al. 2000).

Let $\Upsilon \in \mathsf{S}^2$ be the direction of gravity expressed in the body fixed frame, i.e. $\Upsilon = R^T e_3$. Suppose the fixed value of the momentum map is given by $\nu$. The reduced Euler-Lagrange equations for the 3D pendulum are given by

$$\ddot{\Upsilon} = -\|\dot{\Upsilon}\|^2 \Upsilon + \Upsilon \times \Sigma, \tag{2.72}$$

where the vector $\Sigma \in \mathbb{R}^3$, constants $c, b, \Upsilon \in \mathbb{R}$ are defined as

$$\Sigma = b\dot{\Upsilon} + J^{-1}\left[(J(\dot{\Upsilon} \times \Upsilon) - bJ\Upsilon) \times ((\dot{\Upsilon} \times \Upsilon) - b\Upsilon) + \Upsilon^2 J\Upsilon \times \Upsilon - mg\Upsilon \times \rho - c\dot{\Upsilon}\right], \tag{2.73}$$

$$c = \Upsilon\left\{\mathrm{tr}[J] - 2\frac{\|J\Upsilon\|^2}{\Upsilon \cdot J\Upsilon}\right\}, \quad b = \frac{J\Upsilon \cdot (\dot{\Upsilon} \times \Upsilon)}{\Upsilon \cdot J\Upsilon}, \quad \Upsilon = \frac{\nu}{\Upsilon \cdot J\Upsilon}. \tag{2.74}$$

### 2.3.3 3D Pendulum with an Internal Degree of Freedom

A 3D pendulum is a rigid body supported by a frictionless pivot point acting under a uniform gravitational potential. An internal degree of freedom is modeled as a single mass particle that is constrained to move along a linear slot fixed in the pendulum body. We assume that the mass particle is connected to a linear spring.

*Configuration manifold.* We define three frames; a reference frame, a body fixed frame for the 3D pendulum whose origin is located at the pivot point, and a slot frame fixed to the pendulum body. The origin of the slot frame is located at the point along the slot whose distance $d$ to the mass center is minimum. The first axis is aligned to the slot, the second axis is aligned to the mass center, and the third axis is orthogonal to the first and the second axis. Define

| | |
|---|---|
| $R \in \mathsf{SO}(3)$ | 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 |
| $Q \in \mathsf{SO}(3)$ | Rotation matrix from the slot frame to the body fixed frame |
| $d \in \mathbb{R}$ | Minimum distance from the slot to the mass center |
| $x \in \mathbb{R}$ | Displacement of the mass particle along the slot |
| $\rho_c \in \mathbb{R}^3$ | Vector from the pivot to the mass center of the pendulum |
| $\rho_x \in \mathbb{R}^3$ | Vector from the pivot to the mass particle |
| $m \in \mathbb{R}$ | Mass of the pendulum |
| $m_x \in \mathbb{R}$ | Mass of the particle |
| $\kappa \in \mathbb{R}$ | spring constant |

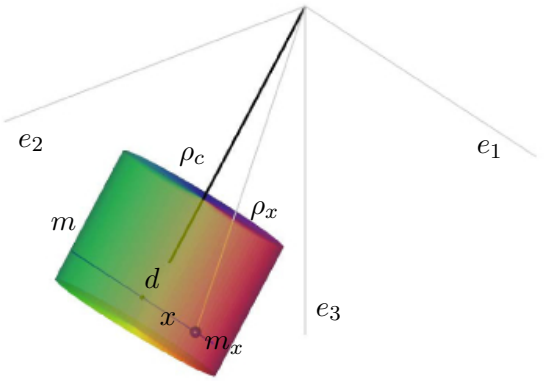



Figure 2.3: 3D Pendulum with an internal degree of freedom

The configuration manifold is $\mathsf{SO}(3) \times \mathbb{R}$. We identify $\mathsf{TSO}(3)$ with $\mathsf{SO}(3) \times \mathfrak{so}(3)$ by left trivialization, and the Lie algebra $\mathfrak{so}(3)$ is identified with $\mathbb{R}^3$ by an isomorphism $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$. We define an inner product on $\mathfrak{so}(3)$ using the standard inner product on $\mathbb{R}^3$ as $\left\langle \hat{\Omega}_1, \hat{\Omega}_2 \right\rangle = \frac{1}{2}\mathrm{tr}\left[\hat{\Omega}_1^T \hat{\Omega}_2\right] = \Omega_1 \cdot \Omega_2$. This defines an inner product on $\mathsf{TSO}(3)$ by the left-trivialization as $\langle X,\ Y \rangle = \langle \mathsf{T}_R\mathsf{L}_{R^{-1}} \cdot X, \mathsf{T}_R\mathsf{L}_{R^{-1}} \cdot Y \rangle$ for $X, Y \in \mathsf{T}_R\mathsf{SO}(3)$. The dual space $\mathsf{T}^*\mathsf{SO}(3)$ is identified with $\mathsf{TSO}(3)$ using this inner product. Let $\mathbf{J} : \mathfrak{so}(3) \to \mathfrak{so}(3)^*$ be defined as $\mathbf{J}(\hat{\Omega}) = J_d\hat{\Omega} + \hat{\Omega}J_d = \widehat{J\Omega}$. The inertia matrix is denoted by $J \in \mathbb{R}^{3\times3}$, and the non-standard inertia matrix is defined by $J_d = \frac{1}{2}\mathrm{tr}[J]\, I_{3\times3} - J \in \mathbb{R}^{3\times3}$. This induces a metric on $\mathfrak{so}(3)$ as $\langle\langle \Omega_1, \Omega_2 \rangle\rangle = \left\langle \mathbf{J}(\hat{\Omega}_1), \hat{\Omega}_2 \right\rangle$. This defines a metric on $\mathsf{TSO}(3)$ by left trivialization. The $\mathrm{ad}$ operator on $\mathsf{SO}(3) \times \mathbb{R}$ is given by $\mathrm{ad}_{(\Omega,\dot{x})}(\Omega', \dot{x}') = (\hat{\Omega}\Omega', 0)$, $\mathrm{ad}^*_{(\Omega,\dot{x})}(\Omega', \dot{x}') = (-\hat{\Omega}\Omega', 0)$.

*Lagrangian.* The rotation matrix $Q \in \mathsf{SO}(3)$ defines the orientation of the slot with respect to the body fixed frame: its first column denotes the direction of the slot in the body fixed frame, and its second column denotes the direction from the mass center of the pendulum to the origin of the slot frame.

Let $r \in \mathbb{R}^3$ be $r = [x; d; 0] \in \mathbb{R}^3$. The vector from the mass center of the pendulum to the mass particle is represented by $Qr$ in the body fixed frame. Therefore, the vector from the pivot to the particle in the body fixed frame is given by

$$\rho_x = \rho_c + Qr,$$

where $\rho_c \in \mathbb{R}^3$ is the vector from the pivot to the mass center expressed in the body fixed frame. Note that $\dot{\rho}_x = Qe_1\dot{x}$ since $\rho_c$, $Q$, $d$ are fixed quantities. Thus, the velocity of the mass particle in the inertial frame is given by

$$\frac{d}{dt}(R\rho_x) = R\dot{\rho}_x + \dot{R}\rho_x = RQe_1\dot{x} + R\hat{\Omega}\rho_x.$$

The kinetic energy is composed of the rotational kinetic energy of the 3D pendulum and the translational kinetic energy of the mass particle.

$$\begin{aligned} T(\Omega, \dot{x}) &= \frac{1}{2}\langle\langle \Omega, \Omega \rangle\rangle + \frac{1}{2} m_x \left\| Q e_1 \dot{x} + \hat{\Omega} \rho_x \right\|^2 \\ &= \frac{1}{2}\Omega^T J \Omega + \frac{1}{2} m_x (\dot{x}^2 - \rho_x^T \hat{\Omega}^2 \rho_x + 2 e_1^T Q^T \hat{\Omega} \rho_x \dot{x}). \end{aligned}$$

The potential energy consists of the gravitational potential energy and the potential energy for the linear spring,

$$U(R, x) = -mg e_3^T R \rho_c - m_x g e_3^T R \rho_x + \frac{1}{2}\kappa x^2.$$

Therefore, the Lagrangian $L : (\mathsf{SO(3)} \times \mathbb{R}) \times (\mathbb{R}^3 \times \mathbb{R}) \to \mathbb{R}$ is given by

$$\begin{aligned} L(R, x, \Omega, \dot{x}) = &\frac{1}{2}\Omega^T J \Omega + \frac{1}{2} m_x (\dot{x}^2 - \rho_x^T \hat{\Omega}^2 \rho_x + 2 e_1^T Q^T \hat{\Omega} \rho_x \dot{x}) \\ &+ mg e_3^T R \rho_c + m_x g e_3^T R \rho_x - \frac{1}{2}\kappa x^2. \end{aligned} \tag{2.75}$$

*Euler-Lagrange equations.* The derivatives of the Lagrangian are given by

$$\begin{aligned} \mathbf{D}_\Omega L \cdot \delta\Omega &= (J\Omega - m_x \hat{\rho}_x^2 \Omega + m_x \dot{x} \hat{\rho}_x Q e_1) \cdot \delta\Omega, \\ -\mathrm{ad}^*_\Omega \cdot \mathbf{D}_\Omega L &= \hat{\Omega}(J\Omega - m_x \hat{\rho}_x^2 \Omega + m_x \dot{x} \hat{\rho}_x Q e_1), \\ (\mathsf{T}_e^* L_R \cdot \mathbf{D}_R L) \cdot \hat{\eta} &= (mg \hat{\rho}_c R^T e_3 + m_x g \hat{\rho}_x R^T e_3) \cdot \eta, \\ \mathbf{D}_x L \cdot \delta x &= (-m_x \rho_x^T \hat{\Omega}^2 Q e_1 + m_x g e_3^T R Q e_1 - \kappa x) \cdot \delta x, \\ \mathbf{D}_{\dot{x}} L \cdot \delta\dot{x} &= (m_x \dot{x} + m_x e_1^T Q^T \hat{\Omega} \rho_x) \cdot \delta\dot{x}. \end{aligned}$$

Substituting these expressions into (2.9), we obtain the Euler-Lagrange equations.

$$\begin{aligned} &(J - m_x \hat{\rho}_x^2)\dot{\Omega} + m_x \hat{\rho}_x Q e_1 \ddot{x} - m_x(\widehat{Qe_1}\hat{\rho}_x + \hat{\rho}_x \widehat{Qe_1})\Omega \dot{x} \\ &\quad + \hat{\Omega}(J\Omega - m_x \hat{\rho}_x^2 \Omega + m_x \dot{x} \hat{\rho}_x Q e_1) - (mg \hat{\rho}_c R^T e_3 + m_x g \hat{\rho}_x R^T e_3) = 0, \end{aligned} \tag{2.76}$$

$$\dot{R} = R\hat{\Omega}, \tag{2.77}$$

$$m_x \ddot{x} - m_x e_1^T Q^T \hat{\rho}_x \dot{\Omega} + m_x \rho_x^T \hat{\Omega}^2 Q e_1 - m_x g e_3^T R Q e_1 + \kappa x = 0. \tag{2.78}$$

*Legendre Transformations.* From (2.14), the Legendre transformation is given by

$$\begin{bmatrix} p_\Omega \\ p_x \end{bmatrix} = \begin{bmatrix} J - m_x \hat{\rho}_x^2 & m_x \hat{\rho}_x Q e_1 \\ (m_x \hat{\rho}_x Q e_1)^T & m_x \end{bmatrix} \begin{bmatrix} \Omega \\ \dot{x} \end{bmatrix}. \tag{2.79}$$

From Corollary 2.2, this yields the Hamilton's equations.

$$\dot{p}_\Omega + \Omega \times p_\Omega - (mg \hat{\rho}_c R^T e_3 + m_x g \hat{\rho}_x R^T e_3) = 0, \tag{2.80}$$

$$\dot{p}_x + m_x \rho_x^T \hat{\Omega}^2 Q e_1 - m_x g e_3^T R Q e_1 + \kappa x = 0. \tag{2.81}$$

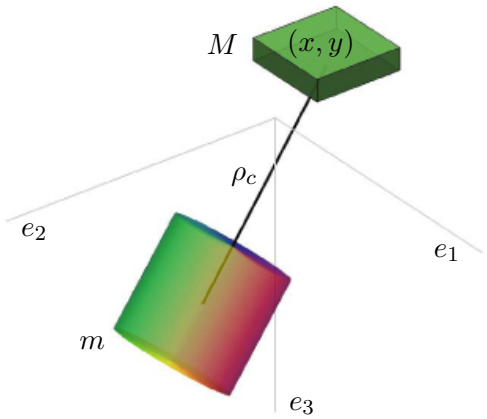



Figure 2.4: 3D Pendulum on a cart

### 2.3.4 3D Pendulum on a Cart

Consider a 3D pendulum whose pivot is attached to a cart moving on a horizontal plane.

*Configuration manifold.* We define two frames; a reference frame and a body fixed frame for the 3D pendulum whose origin is located at the moving pivot point. Define

| | |
|---|---|
| $x \in \mathbb{R}$ | 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 |
| $R \in \mathsf{SO}(3)$ | 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 |
| $\rho_c \in \mathbb{R}^3$ | 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 $\mathsf{SO}(3) \times \mathbb{R}^2$. The tangent bundle $\mathsf{T}(\mathsf{SO}(3) \times \mathbb{R}^2)$ is identified with $(\mathsf{SO}(3) \times \mathbb{R}^2) \times (\mathbb{R}^3 \times \mathbb{R}^2)$ by using the left trivialization presented in Section 2.3.3. The $\mathrm{ad}$ operator on $\mathsf{SO}(3) \times \mathbb{R}^2$ is given by $\mathrm{ad}_{(\Omega,\dot{x},\dot{y})}(\Omega', \dot{x}', \dot{y}') = (\hat{\Omega}\Omega', 0, 0)$, $\mathrm{ad}^*_{(\Omega,\dot{x},\dot{y})}(\Omega', \dot{x}', \dot{y}') = (-\hat{\Omega}\Omega', 0, 0)$.

*Lagrangian.* The Lagrangian $L : (\mathsf{SO}(3) \times \mathbb{R}^2) \times (\mathbb{R}^3 \times \mathbb{R}^2) \to \mathbb{R}$ is the difference between the kinetic energy $T : (\mathsf{SO}(3) \times \mathbb{R}^2) \times (\mathbb{R}^3 \times \mathbb{R}^2) \to \mathbb{R}$ and the potential $U : \mathsf{SO}(3) \times \mathbb{R}^2 \to \mathbb{R}$.

$$L(R, x, y, \Omega, \dot{x}, \dot{y}) = T(R, x, y, \Omega, \dot{x}, \dot{y}) - U(R, x, y). \tag{2.82}$$

The kinetic energy of the cart is given by $T_{cart} = \frac{1}{2}M(\dot{x}^2 + \dot{y}^2)$. Let $\rho \in \mathbb{R}^3$ be the vector from the mass center of the pendulum to a mass element represented in the body fixed frame. The vector

to the mass element from the origin of the reference frame is given by $xe_1 + ye_2 + R(\rho_c + \rho)$. Thus, the kinetic energy of the pendulum can be written as

$$\begin{aligned} T_{pend} &= \frac{1}{2}\int_{\mathcal{B}} \left\| \dot{x}e_1 + \dot{y}e_2 + \dot{R}(\rho_c + \rho) \right\|^2 \, dm(\rho) \\ &= \frac{1}{2}\int_{\mathcal{B}} \left( \dot{x}^2 + \dot{y}^2 + \mathrm{tr}\left[ \hat{\Omega}(\rho_c + \rho)(\rho_c + \rho)^T \hat{\Omega}^T \right] + 2\dot{x}e_1^T R\hat{\Omega}(\rho_c + \rho) + 2\dot{y}e_2^T R\hat{\Omega}(\rho_c + \rho) \right) \, dm(\rho). \end{aligned}$$

We have $\int_{\mathcal{B}} \rho \, dm = 0$ from the definition of the mass center. Define the nonstandard moment of inertia matrix with respect to the pivot as $J_d = \int (\rho_c + \rho)(\rho_c + \rho)^T \, dm$. The kinetic energy of the pendulum is given by

$$T_{pend} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}\mathrm{tr}\left[ \hat{\Omega} J_d \hat{\Omega}^T \right] + m\dot{x}e_1^T R\hat{\Omega}\rho_c + m\dot{y}e_2^T R\hat{\Omega}\rho_c. \tag{2.83}$$

Or equivalently, the second term of the above equation can be written as $\frac{1}{2}\Omega^T J\Omega$ for the standard moment of inertia matrix $J = \mathrm{tr}[J_d]\, I - J_d$. The total kinetic energy is given by $T = T_{cart} + T_{pend}$. The gravitational potential energy of the pendulum is $U(R) = -mge_3^T Rd$. Thus, the Lagrangian is given by

$$L(R, \Omega, \dot{x}, \dot{y}) = \frac{1}{2}(M + m)(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}\Omega^T J\Omega + m\dot{x}e_1^T R\hat{\Omega}\rho_c + m\dot{y}e_2^T R\hat{\Omega}\rho_c + mge_3^T R\rho_c. \tag{2.84}$$

*Euler-Lagrange equations.* We have

$$\begin{aligned} \mathbf{D}_\Omega L \cdot \delta\Omega &= (J\Omega + m\dot{x}\hat{\rho}_c R^T e_1 + m\dot{y}\hat{\rho}_c R^T e_2) \cdot \delta\Omega, \\ -\mathrm{ad}^*_\Omega \cdot D_\Omega L &= \hat{\Omega}(J\Omega + m\dot{x}\hat{\rho}_c R^T e_1 + m\dot{y}\hat{\rho}_c R^T e_2), \\ (\mathsf{T}^*_e \mathsf{L}_R \cdot \mathbf{D}_R L) \cdot \hat{\eta} &= (m\dot{x}\widehat{\hat{\Omega}\rho_c} R^T e_1 + m\dot{y}\widehat{\hat{\Omega}\rho_c} R^T e_2 + mg\hat{\rho}_c R^T e_3) \cdot \eta, \\ \mathbf{D}_{(\dot{x},\dot{y})} L \cdot (\delta\dot{x}, \delta\dot{y}) &= ((M + m)\dot{x} + me_1^T R\hat{\Omega}\rho_c) \cdot \delta\dot{x} + ((M + m)\dot{y} + me_2^T R\hat{\Omega}\rho_c) \cdot \delta\dot{y}, \\ \mathbf{D}_{(x,y)} L \cdot (\delta x, \delta y) &= 0. \end{aligned}$$

Substituting these into (2.9) and (2.10), we obtain the Euler-Lagrange equations.

$$J\dot{\Omega} + m\ddot{x}\hat{\rho}_c R^T e_1 + m\ddot{y}\hat{\rho}_c R^T e_2 + \hat{\Omega}J\Omega = mg\hat{\rho}_c R^T e_3, \tag{2.85}$$

$$\dot{R} = R\hat{\Omega}, \tag{2.86}$$

$$(M + m)\ddot{x} - me_1^T R\hat{\rho}_c\dot{\Omega} + me_1^T R\hat{\Omega}^2\rho_c = 0, \tag{2.87}$$

$$(M + m)\ddot{y} - me_2^T R\hat{\rho}_c\dot{\Omega} + me_2^T R\hat{\Omega}^2\rho_c = 0. \tag{2.88}$$

In a matrix form, these can be written as

$$\begin{bmatrix} J & m\hat{\rho}_c R^T e_1 & m\hat{\rho}_c R^T e_2 \\ (m\hat{\rho}_c R^T e_1)^T & M + m & 0 \\ (m\hat{\rho}_c R^T e_2)^T & 0 & M + m \end{bmatrix} \begin{bmatrix} \dot{\Omega} \\ \ddot{x} \\ \ddot{y} \end{bmatrix} + \begin{bmatrix} \Omega \times J\Omega \\ me_1^T R\hat{\Omega}^2\rho_c \\ me_2^T R\hat{\Omega}^2\rho_c \end{bmatrix} = \begin{bmatrix} mg\hat{\rho}_c R^T e_3 \\ 0 \\ 0 \end{bmatrix}. \tag{2.89}$$

*Legendre Transformations.* The Legendre transformation is given by

$$\begin{bmatrix} p_\Omega \\ p_x \\ p_y \end{bmatrix} = \begin{bmatrix} J & m\hat{\rho}_c R^T e_1 & m\hat{\rho}_c R^T e_2 \\ (m\hat{\rho}_c R^T e_1)^T & M+m & 0 \\ (m\hat{\rho}_c R^T e_2)^T & 0 & M+m \end{bmatrix} \begin{bmatrix} \dot{\Omega} \\ \ddot{x} \\ \ddot{y} \end{bmatrix}. \tag{2.90}$$

From Corollary 2.2, this yields the Hamilton's equations.

$$\dot{p}_\Omega + \Omega \times p_\Omega - (m\dot{x}\widehat{\hat{\Omega}\rho_c}R^T e_1 + m\dot{y}\widehat{\hat{\Omega}\rho_c}R^T e_2 + mg\hat{\rho}_c R^T e_3) = 0, \tag{2.91}$$

$$\dot{p}_x = 0, \tag{2.92}$$

$$\dot{p}_y = 0. \tag{2.93}$$

### 2.3.5 Single Rigid Body

Consider a rigid body acting under a potential that is dependent on the attitude and the position of the body.

*Configuration manifold.* Consider a reference frame, and a body fixed frame that is attached to the rigid body. We assume that the origin of the body fixed frame is located at the mass center. The group of rigid transformation on $\mathbb{R}^3$ is defined as the set of mappings $g : \mathbb{R}^3 \to \mathbb{R}^3$ of the form $g(p) = Ry+x$, where $R \in \mathsf{SO(3)}$ and $x, y \in \mathbb{R}^3$. An element of $\mathsf{SE(3)}$ is written as $(R, x) \in \mathsf{SE(3)}$, and $\mathsf{SE(3)}$ is embedded in the general linear group $\mathsf{GL}(4, \mathbb{R})$ using homogeneous coordinates

$$g = \begin{bmatrix} R & x \\ 0 & 1 \end{bmatrix}.$$

The rotation matrix $R \in \mathsf{SO(3)}$ represents the linear transformation from the body fixed frame to the reference frame, and the vector $x \in \mathbb{R}^3$ represents the location of the origin of the body fixed frame. The Lie algebra, denoted by $\mathfrak{se}(3)$, is isomorphic to $\mathbb{R}^3 \oplus \mathbb{R}^3$ via the mapping $\hat{\cdot} : \mathbb{R}^6 \to \mathfrak{se}(3)$;

$$[\Omega; V]\hat{} = \begin{bmatrix} \hat{\Omega} & V \\ 0 & 0 \end{bmatrix},$$

where $[\Omega;\ V] \in \mathbb{R}^6$ and $\hat{\Omega} \in \mathfrak{so}(3)$. We use the same notation $\hat{\cdot}$ to denote the Lie algebra isomorphism of $\mathfrak{so}(3)$ and $\mathfrak{se}(3)$.

We define an inner product on the Lie algebra $\mathfrak{se}(3)$ using the standard inner product on $\mathbb{R}^3$ $\langle \xi_1,\ \xi_2 \rangle = V_1 \cdot V_2 - \frac{1}{2}\text{tr}\left[\hat{\Omega}_1\hat{\Omega}_2\right] = V_1 \cdot V_2 + \Omega_1 \cdot \Omega_2$ where $\xi_i = (\Omega_i, V_i) \in \mathfrak{se}(3)$ for $i \in \{1, 2\}$. This defines an inner product on $\mathsf{TSE(3)}$ by left-trivialization as $\langle X,\ Y \rangle = \left\langle \mathsf{T}_g\mathsf{L}_{g^{-1}}X,\ \mathsf{T}_g\mathsf{L}_{g^{-1}}Y \right\rangle$ for $X, Y \in \mathsf{T}_g\mathsf{SE(3)}$. The dual space $\mathsf{T}^*\mathsf{SE(3)}$ is identified with $\mathsf{TSE(3)}$ using this inner product.

Let $\mathbf{J} : \mathfrak{se}(3) \to \mathfrak{se}(3)^*$ be defined as

$$\mathbf{J}((\Omega, V)) = \begin{bmatrix} \widehat{J\Omega} & mV \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} J_d\hat{\Omega} + \hat{\Omega}J_d & mV \\ 0 & 0 \end{bmatrix}$$

for the mass $m \in \mathbb{R}^3$, and the inertia matrix $J \in \mathbb{R}^{3\times3}$ of the rigid body. The non-standard inertia matrix is given by $J_d = \frac{1}{2}\mathrm{tr}[J]\, I_{3\times3} - J \in \mathbb{R}^{3\times3}$. This induces a metric on $\mathfrak{se}(3)$ as $\langle\!\langle \xi_1, \xi_2 \rangle\!\rangle_{\mathfrak{se}(3)} = \langle \mathbf{J}(\xi_1),\ \xi_2 \rangle$. The corresponding metric on $\mathsf{TSE(3)}$ is obtained by left-trivialization as $\langle\!\langle X, Y \rangle\!\rangle_{\mathsf{TSE(3)}} = \langle\!\langle \mathsf{T}_g\mathsf{L}_{g^{-1}} X, \mathsf{T}_g\mathsf{L}_{g^{-1}} Y \rangle\!\rangle_{\mathfrak{se}(3)}$ for $X, Y \in \mathsf{T}_g\mathsf{SE(3)}$.

The $\mathrm{ad}$ operator for $\mathfrak{se}(3)$ can be written in a matrix form as

$$\mathrm{ad}_{(\Omega,V)} = \begin{bmatrix} \hat{\Omega} & 0 \\ \hat{V} & \hat{\Omega} \end{bmatrix}, \quad \mathrm{ad}^*_{(\Omega,V)} = \begin{bmatrix} -\hat{\Omega} & -\hat{V} \\ 0 & -\hat{\Omega} \end{bmatrix}. \tag{2.94}$$

*Lagrangian.* The Lagrangian $\mathsf{SE(3)} \times \mathfrak{se}(3) \to \mathbb{R}$ is given by

$$L(g, \xi) = \frac{1}{2}\langle\!\langle \xi, \xi \rangle\!\rangle - U(g) \tag{2.95}$$

for a configuration dependent potential $U : \mathsf{SE(3)} \to \mathbb{R}$.

*Euler-Lagrange equations.* Derivatives of the Lagrangian are given by

$$\begin{aligned} \mathbf{D}_\xi L(g,\xi) \cdot \xi &= \langle\!\langle \xi, \delta\xi \rangle\!\rangle = \langle \mathbf{J}(\xi), \delta\xi \rangle, \\ -\mathrm{ad}^*_\xi \cdot \mathbf{D}_\xi L(g,\xi) &= \begin{bmatrix} \hat{\Omega} & \hat{V} \\ 0 & \hat{\Omega} \end{bmatrix} \begin{bmatrix} J\Omega \\ mV \end{bmatrix} = \begin{bmatrix} \hat{\Omega}J\Omega & m\hat{\Omega}V \\ 0 & 0 \end{bmatrix}, \\ \mathbf{D}_g L(g,\xi) \cdot \delta g &= \begin{bmatrix} \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R} & -R^T \frac{\partial U}{\partial x} \\ 0 & 0 \end{bmatrix} \cdot g^{-1}\delta g. \end{aligned} \tag{2.96}$$

Substituting these equations into (2.9), we obtain the continuous equation of motion for a rigid body on $\mathsf{SE(3)}$ in homogeneous coordinates as

$$\begin{bmatrix} \widehat{J\dot{\Omega}} & m\dot{V} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \hat{\Omega}J\Omega & m\hat{\Omega}V \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R} & -R^T \frac{\partial U}{\partial x} \\ 0 & 0 \end{bmatrix} = 0,$$
$$\begin{bmatrix} \dot{R} & \dot{x} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} R & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\Omega} & V \\ 0 & 0 \end{bmatrix}.$$

This can be written as

$$J\dot{\Omega} + \Omega \times J\Omega = M, \tag{2.97}$$
$$m\dot{V} + m\Omega \times V = -R^T \frac{\partial U}{\partial x}, \tag{2.98}$$
$$\dot{R} = R\hat{\Omega}, \tag{2.99}$$
$$\dot{x} = RV, \tag{2.100}$$

where the moment due to the potential $M \in \mathbb{R}^3$ is determined by

$$\hat{M} = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}. \tag{2.101}$$

*Legendre Transformation.* From (2.14), the Legendre transformation for the single rigid body is given by $(\Pi, \Gamma) = (J\Omega, mV)$, which represents the angular momentum and the linear momentum of the rigid body, represented with respect to the body fixed frame. This yields the Hamilton's equations.

$$\dot{\Pi} + \Omega \times \Pi = M, \tag{2.102}$$

$$\dot{\Gamma} + \Omega \times \Gamma = -R^T \frac{\partial U}{\partial x}, \tag{2.103}$$

$$\dot{R} = R\widehat{J^{-1}\Gamma}, \tag{2.104}$$

$$\dot{x} = \frac{1}{m} R\Gamma. \tag{2.105}$$

### 2.3.6 Full Body Problem

A full body problem deals with the dynamics of non-spherical rigid bodies in space interacting under their mutual potential. Since the mutual 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.

*Euler-Lagrange Equations.* The configuration manifold of full $n$ body problem is $(\mathsf{SE}(3))^n$. Since the dynamics of each body is only coupled through the mutual potential, the development for a single rigid body, presented in Section 2.3.5, is readily extended to the full $n$ body problem to obtain the Euler-Lagrange equations for the full $n$-body problem.

$$J_i\dot{\Omega}_i + \Omega_i \times J_i\Omega_i = M_i, \tag{2.106}$$

$$m_i\dot{V}_i + m_i\Omega_i \times V_i = -R_i^T \frac{\partial U}{\partial x_i}, \tag{2.107}$$

where the subscript $i$ denotes variables for the $i$-th rigid body for $i \in \{1, \dots n\}$.

*Reduced Euler-Lagrange Equations.* Suppose that the potential energy of the full body problem is dependent only on the relative attitudes between rigid bodies. Then, the Lagrangian is invariant under the action of $\mathsf{SE}(3)$, and the configuration manifold can be reduced to a quotient space $(\mathsf{SE}(3))^{n-1}$. According to Noether's theorem, the total linear momentum and the total angular momentum are preserved. For example, the mutual gravitational potential depends on the relative location and the relative attitude of the rigid bodies. In this case, it is desirable to write the equations of motion in the body fixed frame of one rigid body.

Reduced Euler-Lagrange equations for a full two body problem have been developed by Lee et al. (2007b,c). Here, we present the resulting reduced Euler-Lagrange equations for the full two body problem.

The reduced variables are defined with respect to the body fixed frame of the second rigid body.

$$X = R_2^T(x_1 - x_2), \quad R = R_2^T R_1,$$

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 $R \in \mathsf{SO(3)}$ is the relative attitude of the first body with respect to the second body. The corresponding linear and angular velocities are also defined as

$$V = R_2^T(\dot{x}_1 - \dot{x}_2), \quad \Omega = R\Omega_1,$$

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. We also define the velocity of the second body expressed in the second body-fixed frame as $V_2 = R_2^T \dot{x}_2$. 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.

The Lagrangian of the full two body problem can be written in terms of these reduced variables as

$$\begin{aligned} L(R, X, \Omega, V, \Omega_2, V_2) = \frac{1}{2}m_1 \left\|V_1 + V_2\right\|^2 + \frac{1}{2}m_2 \left\|V_2\right\|^2 + \frac{1}{2}\mathrm{tr}\left[\hat{\Omega}J_{d_R}\hat{\Omega}^T\right] + \frac{1}{2}\mathrm{tr}\left[\hat{\Omega}_2 J_{d_2}\hat{\Omega}_2^T\right] \\ - U(R, X). \end{aligned}$$

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

$$\delta R = \delta R_2^T R_1 + R_2^T \delta R_1 = -\hat{\eta}_2 R + \hat{\eta} R,$$

where $\eta = R\hat{\eta}R^T$. The variations of other reduced variables can be obtained in a similar way.

$$\begin{aligned} \delta X &= \chi - \eta_2 X, \\ \delta\hat{\Omega} &= \dot{\eta} - \hat{\Omega}\eta + \eta\hat{\Omega} + \hat{\Omega}\eta_2 - \eta_2\hat{\Omega} + \hat{\Omega}_2\eta - \eta\hat{\Omega}_2, \\ \delta V &= \dot{\chi} + \hat{\Omega}_2\chi - \eta_2 V, \\ \delta\hat{\Omega}_2 &= \dot{\eta}_2 + \hat{\Omega}_2\eta_2 - \eta_2\hat{\Omega}_2, \\ \delta V_2 &= \dot{\chi}_2 + \hat{\Omega}_2\chi_2 - \eta_2 V_2, \end{aligned}$$

where $\chi, \chi_2 \in \mathbb{R}^3$. By taking the variation of the reduced Lagrangian using these constrained variations, we obtain the reduced Euler-Lagrange equations for the full two body problem as follows.

$$\dot{V} + \Omega_2 \times V = -\frac{1}{m}\frac{\partial U}{\partial X}, \tag{2.108}$$

$$\dot{(J_R\Omega)} + \Omega_2 \times J_R\Omega = -M, \tag{2.109}$$

$$J_2\dot{\Omega}_2 + \Omega_2 \times J_2\Omega_2 = X \times \frac{\partial U}{\partial X} + M, \tag{2.110}$$

$$\dot{X} + \Omega_2 \times X = V, \tag{2.111}$$

$$\dot{R} = S(\Omega)R - S(\Omega_2)R, \tag{2.112}$$

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

$$M = r_1 \times u_{r_1} + r_2 \times u_{r_2} + r_3 \times u_{r_3}, \tag{2.113}$$

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

### 2.3.7 Two Rigid Bodies Connected by a Ball Joint

We consider two rigid bodies connected with a ball joint. We assume that the ball joint has three rotational degrees of freedom. The relative equilibria structure of this rigid body dynamics has been studied by Wang (1990).

*Configuration manifold.* We define three frames; a reference frame, and two body fixed frames. Define

| | |
|---|---|
| $x \in \mathbb{R}^3$ | Position of the ball joint in a reference frame |
| $R_i \in \mathsf{SO(3)}$ | 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 $\mathsf{SO(3)} \times \mathsf{SO(3)} \times \mathbb{R}^3$. The tangent bundle $\mathsf{T}(\mathsf{SO(3)} \times \mathsf{SO(3)} \times \mathbb{R}^3)$ is identified with $(\mathsf{SO(3)} \times \mathsf{SO(3)} \times \mathbb{R}^3) \times (\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3)$ by the left trivialization presented in Section 2.3.3.

*Lagrangian.* Let $\rho_i \in \mathbb{R}^3$ be the vector from the mass center of the $i$-th body to a mass element expressed in the $i$-th body fixed frame. The vector to the mass element from the origin of the reference frame is given by $x + R_i(d_i + \rho_i)$. Thus, the kinetic energy of the $i$-th body is given by

$$\begin{aligned} T_i(R_i, x_i, \Omega_i) &= \frac{1}{2} \int_{\mathcal{B}_i} \left\| \dot{x} + \dot{R}_i(d_i + \rho_i) \right\|^2 dm(\rho_i) \\ &= \frac{1}{2} m_i \dot{x} \cdot \dot{x} + \frac{1}{2} \mathrm{tr}\left[ \hat{\Omega}_i J_{d_i} \hat{\Omega}_i^T \right] + m_i \dot{x} \cdot R_i \hat{\Omega}_i d_i \\ &= \frac{1}{2} m_i \dot{x} \cdot \dot{x} + \frac{1}{2} \Omega_i \cdot J_i \Omega_i + m_i \dot{x} \cdot R_i \hat{\Omega}_i d_i, \end{aligned}$$

where $J_{d_i} = m_i d_i d_i^T + \int_{\mathcal{B}_i} \rho_i \rho_i^T \, dm(\rho_i) \in \mathbb{R}^{3\times 3}$, and $J_i = m_i \hat{d}_i^T \hat{d}_i + \int_{\mathcal{B}_i} \hat{\rho}_i^T \hat{\rho}_i \, dm(\rho_i)$. It can be shown that $J_{d_i} = \frac{1}{2}\mathrm{tr}[J_i]\, I_{3\times 3} - J_i$. Using this expression, the Lagrangian of the connected two

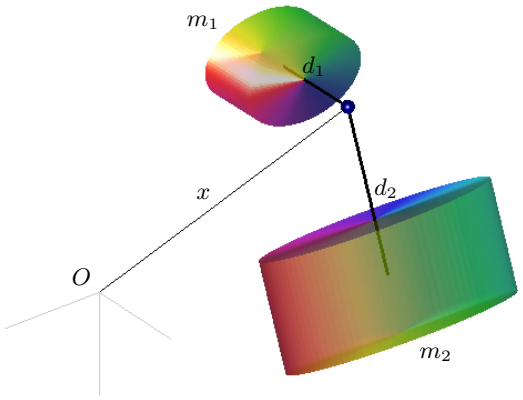



Figure 2.5: Two rigid bodies connected by a ball joint

rigid bodies $L : (\mathsf{SO}(3) \times \mathsf{SO}(3) \times \mathbb{R}^3) \times (\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3) \to \mathbb{R}$ is given by

$$\begin{aligned} L(R_1, R_2, x, \Omega_1, \Omega_2, \dot{x}) = \frac{1}{2}(m_1 + m_2)\dot{x} \cdot \dot{x} + \frac{1}{2}\Omega_1 \cdot J_1\Omega_1 + \frac{1}{2}\Omega_2 \cdot J_2\Omega_2 \\ + \dot{x} \cdot (m_1 R_1 \hat{\Omega}_1 d_1 + m_2 R_2 \hat{\Omega}_2 d_2) - U(R_1, R_2, x) \end{aligned} \tag{2.114}$$

for a potential $U : \mathsf{SO}(3) \times \mathsf{SO}(3) \times \mathbb{R}^3 \to \mathbb{R}$ that depends on the configuration of the connected rigid bodies.

*Euler-Lagrange equations.* Derivatives of the Lagrangian are given by

$$(\mathbf{D}_{(\Omega_1,\Omega_2)}L) \cdot (\delta\Omega_1, \delta\Omega_2) = \left(J_1\Omega_1 + m_1\hat{d}_1 R_1^T \dot{x}\right) \cdot \delta\Omega_1 + \left(J_2\Omega_2 + m_2\hat{d}_2 R_2^T \dot{x}\right) \cdot \delta\Omega_2,$$

$$\mathrm{ad}^*_{(\Omega_1,\Omega_2)} \cdot \mathbf{D}_{(\Omega_1,\Omega_2)}L = \left(-\hat{\Omega}_1(J_1\Omega_1 + m_1\hat{d}_1 R_1^T \dot{x}), -\hat{\Omega}_2(J_2\Omega_2 + m_2\hat{d}_2 R_2^T \dot{x})\right),$$

$$(\mathsf{T}^*_e\mathsf{L}_{(R_1,R_2)} \cdot \mathbf{D}_{(R_1,R_2)}L) \cdot (\xi_1, \xi_2) = \left(m_1\widehat{\hat{\Omega}_1 d_1} R_1^T \dot{x} + M_1\right) \cdot \xi_1 + \left(m_2\widehat{\hat{\Omega}_2 d_2} R_2^T \dot{x} + M_2\right) \cdot \xi_2,$$

$$\mathbf{D}_{\dot{x}}L = (m_1 + m_2)\dot{x} - m_1 R_1 \hat{d}_1 \Omega_1 - m_2 R_2 \hat{d}_2 \Omega_2,$$

$$\mathbf{D}_x L = -\frac{\partial U}{\partial x},$$

where $M_i \in \mathbb{R}^3$ is obtained by the relationship, $\hat{M}_i = \frac{\partial U}{\partial R_i}^T R_i - R_i^T \frac{\partial U}{\partial R_i}$ for $i = 1, 2$.

Substituting these into (2.9) and (2.10), we obtain the Euler-Lagrange equations.

$$J_i\dot{\Omega}_i + m_i\hat{d}_i R_i^T \ddot{x} + \hat{\Omega}_i J_i \Omega_i - M_i = 0, \tag{2.115}$$

$$\dot{R}_i = R_i\hat{\Omega}_i, \tag{2.116}$$

$$(m_1 + m_2)\ddot{x} - m_1 R_1 \hat{\Omega}_1^2 d_1 + m_2 R_2 \hat{\Omega}_2^2 d_2 - m_1 R_1 \hat{d}_1 \dot{\Omega}_1 - m_2 R_2 \hat{d}_2 \dot{\Omega}_2 + \frac{\partial U}{\partial x} = 0 \tag{2.117}$$

for $i = 1, 2$. In a matrix form, these can be written as

$$\begin{bmatrix} J_1 & 0 & m_1\hat{d}_1 R_1^T \\ 0 & J_2 & m_2\hat{d}_2 R_2^T \\ (m_1\hat{d}_1 R_1^T)^T & (m_2\hat{d}_2 R_2^T)^T & (m_1+m_2)I_{3\times 3} \end{bmatrix} \begin{bmatrix} \dot{\Omega}_1 \\ \dot{\Omega}_2 \\ \ddot{x} \end{bmatrix} + \begin{bmatrix} \Omega_1 \times J_1\Omega_1 \\ \Omega_2 \times J_2\Omega_2 \\ m_1 R_1\hat{\Omega}_1^2 d_1 + m_2 R_2\hat{\Omega}_2^2 d_2 \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \\ -U_x \end{bmatrix}. \tag{2.118}$$

*Legendre Transformation.* The Legendre transformation is given by

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} J_1 & 0 & m_1\hat{d}_1 R_1^T \\ 0 & J_2 & m_2\hat{d}_2 R_2^T \\ (m_1\hat{d}_1 R_1^T)^T & (m_2\hat{d}_2 R_2^T)^T & (m_1+m_2)I_{3\times 3} \end{bmatrix} \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \dot{x} \end{bmatrix}. \tag{2.119}$$

From Corollary 2.2, this yields the Hamilton's equations.

$$\dot{p}_1 + \Omega_1 \times p_1 = m_1\widehat{\hat{\Omega}_1 d_1} R_1^T \dot{x} + M_1, \tag{2.120}$$

$$\dot{p}_2 + \Omega_2 \times p_2 = m_2\widehat{\hat{\Omega}_2 d_2} R_2^T \dot{x} + M_2, \tag{2.121}$$

$$\dot{p}_3 = -\frac{\partial U}{\partial x}. \tag{2.122}$$

## 2.4 Examples of Mechanical Systems on Two-Spheres

In Section 2.2, we have developed Lagrangian mechanics on a product of two-spheres. This result can be applied to any mechanical system whose Lagrangian is expressed as a difference between a kinetic energy with constant inertia terms, and a configuration dependent potential.

In this section, we apply the general theory developed in Section 2.2 to the following mechanical systems that evolve on a product of two-spheres. For each example, a mathematical model is defined, and the corresponding expression for Lagrangian is derived. Euler-Lagrange equations are obtained from Proposition 2.4 and Corollary 2.3.

| Section | Mechanical System |
|---|---|
| 2.4.1 | Double Spherical Pendulum |
| 2.4.2 | $n$-body Problem on a Sphere |
| 2.4.3 | Interconnection of Spherical Pendula |
| 2.4.4 | Pure Bending of Elastic Rod |
| 2.4.5 | Spatial Array of Magnetic Dipoles |
| 2.4.6 | Molecular Dynamics on a Sphere |

### 2.4.1 Double Spherical Pendulum

A double spherical pendulum is defined by two mass particles serially connected to frictionless spherical joints by rigid massless links acting under a uniform gravitational potential. The dynamics of a double spherical pendulum has been studied in Marsden et al. (1993).

Let the masses and the lengths of the pendula 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 $q_1 \in \mathsf{S}^2$ represents the direction from the pivot to the first mass, and the vector $q_2 \in \mathsf{S}^2$ represents the direction from the first mass to the second mass. The kinetic energy is given by

$$\begin{aligned} T(\dot{q}_1, \dot{q}_2) &= \frac{1}{2} m_1 \left\| l_1 \dot{q}_1 \right\|^2 + \frac{1}{2} m_2 \left\| l_1 \dot{q}_1 + l_2 \dot{q}_2 \right\|^2 \\ &= \frac{1}{2}(m_1 + m_2) l_1^q (\dot{q}_1 \cdot \dot{q}_1) + m_2 l_1 l_2 (\dot{q}_1 \cdot \dot{q}_2) + \frac{1}{2} m_2 l_2^2 (\dot{q}_2 \cdot \dot{q}_2). \end{aligned}$$

Thus, 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 $U(q_1, q_2) = -(m_1 + m_2) g l_1 e_3 \cdot q_1 - m_2 g l_2 e_3 \cdot q_2$.

Substituting these into (2.35)–(2.36), the equations of motion for the double spherical pendulum are given by

$$\begin{bmatrix} (m_1 + m_2) l_1^2 I_{3\times 3} & -m_2 l_1 l_2 \hat{q}_1 \hat{q}_2 \\ -m_2 l_1 l_2 \hat{q}_2 \hat{q}_1 & m_2 l_2^2 I_{3\times 3} \end{bmatrix} \begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \end{bmatrix} = \begin{bmatrix} m_2 l_1 l_2 (\omega_2 \cdot \omega_2) \hat{q}_1 q_2 + (m_1 + m_2) g l_1 \hat{q}_1 e_3 \\ m_2 l_1 l_2 (\omega_1 \cdot \omega_1) \hat{q}_2 q_1 + m_2 g l_2 \hat{q}_2 e_3 \end{bmatrix}, \tag{2.123}$$

$$\dot{q}_1 = \omega_1 \times q_1, \quad \dot{q}_2 = \omega_2 \times q_2, \tag{2.124}$$

which are more compact than if the equations were 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.

### 2.4.2 $n$-body Problem on a 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 (see Hairer et al. 2003). Let $m_i \in \mathbb{R}$ and $q_i \in \mathsf{S}^2$ be the mass and the position vector of the $i$-th particle, respectively. The kinetic energy is given by

$$T(\dot{q}_1, \ldots, \dot{q}_n) = \sum_{i=1}^{n} \frac{1}{2} m_i (\dot{q}_i \cdot \dot{q}_i).$$

Thus, the $i, j$-th element of the inertia matrix is $M_{ij} = m_i$ when $i = j$, and $M_{ij} = 0$ otherwise.

In Kozlov and Harin (1992), the following expression for the potential is introduced as an analogue of a gravitational potential,

$$U(q_1, \ldots, q_n) = -\frac{\gamma}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \frac{q_i \cdot q_j}{\sqrt{1 - (q_i \cdot q_j)^2}}$$

for a constant $\gamma$. Substituting these into (2.32), the equations of motion for the $n$-body problem on

a sphere are given by

$$m_i \ddot{q}_i = -m_i(\dot{q}_i \cdot \dot{q}_i)q_i - q_i \times \big(q_i \times \gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{q_j}{(1-(q_i \cdot q_j)^2)^{3/2}}\big) \tag{2.125}$$

for $i \in \{1, \ldots, n\}$.

### 2.4.3 Interconnection of Spherical Pendula

We now 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 common horizontal plane, and some pairs of pendula 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 $q_i \in \mathsf{S}^2$ represents the direction from the $i$-th pivot to the $i$-th mass. The kinetic energy is given by

$$T(\dot{q}_1, \ldots, \dot{q}_n) = \sum_{i=1}^{n} \frac{1}{2} m_i l_i^2 (\dot{q}_i \cdot \dot{q}_i).$$

Thus, the inertia matrix is given by $M_{ij} = m_i l_i^2$ when $i = j$, and $M_{ij} = 0$ otherwise.

Let $\Xi$ 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 direction along the 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

$$U(q_1, \ldots q_n) = -\sum_{i=1}^{n} m_i g l_i q_i \cdot e_3 + \sum_{(i,j)\in\Xi} \frac{1}{2}\kappa_{ij} \left( \left\| r_{ij} + \frac{1}{2} l_j q_j - \frac{1}{2} l_i q_i \right\| - \|r_{ij}\| \right)^2.$$

Substituting these into (2.34)–(2.35), the equations of motion for the interconnection of spherical pendula are given by

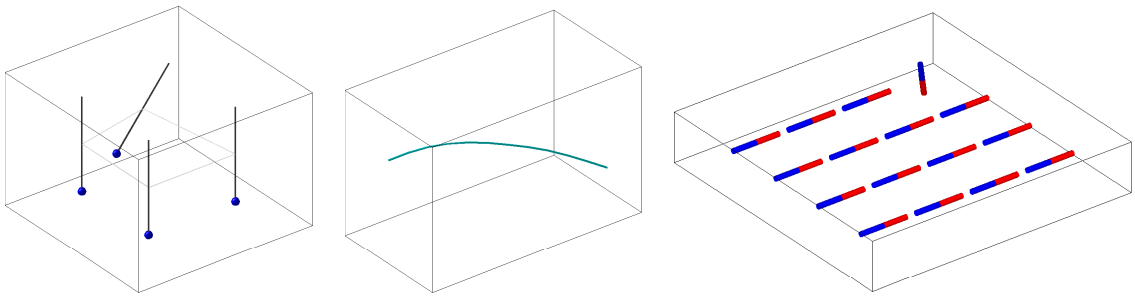
$$m_i l_i^2 \dot{\omega}_i = -q_i \times \frac{\partial U}{\partial q_i}, \tag{2.126}$$

$$\dot{q}_i = \omega_i \times q_i \tag{2.127}$$

for $i \in \{1, \ldots, n\}$.

### 2.4.4 Pure Bending of an Elastic Rod

Consider a pure bending motion of a slender elastic rod. We approximate the elastic rod by $n + 1$ slender rigid rod elements that are serially connected by spherical joints. Each rigid rod element is modeled as a line along which the mass is uniformly distributed. We assume that the 'zeroth' rod is fixed to a rigid wall. The joint has two rotational degrees of freedom; the tip of the $i$-th rod lies on



(a) Interconnection of spherical pendula (b) Pure bending of an elastic rod (c) Spatial array of magnetic dipoles

Figure 2.6: Examples of mechanical systems on two-spheres

a sphere centered at the $i$-th joint. Let $q_i \in \mathsf{S}^2$ be the direction of the $i$-th rod in the inertial frame, and let $m_i, l_i \in \mathbb{R}$ be the mass and the length of the $i$-th rod. The configuration manifold is $(\mathsf{S}^2)^n$.

The vector from the first joint to the $i$-th joint is given by $l_1 q_1 + \cdots + l_{i-1} q_{i-1} = \sum_{j=1}^{i-1} l_j q_j$. Let $s_i \in [0, l_i]$ be the distance from the $i$-th joint to a mass element $dm_i$ in the $i$-th rod. Since the mass is uniformly distributed, we have $dm_i = \frac{m_i}{l_i} ds_i$. The kinetic energy of the $i$-th rod is given by

$$\begin{aligned} T_i &= \frac{1}{2}\frac{m_i}{l_i}\int_0^{l_i} \Big\| \sum_{j=1}^{i-1} l_j \dot{q}_j + s_i \dot{q}_i \Big\|^2 ds_i \\ &= \frac{1}{6} m_i l_i^2 (\dot{q}_i \cdot \dot{q}_i) + \frac{1}{2} m_i l_i \sum_{j=1}^{i-1} l_j \dot{q}_j \cdot \dot{q}_i + \frac{1}{2} m_i \Big\| \sum_{j=1}^{i-1} l_j \dot{q}_j \Big\|^2. \end{aligned}$$

Using this, the total kinetic energy can be written as

$$T = \frac{1}{2} \sum_{i,j=1}^{n} M_{ij} \dot{q}_i \cdot \dot{q}_j, \tag{2.128}$$

where constants $M_{ij}$ for $j < i$ are defined as

$$M_{ii} = \frac{1}{3} m_i l_i^2 + \sum_{p=i+1}^{n} m_p l_i^2, \tag{2.129}$$

$$M_{ij} = \frac{1}{2}\left[ \left( \sum_{p=i+1}^{n} 2 m_p l_j l_i \right) + m_i l_j l_i \right]. \tag{2.130}$$

The potential energy is composed of a gravitational potential and a strain energy. The vector from the first joint to the mass center of the $i$-th rod is given by $l_1 q_1 + \cdots + l_{i-1} q_{i-1} + \frac{1}{2} l_i q_i$. Thus, the gravitational potential is given by

$$U_g(q_1, \ldots, q_n) = \sum_{i=1}^{n} \left[ -m_i g \left( \sum_{j=1}^{i-1} q_j l_j + \frac{1}{2} l_i q_i \right) \cdot e_3 \right]. \tag{2.131}$$

The strain potential energy for pure bending of an elastic beam is given by

$$U_\epsilon = \int_0^l \frac{EI}{2R^2} ds,$$

where $E$ is Young's modulus, $I$ is sectional area moment, and $R$ is the radius of curvature. We assume that the radius of curvature is constant along the $i$-th rod, and is approximated by $R_i = \frac{l_i}{\sin\theta_i/2}$, where $\theta_i$ denotes the angle between the $i$-th rod and the $i-1$-th rod. The strain potential energy is approximated as

$$U_\epsilon(q_1, \ldots, q_n) = \sum_{i=1}^{n} \frac{EI \sin^2(\theta_i/2)}{2l_i^2} l_i = \sum_{i=1}^{n} \frac{EI}{4l_i^2}(1 - \cos\theta_i) = \sum_{i=1}^{n} \frac{EI}{4l_i^2}(1 - q_{i-1} \cdot q_i). \quad (2.132)$$

Therefore, the total potential energy is given by

$$U(q_1, \ldots, q_n) = \sum_{i=1}^{n} \left[ -m_i g \left( \sum_{j=1}^{i-1} q_j l_j + \frac{1}{2} l_i q_i \right) \cdot e_3 + \frac{EI}{4l_i^2}(1 - q_{i-1} \cdot q_i) \right]. \quad (2.133)$$

Substituting these expressions for the inertia matrix $M_{ij}$ and the potential energy into (2.34)–(2.35), we obtain the equations of motion for the finite element model of the pure bending motion of the elastic rod.

$$M_{ii}\dot{\omega}_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} (M_{ij} q_i \times (q_j \times \dot{\omega}_j) + M_{ij}(\omega_j \cdot \omega_j) q_i \times q_j) - q_i \times \frac{\partial U}{\partial q_i}, \quad (2.134)$$

$$\dot{q}_i = \omega_i \times q_i \quad (2.135)$$

for $i \in \{1, \ldots, n\}$.

### 2.4.5 Spatial Array of Magnetic Dipoles

We now study the 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. The $n$ magnetic dipoles act under their mutual magnetic field. This can be considered as a simplified model for the dynamics of micro-magnetic particles (see Cheng et al. 2006).

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, and $q_i \in \mathsf{S}^2$ is the direction of the north pole from the pivot point. Thus, the configuration manifold is $(\mathsf{S}^2)^n$. 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 is given by

$$U(q_1,\dots,q_n)=\frac{1}{2}\sum_{\substack{i,j=1\\ j\neq i}}^{n}\frac{\mu\,\nu_i\nu_j}{4\pi\|r_{ij}\|^3}\left[(q_i\cdot q_j)-\frac{3}{\|r_{ij}\|^2}(q_i\cdot r_{ij})(q_j\cdot r_{ij})\right],$$

where $\mu$ is the permeability constant.

Substituting these into (2.34)–(2.35), the equations of motion of the spatial array of magnetic dipoles are given by

$$\frac{1}{12}m_i l_i^2\dot{\omega}_i=-q_i\times\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{\mu\,\nu_i\nu_j}{4\pi\|r_{ij}\|^3}\left[q_j-\frac{3}{\|r_{ij}\|^2}r_{ij}(q_j\cdot r_{ij})\right], \tag{2.136}$$

$$\dot{q}_i=\omega_i\times q_i \tag{2.137}$$

for $i\in\{1,\dots,n\}$.

### 2.4.6 Molecular Dynamics on a Sphere

We now study molecular dynamics on $\mathsf{S}^2$. Each molecule is modeled as a particle moving on $\mathsf{S}^2$. 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 $q_i\in\mathsf{S}^2$ 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 often used in molecular dynamics (see Lennard-Jones 1931)

$$U(q_1,\dots,q_n)=\frac{1}{2}\sum_{\substack{i,j=1\\ j\neq i}}^{n}4\epsilon\left[\left(\frac{\sigma}{\|q_i-q_j\|}\right)^{12}-\left(\frac{\sigma}{\|q_i-q_j\|}\right)^{6}\right],$$

where the first term models repulsion between nearby molecules according to the Pauli principle, and the second term models attraction at long distances generated by Van der Waals forces. The constants $\epsilon$ and $\sigma$ are molecular constants; $\epsilon$ is proportional to the strength of the mutual potential, and $\sigma$ characterizes inter-molecular forces.

Substituting these into (2.32), the equations of motion for molecular dynamics on a sphere are given by

$$m_i\ddot{q}_i=-m_i(\dot{q}_i\cdot\dot{q}_i)q_i-q_i\times\Big(q_i\times\sum_{\substack{j=1\\ j\neq i}}^{n}4\epsilon\frac{q_i-q_j}{\|q_i-q_j\|}\left[\frac{12\sigma^{12}}{\|q_i-q_j\|^{13}}-\frac{6\sigma^6}{\|q_i-q_j\|^{7}}\right]\Big) \tag{2.138}$$

for $i\in\{1,\dots,n\}$.

## 2.5 Conclusions

In Section 2.1, we have developed Euler-Lagrange equations for dynamic systems on an arbitrary Lie group. The symplectic property of Lagrangian flow, symmetry, and Lagrange-Routh reduction have been discussed. The Euler-Lagrange equations presented in Proposition 2.1 can be considered as either a generalized form of the Euler-Poincaré equation or a left-trivialized form of Euler-Lagrange equations on a manifold. The tangent bundle of a Lie group $\mathsf{TG}$ is identified with $\mathsf{G} \times \mathfrak{g}$ by left trivialization, and the equations of motion are expressed in terms of Lie group elements and Lie algebra elements. Even if the Lagrangian is not left invariant, this approach is still desirable: since the Lie algebra is a linear vector space at the fixed identity element, there is no need to deal with covariant derivatives or Christoffel symbols.

Since these expressions are coordinate-free, they are independent of a specific choice of local coordinates, and they completely avoid any singularity, ambiguity, and confusion associated with local coordinates. They provides a general framework that can be uniformly applied to dynamics of multiple rigid bodies that evolve on a Lie group. In Section 2.3, we have shown that the resulting intrinsic form of the Euler-Lagrange equations are more compact than equations expressed in terms of local coordinates, when applied to dynamics of rigid bodies.

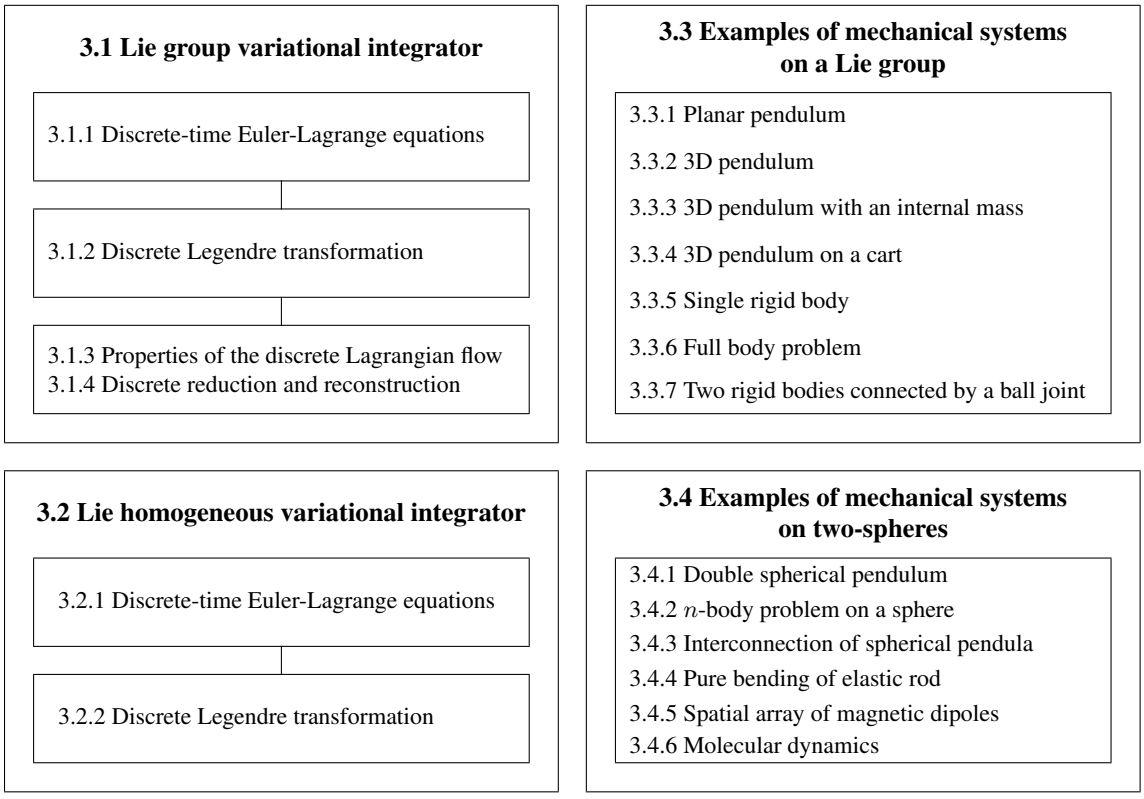
In Section 2.2, we have extended these results to mechanical systems evolving on a product of two-spheres. The variation of a curve on $\mathsf{S}^2$ is expressed in terms of $\mathfrak{so}(3)$ using the fact that the special orthogonal group acts on the two-sphere transitively. Using this property, we have derived a coordinate-free form of the Euler-Lagrange equations on $(\mathsf{S}^2)^n$. Compared with the previous literature, this approach does not require $2n$ angles or $n$ explicit equality constraints.

As shown in Section 2.4, this approach yields a compact form of the equations of motion, and also provides insight into the global dynamics on $(\mathsf{S}^2)^n$. A manifold on which a Lie group acts in a transitively way is referred to as a homogeneous manifold. The key idea of this development can be generalized to an abstract homogeneous manifold.

## CHAPTER 3

# COMPUTATIONAL GEOMETRIC MECHANICS FOR RIGID BODIES ON A LIE GROUP

This chapter deals with computational geometric mechanics for rigid bodies that evolve on a Lie group. The goal is to develop numerical integrators that preserve the geometric properties of the rigid body dynamics. The core idea is constructing computational algorithms from discrete analogues of physical principles, so that the physical properties of the dynamics are preserved naturally by the numerical computations. In particular, we discretize Hamilton's principle with careful consideration for the Lie group structure in order to develop structure-preserving numerical integrators, referred to as Lie group variational integrators or discrete-time Euler-Lagrange equations. Throughout this dissertation, these are viewed as discrete-time mechanical systems. This is in contrast with the perspective that considers numerical integration as an approximation for continuous-time equations.




This chapter has a parallel structure with Chapter 2; it may be considered as a discrete-time

version of geometric mechanics on a Lie group presented in Chapter 2. In Section 3.1, we develop discrete-time Euler-Lagrange equations for dynamic systems evolving on an abstract Lie group, and several properties of the discrete Lagrangian flow are discussed. These results are applied to several rigid body dynamics problems in Section 3.3, and numerical results are presented. The remaining part of this chapter develops computational geometric mechanics on a product of two-spheres: discrete-time Euler-Lagrange equations are developed in Section 3.2, and they are applied to several mechanical systems in Section 3.4.

Throughout this chapter, a subscript $k$ denotes the value of a variable at $t = kh + t_0$ for a fixed time step size $h > 0$, and an integer $N$ is defined such that $t_f - t_0 = Nh$.

## 3.1 Lie Group Variational Integrator

Geometric numerical integration deals with numerical integration methods that preserve geometric properties of the flow of a differential equation, such as invariants, symplecticity, and the structure of a configuration manifold (see Hairer et al. 2000; Leimkuhler and Reich 2004; McLachlan and Quispel 2001).

Numerical methods that conserve energy, momentum, or symplecticity of mechanical systems have been developed (see, for example, LaBudde and Greenspan 1976; Lasagni 1988; Sanz-Serna 1992, 1988; Simo et al. 1992). But the conservation property is often enforced by nonlinear constraints or by a projection onto the manifold defined by the constant conserved quantity.

Alternatively, a discrete-time mechanical system has been developed according to Hamilton's principle by Moser and Veselov (1991); Veselov (1988). The variational view of discrete-time mechanics is further developed by Kane et al. (1999, 2000); Wendlandt and Marsden (1997), and an intrinsic form of discrete-time variational principle is established by Marsden and West (2001). The resulting geometric numerical integrators, referred to as variational integrators, have desirable properties; they are symplectic, momentum preserving, and they exhibit excellent energy conservation property.

For differential equations that evolve on a Lie group, a group element can be updated by the corresponding group action so that the group structure is preserved naturally. This is referred to as a Lie group method (see Iserles et al. 2000). For mechanical systems evolving on a Lie group, a discrete-time Euler-Poincaré equation has been introduced for a left-invariant Lagrangian system by Marsden et al. (1999), with application to the free attitude dynamics of a rigid body. A similar work is presented for the attitude dynamics of an axially symmetric rigid body acting under a gravitational potential in Bobenko and Suris (1999).

In this section, we develop discrete-time Euler-Lagrange equations for a mechanical system evolving on an abstract Lie group $\mathsf{G}$. The Lie group method is explicitly adopted in the context of a variational integrator to construct a unified geometric integrator, referred to as a Lie group variational integrator. It preserves the geometric features of dynamics, such as symplecticity and any momentum map, as well as the geometry of the configuration manifold by automatically remaining

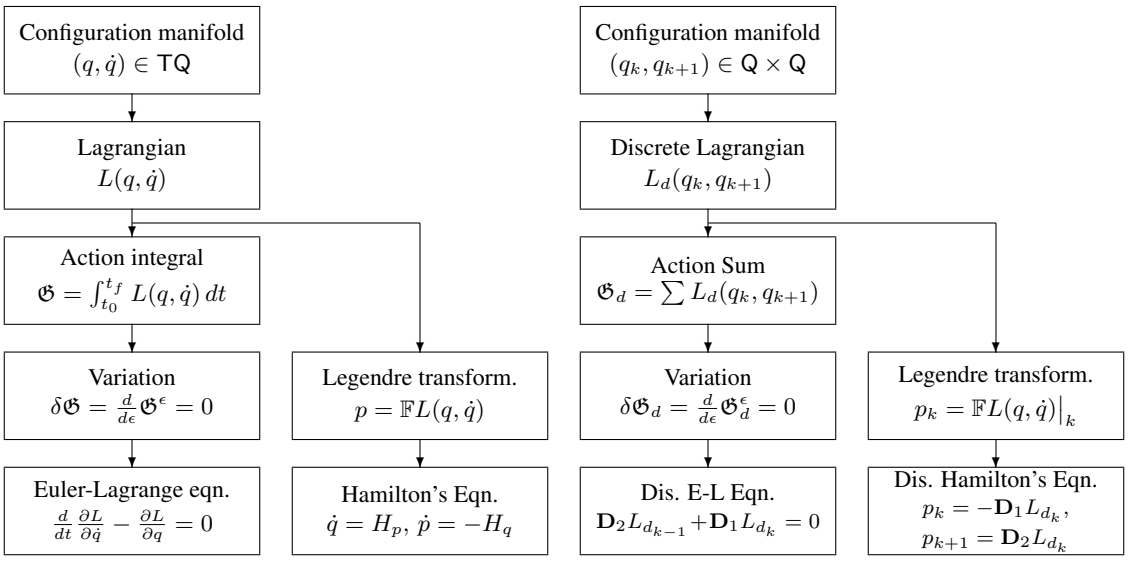



Figure 3.1: Procedures to derive the continuous/discrete Euler-Lagrange equations

on a Lie group. This provides a unified geometric numerical integrator for rigid body systems whose configuration manifold is expressed as a Lie group.

### 3.1.1 Discrete-time Euler-Lagrange Equations

Consider a mechanical system evolving on a Lie group $\mathsf{G}$. The procedures to derive the discrete-time Euler-Lagrange equations on $\mathsf{G}$ are summarized by Figure 3.1: the discrete-time trajectory is derived such that it minimizes the summation of a discrete Lagrangian, called the action sum. The discrete-time Legendre transformation provides an alternative description of mechanical systems, referred to as discrete-time Hamiltonian mechanics. The essential ideas are discretizing Hamilton's principle, where the variations of group elements are expressed in term of the Lie algebra $\mathfrak{g}$ using the exponential map, and updating group elements using group operations.

#### Configuration Manifold and Discrete Lagrangian

Discrete-time Euler-Lagrange equations evolve on $\mathsf{G} \times \mathsf{G}$. Define $f_k \in \mathsf{G}$ such that

$$g_{k+1} = g_k f_k. \tag{3.1}$$

This may be considered as a discrete-time kinematics equation, where the group element $g_{k+1}$ is obtained by a group action of $f_k$ on $g_k$. This is the essential idea of Lie group methods (see Iserles et al. 2000): this guarantees that the discrete-time flow lies on $\mathsf{G}$ without need for additional constraints or projections.

We choose the discrete Lagrangian $L_d : \mathsf{G} \times \mathsf{G} \to \mathbb{R}$ such that it approximates the integral of the Lagrangian along the exact solution of the Euler-Lagrange equations over a single time step, which

is referred to as the exact discrete Lagrangian

$$L_d^{\text{exact}}(g_k, f_k) = \int_0^h L(\tilde{g}(t), \tilde{g}^{-1}(t)\dot{\tilde{g}}(t))\, dt,$$

where $\tilde{g}(t) : [0, h] \to \mathsf{G}$ satisfies the Euler-Lagrange equations (2.9), (2.10) over $[0, h]$ with boundary conditions $\tilde{g}(0) = g_k$, $\tilde{g}(h) = g_k f_k$. The accuracy of the resulting variational integrator is equal to the accuracy of the discrete Lagrangian (see Marsden and West 2001).

### Action Sum

Define the action sum as

$$\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(g_k, f_k). \tag{3.2}$$

Since the discrete Lagrangian approximates the integral of the Lagrangian over one time step, the action sum approximates the action integral. Discrete Hamilton's principle states that this action sum does not vary to the first order for all possible variations of a curve in $\mathsf{G}$.

$$\delta\mathfrak{G}_d = \sum_{k=0}^{N-1} \delta L_d(g_k, f_k) = 0. \tag{3.3}$$

### Variations

Similar to the continuous time case given by (2.4), the variation of a sequence $\{g_k\}_{k=0}^N$ is expressed as

$$g_k^\epsilon = g_k \exp \epsilon\eta_k \tag{3.4}$$

for $k \in \{0, \ldots, N\}$, where $\{\eta_k\}_{k=0}^N$ is a sequence in $\mathfrak{g}$ satisfying $\eta_0 = \eta_N = 0$. The corresponding infinitesimal variation is given by

$$\delta g_k = g_k \eta_k. \tag{3.5}$$

Using (3.1), the infinitesimal variation of $f_k$ is given by

$$\begin{aligned}
\delta f_k &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (g_k^\epsilon)^{-1} g_{k+1}^\epsilon = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \exp(-\epsilon\eta_k) g_k^{-1} g_{k+1} \exp(\epsilon\eta_{k+1}) \\
&= -\mathsf{T}_e\mathsf{R}_{f_k} \cdot \eta_k + \mathsf{T}_e\mathsf{L}_{f_k} \cdot \eta_{k+1} \\
&= -(\mathsf{T}_{f_k}(\mathsf{L}_{f_k} \circ \mathsf{L}_{f_k^{-1}}) \circ \mathsf{T}_e\mathsf{R}_{f_k}) \cdot \eta_k + \mathsf{T}_e\mathsf{L}_{f_k} \cdot \eta_{k+1} \\
&= \mathsf{T}_e\mathsf{L}_{f_k} \cdot \left\{ -\mathrm{Ad}_{f_k^{-1}} \eta_k + \eta_{k+1} \right\}.
\end{aligned} \tag{3.6}$$

**Discrete-time Euler-Lagrange Equations**

The variation of the discrete Lagrangian is given by

$$\delta L_d(g_k, f_k) = \mathbf{D}_{g_k} L_d(g_k, f_k) \cdot \delta g_k + \mathbf{D}_{f_k} L_d(g_k, f_k) \cdot \delta f_k.$$

Using the definition of the cotangent map, we obtain

$$\begin{aligned}
\delta L_d(g_k, f_k) &= \left\langle \mathbf{D}_{g_k} L_d(g_k, f_k),\ (\mathsf{T}_e \mathsf{L}_{g_k} \circ \mathsf{T}_{g_k} \mathsf{L}_{g_k^{-1}}) \cdot \delta g_k \right\rangle \\
&\quad + \left\langle \mathbf{D}_{f_k} L_d(g_k, f_k),\ (\mathsf{T}_e \mathsf{L}_{f_k} \circ \mathsf{T}_{g_k} \mathsf{L}_{f_k^{-1}}) \cdot \delta f_k \right\rangle \\
&= \left\langle \mathsf{T}_e^* \mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k),\ \mathsf{T}_{g_k} \mathsf{L}_{g_k^{-1}} \cdot \delta g_k \right\rangle \\
&\quad + \left\langle \mathsf{T}_e^* \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k),\ \mathsf{T}_{g_k} \mathsf{L}_{f_k^{-1}} \cdot \delta f_k \right\rangle.
\end{aligned}$$

Substituting (3.5) and (3.6) into this, the variation of the Lagrangian is given by

$$\begin{aligned}
\delta L_d(g_k, f_k) &= \langle \mathsf{T}_e^* \mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k),\ \eta_k \rangle + \left\langle \mathsf{T}_e^* \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k),\ -\mathrm{Ad}_{f_k^{-1}} \eta_k + \eta_{k+1} \right\rangle \\
&= \left\langle \mathsf{T}_e^* \mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k) - \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}_e^* \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k)),\ \eta_k \right\rangle \\
&\quad + \langle \mathsf{T}_e^* \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} \mathsf{L}_d(g_k, f_k),\ \eta_{k+1} \rangle.
\end{aligned} \tag{3.7}$$

Therefore, the variation of the action sum is given by

$$\begin{aligned}
\delta \mathfrak{G}_d = \sum_{k=0}^{N-1} &\left\langle \mathsf{T}_e^* \mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k) - \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}_e^* \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k)),\ \eta_k \right\rangle \\
&+ \langle \mathsf{T}_e^* \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} \mathsf{L}_d(g_k, f_k),\ \eta_{k+1} \rangle.
\end{aligned} \tag{3.8}$$

The summation index can be rewritten as

$$\begin{aligned}
\delta \mathfrak{G}_d = &\left\langle \mathsf{T}_e^* \mathsf{L}_{f_{N-1}} \cdot \mathbf{D}_{f_{N-1}} L_d(g_{N-1}, f_{N-1}),\ \eta_N \right\rangle \\
&+ \left\langle \mathsf{T}_e^* \mathsf{L}_{g_0} \cdot \mathbf{D}_{g_0} L_d(g_0, f_0) - \mathrm{Ad}^*_{f_0^{-1}} \cdot (\mathsf{T}_e^* \mathsf{L}_{f_0} \cdot \mathbf{D}_{f_0} L_d(g_0, f_0)),\ \eta_0 \right\rangle \\
&+ \sum_{k=1}^{N-1} \left\langle \mathsf{T}_e^* \mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k) - \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}_e^* \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k)),\ \eta_k \right\rangle \\
&\quad + \left\langle \mathsf{T}_e^* \mathsf{L}_{f_{k-1}} \cdot \mathbf{D}_{f_{k-1}} \mathsf{L}_d(g_{k-1}, f_{k-1}),\ \eta_k \right\rangle.
\end{aligned} \tag{3.9}$$

Since $\eta_k = 0$ at $k = 0, N$, the first two terms of the above equation vanish. From discrete Hamilton's principle, $\delta \mathfrak{G}_d = 0$ for all possible variations, which yields the discrete-time Euler-Lagrange equations on $\mathsf{G}$.

**Proposition 3.1** *Consider a mechanical system evolving on a Lie group $\mathsf{G}$. The discrete-time kinematics equation is defined as (3.11), where the group element $g_{k+1}$ is updated by the right group action of $f_k \in \mathsf{G}$ on $g_k$. For the given discrete Lagrangian $L_d(g_k, f_k) : \mathsf{G} \times \mathsf{G} \to \mathbb{R}$, the corre-*

*sponding discrete-time Euler-Lagrange equations are given by*

$$\mathsf{T}_e^*\mathsf{L}_{f_{k-1}} \cdot \mathbf{D}_{f_{k-1}} L_d(g_{k-1}, f_{k-1}) - \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k)) + \mathsf{T}_e^*\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k) = 0, \tag{3.10}$$

$$g_{k+1} = g_k f_k. \tag{3.11}$$

*For given $(g_{k-1}, f_{k-1})$, we obtain $g_k = g_{k-1}f_{k-1}$ from (3.11), and we solve (3.10) to find $f_k$. This yields a discrete-time flow map $(g_{k-1}, f_{k-1}) \to (g_k, f_k)$, and this process is repeated.*

**Remark 3.1** If the discrete Lagrangian is not dependent on $g_k$, then the third term of (3.10) vanishes. The resulting equations are equivalent to the discrete Euler-Poincaré equations (see Marsden et al. 1999). Therefore, (3.10) can be considered as a generalization of the discrete Euler-Poincaré equations.

**Remark 3.2** These equations are obtained using the right group action of $f_k$ on $g_k$ at (3.11). We can develop similar equations using the left group action on $g_k$. This is summarized by the following corollary.

**Corollary 3.1** *Consider a mechanical system evolving on a Lie group $\mathsf{G}$. The discrete-time kinematics equation is defined as (3.13), where the group element $g_{k+1}$ is updated by the left group action of $r_k$ on $g_k$. For the given discrete Lagrangian $L_d(g_k, r_k) : \mathsf{G} \times \mathsf{G} \to \mathbb{R}$ the corresponding discrete-time Euler-Lagrange equations are given by*

$$\mathsf{T}_e^*\mathsf{R}_{r_{k-1}} \cdot \mathbf{D}_{r_{k-1}} L_d(g_{k-1}, r_{k-1}) - \mathrm{Ad}^*_{r_k} \cdot (\mathsf{T}_e^*\mathsf{R}_{r_k} \cdot \mathbf{D}_{r_k} L_d(g_k, r_k)) + \mathsf{T}_e^*\mathsf{R}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, r_k) = 0, \tag{3.12}$$

$$g_{k+1} = r_k g_k. \tag{3.13}$$

### 3.1.2 Discrete Legendre Transformation

Equations (3.10) and (3.11) yield the discrete-time Lagrangian flow map $(g_k, f_k) \to (g_{k+1}, f_{k+1})$. But sometimes, it is more useful to express the discrete-time flow map in the cotangent bundle using the discrete Legendre transformation.

Define discrete Legendre transforms $\mathbb{F}^+L_d, \mathbb{F}^-L_d : \mathsf{G} \times \mathsf{G} \to \mathsf{G} \times \mathfrak{g}^*$ as

$$\mathbb{F}^+L_d(g_k, f_k) = (g_k f_k, \mu_{k+1}),$$
$$\mathbb{F}^-L_d(g_k, f_k) = (g_k, \mu_k),$$

where $\mu_k, \mu_{k+1} \in \mathfrak{g}^*$ are given by

$$\mu_k = -\mathsf{T}_e^*\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_{d_k} + \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_{d_k}), \tag{3.14}$$

$$\mu_{k+1} = \mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_{d_k}. \tag{3.15}$$

These are well-defined, since (3.10) can be expressed in terms of these discrete Legendre transforms as

$$\mathbb{F}^+L_d(g_{k-1}, f_{k-1}) = \mathbb{F}^-L_d(g_k, f_k). \tag{3.16}$$

For a given $(g_k, \mu_k) \in \mathsf{G} \times \mathfrak{g}^*$, the inverse of the negative discrete Legendre transformation gives $(\mathbb{F}^-L_d)^{-1}(g_k, \mu_k) = (g_k, f_k)$, and the positive discrete Legendre transformation of this gives $\mathbb{F}^+L_d(g_k, f_k) = (g_{k+1}, \mu_{k+1})$. We combine these to obtain the discrete-time Hamiltonian flow $\tilde{\mathcal{F}}_{L_d} : \mathsf{G} \times \mathfrak{g}^* \to \mathsf{G} \times \mathfrak{g}^*$ as

$$\tilde{\mathcal{F}}_{L_d} = \mathbb{F}^+L_d \circ (\mathbb{F}^-L_d)^{-1}, \tag{3.17}$$

Using the discrete-time Lagrangian flow map $\mathcal{F}_{L_d}(g_k, f_k) = (g_{k+1}, f_{k+1})$, (3.16) can be written as

$$\mathbb{F}^+L_d = \mathbb{F}^-L_d \circ \mathcal{F}_{L_d}.$$

Using this, the discrete-time Hamiltonian flow map can be alternatively written as

$$\tilde{\mathcal{F}}_{L_d} = \mathbb{F}^{\pm}L_d \circ F_{L_d} \circ (\mathbb{F}^{\pm}L_d)^{-1}. \tag{3.18}$$

The Hamiltonian flow map that corresponds to (3.17) is summarized as follows.

**Corollary 3.2** *Consider a mechanical system evolving on a Lie group* $\mathsf{G}$*. The discrete-time kinematics equation is defined as (3.11), where the group element* $g_{k+1}$ *is updated by the right group action of* $f_k \in \mathsf{G}$ *on* $g_k$*. For the given discrete Lagrangian* $L_d(g_k, f_k) : \mathsf{G} \times \mathsf{G} \to \mathbb{R}$*, the corresponding discrete-time Hamilton's equations are given by*

$$\mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}^*_e\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k}L_{d_k}) = \mu_k + \mathsf{T}^*_e\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k}L_{d_k}, \tag{3.19}$$

$$g_{k+1} = g_k f_k, \tag{3.20}$$

$$\mu_{k+1} = \mathrm{Ad}^*_{f_k} \cdot (\mu_k + \mathsf{T}^*_e\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k}L_{d_k}). \tag{3.21}$$

*For given* $(g_k, \mu_k)$*, we solve (3.19) to find* $f_k$*. Then,* $g_{k+1}$ *and* $\mu_{k+1}$ *are obtained from (3.20) and (3.21), respectively. This yields a discrete-time flow map* $(g_k, \mu_k) \to (g_{k+1}, \mu_{k+1})$*, and this process is repeated.*

### 3.1.3 Properties of the Discrete-time Lagrangian Flow

We show two properties of the discrete-time Lagrangian flow, namely symplecticity and momentum preservation. The subsequent development can be considered as a special form of general properties of discrete Lagrangian flows, applied to a Lie group configuration manifold (see Marsden and West 2001).

**Symplecticity**

Let $\Theta^+_{L_d}, \Theta^-_{L_d}$ be the discrete Lagrangian one-forms on $\mathsf{G} \times \mathsf{G}$ given by

$$\Theta^+_{L_d}(g_k, f_k) \cdot (\delta g_k, \delta f_k) = \left\langle \mathsf{T}^*_e \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_{d_k},\ f_k^{-1}\delta f_k + \mathrm{Ad}_{f_k^{-1}} \cdot g_k^{-1}\delta g_k \right\rangle, \tag{3.22}$$

$$\Theta^-_{L_d}(g_k, f_k) \cdot (\delta g_k, \delta f_k) = -\left\langle \mathsf{T}^*_e \mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_{d_k} - \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}^*_e \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_{d_k}),\ g_k^{-1}\delta g_k \right\rangle. \tag{3.23}$$

From (3.6), we have

$$\eta_{k+1} = f_k^{-1}\delta f_k + \mathrm{Ad}_{f_k^{-1}} \cdot g_k^{-1}\delta g_k.$$

Substituting this into (3.22) and comparing with (3.7), it can be shown that $\mathbf{d}L_d = \Theta^+_{L_d} - \Theta^-_{L_d}$. Since the second-order exterior derivatives of any form is zero, i.e. $\mathbf{d}^2 = 0$, the exterior derivatives of two discrete Lagrangian one-forms are the same $\mathbf{d}\Theta^+_{L_d} = \mathbf{d}\Theta^-_{L_d}$, which is defined to be the discrete Lagrangian symplectic form $\Omega_{L_d}$ on $\mathsf{G} \times \mathsf{G}$.

$$\Omega_{L_d} = \mathbf{d}\Theta^+_{L_d} = \mathbf{d}\Theta^-_{L_d}. \tag{3.24}$$

We define the discrete-time Lagrangian flow map $\mathcal{F}_{L_d} : \mathsf{G} \times \mathsf{G} \to \mathsf{G} \times \mathsf{G}$ as the flow of the discrete Euler-Lagrange equations (3.10), (3.11).

**Proposition 3.2** *The discrete-time Lagrangian flow preserves the discrete Lagrangian two-form as follows*

$$(\mathcal{F}^{N-1}_{L_d})^* \Omega_{L_d} = \Omega_{L_d}. \tag{3.25}$$

**Proof.** Define the solution space $\mathcal{C}_{L_d}$ to be the set of solutions $\{g_k \in \mathsf{G}\}_{k=0}^N$ of (3.10) and (3.11). Since an element of $\mathcal{C}_{L_d}$ is uniquely determined by the initial condition $(g_0, f_0) \in \mathsf{G} \times \mathsf{G}$, we can identify $\mathcal{C}_{L_d}$ with the manifold of initial conditions $\mathsf{G} \times \mathsf{G}$. Define the restricted action map $\hat{\mathfrak{G}}_d : \mathsf{G} \times \mathsf{G} \to \mathbb{R}$ by

$$\hat{\mathfrak{G}}_d(g_0, f_0) = \mathfrak{G}_d(\{g'_k\}_{k=0}^N),$$

where $\{g'_k\}_{k=0}^N \in \mathcal{C}_L$ is the solution of the discrete-time Euler-Lagrange equations with the initial conditions $(g'_0, g'_1) = (g_0, g_0 f_0)$. Since this satisfies (3.10), (3.9) reduces to

$$\mathbf{d}\hat{\mathfrak{G}}_d \cdot w = ((\mathcal{F}^{N-1}_{L_d})^* \Theta^+_{L_d} - \Theta^-_{L_d}) \cdot w \tag{3.26}$$

for any $w = (\delta g_k, \delta f_k) \in \mathsf{TG} \times \mathsf{TG}$. We take a derivative of (3.26). Since exterior derivatives and pull backs commute, we obtain

$$\mathbf{d^2}\hat{\mathfrak{G}}_d = ((\mathcal{F}^{N-1}_{L_d})^* \mathbf{d}\Theta^+_{L_d} - \mathbf{d}\Theta^-_{L_d}).$$

Since $\mathbf{d^2}\hat{\mathfrak{G}}_d = 0$, we obtain (3.25). ∎

**Discrete Noether's theorem**

Consider the action of a Lie group $\mathsf{H}$ on $\mathsf{G}$, $\Phi : \mathsf{H} \times \mathsf{G} \to \mathsf{G}$ introduced in Section 2.1.4. Recall the infinitesimal generator $\zeta_{\mathsf{G}} : \mathsf{G} \to \mathsf{G} \times \mathfrak{g}$ defined as (2.20) for $\zeta \in \mathsf{H}$. Here, we define the infinitesimal generator $\zeta_{\mathsf{G}\times\mathsf{G}} : \mathsf{G} \times \mathsf{G} \to \mathsf{TG} \times \mathsf{TG}$ as

$$\zeta_{\mathsf{G}\times\mathsf{G}}(g_k, f_k) = \left(\mathsf{T}_e\mathsf{L}_{g_k} \cdot \zeta_{\mathsf{G}}(g_k), \mathsf{T}_e\mathsf{L}_{f_k} \cdot (-\mathrm{Ad}_{f_k^{-1}}\zeta_{\mathsf{G}}(g_k) + \zeta_{\mathsf{G}}(g_k f_k))\right). \tag{3.27}$$

We define two discrete Lagrangian momentum map $J^+_{L_d}, J^-_{L_d} : \mathsf{G} \times \mathsf{G} \to \mathfrak{h}^*$ as

$$J^+_{L_d}(g_k, f_k) \cdot \zeta = \Theta^+_{L_d} \cdot \zeta_{\mathsf{G}\times\mathsf{G}}(g_k, f_k), \tag{3.28}$$

$$J^-_{L_d}(g_k, f_k) \cdot \zeta = \Theta^-_{L_d} \cdot \zeta_{\mathsf{G}\times\mathsf{G}}(g_k, f_k). \tag{3.29}$$

**Proposition 3.3** *Suppose that the discrete Lagrangian is invariant under the lifted action, i.e.* $\mathbf{d}L_d \cdot \zeta_{\mathsf{G}\times\mathsf{G}} = 0$ *for any* $\zeta \in \mathfrak{h}$. *Then, the two discrete Lagrangian momentum map are the same,* $J^+_{L_d} = J^-_{L_d}$, *which is denoted by* $J_{L_d} : \mathsf{G} \times \mathsf{G} \to \mathfrak{h}^*$, *and the discrete-time Lagrangian flow preserves the discrete Lagrangian momentum map.*

$$J_{L_d}((\mathcal{F}^{N-1}_{L_d}(g_0, f_0)) = J_{L_d}(g_0, f_0). \tag{3.30}$$

*This is called discrete Noether's theorem.*

**Proof.** Since $\mathbf{d}L_d = \Theta^+_{L_d} - \Theta^-_{L_d}$, we have

$$\mathbf{d}L_d \cdot \zeta_{\mathsf{G}\times\mathsf{G}} = (\Theta^+_{L_d} - \Theta^-_{L_d}) \cdot \zeta_{\mathsf{G}\times\mathsf{G}} = (J^+_{L_d} - J^-_{L_d}) \cdot \zeta,$$

which is equal to zero for any $\zeta \in \mathfrak{h}$ since the discrete Lagrangian is invariant under the lifted action. Thus, $J^+_{L_d} = J^-_{L_d}$.

Since the action is the summation of the discrete Lagrangian, $L_d(g_k, f_k) \cdot \zeta_{\mathsf{G}\times\mathsf{G}}$ implies that $\mathfrak{G}_d \cdot \zeta_{\mathsf{G}\times\mathsf{G}} = 0$. We can restrict it to the solution space to obtain

$$\mathbf{d}\hat{\mathfrak{G}}_d \cdot \zeta_{\mathsf{G}\times\mathsf{G}} = 0.$$

But, from (3.26), we obtain

$$\begin{aligned}\mathbf{d}\hat{\mathfrak{G}}_d \cdot \zeta_{\mathsf{G}\times\mathsf{G}} &= ((\mathcal{F}^{N-1}_{L_d})^*\Theta^+_{L_d} - \Theta^-_{L_d}) \cdot \zeta_{\mathsf{G}\times\mathsf{G}} \\ &= (J^+_{L_d}(\mathcal{F}^{N-1}_{L_d}(g_k, f_k)) - J^-_{L_d}(g_k, f_k)) \cdot \zeta \end{aligned} \tag{3.31}$$

for any $\zeta \in \mathfrak{h}$, which yields (3.30). ∎

### 3.1.4 Discrete Reduction and Reconstruction

In Section 2.1.5, we have discussed that if there is a symmetry in the Lagrangian, the configuration manifold can be reduced to a shape space. Similarly, if the discrete Lagrangian has a symmetry,

discrete-time reduced Euler-Lagrange equations evolving on the shape space can be derived. The discrete-time flow in the original configuration manifold can be reconstructed from the solution of the discrete-time reduced Euler-Lagrange equations.

The procedure for the discrete Lagrange-Routh reduction is similar to that of continuous-time Lagrange-Routh reduction (see Jalnapurkar et al. 2006). In essence, the discrete reduced Lagrangian satisfies a reduced variational principle with a one-form derived from the connection. This yields the discrete-time flow on the shape space. For a given trajectory on the shape space, we reconstruct the flow on the original configuration manifold by finding the horizontal lift of the reduced trajectory and applying the momentum map preservation property.

Discrete-time reduction and reconstruction for the full body problem has been studied in Lee et al. (2007c), and the results are summarized in Section 3.3.6.

## 3.2 Lie Homogeneous Variational Integrator on Two-Spheres

In this section, we develop discrete-time Euler-Lagrange equations for mechanical systems evolving on a product of two-spheres. The goal is to develop geometric numerical integrators that preserve the geometric properties of the Lagrangian/Hamiltonian dynamics as well as the structure of two-spheres.

As discussed in Section 2.2, the special orthogonal group $\mathsf{SO}(3) = \{R \in \mathbb{R}^{3\times 3} \,|\, R^T R = I, \det R = 1\}$ acts on the two-sphere transitively, i.e. for any $q_1, q_2 \in \mathsf{S}^2$, there exists a $R \in \mathsf{SO}(3)$ such that $q_2 = Rq_1$. The essential idea to derive discrete-time Euler-Lagrange equations on two-spheres is to update elements in the two-sphere using the group action for the special orthogonal group $\mathsf{SO}(3)$. Consequently, the discrete flow evolves on the two-spheres without need for constraints or reprojection. This is referred to as a Lie homogeneous variational integrator on two-spheres.

Compared with geometric numerical integrators on $\mathsf{S}^2$ developed by Lewis and Nigam (2003); Lewis and Olver (2001); Munthe-Kaas and Zanna (1997), this approach conserves the geometric properties of dynamic systems as well as the structure of two-spheres, and it does not require local coordinates or explicit equality constraints. The subsequent development can also be generalized to mechanical systems evolving on an abstract homogeneous manifold.

### 3.2.1 Discrete-time Euler-Lagrange Equations

The procedures to derive the discrete-time Euler-Lagrange equations on $(\mathsf{S}^2)^n$ are summarized by Figure 3.1: the discrete-time trajectory is derived such that it minimizes the summation of a discrete Lagrangian, called the action sum. The discrete-time Legendre transformation provides an alternative description of mechanical systems, referred to as discrete-time Hamiltonian mechanics. The essential ideas are discretizing Hamilton's principle, where the variations are expressed in term of the Lie algebra $\mathfrak{so}(3)$ using the exponential map, and updating elements in two-spheres using the group operation of $\mathsf{SO}(3)$.

### Configuration Manifold and Discrete Lagrangian

We consider a mechanical system evolving on a product of $n$ two-spheres $(\mathsf{S}^2)^n$ introduced in Section 2.2. Recalling (2.26), the Lagrangian has the following structure

$$L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) = \frac{1}{2} \sum_{i,j=1}^{n} M_{ij} \dot{q}_i \cdot \dot{q}_j - U(q_1, \ldots, q_n) \tag{3.32}$$

for a constant symmetric positive definite inertia matrix $\{M_{ij}\}_{i,j=1}^n$ and a configuration dependent potential $U : (\mathsf{S}^2)^n \to \mathbb{R}$.

According to the trapezoidal rule, we define a discrete Lagrangian $L_d : (\mathsf{S}^2)^n \times (\mathsf{S}^2)^n \to \mathbb{R}$

$$L_d(q_{1_k}, \ldots, q_{n_k}, q_{1_{k+1}}, \ldots, q_{n_{k+1}}) = \frac{1}{2h} \sum_{i,j=1}^{n} M_{ij}(q_{i_{k+1}} - q_{i_k}) \cdot (q_{j_{k+1}} - q_{j_k}) - \frac{h}{2} U_k - \frac{h}{2} U_{k+1}, \tag{3.33}$$

where $U_k$ denotes the value of the potential at the $k$-th step, i.e. $U_k = U(q_{1_k}, \ldots, q_{n_k})$. Here, we assume that the inertia matrix is constant and the discrete Lagrangian is obtained by using the trapezoidal rule. The following development can be generalized to mechanical systems with a configuration dependent inertia or a general form of the discrete Lagrangian.

### Action Sum

Using the expression for the discrete Lagrangian, the action sum is defined as

$$\begin{aligned} \mathfrak{G}_d &= \sum_{k=0}^{N-1} L_d(q_{1_k}, \ldots, q_{n_k}, q_{1_{k+1}}, \ldots, q_{n_{k+1}}) \\ &= \frac{1}{2h} \sum_{k=0}^{N-1} \sum_{i,j=1}^{n} M_{ij}(q_{i_{k+1}} - q_{i_k}) \cdot (q_{j_{k+1}} - q_{j_k}) - \frac{h}{2} U_k - \frac{h}{2} U_{k+1}. \end{aligned}$$

Since the discrete Lagrangian approximates the integral of the Lagrangian over one time step, the action sum approximates the action integral. Discrete Hamilton's principle states that the variation of the action sum is zero.

### Variations

Similar to the continuous time case given by (2.28), the variation of a discrete-time curve $\{q_{i_k}\}_{k=0}^N$ is expressed as

$$q_{i_k}^{\epsilon} = \exp \epsilon \hat{\eta}_{i_k} q_{i_k} \tag{3.34}$$

for a discrete-time curve $\{\eta_{i_k}\}_{k=0}^N$ on $\mathbb{R}^3$ satisfying $\eta_{i_0} = \eta_{i_N} = 0$ for $i \in \{1, \ldots, n\}$. We assume $\eta_{i_k} \in \mathbb{R}^3$ is constrained to be orthogonal to $q_{i_k}$, i.e. $\eta_{i_k} \cdot q_{i_k} = 0$. The corresponding infinitesimal

variation of $q_{i_k}$ is written as

$$\delta q_{i_k} = \eta_{i_k} \times q_{i_k}. \tag{3.35}$$

**Discrete-time Euler-Lagrange Equations**

The variation of the discrete Lagrangian can be written as

$$\delta L_{d_k} = \frac{1}{h} \sum_{i,j=1}^{n} (\delta q_{i_{k+1}} - \delta q_{i_k}) \cdot M_{ij}(q_{j_{k+1}} - q_{j_k}) - \frac{h}{2} \sum_{i=1}^{n} \left( \delta q_{i_k} \cdot \frac{\partial U_k}{\partial q_{i_k}} + \delta q_{i_{k+1}} \cdot \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}} \right). \tag{3.36}$$

Substituting (3.35) into (3.36), 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

$$\begin{aligned} \delta L_{d_k} = & \frac{1}{h} \sum_{i,j=1}^{n} \left( \eta_{i_{k+1}} \cdot (q_{i_{k+1}} \times M_{ij}(q_{j_{k+1}} - q_{j_k})) - \eta_{i_k} \cdot (q_{i_k} \times M_{ij}(q_{j_{k+1}} - q_{j_k})) \right) \\ & - \frac{h}{2} \sum_{i=1}^{n} \left( \eta_{i_k} \cdot \left( q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right) + \eta_{i_{k+1}} \cdot \left( q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}} \right) \right). \end{aligned} \tag{3.37}$$

Therefore, the variation of the action sum is given by

$$\begin{aligned} \delta \mathfrak{G}_d = & \frac{1}{h} \sum_{k=0}^{N-1} \sum_{i,j=1}^{n} \left( \eta_{i_{k+1}} \cdot (q_{i_{k+1}} \times M_{ij}(q_{j_{k+1}} - q_{j_k})) - \eta_{i_k} \cdot (q_{i_k} \times M_{ij}(q_{j_{k+1}} - q_{j_k})) \right) \\ & - \frac{h}{2} \sum_{k=0}^{N-1} \sum_{i=1}^{n} \left( \eta_{i_k} \cdot \left( q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right) + \eta_{i_{k+1}} \cdot \left( q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}} \right) \right) \\ = & \sum_{i,j=1}^{n} \eta_{i_N} \cdot \left[ \frac{1}{h} q_{i_N} \times M_{ij}(q_{j_N} - q_{j_{N-1}}) - \frac{h}{2} q_{i_N} \times \frac{\partial U_N}{\partial q_{i_N}} \right] \\ & - \sum_{i,j=1}^{n} \eta_{i_0} \cdot \left[ \frac{1}{h} q_{i_0} \times M_{ij}(q_{j_1} - q_{j_0}) - \frac{h}{2} q_{i_0} \times \frac{\partial U_0}{\partial q_{i_0}} \right] \\ & + \sum_{k=1}^{N-1} \sum_{i=1}^{n} \eta_{i_k} \cdot \left[ \frac{1}{h} (q_{i_k} \times \sum_{j=1}^{n} M_{ij}(-q_{j_{k+1}} + 2q_{j_k} - q_{j_{k-1}})) - h q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right]. \end{aligned}$$

Since $\eta_{i_0} = \eta_{i_N} = 0$ for $i \in \{1, \ldots, n\}$, we obtain

$$\delta \mathfrak{G}_d = \sum_{k=1}^{N-1} \sum_{i=1}^{n} \eta_{i_k} \cdot \left[ \frac{1}{h} (q_{i_k} \times \sum_{j=1}^{n} M_{ij}(-q_{j_{k+1}} + 2q_{j_k} - q_{j_{k-1}})) - h q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right]. \tag{3.38}$$

From discrete Hamilton's principle $\delta \mathfrak{G}_d = 0$ for any $\eta_{i_k}$ perpendicular to $q_{i_k}$. Using the same

argument given in (2.30), the discrete-time equations of motion are given by

$$\frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(-q_{j_{k+1}} + 2q_{j_k} - q_{j_{k-1}})) - hq_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} = 0 \tag{3.39}$$

for $i \in \{1, \dots n\}$.

In addition, we require that the unit length of the vector $q_{i_k}$ be preserved. This is achieved by viewing $\mathsf{S}^2$ as a homogeneous manifold. Since the special orthogonal group $\mathsf{SO}(3)$ acts on $\mathsf{S}^2$ transitively, we can define a discrete update map for $q_{i_k}$ as

$$q_{i_{k+1}} = F_{i_k} q_{i_k}$$

for $F_{i_k} \in \mathsf{SO}(3)$. Then, the unit length of the vector $q_i$ is preserved through the discrete-time 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 3.4** *Consider a mechanical system on $(\mathsf{S}^2)^n$ whose Lagrangian is expressed as (2.26). The discrete-time Euler-Lagrange equations are given by*

$$M_{ii} q_{i_k} \times F_{i_k} q_{i_k} + q_{i_k} \times \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{ij}(F_{j_k} - I_{3\times 3}) q_{j_k} = q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_k} - q_{j_{k-1}}) - h^2 q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}}, \tag{3.40}$$

$$q_{i_{k+1}} = F_{i_k} q_{i_k} \tag{3.41}$$

*for $i \in \{1, \dots n\}$. For given $(q_{i_{k-1}}, q_{i_k})$, we solve (3.40) to obtain $F_{i_k} \in \mathsf{SO}(3)$. Then, $q_{i_{k+1}}$ is computed by (3.41). This yields a discrete-time flow map $(q_{i_{k-1}}, q_{i_k}) \to (q_{i_k}, q_{i_{k+1}})$, and this process is repeated.*

### 3.2.2 Discrete Legendre Transformation

The discrete Legendre transformation is given as follows.

$$\begin{aligned} p_{i_k} \cdot \delta q_{i_k} &= -\mathbf{D}_{q_{i_k}} L_{d_k} \cdot \delta q_{i_k} \\ &= \left[ \frac{1}{h} \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k}) + \frac{h}{2} \frac{\partial U_k}{\partial q_{i_k}} \right] \cdot \delta q_{i_k}, \end{aligned}$$

which can be directly obtained from (3.36). This is satisfied for any $\delta q_{i_k} \in \mathsf{T}_{q_{i_k}} \mathsf{S}^2$ perpendicular to $q_{i_k}$. Using the same argument used to derive (2.37), 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

$$p_{i_k} = -\frac{1}{h} q_{i_k} \times (q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{i_k} \times \left( q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right). \tag{3.42}$$

Similarly, we obtain

$$\begin{aligned} p_{i_{k+1}} \cdot \delta q_{i_{k+1}} &= \mathbf{D}_{q_{i_{k+1}}} L_{d_k} \cdot \delta q_{i_{k+1}} \\ &= \left[ \frac{1}{h} \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k}) - \frac{h}{2} \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}} \right] \cdot \delta q_{i_{k+1}}. \end{aligned}$$

Since $p_{i_{k+1}}$ is perpendicular to $q_{i_{k+1}}$, it is given by

$$p_{i_{k+1}} = -\frac{1}{h} q_{i_{k+1}} \times (q_{i_{k+1}} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) + \frac{h}{2} q_{i_{k+1}} \times \left( q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}} \right). \quad (3.43)$$

This yields the discrete-time Hamilton's equations as follows.

**Corollary 3.3** *Consider a mechanical system on $(\mathsf{S}^2)^n$ whose Lagrangian is expressed as (2.26). The discrete-time Hamilton's equations are given by*

$$p_{i_k} = -\frac{1}{h} q_{i_k} \times (q_{i_k} \times \sum_{j=1}^{n} M_{ij}(F_{j_k} - I_{3\times 3}) q_{j_k}) - \frac{h}{2} q_{i_k} \times \left( q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right), \quad (3.44)$$

$$q_{i_{k+1}} = F_{i_k} q_{i_k}, \quad (3.45)$$

$$p_{i_{k+1}} = -\frac{1}{h} q_{i_{k+1}} \times (q_{i_{k+1}} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) + \frac{h}{2} q_{i_{k+1}} \times \left( q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}} \right) \quad (3.46)$$

*for $i \in \{1, \ldots, n\}$. For given $(q_{i_k}, p_{i_k})$, we solve (3.44) to obtain $F_{i_k} \in \mathsf{SO}(3)$. Then, $q_{i_{k+1}}$ and $p_{i_{k+1}}$ are computed by (3.45) and (3.46), respectively. This yields a discrete-time flow map $(q_{i_k}, p_{i_k}) \to (q_{i_{k+1}}, p_{i_{k+1}})$, 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 (3.44) to (2.37), substituting $\dot{q}_{i_k} = \omega_{i_k} \times q_{i_k}$, and rearranging, we obtain

$$q_{i_k} \times \left[ M_{ii}\omega_{i_k} + (q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(\omega_{j_k} \times q_{j_k})) - \frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right] = 0.$$

Since the expression in the brackets is orthogonal to $q_{i_k}$, the left side is equal to zero if and only if the expression in the brackets is zero. Thus,

$$M_{ii}\omega_{i_k} + (q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(\omega_{j_k} \times q_{j_k})) = \frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) + \frac{h}{2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}}. \quad (3.47)$$

This provides a relationship between $(q_{i_k}, \omega_{i_k})$ and $(q_{i_k}, q_{i_{k+1}})$. Comparing this with (3.39), we obtain

$$M_{ii}\omega_{i_k} + (q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(\omega_{j_k} \times q_{j_k})) = \frac{1}{h}(q_{i_k} \times \sum_{j=1}^{n} M_{ij}(q_{j_k} - q_{j_{k-1}})) - \frac{h}{2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}}, \quad (3.48)$$

which provides a relationship between $(q_{i_k}, \omega_{i_k})$ and $(q_{i_{k-1}}, q_{i_k})$. Equations (3.47) and (3.48) provide a discrete-time flow map in terms of the angular velocity; for a given $(q_{i_k}, \omega_{i_k})$, we find $(q_{i_k}, q_{i_{k+1}})$ by using (3.47). Substituting this into (3.48) expressed at the $k+1$th step, we obtain $(q_{i_{k+1}}, \omega_{i_{k+1}})$. This procedure is summarized as follows.

**Corollary 3.4** *The discrete-time equations of motion given by (3.40) and (3.41) can be written in terms of the angular velocity as*

$$\begin{aligned} M_{ii} q_{i_k} \times F_{i_k} q_{i_k} + q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(F_{j_k} - I_{3\times 3}) q_{j_k} \\ = M_{ii} h \omega_{i_k} - (q_{i_k} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(q_{j_k} \times h\omega_{j_k})) - \frac{h^2}{2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}}, \end{aligned} \quad (3.49)$$

$$q_{i_{k+1}} = F_{i_k} q_{i_k}, \quad (3.50)$$

$$\begin{aligned} M_{ii}\omega_{i_{k+1}} - (q_{i_{k+1}} \times \sum_{\substack{j=1 \\ j\neq i}}^{n} M_{ij}(q_{j_{k+1}} \times \omega_{j_{k+1}})) \\ = \frac{1}{h}(q_{i_{k+1}} \times \sum_{j=1}^{n} M_{ij}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}} \end{aligned} \quad (3.51)$$

*for $i \in \{1, \ldots, n\}$. Equivalently, (3.51) can be written in a matrix form as*

$$\begin{aligned} &\begin{bmatrix} M_{11} I_{3\times 3} & -M_{12}\hat{q}_{1_{k+1}}\hat{q}_{2_{k+1}} & \cdots & -M_{1n}\hat{q}_1\hat{q}_{n_{k+1}} \\ -M_{21}\hat{q}_{2_{k+1}}\hat{q}_{1_{k+1}} & M_{22} I_{3\times 3} & \cdots & -M_{2n}\hat{q}_{2_{k+1}}\hat{q}_{n_{k+1}} \\ \vdots & \vdots & & \vdots \\ -M_{n1}\hat{q}_{n_{k+1}}\hat{q}_{1_{k+1}} & -M_{n2}\hat{q}_{n_{k+1}}\hat{q}_{2_{k+1}} & \cdots & M_{nn} I_{3\times 3} \end{bmatrix} \begin{bmatrix} \omega_{1_{k+1}} \\ \omega_{2_{k+1}} \\ \vdots \\ \omega_{n_{k+1}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{h}(q_{1_{k+1}} \times \sum_{j=1}^{n} M_{1j}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{1_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{1_{k+1}}} \\ \frac{1}{h}(q_{2_{k+1}} \times \sum_{j=1}^{n} M_{2j}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{2_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{2_{k+1}}} \\ \vdots \\ \frac{1}{h}(q_{n_{k+1}} \times \sum_{j=1}^{n} M_{nj}(q_{j_{k+1}} - q_{j_k})) - \frac{h}{2} q_{n_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{n_{k+1}}} \end{bmatrix}. \end{aligned} \quad (3.52)$$

*For a given $(q_{i_k}, \omega_{i_k})$, we solve (3.49) to obtain $F_{i_k} \in \mathsf{SO(3)}$. Then, $q_{i_{k+1}}$ and $\omega_{i_{k+1}}$ are computed*

*by (3.50) and (3.52), respectively. This yields a discrete-time flow map in terms of the angular velocity* $(q_{i_k}, \omega_{i_k}) \to (q_{i_{k+1}}, \omega_{i_{k+1}})$*, and this process is repeated.*

The discrete-time Euler-Lagrange equations given by (3.40) and (3.49) are implicit equations: we need to solve an implicit equation at each time step, in order to find the relative update represented by the rotation matrix $F_{i_k} \in \mathsf{SO}(3)$. A computational approach to solve the implicit equations is presented in Section 3.4.7, where it is shown that the discrete-time Euler-Lagrange equations become explicit when the inertia matrix is diagonal, i.e. $M_{ij} = 0$ for $i \neq j$. The explicit form of the discrete-time Euler-Lagrange equations for mechanical systems evolving on a product of two-spheres is summarized as follows.

**Corollary 3.5** *Consider a mechanical system on* $(\mathsf{S}^2)^n$ *whose Lagrangian is expressed as (2.26) where* $M_{ij} = 0$ *for* $i \neq j$*, i.e. the dynamics are coupled only though the potential energy. The explicit discrete-time equations of motion are given by*

$$q_{i_{k+1}} = \left( h\omega_{i_k} - \frac{h^2}{2M_{ii}} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right) \times q_{i_k} + \left( 1 - \left\| h\omega_{i_k} - \frac{h^2}{2M_{ii}} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right\|^2 \right)^{1/2} q_{i_k}, \quad (3.53)$$

$$\omega_{i_{k+1}} = \omega_{i_k} - \frac{h}{2M_{ii}} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} - \frac{h}{2M_{ii}} q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}} \quad (3.54)$$

*for* $i \in \{1, \ldots, n\}$.

## 3.3 Examples of Mechanical Systems on a Lie Group

In Section 3.1, we have developed discrete-time Lagrangian mechanics on an abstract Lie group. Since the configuration manifold of dynamics of rigid bodies is a Lie group, the results provide a unified framework that can be applied to various rigid body dynamics.

In this section, we apply the general theory developed in Section 3.1 to the following rigid body dynamics discussed in Section 2.3. For each example, a discrete Lagrangian is chosen, and Euler-Lagrange equations and Legendre transformations are obtained, followed by numerical results.

| Section | Mechanical System | $\mathsf{G}$ |
|---|---|---|
| 3.3.1 | Planar pendulum | $\mathsf{SO}(2)$ |
| 3.3.2 | 3D pendulum | $\mathsf{SO}(3)$ |
| 3.3.3 | 3D pendulum with an internal degree of freedom | $\mathsf{SO}(3) \times \mathbb{R}$ |
| 3.3.4 | 3D pendulum on a cart | $\mathsf{SO}(3) \times \mathbb{R}^2$ |
| 3.3.5 | Single rigid body | $\mathsf{SE}(3)$ |
| 3.3.6 | Full body problem | $(\mathsf{SE}(3))^n$ |
| 3.3.7 | Two rigid bodies connected by a ball joint | $\mathsf{SO}(3) \times \mathsf{SO}(3) \times \mathbb{R}^3$ |

### 3.3.1 Planar Pendulum

Consider the planar pendulum model presented in Section 2.3.1.

*Configuration Manifold.* The configuration manifold is $\mathsf{SO}(2)$, and the group action for $\mathsf{SO}(2)$ is matrix multiplication. Thus, the discrete update map (3.11) can be written as

$$R_{k+1} = R_k F_k,$$

for $F_k \in \mathsf{SO}(2)$. The adjoint operator $\mathrm{Ad}_R$ for $R \in \mathsf{SO}(2)$ is the identity on $\mathfrak{so}(2)$.

*Discrete Lagrangian.* Recalling (2.41), the Lagrangian of the planar pendulum is given by

$$L(R, \hat{\Omega}) = \frac{1}{4} m l^2 \mathrm{tr}\left[\hat{\Omega}^T \hat{\Omega}\right] + m g e_2^T R \rho. \tag{3.55}$$

From the attitude kinematics equation (2.43), the angular velocity is approximated by

$$\hat{\Omega}_k \approx \frac{1}{h} R_k^T (R_{k+1} - R_k) = \frac{1}{h}(F_k - I).$$

According to the trapezoidal rule, the discrete Lagrangian is chosen as

$$\begin{aligned} L_d(R_k, F_k) &= \frac{1}{4h} m l^2 \mathrm{tr}\big[(F_k - I_{2\times 2})^T (F_k - I_{2\times 2})\big] + \frac{h}{2} m g l e_2^T R_k e_2 + \frac{h}{2} m g l e_2^T R_k F_k e_2 \\ &= \frac{1}{2h} m l^2 \mathrm{tr}[(I_{2\times 2} - F_k)] + \frac{h}{2} m g l e_2^T R_k e_2 + \frac{h}{2} m g l e_2^T R_k F_k e_2. \end{aligned}$$

*Discrete Euler-Lagrange Equations.* The variation of the discrete Lagrangian is given by

$$\begin{aligned}\delta L(R_k, F_k) &= -\frac{1}{2h}ml^2\mathrm{tr}[\delta F_k] + \frac{h}{2}mgle_2^T R_k\hat{\eta}_k e_2 + \frac{h}{2}mgle_2^T R_k\hat{\eta}_k F_k e_2 + \frac{h}{2}mgle_2^T R_k\delta F_k e_2\\ &= -\frac{1}{2h}ml^2\mathrm{tr}\big[F_k(F_k^T\delta F_k)\big] + \frac{h}{2}mgle_2^T R_k\hat{\eta}_k e_2 + \frac{h}{2}mgle_2^T R_k F_k(F_k^T\hat{\eta}_k F_k)e_2\\ &\quad + \frac{h}{2}mgle_2^T R_k F_k(F_k^T\delta F_k)e_2,\end{aligned}$$

Since $F\hat{\eta}F^T = \hat{\eta}$ for any $F \in \mathrm{SO}(2)$ and $\hat{\eta} \in \mathfrak{so}(2)$, $F_k^T\delta F_k$ is skew-symmetric, we obtain

$$\begin{aligned}\delta L(R_k, F_k) &= -\frac{1}{4h}ml^2\mathrm{tr}\big[(F_k - F_k^T)(F_k^T\delta F_k)\big] - \frac{h}{2}mgle_2^T(R_k + R_{k+1})e_1\cdot\eta_k\\ &\quad - \frac{h}{2}mgle_2^T R_{k+1}e_1\cdot(F_k^T\delta F_k)^\wedge\\ &= \left\langle \frac{1}{2h}ml^2(F_k - F_k^T) - \frac{h}{2}mgle_2^T\widehat{R_{k+1}e_1},\ F_k^T\delta F_k\right\rangle\\ &\quad - \left\langle \frac{h}{2}mgl(\widehat{e_2^T R_k e_1} + e_2^T\widehat{R_{k+1}e_1}),\ \hat{\eta}_k\right\rangle,\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}\mathsf{T}_e^*\mathsf{L}_{F_k}\cdot\mathbf{D}_{F_k}L_d(R_k, F_k) &= \frac{1}{2h}ml^2(F_k - F_k^T) - \frac{h}{2}mgle_2^T\widehat{R_{k+1}e_1},\\ \mathsf{T}_e^*\mathsf{L}_{R_k}\cdot\mathbf{D}_{R_k}L_d(R_k, F_k) &= -\frac{h}{2}mgl(\widehat{e_2^T R_k e_1} + e_2^T\widehat{R_{k+1}e_1}).\end{aligned}$$

Substituting these equations into (3.10), we obtain the discrete-time Euler-Lagrange equations for the planar pendulum as

$$\frac{1}{2h}ml^2(F_k - F_k^T) - \frac{1}{2h}ml^2(F_{k+1} - F_{k+1}^T) - hmgle_2^T\widehat{R_{k+1}e_1} = 0, \tag{3.56}$$

$$R_{k+1} = R_k F_k. \tag{3.57}$$

*Discrete Legendre Transformation.* From (3.14) and (3.15), the discrete Legendre transformations are given by

$$\begin{aligned}\hat{\Pi}_k &= \frac{1}{2h}ml^2(F_k - F_k^T) + \frac{h}{2}mgle_2^T\widehat{R_k e_1},\\ \hat{\Pi}_{k+1} &= \frac{1}{2h}ml^2(F_k - F_k^T) - \frac{h}{2}mgle_2^T\widehat{R_{k+1}e_1},\end{aligned}$$

where $\Pi_k = ml^2\Omega \in \mathfrak{so}(2)^* \simeq \mathbb{R}^*$. These yield the discrete-time Hamilton's equations as

$$\frac{1}{2h}ml^2(F_k - F_k^T) = \hat{\Pi}_k - \frac{h}{2}mgle_2^T\widehat{R_k e_1}, \tag{3.58}$$

$$R_{k+1} = R_k F_k, \tag{3.59}$$

$$\hat{\Pi}_{k+1} = \hat{\Pi}_k - \frac{h}{2}mgle_2^T\widehat{R_k e_1} - \frac{h}{2}mgle_2^T\widehat{R_{k+1}e_1}. \tag{3.60}$$

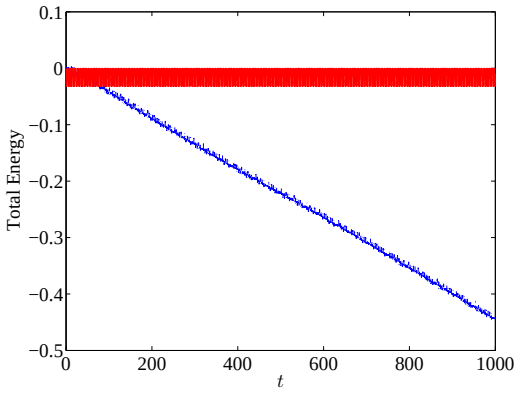



Figure 3.2: Numerical simulation for a planar pendulum: computed total energy (LGVI: red, solid, RK4(5): blue, dotted)

For a given $(R_k, \hat{\Pi}_k)$, we solve (3.58) to obtain $F_k$. Using this, $(R_{k+1}, \hat{\Pi}_{k+1})$ is obtained by (3.59) and (3.60). This yields a discrete-time flow map $(R_k, \hat{\Pi}_k) \to (R_{k+1}, \hat{\Pi}_{k+1})$. If we parameterize $R_k$ using $\theta_k$, (3.58) and (3.59) are equivalent to

$$\theta_{k+1} = \theta_k + \sin^{-1}\left(\frac{h}{ml^2}\Pi_k - \frac{h^2 g}{2l} e_2^T R_k e_1\right). \tag{3.61}$$

*Numerical Results.* We compare the computational properties of the discrete-time equations of motion given by (3.58)–(3.60) 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-time equations of motion is chosen such that the CPU times are identical. Figure 3.2 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 $1.0835 \times 10^{-2}\,\mathrm{Nm}$. 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.

### 3.3.2 3D Pendulum

Consider the 3D pendulum model presented in Section 2.3.2.

*Configuration manifold.* The configuration manifold for the 3D pendulum is the special orthogonal group, $\mathsf{SO(3)}$. A rotation matrix $R \in \mathsf{SO(3)}$ is a linear transformation from a representation of a vector in the body fixed frame into a representation of the vector in the inertial frame.

We define $F_k \in \mathsf{SO(3)}$ as $F_k = R_k^T R_{k+1}$. Thus, we have the discrete-time attitude kinematics equation as

$$R_{k+1} = R_k F_k. \tag{3.62}$$

The rotation matrix $F_k$ represents the relative attitude update between two integration steps, and by

requiring that $F_k$ lies on $\mathsf{SO(3)}$ we guarantee that the discrete flow $R_k$ for $k \in \{0, \ldots, N\}$ evolves on $\mathsf{SO(3)}$ automatically.

The adjoint operator is as follows.

$$\mathrm{Ad}_F \hat{\eta} = F\hat{\eta}F^T = \widehat{F\eta}, \quad \mathrm{Ad}^*_F \hat{\eta} = F^T\hat{\eta}F = \widehat{F^T\eta} \tag{3.63}$$

for $F \in \mathsf{SO(3)}$, $\hat{\eta} \in \mathfrak{so}(3)$.

*Discrete Lagrangian.* Recall that the Lagrangian $L : \mathsf{SO(3)} \times \mathfrak{so}(3) \to \mathbb{R}$ of the attitude dynamics of the 3D pendulum is given by

$$L(R, \hat{\Omega}) = \frac{1}{2}\mathrm{tr}\left[\hat{\Omega} J_d \hat{\Omega}^T\right] - U(R) \tag{3.64}$$

for an attitude dependent gravitational potential $U : \mathsf{SO(3)} \to \mathbb{R}$.

Using the kinematics equations (2.48), (3.62), $\hat{\Omega}_k$ can be approximated as

$$\hat{\Omega}_k = R_k^T \dot{R}_k \approx \frac{1}{h} R_k^T (R_{k+1} - R_k) = \frac{1}{h}(F_k - I). \tag{3.65}$$

From the trapezoidal rule, we choose the following form of the discrete Lagrangian $L_d : \mathsf{SO(3)} \times \mathsf{SO(3)} \to \mathbb{R}$.

$$\begin{aligned} L_d(R_k, F_k) &= \frac{h}{2} L(R_k, \frac{1}{h}(F_k - I)) + \frac{h}{2} L(R_{k+1}, \frac{1}{h}(F_k - I)) \\ &= \frac{1}{2h}\mathrm{tr}\left[(F_k - I) J_d (F_k - I)^T\right] - \frac{h}{2}U(R_k) - \frac{h}{2}U(R_{k+1}) \\ &= \frac{1}{2h}\mathrm{tr}\left[F_k J_d F_k^T - F_k J_d - J_d F_k^T + J_d\right] - \frac{h}{2}U(R_k) - \frac{h}{2}U(R_{k+1}). \end{aligned}$$

Since $\mathrm{tr}\left[F_k J_d F_k^T\right] = \mathrm{tr}\left[J_d F_k^T F_k\right] = \mathrm{tr}[J_d]$, and $\mathrm{tr}\left[J_d F_k^T\right] = \mathrm{tr}[F_k J_d]$, the discrete Lagrangian can be written as

$$L_d(R_k, F_k) = \frac{1}{h}\mathrm{tr}[(I - F_k) J_d] - \frac{h}{2}U(R_k) - \frac{h}{2}U(R_k F_k). \tag{3.66}$$

*Discrete-time Euler-Lagrange Equations.* We first find expressions for the derivatives of the discrete Lagrangian. Let $T_d : \mathsf{SO(3)} \to \mathbb{R}$ be the first term of the discrete Lagrangian

$$T_d(F_k) = \frac{1}{h}\mathrm{tr}[(I - F_k) J_d]\,.$$

The derivative of $T_d(F_k)$ with respect to $F_k$ is given by

$$\begin{aligned} \mathbf{D}_{F_k} T_d(F_k) \cdot \delta F_k &= -\frac{1}{h}\mathrm{tr}[\delta F_k J_d] \cdot \delta F_k = -\frac{1}{h}\mathrm{tr}\left[F_k F_k^T \delta F_k J_d\right] \\ &= -\frac{1}{h}\mathrm{tr}\left[(F_k^T \delta F_k)(J_d F_k)\right], \end{aligned} \tag{3.67}$$

where we use the property of the trace $\mathrm{tr}[AB] = \mathrm{tr}[BA]$ for any $A, B \in \mathbb{R}^{n\times n}$. The following identity is satisfied for any $x \in \mathbb{R}^3, B \in \mathbb{R}^{3\times 3}$, since $\mathrm{tr}[\hat{x}B] = -\mathrm{tr}\big[B^T\hat{x}\big] = -\mathrm{tr}\big[\hat{x}B^T\big]$:

$$\mathrm{tr}[\hat{x}B] = \frac{1}{2}\mathrm{tr}\big[\hat{x}(B - B^T)\big] = -\left\langle B - B^T,\, \hat{x}\right\rangle. \tag{3.68}$$

Since $F_k^T\delta F_k$ is skew symmetric, (3.67) can be written as

$$\mathbf{D}_{F_k}T_d(F_k)\cdot\delta F_k = -\frac{1}{h}\mathrm{tr}[\delta F_k J_d] = \frac{1}{h}\left\langle J_dF_k - F_k^TJ_d,\, F_k^T\delta F_k\right\rangle. \tag{3.69}$$

As in (2.60), the derivative of the potential is given by

$$\mathbf{D}_{R_k}U(R_k)\cdot\delta R_k = -\left\langle \hat{M}_k,\, R_k^T\delta R_k\right\rangle = \frac{1}{2}\mathrm{tr}\Big[\hat{M}_kR_k^T\delta R_k\Big], \tag{3.70}$$

where $M_k$ is determined by $\hat{M}_k = \frac{\partial U_k}{\partial R_k}R_k^T - R_k\frac{\partial U_k}{\partial R_k}^T$. Therefore, we have

$$\begin{aligned}\mathbf{D}_{R_k}U(R_kF_k)\cdot\delta R_k &= \frac{1}{2}\mathrm{tr}\Big[\hat{M}_{k+1}F_k^TR_k^T\delta R_kF_k\Big] = \frac{1}{2}\mathrm{tr}\Big[(R_k^T\delta R_k)(F_k\hat{M}_{k+1}F_k^T)\Big]\\ &= -\left\langle \widehat{F_k\hat{M}_{k+1}},\, R_k^T\delta R_k\right\rangle.\end{aligned} \tag{3.71}$$

Similarly, we also have

$$\mathbf{D}_{F_k}U(R_kF_k)\cdot\delta F_k = -\left\langle \hat{M}_{k+1},\, F_k^T\delta F_k\right\rangle. \tag{3.72}$$

From (3.69), (3.70), (3.71), and (3.72), the derivatives of the discrete Lagrangian are given by

$$\begin{aligned}\mathbf{D}_{F_k}L_d(R_k,F_k)\cdot\delta F_k &= \mathbf{D}_{F_k}T_d(F_k)\cdot\delta F_k - \frac{h}{2}\mathbf{D}_{F_k}U(R_kF_k)\cdot\delta F_k\\ &= \left\langle \frac{1}{h}(J_dF_k - F_k^TJ_d) + \frac{h}{2}\hat{M}_{k+1},\, F_k^T\delta F_k\right\rangle,\\ \mathbf{D}_{R_k}L_d(R_k,F_k)\cdot\delta R_K &= -\frac{h}{2}\mathbf{D}_{R_k}U(R_k)\cdot\delta R_k - \frac{h}{2}\mathbf{D}_{R_k}U(R_kF_k)\cdot\delta R_k\\ &= \frac{h}{2}\left\langle \hat{M}_k + \widehat{F_k\hat{M}_{k+1}},\, R_k^T\delta R_k\right\rangle.\end{aligned}$$

Therefore, we obtain

$$\mathsf{T}_e^*\mathsf{L}_{F_k}\cdot\mathbf{D}_{F_k}L_d(R_k,F_k) = \frac{1}{h}(J_dF_k - F_k^TJ_d) + \frac{h}{2}\hat{M}_{k+1}, \tag{3.73}$$

$$\mathrm{Ad}^*_{F_{k+1}^T}\mathsf{T}_e^*\mathsf{L}_{F_{k+1}}\cdot\mathbf{D}_{F_{k+1}}L_d(R_{k+1},F_{k+1}) = \frac{1}{h}(F_{k+1}J_d - J_dF_{k+1}^T) + \frac{h}{2}\widehat{F_{k+1}\hat{M}_{k+2}}, \tag{3.74}$$

$$\mathsf{T}_e^*\mathsf{L}_{R_{k+1}}\cdot\mathbf{D}_{R_{k+1}}L_d(R_{k+1},F_{k+1}) = \frac{h}{2}(\hat{M}_{k+1} + \widehat{F_{k+1}\hat{M}_{k+2}}). \tag{3.75}$$

Substituting these into (3.10) and (3.11), the discrete-time Euler-Lagrange equations for the 3D

pendulum are given by

$$\frac{1}{h}\left(F_{k+1}J_d - J_dF_{k+1}^T - J_dF_k + F_k^T J_d\right) = h\hat{M}_{k+1}, \tag{3.76}$$

$$R_{k+1} = R_kF_k, \tag{3.77}$$

where $M_k = mg\rho_c \times R_k^T e_3$. For given $(R_k, R_{k+1})$ and $F_k = R_k^T R_{k+1}$, we solve the implicit equation (3.76) to find $F_{k+1}$. Then, $R_{k+2}$ is obtained from (3.77). This yields a discrete-time flow map $(R_k, R_{k+1}) \to (R_{k+1}, R_{k+2})$, and this procedure is repeated.

*Discrete Legendre Transformation.* From (3.14), (3.15), the discrete Legendre transformation is given by

$$\begin{aligned}\hat{\Pi}_k &= -\mathsf{T}_e^*\mathsf{L}_{R_k}\cdot\mathbf{D}_{R_k}L_d(R_k,F_k) + \mathrm{Ad}^*_{F_k^T}\cdot\mathsf{T}_e^*\mathsf{L}_{F_k}\cdot\mathbf{D}_{F_k}L_d(R_k,F_k)\\ &= \frac{1}{h}(F_kJ_d - J_dF_k^T) - \frac{h}{2}\hat{M}_k, \qquad (3.78)\\ \hat{\Pi}_{k+1} &= \mathsf{T}_e^*\mathsf{L}_{F_k}\cdot\mathbf{D}_{F_k}L_d(R_k,F_k)\\ &= \frac{1}{h}(J_dF_k - F_k^TJ_d) + \frac{h}{2}\hat{M}_{k+1}. \qquad (3.79)\end{aligned}$$

Combining these, we have

$$\hat{\Pi}_{k+1} = \frac{1}{h}F_k^T(F_kJ_d - J_dF_k^T)F_k + \frac{h}{2}\hat{M}_{k+1},$$
$$\hat{\Pi}_{k+1} = F_k^T\hat{\Pi}_kF_k + \frac{h}{2}\hat{M}_k + \frac{h}{2}\hat{M}_{k+1}.$$

Using the fact that $F^T\hat{\Pi}F = \widehat{F^T\Pi}$ for any $\Pi \in \mathbb{R}^3$ and $F \in \mathsf{SO(3)}$, we obtain equivalent equations in a vector form. In summary, the discrete-time Hamilton's equations are given by

$$h\big(\Pi_k + \frac{h}{2}M_k\big)^\wedge = F_kJ_d - J_dF_k^T, \tag{3.80}$$

$$R_{k+1} = R_kF_k, \tag{3.81}$$

$$\Pi_{k+1} = F_k^T\Pi_k + \frac{h}{2}F_k^TM_k + \frac{h}{2}M_{k+1}. \tag{3.82}$$

For given $(R_k, \Pi_k)$, we solve the implicit equation (3.80) to find $F_k$. Then, $R_{k+1}$ is obtained from (3.81), and $\Pi_{k+1}$ is obtained from (3.82). This yields a discrete map $(R_k, \Pi_k) \to (R_{k+1}, \Pi_{k+1})$, and this procedure is repeated.

The discrete-time equations given by (3.76) and (3.80) are implicit equations: we need to solve an implicit equation at each time step in order to find the relative update represented by the rotation matrix $F_k \in \mathsf{SO(3)}$. A computational approach for these implicit equations is presented in Section 3.3.8.

**Remark 3.3** In this section, we have developed discrete-time Euler-Lagrange equations and discrete-time Hamilton's equations for the 3D pendulum by substituting the discrete Lagrangian given by

(3.66) into the discrete-time Euler-Lagrange equations on a general Lie group (3.10). Alternatively, the discrete Euler-Lagrange equations for the 3D pendulum can be derived by following the procedure presented in Section 3.1.1. The corresponding discrete Hamilton's principle on $\mathsf{SO(3)}$ has been studied in Lee et al. (2005b).

*Symmetry.* Recall that the symmetry action of the 3D pendulum $\Phi : \mathsf{S}^1 \times \mathsf{SO(3)} \to \mathsf{SO(3)}$ is given by

$$\Phi(\theta, R) = \exp_{\mathsf{SO(3)}}(\theta \hat{e}_3) R, \tag{3.83}$$

which represents the rotation of the 3D pendulum about the gravity direction $e_3$.

We first find expressions for the infinitesimal generators. We identify $\mathfrak{h} = \mathfrak{so}(2)$ with $\mathbb{R}$. From (2.20), for $\zeta \in \mathbb{R}$, the infinitesimal generator $\zeta_{\mathsf{SO(3)}} : \mathsf{SO(3)} \to \mathsf{SO(3)} \times \mathfrak{so}(3)$ is given by

$$\begin{aligned}
\zeta_{\mathsf{SO(3)}}(R) &= \phi_L \circ \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \Phi_{\exp_{\mathsf{SO(2)}} \epsilon\zeta}(R) = \phi_L \circ \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \exp_{\mathsf{SO(3)}}(\epsilon\zeta\hat{e}_3) R \\
&= \phi_L \circ (R,\, \zeta \hat{e}_3 R) = (R,\, \zeta \widehat{R^T e_3}).
\end{aligned} \tag{3.84}$$

From the definition (3.27), the infinitesimal generator $\zeta_{\mathsf{SO(3)}\times\mathsf{SO(3)}} : \mathsf{SO(3)} \times \mathsf{SO(3)} \to \mathsf{TSO(3)} \times \mathsf{TSO(3)}$ is given by

$$\begin{aligned}
\zeta_{\mathsf{SO(3)}\times\mathsf{SO(3)}}(R_k, F_k) &= \left( \mathsf{T}_e \mathsf{L}_{R_k} \cdot \zeta_{\mathsf{SO(3)}}(R_k),\, \mathsf{T}_e \mathsf{L}_{F_k} \cdot (-\mathrm{Ad}_{F_k^T} \zeta_{\mathsf{SO(3)}}(R_k) + \zeta_{\mathsf{SO(3)}}(R_k F_k)) \right) \\
&= \zeta \left( R_k \widehat{R_k^T e_3},\, F_k (R_{k+1}^T e_3 - F_k^T R_k^T e_3)^\wedge \right).
\end{aligned} \tag{3.85}$$

Now we show that the discrete Lagrangian is infinitesimally invariant under the symmetry action. From (3.22), (3.23), the discrete Lagrangian one-forms on $\mathsf{SO(3)} \times \mathsf{SO(3)}$ are given by

$$\begin{aligned}
\Theta^+_{L_d}(R_k, F_k) \cdot (\delta R_k, \delta F_k) &= \left\langle \mathsf{T}^*_e \mathsf{L}_{F_k} \cdot \mathbf{D}_{F_k} L_{d_k},\, F_k^T \delta F_k + \mathrm{Ad}_{F_k^T} \cdot R_k^T \delta R_k \right\rangle \\
&= \left\langle \frac{1}{h}(J_d F_k - F_k^T J_d) + \frac{h}{2} \hat{M}_{k+1},\, F_k^T \delta F_k + F_k^T R_k^T \delta R_k F_k \right\rangle, \\
\Theta^-_{L_d}(R_k, F_k) \cdot (\delta R_k, \delta F_k) &= -\left\langle \mathsf{T}^*_e \mathsf{L}_{R_k} \cdot \mathbf{D}_{R_k} L_{d_k} - \mathrm{Ad}^*_{F_k^T} \cdot (\mathsf{T}^*_e \mathsf{L}_{F_k} \cdot \mathbf{D}_{F_k} L_{d_k}),\, R_k^T \delta R_k \right\rangle \\
&= \left\langle \frac{1}{h}(F_k J_d - J_d F_k^T) - \frac{h}{2} \hat{M}_k,\, R_k^T \delta R_k \right\rangle.
\end{aligned}$$

Since $L_{d_k} = \Theta^+_{L_d} - \Theta^-_{L_d}$, we obtain

$$\begin{aligned}
L_d(R_k, F_k) \cdot \zeta_{\mathsf{SO(3)}\times\mathsf{SO(3)}} &= (\Theta^+_{L_d} - \Theta^-_{L_d}) \cdot \zeta_{\mathsf{SO(3)}\times\mathsf{SO(3)}} \\
&= \left\langle \frac{1}{h}(J_d F_k - F_k^T J_d) + \frac{h}{2} \hat{M}_{k+1},\, \zeta \widehat{R_{k+1}^T e_3} \right\rangle - \left\langle \frac{1}{h}(F_k J_d - J_d F_k^T) - \frac{h}{2} \hat{M}_k,\, \zeta \widehat{R_k^T e_3} \right\rangle.
\end{aligned}$$

Since $M_k = mg\rho_c \times R_k^T e_3$, we have $\left\langle \hat{M}_k, \widehat{R_k^T e_3} \right\rangle = 0$, $\left\langle \hat{M}_{k+1}, \widehat{R_{k+1}^T e_3} \right\rangle = 0$. Thus, we obtain

$$\begin{aligned} L_d(R_k, F_k) \cdot \zeta_{\mathsf{SO(3)}\times\mathsf{SO(3)}} &= \left\langle \frac{1}{h}(J_d F_k - F_k^T J_d),\, \zeta\widehat{R_{k+1}^T e_3} \right\rangle - \left\langle \frac{1}{h}(F_k J_d - J_d F_k^T),\, \zeta\widehat{R_k^T e_3} \right\rangle \\ &= \left\langle \frac{1}{h}(J_d F_k - F_k^T J_d),\, \zeta\widehat{R_{k+1}^T e_3} \right\rangle - \left\langle \frac{1}{h}F_k^T(F_k J_d - J_d F_k^T)F_k,\, \zeta F_k^T\widehat{R_k^T e_3}F_k \right\rangle \\ &= \left\langle \frac{1}{h}(J_d F_k - F_k^T J_d),\, \zeta\widehat{R_{k+1}^T e_3} \right\rangle - \left\langle \frac{1}{h}(J_d F_k - F_k^T J_d),\, \zeta\widehat{(F_k R_k)^T e_3} \right\rangle = 0. \end{aligned}$$

Therefore, the discrete Lagrangian is infinitesimally invariant under the symmetry action.

We now find an expression for the discrete momentum map. According to (3.28), the momentum map $J_{L_d}^+ : \mathsf{SO(3)} \times \mathsf{SO(3)} \to \mathbb{R}^*$ is given by

$$\begin{aligned} J_{L_d}^+(R_k, F_k) \cdot \zeta &= \Theta_{L_d}^+ \cdot \zeta_{\mathsf{G}\times\mathsf{G}}(R_k, F_k) \\ &= \left\langle \frac{1}{h}(J_d F_k - F_k^T J_d) + \frac{h}{2}\hat{M}_{k+1},\, \zeta\widehat{R_{k+1}^T e_3} \right\rangle \\ &= \left\langle \hat{\Pi}_{k+1},\, \zeta\widehat{R_{k+1}^T e_3} \right\rangle = \zeta e_3^T R_{k+1}\Pi_{k+1}, \end{aligned} \tag{3.86}$$

where we use the discrete Legendre transformation (3.79). Thus, $J_{L_d}^+(R_k, F_k) = e_3^T R_{k+1}\Pi_{k+1}$, which represents the angular momentum of the 3D pendulum about the gravity direction. This is preserved by the discrete-time flow according to the discrete Noether's theorem.

**Remark 3.4** The preservation of the angular momentum about the gravity direction can be directly shown from (3.82). Multiplying the left and right sides of (3.82) by $e_3^T R_{k+1}$, we obtain

$$\begin{aligned} e_3^T R_{k+1}\Pi_{k+1} &= e_3^T R_{k+1}F_k^T\Pi_k + \frac{h}{2}e_3^T R_{k+1}F_k^T M_k + \frac{h}{2}e_3^T R_{k+1}M_{k+1} \\ &= e_3^T R_k F_k \Pi_k + \frac{h}{2}e_3^T R_k M_k + \frac{h}{2}e_3^T R_{k+1}M_{k+1}. \end{aligned}$$

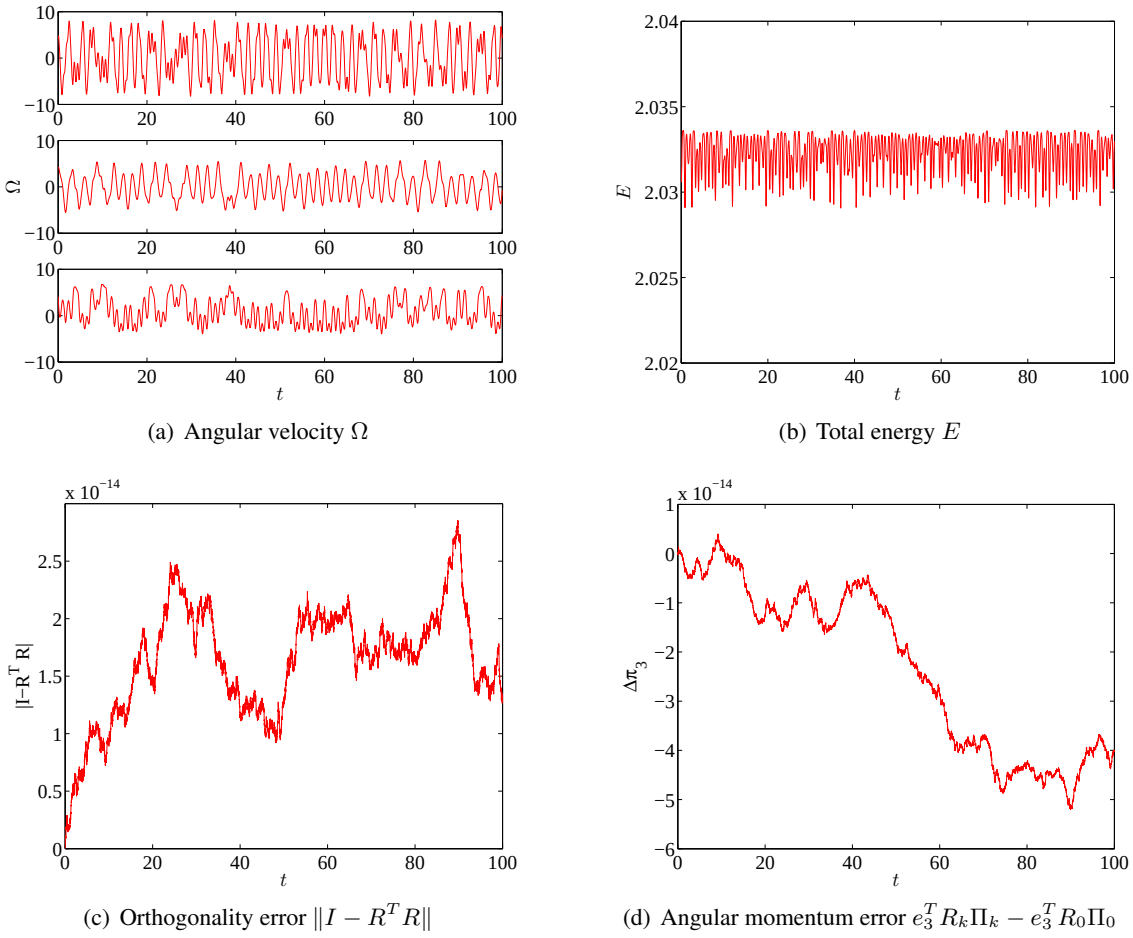
Since $e_3^T R_k M_k = 0$, $e_3^T R_{k+1}M_{k+1} = 0$, this reduces to

$$e_3^T R_{k+1}\Pi_{k+1} = e_3^T R_k F_k \Pi_k,$$

which shows conservation of the momentum map by the discrete-time flow. This is a more concise proof. The preceding development shows a formal application of the abstract discrete Noether's theorem to the 3D pendulum model.

*Numerical Results.* We compute the flow of discrete-time Hamilton's equations. The physical constants for the 3D pendulum are chosen as

$$m = 1\,\mathrm{kg}, \quad \rho_c = [0, 0, 0.3]\,\mathrm{m}, \quad J = \mathrm{diag}[0.13, 0.28, 0.17]\,\mathrm{kgm}^2.$$




(a) Angular velocity $\Omega$

(b) Total energy $E$

(c) Orthogonality error $\|I - R^T R\|$

(d) Angular momentum error $e_3^T R_k \Pi_k - e_3^T R_0 \Pi_0$

Figure 3.3: Numerical simulation for a 3D pendulum

Initial conditions are chosen as

$$R_0 = I, \qquad \Omega_0 = [4.14, 4.14, 4.14]\,\mathrm{rad/s}.$$

The simulation time and step size are $t_f = 100$ seconds, $h = 0.01$ seconds.

Figure 3.3 shows responses of angular velocity of the pendulum, computed total energy, orthogonality error, and the deviation of the angular momentum along the gravity direction. As shown by Figure 3.3(b), the computed total energy of the Lie group variational integrator oscillates near the initial value, but there is no increasing or decreasing drift for long time periods. The orthogonality error of rotation matrices and the error in conservation of the angular momentum about the gravity direction remain at machine precision level.

### 3.3.3 3D Pendulum with an Internal Degree of Freedom

Consider the 3D pendulum with an internal degree of freedom presented in Section 2.3.3.

*Configuration Manifold.* The configuration manifold is $\mathsf{SO(3)} \times \mathbb{R}$. The group action defines the discrete-time update map as

$$\begin{aligned}(R_{k+1}, x_{k+1}) &= (R_k, x_k)(F_k, \Delta x_k) \\ &= (R_k F_k, x_k + \Delta x_k)\end{aligned} \tag{3.87}$$

for $(F_k, \Delta x_k) \in \mathsf{SO(3)} \times \mathbb{R}$. The $\mathrm{Ad}$ operator also has the product structure.

$$\mathrm{Ad}_{(F,\Delta x)}(\hat{\eta}, x') = (F\hat{\eta}F^T, x') = (\widehat{F\eta}, x'), \quad \mathrm{Ad}^*_{(F,\Delta x)}(\hat{\eta}, x') = (F^T\hat{\eta}F, x') = (\widehat{F^T\eta}, x').$$

*Discrete Lagrangian.* Recalling (2.75), the Lagrangian is given by

$$\begin{aligned}L(R, x, \Omega, \dot{x}) = \frac{1}{2}\Omega^T J\Omega + \frac{1}{2}m_x(\dot{x}^2 - \rho_x^T\hat{\Omega}^2\rho_x + 2e_1^T Q^T\hat{\Omega}\rho_x\dot{x}) \\ + mge_3^T R\rho_c + m_x ge_3^T R\rho_x - \frac{1}{2}\kappa x^2.\end{aligned}$$

The discrete Lagrangian is chosen as

$$\begin{aligned}L_d(R_k, x_k, F_k, \Delta x_k) = \frac{1}{h}\mathrm{tr}[(I - F_k)J_d] + \frac{m_x}{2h}\Delta x_k^2 - \frac{m_x}{2h}\mathrm{tr}\big[\rho_{x_k}\rho_{x_k}^T(F_k - I)^2\big] \\ - \frac{m_x}{h}\mathrm{tr}\big[Qe_1\Delta x_k\rho_{x_k}^T(F_k - I)\big] + \frac{h}{2}mge_3^T R_k\rho_c + \frac{h}{2}m_x ge_3^T R_k\rho_{x_k} \\ + \frac{h}{2}mge_3^T R_k F_k\rho_c + \frac{h}{2}m_x ge_3^T R_k F_k\rho_{x_{k+1}} - \frac{h}{4}\kappa x_k^2 - \frac{h}{4}\kappa(x_k + \Delta x_k)^2.\end{aligned} \tag{3.88}$$

We find expressions for the derivatives of the discrete Lagrangian. The derivative of the discrete Lagrangian with respect to $F_k$ is given by

$$\begin{aligned}\mathbf{D}_{F_k}L_{d_k}\cdot\delta F_k &= -\frac{1}{h}\mathrm{tr}[\delta F_k J_d] - \frac{m_x}{2h}\mathrm{tr}\big[\rho_{x_k}\rho_{x_k}^T(F_k\delta F_k + \delta F_k F_k - 2\delta F_k)\big] \\ &\quad - \frac{m_x}{h}\mathrm{tr}\big[Qe_1\Delta x_k\rho_{x_k}^T\delta F_k\big] + \frac{h}{2}mge_3^T R_k\delta F_k\rho_c + \frac{h}{2}m_x ge_3^T R_k\delta F_k\rho_{x_{k+1}} \\ &= -\frac{1}{h}\mathrm{tr}[\delta F_k A_k] + \frac{h}{2}mge_3^T R_k\delta F_k\rho_c + \frac{h}{2}m_x ge_3^T R_k\delta F_k\rho_{x_{k+1}},\end{aligned}$$

where $A_k \in \mathbb{R}^{3\times 3}$ is defined as

$$A_k = J_d + \frac{m_x}{2}(\rho_{x_k}\rho_{x_k}^T F_k + F_k\rho_{x_k}\rho_{x_k}^T - 2\rho_{x_k}\rho_{x_k}^T) + m_x Qe_1\Delta x_k\rho_{x_k}^T.$$

This can be written as

$$\mathbf{D}_{F_k} L_{d_k} \cdot \delta F_k = -\frac{1}{h}\mathrm{tr}\big[(F_k^T \delta F_k) A_k F_k\big] + \frac{h}{2} m g e_3^T R_{k+1} (F_k^T \delta F_k) \rho_c + \frac{h}{2} m_x g e_3^T R_{k+1} (F_k^T \delta F_k) \rho_{x_{k+1}}. \tag{3.89}$$

Since $F_k^T \delta F_k$ is skew symmetric, using (3.68), the first term of (3.89) can be written as

$$-\frac{1}{h}\mathrm{tr}\big[(F_k^T \delta F_k) A_k F_k\big] = \frac{1}{h}\left\langle A_k F_k - F_k^T A_k^T, F_k^T \delta F_k \right\rangle.$$

Since $\hat{x} y = -\hat{y} x$ for any $x, y \in \mathbb{R}^3$, the last two terms of (3.89) can be written as

$$\begin{aligned} &\frac{h}{2} m g e_3^T R_{k+1} (F_k^T \delta F_k) \rho_c + \frac{h}{2} m_x g e_3^T R_{k+1} (F_k^T \delta F_k) \rho_{x_{k+1}} \\ &\qquad = -\frac{h}{2} m g e_3^T R_{k+1} \hat{\rho}_c (F_k^T \delta F_k)^\vee - \frac{h}{2} m_x g e_3^T R_{k+1} \hat{\rho}_{x_{k+1}} (F_k^T \delta F_k)^\vee \\ &\qquad = \frac{h}{2}\left\langle m g \hat{\rho}_c R_{k+1}^T e_3 + m_x g \hat{\rho}_{x_{k+1}} R_{k+1}^T e_3,\ (F_k^T \delta F_k)^\vee \right\rangle. \end{aligned} \tag{3.90}$$

Therefore, we obtain

$$\mathsf{T}_e^* \mathsf{L}_{F_k} \cdot \mathbf{D}_{F_k} L_{d_k} = \frac{1}{h}(A_k F_k - F_k^T A_k^T) + \left(\frac{h}{2} m g \hat{\rho}_c R_{k+1}^T e_3 + \frac{h}{2} m_x g \hat{\rho}_{x_{k+1}} R_{k+1}^T e_3\right)^\wedge. \tag{3.91}$$

The $\mathrm{Ad}^*$ operation gives

$$\begin{aligned} \mathrm{Ad}^*_{F_k^T} \cdot (\mathsf{T}_e^* \mathsf{L}_{F_k} \cdot \mathbf{D}_{F_k} L_{d_k}) = &\frac{1}{h}(F_k A_k - A_k^T F_k^T) \\ &+ \left(\frac{h}{2} m g F_k \hat{\rho}_c R_{k+1}^T e_3 + \frac{h}{2} m_x g F_k \hat{\rho}_{x_{k+1}} R_{k+1}^T e_3\right)^\wedge. \end{aligned} \tag{3.92}$$

The derivative of the discrete Lagrangian with respect to $R_k$ is given by

$$\begin{aligned} \mathbf{D}_{R_k} L_{d_k} \cdot \delta R_k &= \frac{h}{2} m g e_3^T \delta R_k \rho_c + \frac{h}{2} m_x g e_3^T \delta R_k \rho_{x_k} + \frac{h}{2} m g e_3^T \delta R_k F_k \rho_c + \frac{h}{2} m_x g e_3^T \delta R_k F_k \rho_{x_{k+1}} \\ &= \frac{h}{2} m g e_3^T R_k (R_k^T \delta R_k) \rho_c + \frac{h}{2} m_x g e_3^T R_k (R_k^T \delta R_k) \rho_{x_k} \\ &\quad + \frac{h}{2} m g e_3^T R_k (R_k^T \delta R_k) F_k \rho_c + \frac{h}{2} m_x g e_3^T R_k (R_k^T \delta R_k) F_k \rho_{x_{k+1}}. \end{aligned}$$

By following the same procedure used to derive (3.90), we obtain

$$\mathsf{T}_e^* \mathsf{L}_{R_k} \cdot \mathbf{D}_{R_k} L_{d_k} = \frac{h}{2} m g \hat{\rho}_c R_k^T e_3 + \frac{h}{2} m_x g \hat{\rho}_{x_k} R_k^T e_3 + \frac{h}{2} m g F_k \hat{\rho}_c R_{k+1}^T e_3 + \frac{h}{2} m_x g F_k \hat{\rho}_{x_k} R_{k+1}^T e_3. \tag{3.93}$$

The derivative of the discrete Lagrangian with respect to $\Delta x_k$ is given by

$$\mathbf{D}_{\Delta x_k} L_{d_k} = \frac{m_x}{h}\Delta x_k - \frac{m_x}{h}\mathrm{tr}\big[Q e_1 \rho_{x_k}^T (F_k - I)\big] + \frac{h}{2} m_x g e_3^T R_{k+1} Q e_1 - \frac{h}{2}\kappa x_{k+1}. \tag{3.94}$$

The $\mathrm{Ad}^*$ operation is the identity on $\mathbb{R}$. Since $\delta\rho_{x_k} = Qe_1\delta x_k$, the derivative of the discrete Lagrangian with respect to $x_k$ is given by

$$\begin{aligned}\mathbf{D}_{x_k}L_{d_k} = &-\frac{m_x}{2h}\mathrm{tr}\big[(Qe_1\rho_{x_k}^T + \rho_x e_1^T Q^T)(F_k - I)^2 + 2\Delta x_k(F_k - I)\big] \\ &+ \frac{h}{2}m_x g e_3^T R_k Q e_1 + \frac{h}{2}m_x g e_3^T R_{k+1} Q e_1 - \frac{h}{2}\kappa x_k - \frac{h}{2}\kappa x_{k+1}. \end{aligned} \tag{3.95}$$

*Discrete-time Euler-Lagrange Equations.* We substitute (3.91)-(3.95) to (3.10) for $g_k = (R_k, x_k)$ and $f_k = (F_k, \Delta x_k)$ to obtain the discrete-time Euler-Lagrange equations

$$\begin{aligned}&\frac{1}{h}(A_{k-1}F_{k-1} - F_{k-1}^T A_{k-1}^T - F_k A_k + A_k^T F_k^T)^\vee \\ &\qquad + hmg\hat{\rho}_c R_k^T e_3 + hm_x g\hat{\rho}_{x_k} R_k^T e_3 = 0,\end{aligned} \tag{3.96}$$

$$A_k = J_d + \frac{m_x}{2}(\rho_{x_k}\rho_{x_k}^T F_k + F_k\rho_{x_k}\rho_{x_k}^T - 2\rho_{x_k}\rho_{x_k}^T) + m_x Q e_1 \Delta x_k \rho_{x_k}^T, \tag{3.97}$$

$$\begin{aligned}-\frac{m_x}{h}(x_{k+1} - 2x_k + x_{k-1}) &- \frac{m_x}{h}\mathrm{tr}\Big[Qe_1\rho_{x_{k-1}}^T(F_{k-1} - I)\Big] + \frac{m_x}{h}\mathrm{tr}\big[Qe_1\rho_{x_k}^T(F_k - I)\big] \\ &- \frac{m_x}{2h}\mathrm{tr}\big[(Qe_1\rho_{x_k}^T + \rho_x e_1^T Q^T)(F_k - I)^2 + 2(x_{k+1} - x_k)(F_k - I)\big] \\ &+ hm_x g e_3^T R_k Q e_1 - h\kappa x_k = 0.\end{aligned} \tag{3.98}$$

*Discrete-time Legendre Transformation.* The discrete Legendre transformation yields the following discrete-time Hamilton's equations.

$$p_{\Omega_k} = \frac{1}{h}(F_k A_k - A_k^T F_k^T)^\vee - \frac{h}{2}mg\hat{\rho}_c R_k^T e_3 - \frac{h}{2}m_x g\hat{\rho}_{x_k} R_k^T e_3, \tag{3.99}$$

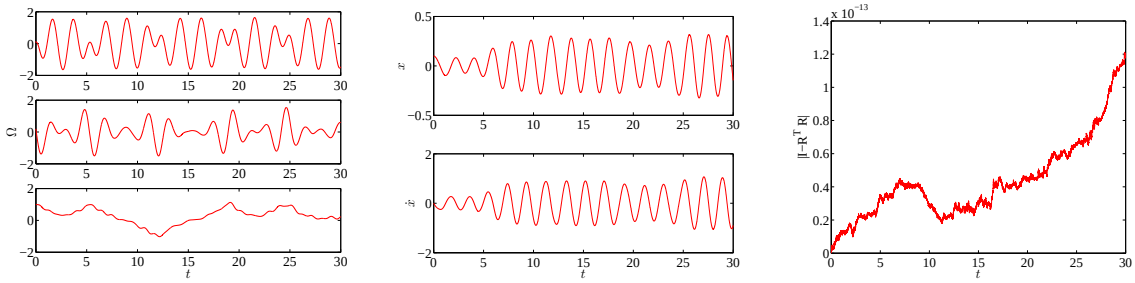
$$\begin{aligned}p_{x_k} = &\frac{m_x}{h}\Delta x_k - \frac{m_x}{h}\mathrm{tr}\big[Qe_1\rho_{x_k}^T(F_k - I)\big] \\ &+ \frac{m_x}{2h}\mathrm{tr}\big[(Qe_1\rho_{x_k}^T + \rho_x e_1^T Q^T)(F_k - I)^2 + 2\Delta x_k(F_k - I)\big] - \frac{h}{2}m_x g e_3^T R_k Q e_1 + \frac{h}{2}\kappa x_k,\end{aligned} \tag{3.100}$$

$$A_k = J_d + \frac{m_x}{2}(\rho_{x_k}\rho_{x_k}^T F_k + F_k\rho_{x_k}\rho_{x_k}^T - 2\rho_{x_k}\rho_{x_k}^T) + m_x Q e_1 \Delta x_k \rho_{x_k}^T, \tag{3.101}$$

$$p_{\Omega_{k+1}} = \frac{1}{h}(A_k F_k - F_k^T A_k)^\vee + \frac{h}{2}mg\hat{\rho}_c R_{k+1}^T e_3 + \frac{h}{2}m_x g\hat{\rho}_{x_{k+1}} R_{k+1}^T e_3, \tag{3.102}$$

$$p_{x_{k+1}} = \frac{m_x}{h}\Delta x_k - \frac{m_x}{h}\mathrm{tr}\big[Qe_1\rho_{x_k}^T(F_k - I)\big] + \frac{h}{2}m_x g e_3^T R_{k+1} Q e_1 - \frac{h}{2}\kappa x_{k+1}. \tag{3.103}$$

For given $(R_k, x_k, \Omega_k, \dot{x}_k)$, we find $(p_{\Omega_k}, p_{x_k})$ by (2.79). We use a fixed point iteration to determine $F_k$. For an initial guess of $F_k$, the corresponding $\Delta x_k$ is determined from (3.100), and $A_k$ is determined from (3.101). Then, we can find $F_k$ by solving (3.99). This is repeated until $F_k$ converges. $(R_{k+1}, x_{k+1})$ are obtained from (3.87), and $(p_{\Omega_{k+1}}, p_{x_{k+1}})$ are obtained from (3.102), (3.103). The velocities $(\Omega_{k+1}, \dot{x}_{k+1})$ can be determined by (2.79). This yields a discrete-time flow map $(R_k, x_k, \Omega_k, \dot{x}_k) \to (R_{k+1}, x_{k+1}, \Omega_{k+1}, \dot{x}_{k+1})$.



(a) Angular velocity $\Omega$ (b) Position and velocity of the particle $x, \dot{x}$ (c) Orthogonality error $\|I - R^T R\|$

Figure 3.4: Numerical simulation for a 3D pendulum with an internal degree of freedom

*Numerical Results.* We compute the flow of discrete-time Hamilton's equations. The physical constants for the 3D pendulum with an internal degree of freedom are chosen as

$$m = 1\,\mathrm{kg}, \quad m_x = 0.1\,\mathrm{kg}, \quad \rho_c = [0, 0, 1]\,\mathrm{m}, \quad d = 0.15\,\mathrm{m}, \quad \kappa = 1.1\,\mathrm{N/m},$$
$$J = \mathrm{diag}[1.03, 1.06, 0.05]\,\mathrm{kgm}^2, \quad Q = \begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.00 & 0.25 & 0.96 \\ 0.00 & 0.96 & -0.25 \end{bmatrix}.$$

Initial conditions are chosen as

$$x_0 = 0.10\,\mathrm{m}, \quad \dot{x}_0 = 0.00\,\mathrm{m/s},$$
$$R_0 = \exp(\frac{\pi}{6}\hat{e}_2), \qquad \Omega_0 = [0.1, 0.2, 1.0]\,\mathrm{rad/s}.$$

Figure 3.4 shows responses of the angular velocity of the pendulum, the position of the particle with respect to the 3D pendulum, and the velocity of the particle. The computed orthogonality error of the rotation matrix is also shown.

### 3.3.4 3D Pendulum on a Cart

Consider the 3D pendulum whose pivot is attached to a cart moving on a horizontal plane, presented in Section 2.3.4.

*Configuration Manifold.* The configuration manifold is $\mathsf{SO}(3) \times \mathbb{R}^2$. The group action defines the discrete-time update map (3.11) as

$$\begin{aligned}(R_{k+1}, x_{k+1}, y_{k+1}) &= (R_k, x_k, y_k)(F_k, \Delta x_k, \Delta y_k) \\ &= (R_k F_k, x_k + \Delta x_k, y_k + \Delta y_k),\end{aligned}$$

for $(F_k, \Delta x_k, \Delta y_k) \in \mathsf{SO}(3) \times \mathbb{R}^2$. The adjoint operator also has a product structure as follows

$$\mathrm{Ad}_{(F,\Delta x,\Delta y)}(\hat{\eta}, x', y') = (F\hat{\eta}F^T, x', y'), \quad \mathrm{Ad}^*_{(F,\Delta x,\Delta y)}(\hat{\eta}, x', y') = (F^T\hat{\eta}F, x', y').$$

*Discrete Lagrangian.* Recalling (2.84), the Lagrangian is given by

$$L(R,\Omega,\dot{x},\dot{y}) = \frac{1}{2}(M+m)(\dot{x}^2+\dot{y}^2) + \frac{1}{2}\Omega^T J\Omega + m\dot{x}e_1^T\dot{R}\rho_c + m\dot{y}e_2^T\dot{R}\rho_c + mge_3^T R\rho_c.$$

The discrete Lagrangian is chosen as

$$\begin{aligned} L_d(R_k,x_k,y_k,F_k,\Delta x_k,\Delta y_k) &= \frac{1}{2h}(M+m)((\Delta x_k)^2+(\Delta y_k)^2) + \frac{1}{h}\mathrm{tr}[(I-F_k)J_d] \\ &+ \frac{m}{h}\Delta x_k e_1^T R_k(F_k-I)\rho_c + \frac{m}{h}\Delta y_k e_2^T F_k(R_k-I)\rho_c + \frac{h}{2}mge_3^T R_k\rho_c + \frac{h}{2}mge_3^T R_k F_k\rho_c. \end{aligned} \tag{3.104}$$

We find the expressions for the derivatives of the discrete Lagrangian. We have

$$\begin{aligned} \mathbf{D}_{F_k}L_{d_k}\cdot\delta F_k &= -\frac{1}{h}\mathrm{tr}[\delta F_k J_d] + \frac{m}{h}\Delta x_k e_1^T R_k\delta F_k\rho_c + \frac{m}{h}\Delta y_k e_2^T R_k\delta F_k\rho_c + \frac{h}{2}mge_3^T R_k\delta F_k\rho_c \\ &= -\frac{1}{h}\mathrm{tr}\big[F_k F_k^T\delta F_k J_d\big] + \frac{m}{h}\Delta x_k e_1^T R_{k+1}F_k^T\delta F_k\rho_c + \frac{m}{h}\Delta y_k e_2^T R_{k+1}F_k^T\delta F_k\rho_c \\ &\quad + \frac{h}{2}mge_3^T R_{k+1}F_k^T\delta F_k\rho_c. \end{aligned} \tag{3.105}$$

Since $F_k^T\delta F_k$ is skew-symmetric, and using (3.68), the first term of (3.105) can be written as

$$-\frac{1}{h}\mathrm{tr}\big[F_k(F_k^T\delta F_k)J_d\big] = -\frac{1}{h}\mathrm{tr}\big[(F_k^T\delta F_k)J_dF_k\big] = \frac{1}{h}\left\langle J_dF_k - F_k^TJ_d, F_k^T\delta F_k\right\rangle.$$

We use the identity: for any $x,y,z\in\mathbb{R}^3$,

$$y^T\hat{x}z = -\mathrm{tr}\big[yz^T\hat{x}\big] = \left\langle yz^T - zy^T,\ \hat{x}\right\rangle. \tag{3.106}$$

Using (3.106), the second term of (3.105) can be written as

$$\frac{m}{h}\Delta x_k e_1^T R_{k+1}F_k^T\delta F_k\rho_c = \frac{m}{h}\Delta x_k\left\langle R_{k+1}^T e_1\rho_c^T - \rho_c e_1^T R_{k+1},\ F_k^T\delta F_k\right\rangle.$$

We apply the same identity to the remaining terms of (3.105) to obtain

$$\begin{aligned} \mathsf{T}_e^*\mathsf{L}_{F_k}\cdot\mathbf{D}_{F_k}L_{d_k} &= \frac{1}{h}(J_dF_k - F_k^TJ_d) + \frac{m}{h}\Delta x_k(R_{k+1}^Te_1\rho_c^T - \rho_ce_1^TR_{k+1}) \\ &\quad + \frac{m}{h}\Delta y_k(R_{k+1}^Te_2\rho_c^T - \rho_ce_2^TR_{k+1}) + \frac{h}{2}mg(R_{k+1}^Te_3\rho_c^T - \rho_ce_3^TR_{k+1}). \end{aligned} \tag{3.107}$$

Similarly, we obtain

$$\begin{aligned} \mathsf{T}_e^*\mathsf{L}_{R_k}\cdot\mathbf{D}_{R_k}L_{d_k} &= \frac{m}{h}\Delta x_k(R_k^Te_1\rho_c^T(F_k^T-I) - (F_k-I)\rho_ce_1R_k) \\ &\quad + \frac{m}{h}\Delta y_k(R_k^Te_2\rho_c^T(F_k^T-I) - (F_k-I)\rho_ce_2R_k) \\ &\quad + \frac{h}{2}mg(R_k^Te_3\rho_c^T - \rho_ce_3^TR_k) + \frac{h}{2}mg(R_k^Te_3\rho_c^TF_k^T - F_k\rho_ce_3^TR_k). \end{aligned} \tag{3.108}$$

We have $\mathrm{Ad}^*_{F_k^{-1}}(\mathsf{T}^*_e\mathsf{L}_{F_k}\cdot\mathbf{D}_{F_k}L_{d_k}) = F_k(\mathsf{T}^*_e\mathsf{L}_{F_k}\cdot\mathbf{D}_{F_k}L_{d_k})F_k^T$. The derivatives of the discrete Lagrangian with respect to $\Delta x_k$, $\Delta y_k$ are given by

$$\mathbf{D}_{\Delta x_k}L_{d_k} = \frac{1}{h}(M+m)\Delta x_k + \frac{m}{h}e_1^T(R_{k+1}-R_k)\rho_c, \tag{3.109}$$

$$\mathbf{D}_{\Delta y_k}L_{d_k} = \frac{1}{h}(M+m)\Delta y_k + \frac{m}{h}e_2^T(R_{k+1}-R_k)\rho_c. \tag{3.110}$$

*Discrete-time Euler-Lagrange Equations.* Substituting (3.107)-(3.110) into (3.10), we obtain the discrete Euler-Lagrange equations for the 3D pendulum on a cart.

$$\frac{1}{h}(M+m)(x_{k+1}-2x_k+x_{k-1}) + \frac{m}{h}e_1^T(R_{k+1}-2R_k+R_{k-1})\rho_c = 0, \tag{3.111}$$

$$\frac{1}{h}(M+m)(y_{k+1}-2y_k+y_{k-1}) + \frac{m}{h}e_2^T(R_{k+1}-2R_k+R_{k-1})\rho_c = 0, \tag{3.112}$$

$$\begin{aligned}\frac{1}{h}(F_kJ_d - J_dF_k^T - J_dF_{k-1} + F_{k-1}^TJ_d) &= -\frac{m}{h}(x_{k+1}-2x_k+x_{k-1})(R_k^Te_1\rho_c^T - \rho_ce_1^TR_k)\\ &\quad - \frac{m}{h}(y_{k+1}-2y_k+y_{k-1})(R_k^Te_2\rho_c^T - \rho_ce_2^TR_k) + hmg(R_k^Te_3\rho_c^T - \rho_ce_3^TR_k).\end{aligned} \tag{3.113}$$

Using the property $\widehat{x\times y} = yx^T - xy^T$, the third equation can be written as

$$\begin{aligned}&\frac{1}{h}(F_kJ_d - J_dF_k^T - J_dF_{k-1} + F_{k-1}^TJ_d)^\vee\\ &\quad= -\frac{m}{h}(x_{k+1}-2x_k+x_{k-1})\hat\rho_cR_k^Te_1 - \frac{m}{h}(y_{k+1}-2y_k+y_{k-1})\hat\rho_cR_k^Te_2 + hmg\hat\rho_cR_k^Te_3.\end{aligned} \tag{3.114}$$

*Discrete-time Legendre Transformation.* The discrete Legendre transformation yields the following discrete-time Hamilton's equations.

$$p_{x_k} = \frac{1}{h}(M+m)(x_{k+1}-x_k) + \frac{m}{h}e_1(R_{k+1}-R_k)\rho_c, \tag{3.115}$$

$$p_{y_k} = \frac{1}{h}(M+m)(y_{k+1}-y_k) + \frac{m}{h}e_2(R_{k+1}-R_k)\rho_c, \tag{3.116}$$

$$\begin{aligned}\hat p_{\Omega_k} &= \frac{1}{h}(F_kJ_d - J_dF_k^T)\\ &\quad + \left\{\frac{m}{h}(x_{k+1}-x_k)\hat\rho_cR_k^Te_1 + \frac{m}{h}(y_{k+1}-y_k)\hat\rho_cR_k^Te_2 - \frac{h}{2}mg\hat\rho_cR_k^Te_3\right\}^\wedge,\end{aligned} \tag{3.117}$$

$$R_{k+1} = R_kF_k, \tag{3.118}$$

$$p_{x_{k+1}} = p_{x_k}, \tag{3.119}$$

$$p_{y_{k+1}} = p_{y_k}, \tag{3.120}$$

$$\begin{aligned}\hat p_{\Omega_{k+1}} &= \frac{1}{h}(J_dF_k - F_k^TJ_d)\\ &\quad + \left\{\frac{m}{h}(x_{k+1}-x_k)\hat\rho_cR_{k+1}^Te_1 + \frac{m}{h}(y_{k+1}-y_k)\hat\rho_cR_{k+1}^Te_2 + \frac{h}{2}mg\hat\rho_cR_{k+1}^Te_3\right\}^\wedge.\end{aligned} \tag{3.121}$$

For given $(R_k, x_k, y_k, \Omega_k, \dot{x}_k, \dot{y}_k)$, we find $(p_{\Omega_k}, p_{x_k}, p_{y_k})$ by (2.90). We use a fixed point iteration to determine $R_{k+1}$. For an initial guess of $R_{k+1}$, the corresponding $x_{k+1}, y_{k+1}$ are obtained from (3.115), (3.116). Then, we can find $F_k$ by solving (3.117). The new value for $R_{k+1}$ is given by (3.118). This is repeated until $R_{k+1}$ converges. Then, $x_{k+1}, y_{k+1}$ are obtained from (3.115), (3.116), and $(p_{\Omega_{k+1}}, p_{x_{k+1}}, p_{y_{k+1}})$ are obtained by (3.119), (3.120), and (3.121). The velocities $(\Omega_{k+1}, \dot{x}_{k+1}, \dot{y}_{k+1})$ can be obtained from (2.90). This yields a discrete-time flow map $(R_k, x_k, y_k, \Omega_k, \dot{x}_k, \dot{y}_k) \to (R_{k+1}, x_{k+1}, y_{k+1}, \Omega_{k+1}, \dot{x}_{k+1}, \dot{y}_{k+1})$.

*Numerical Results.* We compare the computational properties of the discrete-time equations of motion given by (3.115)-(3.121) with a 4(5)-th variables step size Runge-Kutta method. The physical constants for the 3D pendulum on a cart are chosen as

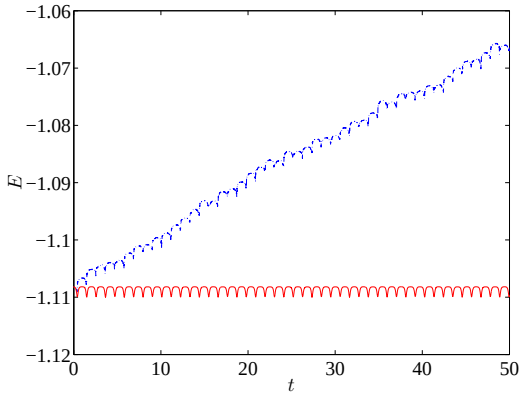
$$M = m = 1\,\mathrm{kg}, \quad \rho_c = [0.25, 0.25, 1]\,\mathrm{m}, \quad J = \begin{bmatrix} 1.09 & -0.06 & -0.25 \\ -0.05 & 1.10 & -0.25 \\ -0.25 & -0.25 & 0.15 \end{bmatrix} \mathrm{kgm}^2.$$

Initial conditions are chosen such that the mass center of the system is located at the origin, and the total linear momentum is zero.

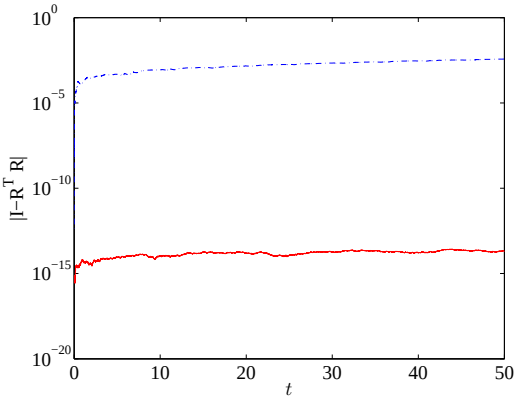
$$x_0 = -0.125\,\mathrm{m}, \quad \dot{x}_0 = 0.525\,\mathrm{m/s}, \quad y_0 = 0.5\,\mathrm{m}, \quad \dot{y}_0 = -0.0125\,\mathrm{m/s},$$
$$R_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \Omega_0 = [0.1, 0.2, 5.0]\,\mathrm{rad/s}.$$

Figure 3.5(a) shows the computed total energy for $50$ seconds. The Lie group variational integrator preserves the total energy well. But there is a notable dissipation of the computed total energy using the Runge-Kutta method. Similar characteristics are observed for the orthogonality error of the rotation matrices in Figure 3.5(b).

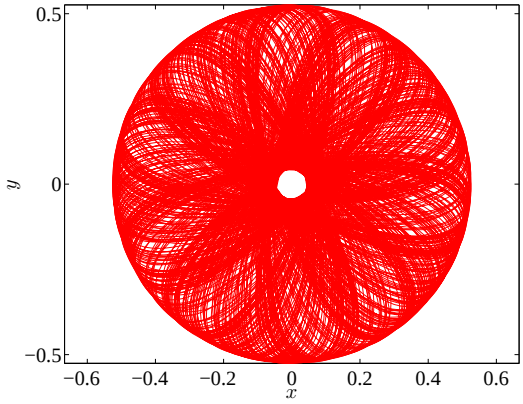
It can be shown that the horizontal component of the total linear momentum is conserved. As a result, the mass center moves along the vertical $e_3$ axis, and the horizontal location of the mass center of the system is fixed. Figure 3.5(c)-3.5(d) show the trajectory of the cart on the horizontal plane, and Figure 3.5(e)-3.5(g) show the initial configuration and the configuration at $t = 700$ sec. It is interesting to observe that the computed mass center of the Runge-Kutta method drifts in the horizontal plane, as it does not preserve the total linear momentum properly.



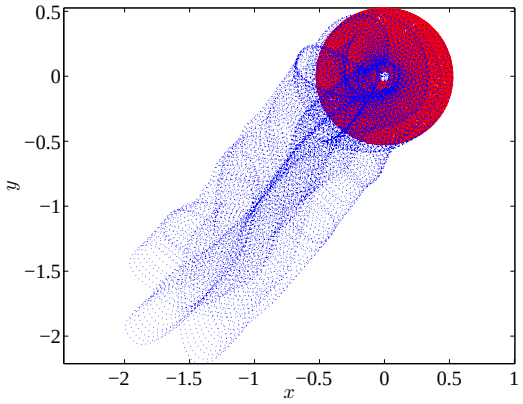

(a) Computed total energy



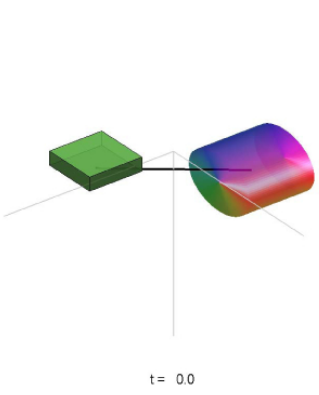

(b) Orthogonality error $\|I - R^T R\|$



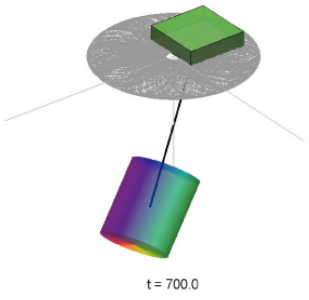

(c) Trajectory of the cart (LGVI)



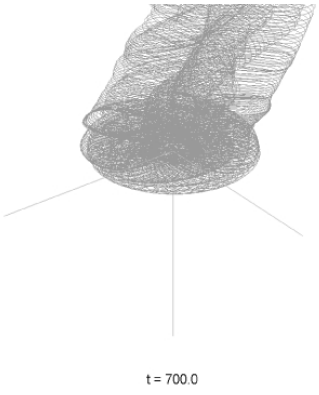

(d) Trajectory of the cart (RK4(5))




(e) Initial configuration




(f) Configuration at $t = 700$s (LGVI)




(g) Configuration at $t = 700$s (RK4(5))

Figure 3.5: Numerical simulation for a 3D pendulum on a cart (LGVI: red, solid, RK4(5): blue, dotted)

### 3.3.5 Single Rigid Body

Consider a rigid body acting under a potential that is dependent of the attitude and the position of the body, presented in Section 2.3.5.

*Configuration Manifold.* Let $g_k = (R_k, x_k), f_k = (F_k, Y_k) \in \mathsf{SE}(3)$ for $R_k, F_k \in \mathsf{SO}(3)$, $x_k, Y_k \in \mathbb{R}^3$. The group action of $\mathsf{SE}(3)$ is matrix multiplication in homogeneous coordinates. The discrete-time kinematics equation (3.11) can be written as

$$\begin{bmatrix} R_{k+1} & x_{k+1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_k & x_k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} F_k & Y_k \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_k F_k & x_k + R_k Y_k \\ 0 & 1 \end{bmatrix}. \tag{3.122}$$

The adjoint operator for $\mathsf{SE}(3)$ can be written in a matrix form as

$$\mathrm{Ad}_{(R,x)} = \begin{bmatrix} R & 0 \\ \hat{x}R & R \end{bmatrix}, \quad \mathrm{Ad}^*_{(R,x)} = \begin{bmatrix} R^T & -R^T\hat{x} \\ 0 & R^T \end{bmatrix}. \tag{3.123}$$

*Discrete Lagrangian.* The discrete Lagrangian is chosen as

$$L_d(g_k, f_k) = T(f_k) - \frac{h}{2}U(g_k) - \frac{h}{2}U(g_k f_k), \tag{3.124}$$

where the discrete kinetic energy term is defined as

$$T(f_k) = \frac{1}{h}\langle\!\langle f_k - e, f_k - e \rangle\!\rangle = \frac{m}{2h} Y_k \cdot Y_k + \frac{1}{h}\mathrm{tr}[(I_{3\times 3} - F_k)J_d]\,.$$

Thus, $\delta T(f_k)$ is given by

$$\begin{aligned} \delta T(f_k) &= \frac{m}{h} Y_k \cdot \delta Y_k + \frac{1}{h}\mathrm{tr}[-\delta F_k J_d] \\ &= \frac{m}{h} Y_k \cdot \delta Y_k + \frac{1}{h}\mathrm{tr}\left[-F_k^T \delta F J_d F_k\right]. \end{aligned}$$

Since $F_k^T \delta F_k$ is skew-symmetric, this can be rewritten as

$$\begin{aligned} \delta T(f_k) &= \frac{m}{h} F_k^T Y_k \cdot F_k^T \delta Y_k - \frac{1}{2h}\mathrm{tr}\left[F_k^T \delta F (J_d F_k - F_k^T J_d)\right] \\ &= \left\langle \begin{bmatrix} \frac{1}{h}(J_d F_k - F_k^T J_d) & \frac{m}{h} F_k^T Y_k \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} F_k^T \delta F_k & F_k^T \delta Y_k \\ 0 & 0 \end{bmatrix} \right\rangle. \end{aligned} \tag{3.125}$$

Using (2.96) and (2.62), the derivatives of the potential term can be written as

$$\begin{aligned} &-\mathbf{D}_{g_k}(U(g_k) + U(g_k f_k)) \cdot \delta g_k - \mathbf{D}_{f_k} U(g_k f_k) \cdot \delta f_k \\ &\quad = \begin{bmatrix} \hat{M}_k & -R_k^T \frac{\partial U_k}{\partial x_k} \\ 0 & 0 \end{bmatrix} \cdot \eta_k + \begin{bmatrix} \hat{M}_{k+1} & -R_{k+1}^T \frac{\partial U_{k+1}}{\partial x_{k+1}} \\ 0 & 0 \end{bmatrix} \cdot (\mathrm{Ad}_{f_k^{-1}} \eta_k + f_k^{-1} \delta f_k). \end{aligned} \tag{3.126}$$

From (3.125) and (3.126), $\mathsf{T}^*_e\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k)$ is given by

$$\mathsf{T}^*_e\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k) = \begin{bmatrix} \frac{1}{h}(J_d F_k - F_k^T J_d) & \frac{m}{h} F_k^T Y_k \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \frac{h}{2}\hat{M}_{k+1} & -\frac{h}{2} R_{k+1}^T \frac{\partial U_{k+1}}{\partial x_{k+1}} \\ 0 & 0 \end{bmatrix}. \tag{3.127}$$

Using (3.123), $\mathrm{Ad}^*_{f_k^{-1}}(\mathsf{T}^*_e\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k))$ is computed in $\mathbb{R}^6$ as

$$\begin{aligned} &\mathrm{Ad}^*_{f_k^{-1}}(\mathsf{T}^*_e\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k)) \\ &\quad = \begin{bmatrix} F_k & \widehat{F_k F_k^T Y_k} \\ 0 & F_k \end{bmatrix} \begin{bmatrix} \frac{1}{h}\check{}(J_d F_k - F_k^T J_d) \\ \frac{m}{h} F_k^T Y_k \end{bmatrix} + \mathrm{Ad}^*_{f_k^{-1}} \begin{bmatrix} \frac{h}{2}\hat{M}_{k+1} & \frac{h}{2} - R_{k+1}^T \frac{\partial U_{k+1}}{\partial x_{k+1}} \\ 0 & 0 \end{bmatrix} \\ &\quad = \begin{bmatrix} \frac{1}{h}\check{}(F_k J_d - J_d F_k^T) \\ \frac{m}{h} Y_k \end{bmatrix} + \mathrm{Ad}^*_{f_k^{-1}} \begin{bmatrix} \frac{h}{2}\hat{M}_{k+1} & -\frac{h}{2} R_{k+1}^T \frac{\partial U_{k+1}}{\partial x_{k+1}} \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{3.128}$$

An expression for $\mathsf{T}^*_e\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k)$ is obtained from (3.126).

$$\mathsf{T}^*_e\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k) = \begin{bmatrix} \frac{h}{2}\hat{M}_k & -\frac{h}{2} R_k^T \frac{\partial U_k}{\partial x_k} \\ 0 & 0 \end{bmatrix} + \mathrm{Ad}^*_{f_k^{-1}} \begin{bmatrix} \frac{h}{2}\hat{M}_{k+1} & -\frac{h}{2} R_{k+1}^T \frac{\partial U_{k+1}}{\partial x_{k+1}} \\ 0 & 0 \end{bmatrix}. \tag{3.129}$$

*Discrete-time Euler-Lagrange Equations.* Substituting (3.127), (3.128), and (3.129) into (3.10),

$$J_d F_{k-1} - F_{k-1}^T J_d - (F_k J_d - J_d F_k^T) + h^2 \hat{M}_k = 0, \tag{3.130}$$

$$\frac{m}{h^2} F_{k-1}^T Y_{k-1} - \frac{m}{h^2} Y_k - R_k^T \frac{\partial U_k}{\partial x_k} = 0. \tag{3.131}$$

Since the left-trivialization is used, these equations are expressed in the body fixed frame. From (3.122), $Y_k = R_k^T(x_{k+1} - x_k)$. Substituting this into the above equation, we obtain

$$\frac{m}{h^2} F_{k-1}^T R_{k-1}^T (x_k - x_{k-1}) - \frac{m}{h^2} R_k^T (x_{k+1} - x_k) - R_k^T \frac{\partial U_k}{\partial x_k} = 0.$$

Multiplying this by $R_k$, and rearranging, we obtain

$$\frac{m}{h^2}(x_{k+1} - 2x_k + x_{k-1}) = -\frac{\partial U_k}{\partial x_k}. \tag{3.132}$$

Therefore, the discrete-time Euler-Lagrange equations are given by (3.130) and (3.132).

*Discrete Legendre Transformation.* Let $(\Pi_k, \Gamma_k) \in \mathfrak{se}(3)^*$ be the angular momentum and the linear momentum of the rigid body, represented with respect to the body fixed frame. From (3.14) and (3.15), the discrete Legendre transformations are given by

$$\begin{bmatrix} \hat{\Pi}_k & \Gamma_k \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{h}(F_k J_d - J_d F_k^T) - \frac{h}{2}\hat{M}_k & \frac{m}{h} Y_k + \frac{h}{2} R_k^T \frac{\partial U_k}{\partial x_k} \\ 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} \hat{\Pi}_{k+1} & \Gamma_{k+1} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{h}(J_d F_k - F_k^T J_d) & \frac{m}{h} F_k^T Y_k \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \frac{h}{2}\hat{M}_{k+1} & -\frac{h}{2} R_{k+1}^T \frac{\partial U_{k+1}}{\partial x_{k+1}} \\ 0 & 0 \end{bmatrix}.$$

It is desirable to express the discrete-time equations for the location and the linear momentum of the rigid body with respect to the reference frame. Let $\gamma \in \mathbb{R}^{3*}$ be the linear momentum represented with respect to the reference frame, i.e. $\gamma = R\Gamma$. The discrete-time Hamilton's equations can be written as

$$h\hat{\Pi}_k + \frac{h^2}{2}\hat{M}_k = J_d F_k - F_k^T J_d, \tag{3.133}$$

$$\Pi_{k+1} = F_k^T \Pi_k + \frac{h}{2} F_k^T M_k + \frac{h}{2} M_{k+1}, \tag{3.134}$$

$$x_{k+1} = x_k + \frac{h}{m}\gamma_k - \frac{h^2}{2m}\frac{\partial U_k}{\partial x_k}, \tag{3.135}$$

$$\gamma_{k+1} = \gamma_k - \frac{h}{2}\frac{\partial U_k}{\partial x_k} - \frac{h}{2}\frac{\partial U_{k+1}}{\partial x_{k+1}}. \tag{3.136}$$

*Numerical Results.* We simulate the dynamics of a dumbbell body acting under a gravitational potential from a fixed spherical body. We assume that the mass of the dumbbell is negligible compared to the mass of the spherical body, and the origin of the reference frame is located at the mass center of the spherical body. The gravitational potential results in a nontrivial coupling between the attitude dynamics and the orbital dynamics. This model is referred to as a restricted full two body problem.

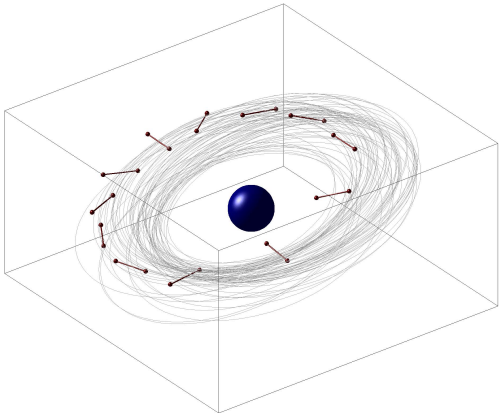
The dumbbell model consists of two equal rigid spheres and a rigid massless connecting rod. This dumbbell rigid body model results in a simple closed form for the mutual gravitational potential given by

$$U(R, x) = -\frac{GMm}{2}\sum_{q=1}^{2} \|x + R\rho_q\|,$$

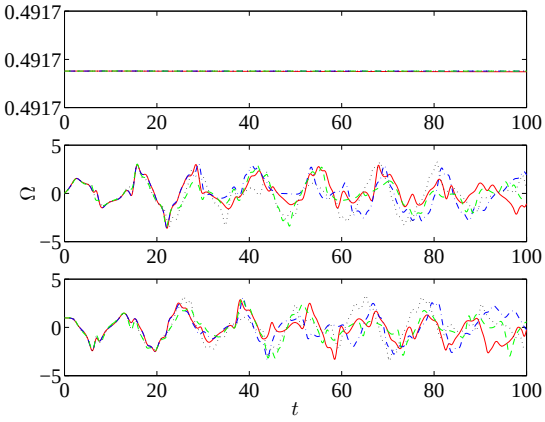
where $G$ is the universal gravitational constant, and $M$ is the mass of the spherical body. The constant $m$ is the mass of the spherical body, and $\rho_q \in \mathbb{R}^3$ is a vector from the origin of the body-fixed frame to the $q$th sphere of the dumbbell in the body-fixed frame.

We compare computational properties of the Lie group variational integrator with Runge-Kutta methods applied to three types of attitude representations, namely, rotation matrices, quaternions, and Euler-angles. The attitude kinematics equation is expressed in terms of quaternions and Euler-angles, and the Runge-Kutta method is applied.

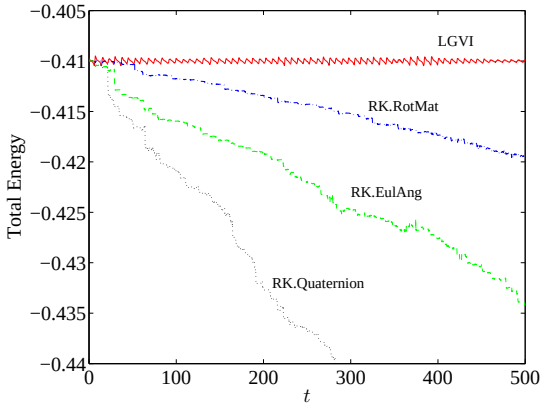
Figure 3.6 shows numerical simulation results for a near-circular orbit, where the trajectory of the dumbbell, the angular velocity response, the computed total energy, and the orthogonality error are presented. The Lie group variational integrator preserves the total energy and the Lie group structure of $\mathsf{SO(3)}$. The mean total energy deviation is $2.5983 \times 10^{-4}$, and the mean orthogonality error is $1.8553 \times 10^{-13}$. But, there is a notable dissipation of the computed total energy and the orthogonality error for the Runge-Kutta method.



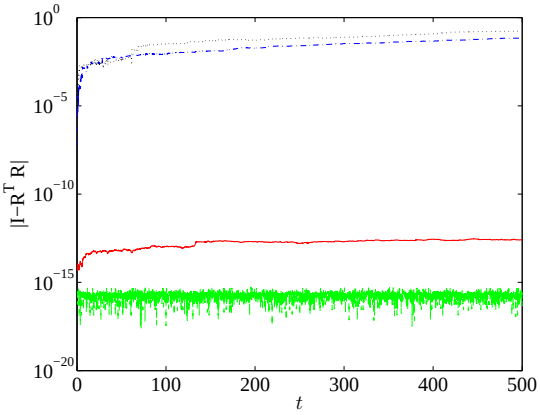
(a) Trajectory of the dumbbell




(b) Angular velocity $\Omega$




(c) Computed Total Energy




(d) Orthogonality Error $\|I - R^T R\|$

Figure 3.6: Numerical simulation for a single rigid body (Lie group variational integrator:LGVI, Runge-Kutta with rotation matrices:RK.RotMat, Runge-Kutta with Euler-angles:RK. EulAng, Runge-Kutta with quaternions:RK.qua

It is interesting to see that the Runge-Kutta method using quaternions, which is generally assumed to have better computational properties than the kinematics equation with rotation matrices, has a larger total energy error and a larger orthogonality error. Since the unit length of the quaternion vector is not preserved in the numerical computations, the orthogonality errors arise when converted to a rotation matrix. This suggests that, even with conventional (non-geometric) integrators, the rotation matrix has better computational properties than quaternions for the rigid body dynamics.

### 3.3.6 Full Body Problem

Consider the full body problem presented in Section 2.3.6.

*Discrete-time Euler-Lagrange Equations.* The configuration manifold of the full $n$ body problem is $(\mathsf{SE}(3))^n$. Since the dynamics of the full body system are only coupled through the mutual potential, the development for a single rigid body, presented in Section 3.3.5, is readily extended to the full $n$ body problem to obtain the discrete-time Euler-Lagrange equations.

$$J_{d_i} F_{i_{k-1}} - F_{i_{k-1}}^T J_{d_i} - (F_{k_i} J_{d_i} - J_{d_i} F_{k_i}^T) + h^2 \hat{M}_{i_k} = 0, \tag{3.137}$$

$$\frac{m_i}{h^2}(x_{i_{k+1}} - 2x_{i_k} + x_{i_{k-1}}) = -\frac{\partial U_k}{\partial x_{i_k}}, \tag{3.138}$$

where the subscript $i$ refers to the $i$-th rigid body for $i \in \{1, \ldots, n\}$. Similarly, discrete-time Hamilton's equations given by (3.133)-(3.136) can be easily extended to the full $n$-body problem.

*Discrete-time Reduced Euler-Lagrange Equations.* When the mutual potential depends only on the relative positions and the relative attitudes of rigid bodies, the Lagrangian is invariant under the action of $\mathsf{SE}(3)$, and the configuration manifold can be reduced to a quotient space $(\mathsf{SE}(3))^{n-1}$.

Discrete-time reduced Euler-Lagrange equations for the full two body problem have been developed in Lee et al. (2007b,c). The procedure to derive discrete-time reduced Euler-Lagrange equations is the same as the procedure presented in Section 3.1. The expressions for the variations should be carefully developed for the reduced variables. Here, we present the resulting discrete-time reduced Euler-Lagrange equations for the full two body problem. The detailed development can be found in Lee et al. (2007c).

The discrete-time reduced Euler-Lagrange equations for the full two body problem are given by

$$F_{2_k} X_{k+1} - 2X_k + F_{2_{k-1}}^T X_{k-1} = -\frac{h^2}{m}\frac{\partial U_k}{\partial X_k}, \tag{3.139}$$

$$F_{k+1} J_{dR_{k+1}} - J_{dR_{k+1}} F_{k+1}^T = F_{2_k}^T \left(F_k J_{dR_k} - J_{dR_k} F_k^T\right) F_{2_k} - h^2 \hat{M}_{k+1}, \tag{3.140}$$

$$F_{2_{k+1}} J_{d_2} - J_{d_2} F_{2_{k+1}}^T = F_{2_k}^T \left(F_{2_k} J_{d_2} - J_{d_2} F_{2_k}^T\right) F_{2_k} + h^2 X_{k+1} \times \frac{\partial U}{\partial X_{k+1}} + h^2 \hat{M}_{k+1}, \tag{3.141}$$

$$R_{k+1} = F_{2_k}^T F_k R_k, \tag{3.142}$$

$$R_{2_{k+1}} = R_{2_k} F_{2_k}. \tag{3.143}$$

The discrete Legendre transformation yields the discrete-time Hamilton's equations.

$$X_{k+1} = F_{2_k}^T \left(X_k + h\frac{\Gamma_k}{m} - \frac{h^2}{2m}\frac{\partial U_k}{\partial X_k}\right), \tag{3.144}$$

$$\Gamma_{k+1} = F_{2_k}^T \left(\Gamma_k - \frac{h}{2}\frac{\partial U_k}{\partial X_k}\right) - \frac{h}{2}\frac{\partial U_{k+1}}{\partial X_{k+1}}, \tag{3.145}$$

$$\Pi_{k+1} = F_{2_k}^T \left(\Pi_k - \frac{h}{2} M_k\right) - \frac{h}{2} M_{k+1}, \tag{3.146}$$

$$\Pi_{2_{k+1}} = F_{2_k}^T \left( \Pi_{2_k} + \frac{h}{2} X_k \times \frac{\partial U}{\partial X_k} + \frac{h}{2} M_k \right) + \frac{h}{2} X_{k+1} \times \frac{\partial U}{\partial X_{k+1}} + \frac{h}{2} M_{k+1}, \tag{3.147}$$

$$R_{k+1} = F_{2_k}^T F_k R_k, \tag{3.148}$$

$$h \left( \Pi_k - \frac{h}{2} M_k \right)^\wedge = F_k J_{dR_k} - J_{dR_k} F_k^T, \tag{3.149}$$

$$h \left( \Pi_{2_k} + \frac{h}{2} X_k \times \frac{\partial U}{\partial X_k} + \frac{h}{2} M_k \right)^\wedge = F_{2_k} J_{d_2} - J_{d_2} F_{2_k}^T, \tag{3.150}$$

where the reduced variables for the linear momentum and the angular momentum are given by $\Gamma = mV$, $\Pi = J_R \Omega = R J_1 \Omega_1$, and $\Pi_2 = J_2 \Omega_2$.

*Numerical Results.* 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 results in 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 $i$th dumbbell, and $\rho_{i_p} \in \mathbb{R}^3$ is a vector from the origin of the body-fixed frame to the $p$th sphere of the $i$th dumbbell in the $i$th 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.

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 the two bodies are near-elliptic orbits. The trajectories of the two dumbbell bodies are shown in Figure 3.7(a).

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 for the full rigid body dynamics. A comparison can be made between the LGVI and other integration methods that 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. (2000) for the continuous equations of motion (2.108)–(2.113), respectively. For the LGVI, the discrete-time equations of motion given by (3.144) through (3.150) 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).

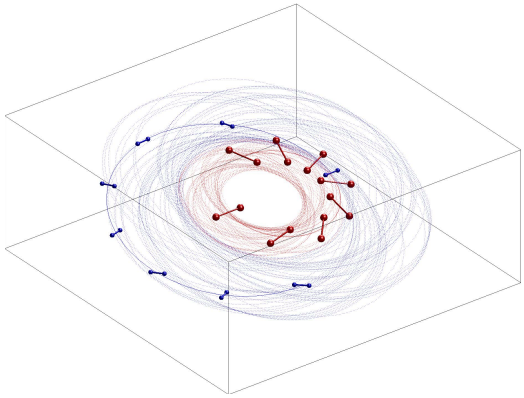
Figure 3.7(b) shows the computed total energy response over 30 seconds with an integration step

size $h = 0.002$ sec. For the LGVI, the total energy is nearly constant, and it has no tendency to drift; while the other integrators fail to preserve the total energy. This can be observed in Figure 3.7(c), 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. Figure 3.7(d) 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 are much less accurate.

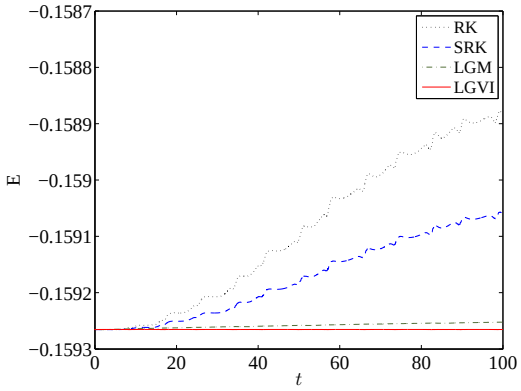
These computational comparisons suggest that for numerical integration of Hamiltonian systems evolving on a Lie group, such as full body problems, it is critical to preserve both the symplectic property and the Lie group structure. For the RK and the SRK, the orthogonality error in the rotation matrix corrupts the attitude of the rigid bodies. The accumulation of this attitude degradation results in 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 geometric 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 Figure 3.7(e), 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 (2.108)–(2.113), 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 Section 3.3.8 is efficient for solving the implicit equation (3.149) and hence it takes less time than the evaluation of (2.108)–(2.113). 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.

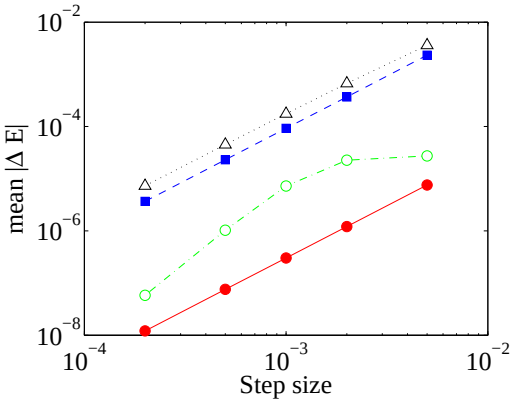
In summary, comparing both Figure 3.7(c) and 3.7(e), we see that the LGVI requires 16 times less CPU time than the LGM, 35 times less CPU time than the RK, and 98 times less CPU time than the SRK for similar total energy error in this computational example for the full body problem.



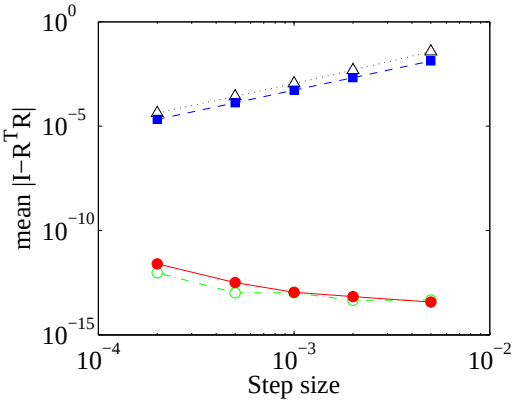
(a) Trajectories of two dumbbell bodies



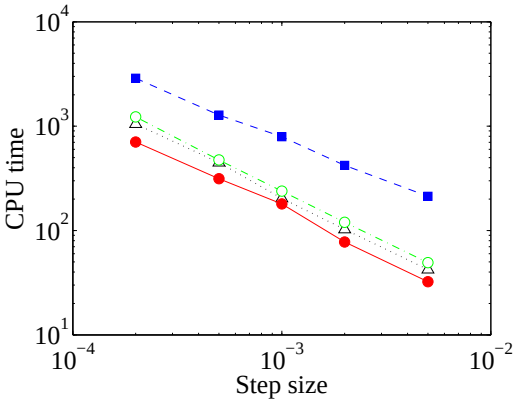

(b) Computed total energy for 30 seconds




(c) Mean total energy error $|E - E_0|$ vs. step size




(d) Mean orthogonality error $\|I - R^T R\|$ vs. step size




(e) CPU time vs. step size

Figure 3.7: Numerical simulation for a full two body problem (Explicit Runge-Kutta:RK, symplectic Runge-Kutta:SRK, Lie group method:LGM, and Lie group variational integrator:LGVI )

### 3.3.7 Two Rigid Bodies Connected by a Ball Joint

Consider the two rigid bodies connected by a ball joint presented in Section 2.3.7.

*Configuration Manifold.* The configuration manifold is $\mathsf{SO(3)} \times \mathsf{SO(3)} \times \mathbb{R}^3$. The discrete update map (3.11) can be written as

$$\begin{aligned}(R_{1_{k+1}}, R_{2_{k+1}}, x_{k+1}) &= (R_{1_k}, R_{2_k}, x_k)(F_{1_k}, F_{2_k}, \Delta x_k) \\ &= (R_{1_k}F_{1_k}, R_{2_k}F_{2_k}, x_k + \Delta x_k)\end{aligned}$$

for $(F_{1_k}, F_{2_k}, \Delta x_k) \in \mathsf{SO(3)} \times \mathsf{SO(3)} \times \mathbb{R}^3$. The adjoint operator also has product structure.

$$\mathrm{Ad}_{(F_1,F_2,x)}(\hat{\eta}_1, \hat{\eta}_2, x') = (F_1\hat{\eta}_1F_1^T, F_2\hat{\eta}_2F_2^T, x), \quad \mathrm{Ad}^*_{(F_1,F_2,x)}(\hat{\eta}_1, \hat{\eta}_2, x') = (F_1^T\hat{\eta}_1F_1, F_2^T\hat{\eta}_2F_2, x).$$

*Discrete Lagrangian.* Recall that, from (2.114), the Lagrangian is given by

$$\begin{aligned}L(R_1, R_2, x, \Omega_1, \Omega_2, \dot{x}) = \frac{1}{2}(m_1 + m_2)\dot{x} \cdot \dot{x} + \frac{1}{2}\Omega_1 \cdot J_1\Omega_1 + \frac{1}{2}\Omega_2 \cdot J_2\Omega_2 \\ + \dot{x} \cdot (m_1R_1\hat{\Omega}_1d_1 + m_2R_2\hat{\Omega}_2d_2) - U(R_1, R_2, x)\end{aligned}$$

for a configuration dependent potential $U : \mathsf{SO(3)} \times \mathsf{SO(3)} \times \mathbb{R}^3 \to \mathbb{R}$.

For $g_k = (R_{1_k}, R_{2_k}, x_k)$ and $f_k = (F_{1_k}, F_{2_k}, \Delta x_k)$, the discrete Lagrangian is chosen as

$$\begin{aligned}L_d(g_k, f_k) = \frac{m_1 + m_2}{2h}\Delta x_k \cdot \Delta x_k + \frac{1}{h}\mathrm{tr}[(I_{3\times3} - F_{1_k})J_{d_1}] + \frac{1}{h}\mathrm{tr}[(I_{3\times3} - F_{2_k})J_{d_2}] \\ + \frac{1}{h}\mathrm{tr}\big[m_1R_{1_k}(F_{1_k} - I_{3\times3})d_1\Delta x_k^T\big] + \frac{1}{h}\mathrm{tr}\big[m_2R_{2_k}(F_{2_k} - I_{3\times3})d_2\Delta x_k^T\big] \\ - hU(g_k).\end{aligned}$$

We find the expressions for the derivatives of the discrete Lagrangian. Using the identity given in (3.68), we obtain

$$\delta(\mathrm{tr}[(I_{3\times3} - F_{i_k})J_{d_i}]) = \mathrm{tr}[-\delta F_{i_k}J_{d_i}] = \mathrm{tr}\big[-F_{i_k}^T\delta F_{i_k}J_{d_i}F_{i_k}\big] = \left\langle F_{i_k}^T\delta F_{i_k}, J_{d_i}F_{i_k} - F_{i_k}^TJ_{i_d}\right\rangle.$$

Similarly, we have

$$\begin{aligned}\delta(\mathrm{tr}\big[m_iR_i(F_{i_k} - I_{3\times3})d_i\Delta x_k^T\big]) = &-\left\langle\hat{\eta}_i, m_i(F_{i_k} - I)d_i\Delta x_k^TR_{i_k} - m_iR_{i_k}^T\Delta x_kd_i^T(F_{i_k}^T - I)\right\rangle \\ &-\left\langle F_{i_k}^T\delta F_{i_k}, m_id_i\Delta x_k^TR_{i_k}F_{i_k} - m_iF_{i_k}^TR_{i_k}^T\Delta x_kd_i^T\right\rangle \\ &+\Delta x_k^Tm_iR_{i_k}(F_{i_k} - I)d_i.\end{aligned}$$

From these, we have

$$\begin{aligned}\mathsf{T}_e^*\mathsf{L}_{F_{i_k}} \cdot \mathbf{D}_{F_{i_k}}L_{d_k} = \frac{1}{h}(J_{d_i} - m_id_i\Delta x_k^TR_{i_k})F_{i_k} - \frac{1}{h}F_{i_k}^T(J_{i_d} - m_iR_{i_k}^T\Delta x_kd_i^T), \\ \mathrm{Ad}^*_{F_{i_k}^{-1}}(\mathsf{T}_e^*\mathsf{L}_{F_{i_k}} \cdot \mathbf{D}_{F_{i_k}}L_{d_k}) = F_{i_k}(\mathsf{T}_e^*\mathsf{L}_{F_{i_k}} \cdot \mathbf{D}_{F_{i_k}}L_{d_k})F_{i_k}^T,\end{aligned}$$

$$\mathsf{T}_e^*\mathsf{L}_{R_{i_k}} \cdot \mathbf{D}_{R_{i_k}} L_{d_k} = \frac{m_i}{h} R_{i_k}^T \Delta x_k d_i^T (F_{i_k}^T - I) - \frac{m_i}{h}(F_{i_k} - I) d_i \Delta x_k^T R_{i_k} - h\hat{M}_{i_k},$$

$$\mathbf{D}_{x_{k+1}} L_{d_k} = \frac{1}{h}(m_1 + m_2)\Delta x_k + \frac{m_1}{h} R_{1_k} \Delta x_k d_1 + \frac{m_2}{h} R_{2_k} \Delta x_k d_2,$$

$$\mathbf{D}_{x_k} L_{d_k} = -\frac{1}{h}(m_1 + m_2)\Delta x_k - \frac{m_1}{h} R_{1_k}(F_{1_k} - I) d_1 - \frac{m_2}{h} R_{2_k}(F_{2_k} - I) d_2 - h\frac{\partial U_k}{\partial x_k}.$$

*Discrete Euler-Lagrange Equations.* Substituting these into (3.10), we obtain the discrete Euler-Lagrange equations for the two rigid bodies connected by a ball joint as

$$A_{i_{k+1}}^T F_{i_{k+1}} - F_{i_{k+1}}^T A_{i_{k+1}} + A_{i_k} F_{i_k} - F_{i_k}^T A_{i_k}^T = h\hat{B}_{i_k} + h^2 \hat{M}_{i_k}, \tag{3.151}$$

$$A_{i_k} = J_{i_d} - m_i R_{i_k}^T (x_{k+1} - x_k) d_i^T, \tag{3.152}$$

$$\hat{B}_{i_k} = \frac{m_i}{h}(F_{i_k} - I) d_i (x_{k+1} - x_k)^T R_{i_k} - \frac{m_i}{h} R_{i_k}^T (x_{k+1} - x_k) d_i^T (F_{i_k}^T - I), \tag{3.153}$$

$$R_{i_{k+1}} = R_{i_k} F_{i_k}, \tag{3.154}$$

$$\begin{aligned} \frac{1}{h}(m_1 + m_2)(x_{k+2} - 2x_{k+1} + x_k) + \frac{m_1}{h}(R_{1_{k+1}}(F_{1_{k+1}} - 2I) + R_{1_k}) d_1 \\ + \frac{m_2}{h}(R_{2_{k+1}}(F_{2_{k+1}} - 2I) + R_{2_k}) d_2 = -h\frac{\partial U_{k+1}}{\partial x_{k+1}} \end{aligned} \tag{3.155}$$

for $i \in \{1, 2\}$. These equation can be solved by a fixed point iteration for $x_{k+2}$. For given $(R_{1_k}, R_{2_k}, F_{1_k}, F_{2_k}, x_k, x_{k+1})$, we guess $x_{k+2}$. Then, $F_{i_{k+1}}$ can be obtained by solving (3.151), and $R_{i_{k+1}}$ is obtained by (3.154). The new value of $x_{k+2}$ is given by (3.155). This procedure is repeated until $x_{k+2}$ converges.

*Discrete Legendre Transformation.* Using the discrete Legendre transformation, the discrete equations of motion in Hamiltonian form are given by

$$\hat{p}_{i_k} = \frac{1}{h}(F_{i_k} J_{d_i} - J_{d_i} F_{i_k}^T) + \frac{m_i}{h}(R_{i_k}^T (x_{k+1} - x_k) d_i^T - d_i (x_{k+1} - x_k)^T R_{i_k}) + h\hat{M}_k, \tag{3.156}$$

$$p_{i_{k+1}} = F_{i_k}^T (p_{i_k} - B_{i_k} - h\hat{M}_k), \tag{3.157}$$

$$\hat{B}_{i_k} = \frac{m_i}{h}(F_{i_k} - I) d_i (x_{k+1} - x_k)^T R_{i_k} - \frac{m_i}{h} R_{i_k}^T (x_{k+1} - x_k) d_i^T (F_{i_k}^T - I), \tag{3.158}$$

$$p_{3_k} = \frac{1}{h}(m_1 + m_2)(x_{k+1} - x_k) + \frac{m_1}{h} R_{1_k}(F_{1_k} - I) d_1 + \frac{m_2}{h} R_{2_k}(F_{2_k} - I) d_2 - h\frac{\partial U_k}{\partial x_k}, \tag{3.159}$$

$$p_{3_{k+1}} = p_{3_k} - h\frac{\partial U_k}{\partial x_k}, \tag{3.160}$$

$$R_{i_{k+1}} = R_{i_k} F_{i_k} \tag{3.161}$$

for $i \in \{1, 2\}$. For given $(R_{1_k}, R_{2_k}, x_k, \Omega_{1_k}, \Omega_{2_k}, \dot{x}_k)$, we find $(p_{1_k}, p_{2_k}, p_{3_k})$ using (2.119). We solve $x_{k+1}$ iteratively: we guess $x_{k+1}$, and solve (3.157) to obtain $F_{i_{k+1}}$. Then, the new value of $x_{k+1}$ is obtained by (3.159). This procedure is repeated until $x_{k+1}$ converges. Then, $p_{i_{k+1}}$

and $p_{3_{k+1}}$ are obtained by (3.157) and (3.160), respectively, and $R_{1_{k+1}}, R_{2_{k+1}}$ are obtained by (3.161). Using (2.119), we find $(\Omega_{1_{k+1}}, \Omega_{2_{k+1}}, \dot{x}_{k+1})$. Thus, this yields a discrete-time flow map $(R_{1_k}, R_{2_k}, x_k, \Omega_{1_k}, \Omega_{2_k}, \dot{x}_k) \to (R_{1_{k+1}}, R_{2_{k+1}}, x_{k+1}, \Omega_{1_{k+1}}, \Omega_{2_{k+1}}, \dot{x}_{k+1})$, and this process is repeated.

*Numerical Results.* We assume there is no potential field, $U \equiv 0$. We choose two elliptic cylinder rigid bodies with different shapes. The properties of the rigid bodies are given by

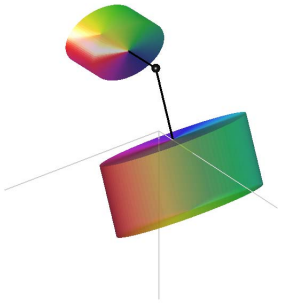
$$m_1 = 2\text{km}, \quad J_1 = \text{diag}[0.1000, 1.1558, 1.2358]\text{kgm}^2, \quad d_1 = [-0.75, 0, 0]\text{m}, \tag{3.162}$$

$$m_1 = 1\text{km}, \quad J_2 = \text{diag}[0.0325, 0.2200, 0.2325]\text{kgm}^2, \quad d_2 = [0.45, 0, 0]\text{m}. \tag{3.163}$$

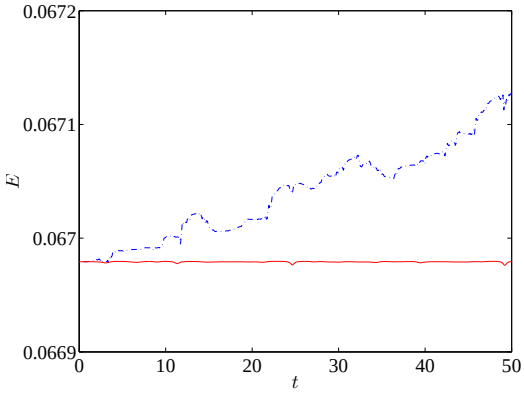
The initial conditions are chosen such that the total linear momentum is zero.

$$R_{1_0} = I, \quad \Omega_{1_0} = [0, 0.5, 0]\text{rad/s}, \quad R_{2_0} = I, \quad \Omega_{2_0} = [1, 0, 0]\text{rad/s},$$
$$x_0 = [0.35, 0, 0]\text{m}, \quad \dot{x}_0 = [0, 0, -0.25]\text{m/s}.$$

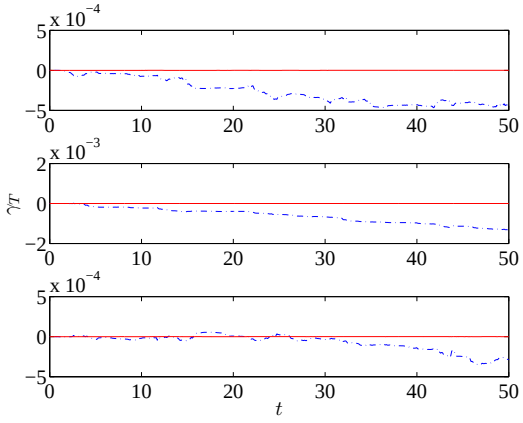
Figure 3.8 shows the computed total energy, total linear momentum, and total angular momentum for the Lie group variational integrator and a 4(5)-th order variable step size Runge-Kutta method. Since there is no potential, the total linear/anguar momentum should be preserved. The Lie group variational integrators preserve the conserved quantities numerically, but there is notable dissipation for the Runge-Kutta method.



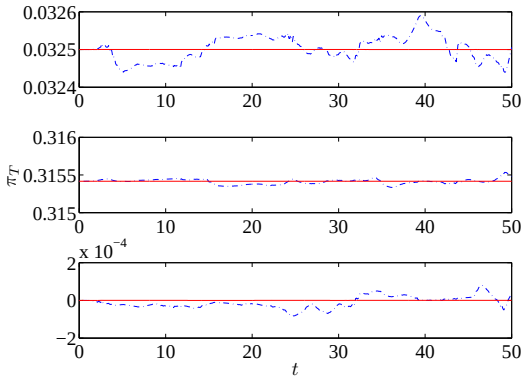
(a) Two rigid bodies connected by a ball joint




(b) Computed total energy




(c) Computed total linear momentum




(d) Computed total angular momentum

Figure 3.8: Numerical simulation for two rigid bodies connect by a ball joint (LGVI: red, solid, RK4(5): blue, dotted)

### 3.3.8 Computational Approach

Most of the discrete-time Euler-Lagrange equations presented in this section include implicit equations. In particular, we need to find the relative update of the group elements represented by the group action of $f_k \in \mathsf{G}$ at each discrete time step. Since the implicit equations are solved repeatedly at each time step, it is important to develop a fast computational algorithm for overall computational efficiency of the presented discrete Euler-Lagrange equations.

The essential idea is expressing the group element $f_k \in \mathsf{G}$ in terms of a Lie algebra element using the exponential map. The exponential map is a local diffeomorphism from the Lie algebra $\mathfrak{g}$ near zero to the Lie group $\mathsf{G}$ near the identity element. Since $f_k$ represents the relative update between two adjacent integration steps, it is close to the identity. Therefore, the implicit equations can be expressed at the Lie algebra level.

This is desirable since the Lie algebra is a linear vector space. The implicit equation is solved numerically in the linear space, and the corresponding group element is obtained by the exponential map. Since the solution of the implicit equation is close to zero in the Lie algebra, the implicit equation is easily solved as the step size is decreased.

In this section, we present a computational approach to solve the implicit equation on $\mathsf{SO}(3)$ appearing in the discrete-time Euler-Lagrange equations for the 3D pendulum. This is the simplest, but nontrivial, form of the implicit equation. The computational approach for other mechanical systems can be obtained by extending the approach.

The discrete-time Euler-Lagrange equations (3.76) and the discrete-time Hamilton equations (3.80) have the following form of the implicit equation; for given $a \in \mathbb{R}^3$ and $J_d \in \mathbb{R}^{3\times 3}$, we need to find $F \in \mathsf{SO}(3)$ satisfying

$$\hat{a} = FJ_d - J_dF^T. \tag{3.164}$$

We rewrite this equation on $\mathbb{R}^3 \simeq \mathfrak{so}(3)$ using the exponential map and the Cayley transformation, to which a Newton iteration method is applied.

#### Exponential map

The exponential map on $\mathsf{SO}(3)$ has a closed form expression, referred to as Rodrigues' formula (see Marsden and Ratiu 1999):

$$F = \exp \hat{f} = I_{3\times 3} + \frac{\sin\|f\|}{\|f\|}\hat{f} + \frac{1-\cos\|f\|}{\|f\|^2}\hat{f}^2 \tag{3.165}$$

for $f \in \mathbb{R}^3$. This physically represents the rotation about the axis $f$ with the rotation angle $\|f\|$.

Substituting (3.165) into (3.164) and using the properties $\hat{f}J_d + J_d\hat{f} = \widehat{Jf}$, $\hat{f}\widehat{Jf} - \widehat{Jf}\hat{f} = \widehat{f \times Jf}$, we obtain

$$\hat{a} = \frac{\sin\|f\|}{\|f\|}\widehat{Jf} + \frac{1-\cos\|f\|}{\|f\|^2}\widehat{f \times Jf}.$$

Thus, (3.80) is converted into an equivalent vector equation in $\mathbb{R}^3$ as

$$0 = -a + \frac{\sin\|f\|}{\|f\|} Jf + \frac{1-\cos\|f\|}{\|f\|^2} f \times Jf \equiv A(f), \tag{3.166}$$

where $A : \mathbb{R}^3 \to \mathbb{R}^3$. We use the Newton method to solve this nonlinear vector equation $0 = A(f)$. We use the following iteration formula iteration

$$f^{(i+1)} = f^{(i)} - \nabla A(f^{(i)})^{-1} A(f^{(i)}), \tag{3.167}$$

where the Jacobian $\nabla A(f)$ in (3.167) can be expressed as

$$\begin{aligned}\nabla A(f) = {} & \frac{\cos\|f\|\,\|f\| - \sin\|f\|}{\|f\|^3} Jff^T + \frac{\sin\|f\|}{\|f\|} J \\ & + \frac{\sin\|f\|\,\|f\| - 2(1-\cos\|f\|)}{\|f\|^4} (f \times Jf) f^T + \frac{1-\cos\|f\|}{\|f\|^2} \left\{ -\widehat{Jf} + \hat{f} J \right\}.\end{aligned}$$

We iterate until $\left\| f^{(i+1)} - f^{(i)} \right\| < \epsilon$ for a given tolerance $\epsilon$, and we find the rotation matrix $F$ using Rodrigues' formula (3.165). The initial guess can be selected as the solution of the linearized vector equation (3.166), which gives $f^{(0)} = J^{-1}a$, or the solution at the previous time step can be used as the initial guess for the current step.

The implicit equation (3.166) expressed in $\mathbb{R}^3$ using the exponential map is well defined, since the rotation matrix $F$ represents the relative attitude update near the identity matrix. This computational approach technically avoids the nonlinear constraints associated with the rotation matrix $F$; the equivalent vector equation is iterated in $\mathbb{R}^3$.

However, the complicated expression for the Jacobian written in terms of trigonometric functions may not be desirable for overall computational efficiency. Thus, we present another computational approach using the Cayley transformation. This approach reduces the computational load using a numerically-efficient local diffeomorphism; the essential idea of solving an equivalent vector equation in $\mathbb{R}^3$ is the same as before.

### Cayley transformation

The Cayley transformation is a local diffeomorphism between $\mathsf{SO}(3)$ and $\mathbb{R}^3$ given by

$$\begin{aligned} F = \mathrm{Cay}(f_c) &= (I + \hat{f}_c)(I - \hat{f}_c)^{-1} = (I - \hat{f}_c)^{-1}(I + \hat{f}_c) \\ &= \frac{1}{1 + f_c^T f_c}((1 - f_c^T f_c)I + 2\hat{f}_c + 2 f_c f_c^T) \end{aligned} \tag{3.168}$$

for $f_c \in \mathbb{R}^3$. This represents a rotation of a rigid body along the direction $f_c / \|f_c\|$ with rotation angle $\theta$ determined by $\|f_c\| = \tan\frac{\theta}{2}$. It can be shown that $\exp\hat{f} = \mathrm{Cay}(\tan\frac{\|f\|}{2}\frac{f}{\|f\|})$. Thus, the Cayley transformation can be considered to be a modification of the exponential map, where the rotation angle is encoded in a different way. But, the Cayley transformation has the numerical advantage that it does not require evaluation of computationally-expensive $\sin$ and $\cos$ terms.

Using the Cayley transformation, we transform the implicit equation (3.164) into an equivalent vector equation. Since the two matrices $(I + \hat{f}_c)$ and $(I - \hat{f}_c)^{-1}$ commute, we can write the Cayley transformation in the following form for convenience.

$$F = \mathrm{Cay}(f_c) = \frac{I + \hat{f}_c}{I - \hat{f}_c}.$$

Substituting this into (3.164), we obtain

$$\hat{a} = \frac{I + \hat{f}_c}{I - \hat{f}_c} J_d - J_d \frac{I - \hat{f}_c}{I + \hat{f}_c}.$$

Multiplying this by $I - \hat{f}_c$ on the left side, by $I + \hat{f}_c$ on the right side, we have

$$(I - \hat{f}_c)\hat{a}(I + \hat{f}_c) = (I + \hat{f}_c)J_d(I + \hat{f}_c) - (I - \hat{f}_c)J_d(I - \hat{f}_c),$$
$$\hat{a} + \widehat{a \times f_c} - \hat{f}_c\hat{a}\hat{f}_c = \hat{f}_c J_d + J_d\hat{f}_c + \hat{f}_c J_d + J_d\hat{f}_c.$$

Using the following identities

$$\widehat{f_c \times a} = \hat{f}_c\hat{a} - \hat{a}\hat{f}_c, \quad \widehat{Jf_c} = J_d\hat{f}_c + \hat{f}_c J_d, \quad \hat{f}_c\hat{a}\hat{f}_c = -(a^T f_c)\hat{f}_c,$$

we obtain

$$(a + a \times f_c + f_c(a^T f_c))^\wedge = 2\widehat{Jf_c}.$$

Therefore, (3.164) is converted into an equivalent vector equation in $\mathbb{R}^3$ as

$$0 = a + a \times f_c + f_c(a^T f_c) - 2Jf_c \equiv A_c(f_c). \tag{3.169}$$

where $A_c : \mathbb{R}^3 \to \mathbb{R}^3$. We use a Newton iteration to solve this nonlinear vector equation $0 = A_c(f)$. We use the following iteration

$$f_c^{(i+1)} = f_c^{(i)} - \nabla A_c(f_c^{(i)})^{-1} A_c(f_c^{(i)}), \tag{3.170}$$

where the Jacobian $\nabla A_c(f_c)$ is given by

$$\nabla A_c(f_c) = \hat{a} + (a^T f_c)I + f_c a^T - 2J. \tag{3.171}$$

We iterate until $\left\| f^{(i+1)} - f^{(i)} \right\| < \epsilon$ for a given tolerance $\epsilon$, and we find the rotation matrix using the Cayley transformation (3.168). The initial guess can be selected as the solution of the linearized vector equation of (3.169), which gives $f_c^{(0)} = (2J - \hat{a})^{-1}a$, or the solution from the previous time step can be used as the initial guess for the current step.

### 3.3.9 Summary of Computational Properties

We have derived discrete-time Euler-Lagrange equations for several rigid body systems evolving on a Lie group, and computational results are presented. Here, we summarize the computational properties of the Lie group variational integrator compared to other numerical integrators for dynamics of rigid bodies.

*Computational Accuracy.* The Lie group variational integrators preserve the geometric properties of mechanical systems on a Lie group. In particular, the computed total energy, momentum map, and deviation from the Lie group configuration manifold are presented. As shown at Figure 3.2, 3.3(b), 3.5(a), 3.6(c), 3.7(b), and 3.8(b), the computed total energy of the Lie group variational integrator oscillates near the initial value, but there is no increasing or decreasing drift for long time periods. This is due to the fact that the numerical solutions of symplectic numerical integrators are the exact solution of a perturbed Hamiltonian (see Hairer 1994). The perturbed value of the Hamiltonian is preserved in the discrete-time flow. This is in contrast to other numerical integrators evaluated in this dissertation, such as the explicit Runge-Kutta method, the symplectic Runge-Kutta method, and (non-symplectic) Lie group methods, where the computed total energy increases (or decreases) linearly with the simulation time.

The Lie group variational integrators preserves the momentum map exactly as discussed in Section 3.1.3. This can be observed in Figure 3.3(d), 3.8(c) and 3.8(d), where the value of the momentum map is preserved up to machine precision. From the numerical simulation of the 3D pendulum on a cart in Figure 3.5(c), the mass center of the system is fixed in space, since the Lie group variational integrator conserves the zero value of the total linear momentum. But the location of the mass center for the explicit Runge-Kutta method drifts, as shown at Figure 3.5(d), since it does not accurately preserve the total linear momentum. Comparing the terminal configurations of both methods in Figure 3.5(f) and Figure 3.5(g), we see that the computational results obtained by the explicit Runge-Kutta method are not reliable.

The structure of the Lie group configuration manifold is well preserved in the Lie group variational integrator. As shown in Figure 3.4(c), 3.5(b), 3.6(d) and 3.7(d), the deviation of the discrete-time flow from the Lie group, measured by the orthogonality error $\left\| I - R^T R \right\|$, remains at machine precision. But interestingly, the orthogonality error increases linearly with respect to the simulation time, and accordingly, the mean orthogonality error increases as the time step decreases for a fixed simulation time. This is due to the accumulation of roundoff errors. For example, since the rotation matrix in $\mathsf{SO}(3)$ is updated by the group action of $\mathsf{SO}(3)$, which is matrix multiplication, the $k$-th rotation matrix is obtained by $k$ matrix multiplications, i.e. $R_k = R_0 F_0 \cdots F_{k-1}$. The roundoff error at each matrix multiplication accumulates, and therefore the orthogonality error increases linearly. The roundoff errors are unavoidable in any numerical computation with a finite digit representation, and they are usually indiscernible. Here we can actually observe the accumulation of roundoff errors, since there is no source for the orthogonality error except the roundoff errors in the numerical computations. The magnitude of the orthogonality error is acceptable for any practical

purpose. In summary, the Lie group variational integrators preserve the Lie group structure up to roundoff errors. The explicit Runge-Kutta method and the symplectic Runge-Kutta method do not preserve the structure of the Lie group. As shown at Figure 3.5(b) and 3.7(d), the orthogonality error for these methods are about $10^{10}$ times larger than the Lie group variational integrators, and consequently, the attitude of the rigid body is not computed accurately.

In summary, the Lie group variational integrators are geometrically accurate numerical integrators in the sense that they preserve all of the geometric properties of the dynamics of rigid bodies, such as symplecticity, momentum map, total energy, and Lie group structure.

*Computational Efficiency.* In Section 3.3.8, computational approaches to solve the implicit equations in the Lie group variational integrator are developed. The essential idea is to write the implicit equations in terms of a Lie algebra element using the exponential map, and numerically solve the transformed implicit equation in a vector space. The corresponding computational requirements are much smaller than for the fully-implicit symplectic Runge-Kutta method. In Section 3.3.6, we have shown that if the discrete Lagrangian is carefully chosen to minimize the number of function evaluations at each step, then the computational requirements of the Lie group variational integrator are comparable to the explicit integrators for the same step size. Therefore, we claim the Lie group variational integrator is *almost explicit*.

For the same level of total energy error, the Lie group variational integrators are faster than the explicit Runge-Kutta method, the symplectic Runge-Kutta method, and the Lie group method by several times. This computational advantage is due to the fact that latter integrators do not preserve all, or one, of the symplecticity and the Lie group structure of the rigid body dynamics.

The RATTLE algorithm is a symplectic numerical integrator for a constrained Hamiltonian system (see Leimkuhler and Reich 2004). By considering the orthogonal structure of the rotation matrix as a nonlinear constraint, one can construct a symplectic numerical integrator that preserves orthogonality using the RATTLE algorithm. A numerical comparison for the 3D pendulum model shows that the Lie group variational integrator is faster than the RATTLE algorithm by 17-60%. This illustrates that it is more efficient to expressing the relative update of group elements in terms of the Lie algebra elements as proposed in this dissertation, than considering the group structure as a nonlinear constraint that should be satisfied at each integration step.

In summary, the Lie group variational integrators have similar computational requirements compared to other explicit numerical integrators for the same step size, and they are substantially more efficient than other numerical integrator for the same level of numerical accuracy.

## 3.4 Examples of Mechanical Systems on Two-Spheres

Several mechanical systems evolving on a product of two-spheres have been introduced in Section 2.4, and the corresponding expressions for inertia matrix, potential, and equations of motion were presented. For the mechanical systems whose Lagrangian is expressed as (2.26), the expressions for the inertia and the potential determine the corresponding discrete-time Euler-Lagrange equations according (3.40), (3.52), (3.53), and (3.54). Therefore, the discrete-time Euler-Lagrange equations for the mechanical systems presented in Section 2.4 have already been obtained. In this section, computational results for the discrete-time Euler-Lagrange equations on two-spheres are presented.

| Section | Mechanical System |
|---|---|
| 3.4.1 | Double Spherical Pendulum |
| 3.4.2 | $n$-body Problem on a Sphere |
| 3.4.3 | Interconnection of Spherical Pendula |
| 3.4.4 | Pure Bending of Elastic Rod |
| 3.4.5 | Spatial Array of Magnetic Dipoles |
| 3.4.6 | Molecular Dynamics on a Sphere |

### 3.4.1 Double Spherical Pendulum

Consider the double spherical pendulum presented in Section 2.4.1. Recall that the inertia matrix is given by $M_{11} = (m_1 + m_2)l_1^2$, $M_{12} = m_2 l_1 l_2$, and $M_{22} = m_2 l_2^2$, and the gravitational potential is written as $U(q_1, q_2) = -(m_1 + m_2)gl_1 e_3 \cdot q_1 - m_2 g l_2 e_3 \cdot q_2$. Substituting these into (3.49)–(3.52), the discrete-time equations of motion for the double spherical pendulum are given by

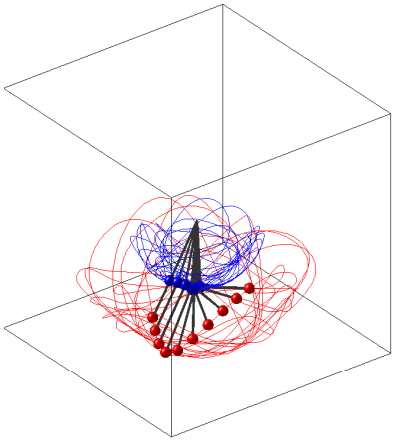
$$\begin{aligned}(m_1 + m_2)l_1^2 q_{1_k} \times F_{1_k} q_{1_k} + q_{1_k} \times m_2 l_1 l_2 (F_{2_k} - I) q_{2_k} \\ = (m_1 + m_2)l_1^2 h\omega_{1_k} - q_{1_k} \times m_2 l_1 l_2 (q_{2_k} \times h\omega_{2_k}) + \frac{h^2}{2} q_{1_k} \times (m_1 + m_2) g l_1 e_3,\end{aligned} \tag{3.172}$$

$$\begin{aligned}m_2 l_2^2 q_{2_k} \times F_{2_k} q_{1_k} + q_{2_k} \times m_2 l_1 l_2 (F_{1_k} - I) q_{1_k} \\ = m_2 l_2^2 h\omega_{2_k} - q_{2_k} \times m_2 l_1 l_2 (q_{1_k} \times h\omega_{1_k}) + \frac{h^2}{2} q_{1_k} \times m_2 g l_2 e_3,\end{aligned} \tag{3.173}$$

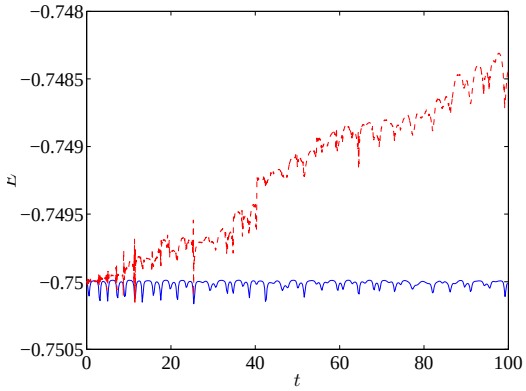
$$q_{1_{k+1}} = F_{1_k} q_{1_k}, \quad q_{2_{k+1}} = F_{2_k} q_{2_k}, \tag{3.174}$$

$$\begin{aligned}&\begin{bmatrix} (m_1 + m_2)l_1^2 I_{3\times 3} & -m_2 l_1 l_2 \hat{q}_{1_{k+1}} \hat{q}_{2_{k+1}} \\ -m_2 l_1 l_2 \hat{q}_{2_{k+1}} \hat{q}_{1_{k+1}} & m_2 l_2^2 I_{3\times 3} \end{bmatrix} \begin{bmatrix} \omega_{1_{k+1}} \\ \omega_{2_{k+1}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{h} q_{1_{k+1}} \times m_2 l_1 l_2 (q_{2_{k+1}} - q_{2_k}) + \frac{h}{2} q_{1_{k+1}} \times (m_1 + m_2) g l_1 e_3 \\ \frac{1}{h} q_{2_{k+1}} \times m_2 l_1 l_2 (q_{1_{k+1}} - q_{1_k}) + \frac{h}{2} q_{2_{k+1}} \times m_2 g l_2 e_3 \end{bmatrix}.\end{aligned} \tag{3.175}$$

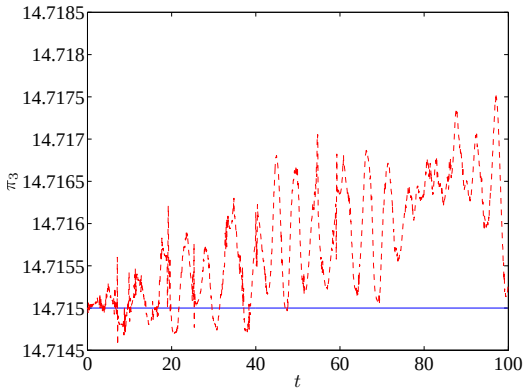
We compare the computational properties of the discrete-time equations of motion with a 4(5)-th order variable step size Runge-Kutta method for (2.124)–(2.123). We choose $m_1 = m_2 = 1\,\mathrm{kg}$, $l_1 = l_2 = 9.81\,\mathrm{m}$. The initial conditions are $q_{1_0} = [0.8660,\ 0,\ 0.5]$, $q_{2_0} = [0,\ 0,\ 1]$, $\omega_{1_0} =$



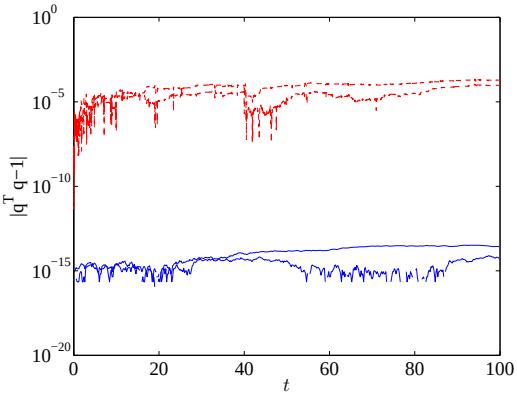
(a) Trajectory of pendulum




(b) Computed total energy




(c) Total angular momentum about the gravity direction



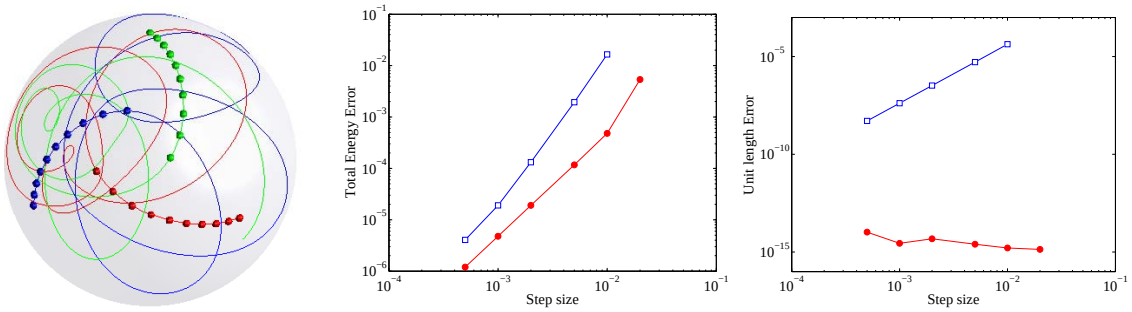

(d) Unit length error

Figure 3.9: Numerical simulation for a double spherical pendulum (RK45: blue, dotted, VI: red, solid)

$[-0.4330,\ 0,\ 0.75]$, $\omega_{2_0} = [0,\ 1,\ 0]\,\mathrm{rad/sec}$. The simulation time is $100\,\mathrm{sec}$, and the step size of the discrete-time equations of motion is $h = 0.01$. Figure 3.9 shows the computed total energy and the configuration manifold errors. The variational integrator preserves the total energy and the structure of $(\mathsf{S}^2)^n$ well for this chaotic motion of the double spherical pendulum. The mean total energy variation is $2.1641 \times 10^{-5}\,\mathrm{Nm}$, and the mean unit length error is $8.8893 \times 10^{-15}$. There is a notable increase of the computed total energy for the Runge-Kutta method, where the mean variation of the total energy is $7.8586 \times 10^{-4}\,\mathrm{Nm}$. The mean deviations of the angular momentum about the gravity direction are $1.0217 \times 10^{-10}\,\mathrm{kgm^2/s}$ for the variational integrator, and $8.0497 \times 10^{-4}$ for the Runge-Kutta method, respectively. The Runge-Kutta method also fails to preserve the structure of $(\mathsf{S}^2)^n$. The mean unit length error is $6.2742 \times 10^{-5}$.

### 3.4.2 $n$-body Problem on a Sphere

Consider the $n$-body problem on a sphere presented in Section 2.4.2. It deals with the motion of $n$ mass particles constrained to lie on a two-sphere, acting under a mutual potential. Recall that the




(a) Trajectory of particles (b) Total energy error v.s. step size (c) Unit length error v.s. step size

Figure 3.10: Numerical simulation for a 3-body problem on sphere (RK2: blue, square, VI: red, circle)

inertia matrix is given by$M_{ij} = m_i$ when $i = j$, and $M_{ij} = 0$ otherwise, and the gravitational potential is given by $U(q_1, \ldots, q_n) = -\frac{\gamma}{2} \sum_{j \neq i} \frac{q_i \cdot q_j}{\sqrt{1-(q_i \cdot q_j)^2}}$. Substituting these into (3.53)–(3.54), the discrete-time equations of motion are given by

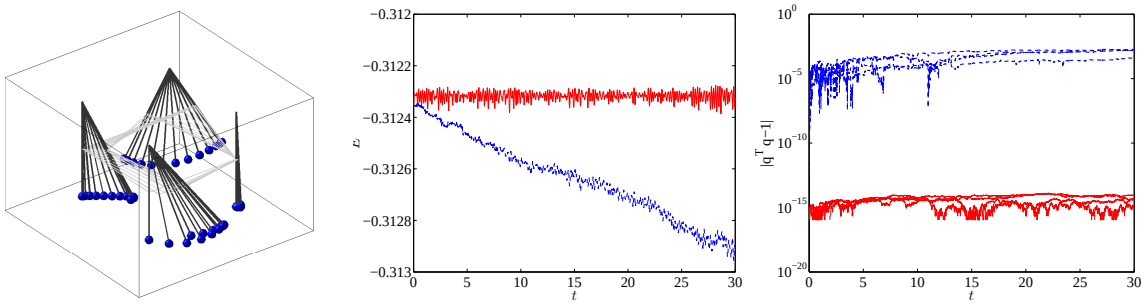
$$q_{i_{k+1}} = \left( h\omega_{i_k} - \frac{h^2}{2m_i} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right) \times q_{i_k} + \left( 1 - \left\| h\omega_{i_k} - \frac{h^2}{2m_i} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right\|^2 \right)^{1/2} q_{i_k}, \quad (3.176)$$

$$\omega_{i_{k+1}} = \omega_{i_k} - \frac{h}{2m_i} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} - \frac{h}{2m_i} q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}}, \quad (3.177)$$

where

$$\frac{\partial U_k}{\partial q_{i_k}} = -\gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{q_{j_k}}{(1 - (q_{i_k} \cdot q_{j_k})^2)^{3/2}}.$$

A two-body problem on the two-sphere under this gravitational potential is studied in Hairer et al. (2003) 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-time equations of motion with a 2-nd order fixed step size Runge-Kutta method for (2.125). 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 $10\,\text{sec}$. Figure 3.10 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 step size by several orders of magnitude. For the variational integrator, the total energy error is reduced by almost 100 times from $1.1717 \times 10^{-4}$ to $1.1986 \times 10^{-6}$ when the step size is reduced by 10 times from $10^{-3}$ to $10^{-4}$, which verifies the second order accuracy numerically.




(a) Motion of pendula (b) Computed total energy (c) Unit length error

Figure 3.11: Numerical simulation for an interconnection of 4 spherical pendula (RK2: blue, dotted, RK2 with projection: black, dashed, VI: red, solid)

### 3.4.3 Interconnection of Spherical Pendula

Consider the interconnection of spherical pendula presented in Section 2.4.3. Recall that the inertia matrix is given by $M_{ij} = m_i l_i^2$ when $i = j$, and $M_{ij} = 0$ otherwise, and the potential is given by

$$U(q_1, \dots q_n) = -\sum_{i=1}^{n} m_i g l_i q_i \cdot e_3 + \sum_{(i,j)\in\Xi} \frac{1}{2}\kappa_{ij} \left( \left\| r_{ij} + \frac{1}{2} l_j q_j - \frac{1}{2} l_i q_i \right\| - \|r_{ij}\| \right)^2.$$

Substituting these into (3.53)–(3.54), the discrete-time equations of motion are given by

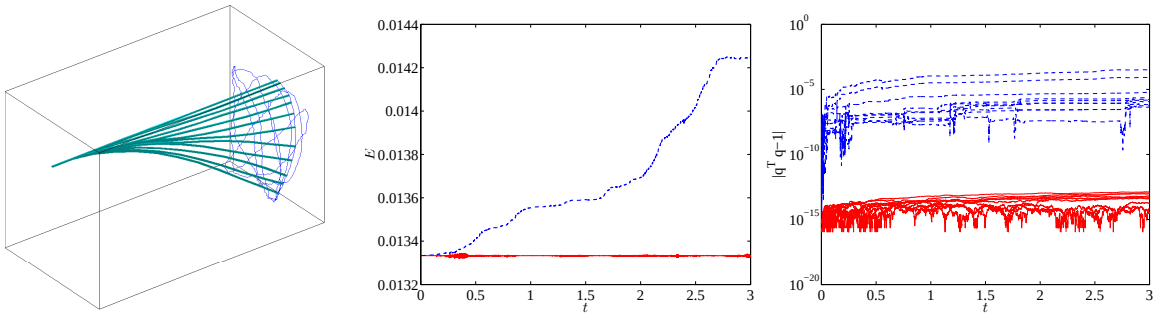
$$q_{i_{k+1}} = \left( h\omega_{i_k} - \frac{h^2}{2m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right) \times q_{i_k} + \left( 1 - \left\| h\omega_{i_k} - \frac{h^2}{2m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right\|^2 \right)^{1/2} q_{i_k}, \tag{3.178}$$

$$\omega_{i_{k+1}} = \omega_{i_k} - \frac{h}{2m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} - \frac{h}{2m_i l_i^2} q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}}. \tag{3.179}$$

We compare the computational properties of the discrete-time equations of motion with a 2-nd order fixed step size explicit Runge-Kutta method for (2.126)–(2.127), and the same Runge-Kutta method with reprojection; at each time step, the vectors $q_{i_k}$ are projected onto $\mathsf{S}^2$ by using normalization.

We choose four interconnected pendula, $n = 4$, and we assume each pendulum has the same mass and length; $m_i = 0.1\,\mathrm{kg}$, $l_i = 0.1\,\mathrm{m}$. A set $\Xi$ is defined such that $(i, j) \in \Xi$ if the $i$-th pendulum and the $j$-th pendulum are connected. We choose $\Xi = \{(1,2), (2,3), (3,4), (4,1)\}$, 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 3.11 shows the computed total energy and the unit length errors. The variational integrator preserves the total energy and the structure of $(\mathsf{S}^2)^n$ well. The mean total energy variation is $3.6171 \times 10^{-5}\,\mathrm{Nm}$, and the mean unit length error is $4.2712 \times 10^{-15}$. For both Runge-Kutta




(a) Deformation of rod (b) Computed total energy (c) Unit length error

Figure 3.12: Numerical simulation for an elastic rod (RK45: blue, dotted, VI: red, solid)

methods, there is a notable increase of the computed total energy. It is interesting to see that the reprojection approach increases the total energy error, even though it preserves the structure of $(\mathsf{S}^2)^n$ accurately. This shows that a standard reprojection method can significantly corrupt numerical trajectories (see Hairer et al. 2000; Lewis and Nigam 2003).

### 3.4.4 Pure Bending of Elastic Rod

Consider the pure bending motion of a slender elastic rod presented in Section 2.4.4. Recall that the moment of inertia matrix is given by

$$M_{ii} = \frac{1}{3} m_i l_i^2 + \sum_{p=i+1}^{n} m_p l_i^2,$$
$$M_{ij} = \frac{1}{2} \left[ \left( \sum_{p=i+1}^{n} 2 m_p l_j l_i \right) + m_i l_j l_i \right]$$

for $1 \leq i < j \leq n$, and the potential is given by

$$U(q_1, \ldots, q_n) = \sum_{i=1}^{n} \left[ -m_i g \left( \sum_{j=1}^{i-1} q_j l_j + \frac{1}{2} l_i q_i \right) \cdot e_3 + \frac{EI}{4 l_i^2} (1 - q_{i-1} \cdot q_i) \right].$$

Substituting these into (3.49)–(3.52), we obtain the discrete-time equations of motion.

We compare the computational properties of the discrete-time equations of motion with a 4(5)-th order variable step size Runge-Kutta method. 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 $3\,\mathrm{sec}$, and the step size is $h = 0.0001$.

Figure 3.12 shows the computed total energy and the unit length errors. The variational integrator preserves the total energy and the structure of $(\mathsf{S}^2)^n$. The mean total energy variation is

$1.4310 \times 10^{-6}\,\mathrm{Nm}$, and the mean unit length error is $2.9747 \times 10^{-14}$. There is a notable dissipation of the computed total energy for the Runge-Kutta method, where the mean variation of the total energy is $3.5244 \times 10^{-4}\,\mathrm{Nm}$. The Runge-Kutta method also fail to preserve the structure of $(\mathsf{S}^2)^n$. The mean unit length error is $1.8725 \times 10^{-5}$.

### 3.4.5 Spatial Array of Magnetic Dipoles

Consider the spatial array of magnetic dipoles presented in Section 2.4.5. Recall that the inertia matrix is given by $M_{ij} = \frac{1}{12} m_i l_i^2$ when $i = j$, and $M_{ij} = 0$ otherwise. The magnetic potential is given by

$$U(q_1, \ldots, q_n) = \frac{1}{2} \sum_{\substack{i,j=1 \\ j \neq i}}^{n} \frac{\mu\, \nu_i \nu_j}{4\pi \|r_{ij}\|^3} \left[ (q_i \cdot q_j) - \frac{3}{\|r_{ij}\|^2} (q_i \cdot r_{ij})(q_j \cdot r_{ij}) \right],$$

Substituting these into (3.53)–(3.54), we obtain the discrete-time Euler-Lagrange equations as

$$q_{i_{k+1}} = \left( h\omega_{i_k} - \frac{6h^2}{m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right) \times q_{i_k} + \left( 1 - \left\| h\omega_{i_k} - \frac{6h^2}{m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right\|^2 \right)^{1/2} q_{i_k}, \quad (3.180)$$

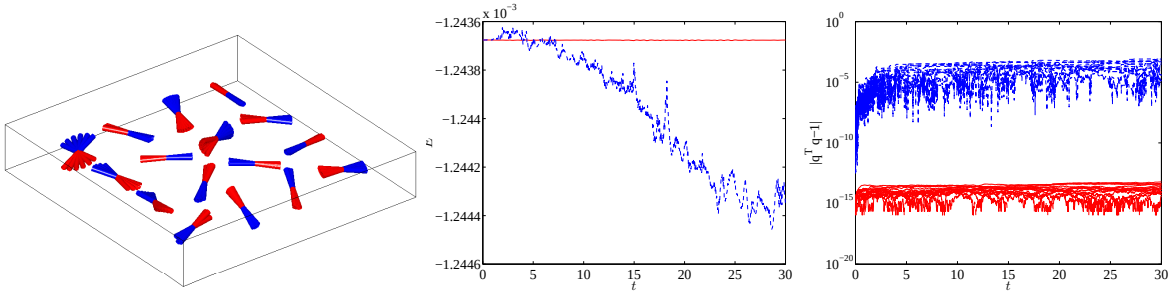
$$\omega_{i_{k+1}} = \omega_{i_k} - \frac{6h}{m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} - \frac{6h}{m_i l_i^2} q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}}, \quad (3.181)$$

where

$$\frac{\partial U_k}{\partial q_{i_k}} = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\mu\, \nu_i \nu_j}{4\pi \|r_{ij}\|^3} \left[ q_j - \frac{3}{\|r_{ij}\|^2} r_{ij} (q_j \cdot r_{ij}) \right].$$

We compare the computational properties of the discrete-time equations of motion with a 4(5)-th order variable step size Runge-Kutta method for (2.136)–(2.137). 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 $4 \times 4$ 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 3.13 shows the computed total energy and the unit length errors. The variational integrator preserves the total energy and the structure of $(\mathsf{S}^2)^n$ well. The mean total energy variation is $8.5403 \times 10^{-10}\,\mathrm{Nm}$, and the mean unit length error is $1.6140 \times 10^{-14}$. There is a notable dissipation of the computed total energy for the Runge-Kutta method, where the mean variation of the total energy is $2.9989 \times 10^{-7}\,\mathrm{Nm}$. The Runge-Kutta method also fail to preserve the structure of $(\mathsf{S}^2)^n$. The mean unit length error is $1.7594 \times 10^{-4}$.




(a) Motion of magnetic dipoles (b) Computed total energy (c) Unit length error

Figure 3.13: Numerical simulation for an array of magnetic dipoles (RK45: blue, dotted, VI: red, solid)

### 3.4.6 Molecular Dynamics on a Sphere

Consider the molecular dynamics model on $\mathsf{S}^2$ presented in Section 2.4.6. Recall that the inertia matrix is $M_{ij} = m_i$ when $i = j$, and $M_{ij} = 0$ otherwise. The Lennard-Jones potential is given by

$$U(q_1, \ldots, q_n) = \frac{1}{2} \sum_{\substack{i,j=1 \\ j \neq i}}^{n} 4\epsilon \left[ \left( \frac{\sigma}{\|q_i - q_j\|} \right)^{12} - \left( \frac{\sigma}{\|q_i - q_j\|} \right)^{6} \right],$$

Substituting these into (3.53)–(3.54), we obtain the discrete-time Euler-Lagrange equations as

$$q_{i_{k+1}} = \left( h\omega_{i_k} - \frac{h^2}{2m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right) \times q_{i_k} + \left( 1 - \left\| h\omega_{i_k} - \frac{h^2}{2m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} \right\|^2 \right)^{1/2} q_{i_k}, \tag{3.182}$$

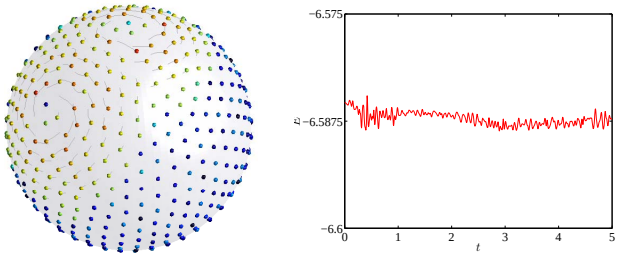
$$\omega_{i_{k+1}} = \omega_{i_k} - \frac{h}{2m_i l_i^2} q_{i_k} \times \frac{\partial U_k}{\partial q_{i_k}} - \frac{h}{2m_i l_i^2} q_{i_{k+1}} \times \frac{\partial U_{k+1}}{\partial q_{i_{k+1}}}, \tag{3.183}$$

where

$$\frac{\partial U_k}{\partial q_{i_k}} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} 4\epsilon \frac{q_i - q_j}{\|q_i - q_j\|} \left[ \frac{12\sigma^{12}}{\|q_i - q_j\|^{13}} - \frac{6\sigma^6}{\|q_i - q_j\|^7} \right].$$

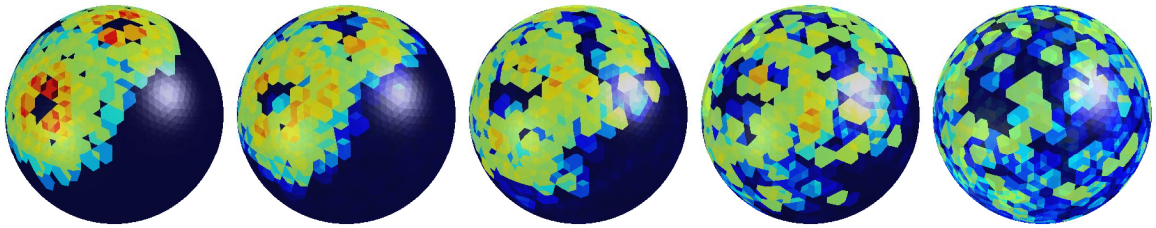
We choose 642 molecules, $n = 642$, and we assume each molecule has the same mass, $m_i = 1$. Initially, molecules are uniformly distributed on a sphere. The strength of the potential is chosen as $\epsilon = 0.01$, and the constant $\sigma$ is chosen such that the inter-molecular force between neighboring molecules is close to zero. The initial velocities are modeled as two vortices separated by $30$ degrees. The simulation time is $5$ s, and the step size is $h = 0.005$.

Trajectories of molecules and the computed total energy are shown in Figure 3.14. The mean deviation of the total energy is $1.8893 \times 10^{-3}$, and the mean unit length error is $5.2623 \times 10^{-15}$. In molecular dynamics simulations, macroscopic quantities such as temperature and pressure are more useful than trajectories of individual molecules. Figure 3.15 shows the change of kinetic energy distributions over time, which measures the temperature (see Allen and Tildesley 1987); the sphere




(a) Initial trajectories (b) Computed total energy

Figure 3.14: Numerical simulation for molecular dynamics on a sphere



(a) $t = 0$ (b) $t = 0.25$ (c) $t = 0.5$ (d) $t = 0.75$ (e) $t = 5$

Figure 3.15: Numerical simulation for molecular dynamics on a sphere: kinetic energy distributions over time

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.

### 3.4.7 Computational Approach

For the discrete equations of motion, we need to solve (3.40) and (3.49) to obtain $F_{i_k} \in \mathsf{SO}(3)$. Here we present a computational approach. The implicit equations given by (3.40) and (3.49) have the following structure.

$$M_{ii} q_i \times F_i q_i + q_i \times \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{ij} (F_j - I_{3\times 3}) q_j = d_i \tag{3.184}$$

for $i \in \{1, \ldots, n\}$, where $M_{ij} \in \mathbb{R}$, $q_i \in \mathsf{S}^2$, $d_i \in \mathbb{R}^3$ are known, and we need to find $F_i \in \mathsf{SO}(3)$. 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. But, the solution is not unique since if $F_i \in \mathsf{SO}(3)$ satisfies (3.184), then $F_i \exp(\theta q_i) \in \mathsf{SO}(3)$ also satisfies (3.184) for any $\theta \in \mathsf{S}^1$. To avoid this ambiguity, we search a 2-dimensional subgroup of $\mathsf{SO}(3)$ in which the solution of (3.184) is unique.

Similar to the computational approach to solve the implicit equation of the Lie group variational integrator on $\mathsf{SO}(3)$ presented in Section 3.3.8, the essential idea is to express the rotation matrix

$F_i$ in terms of $f_i \in \mathbb{R}^3$ using the the Cayley transformation. Using the Cayley transformation, $F_i \in \mathsf{SO(3)}$ can be expressed in terms of $f_i \in \mathbb{R}^3$ as

$$\begin{aligned} F_i &= (I_{3\times 3} + \hat{f}_i)(I_{3\times 3} - \hat{f}_i)^{-1} \\ &= \frac{1}{1 + f_i \cdot f_i}((1 - f_i \cdot f_i)I_{3\times 3} + 2f_i f_i^T + 2\hat{f}_i). \end{aligned}$$

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

$$F_i q_i = \frac{1}{1 + f_i \cdot f_i}((1 - f_i \cdot f_i)q_i + 2\hat{f}_i q_i). \tag{3.185}$$

Thus, we obtain

$$\begin{aligned} q_i \times F_i q_i = \frac{2}{1 + f_i \cdot f_i} q_i \times (f_i \times q_i) = \frac{2}{1 + f_i \cdot f_i} f_i, \\ (F_j - I_{3\times 3})q_j = -\frac{2}{1 + f_j \cdot f_j}(q_j f_j^T + \hat{q}_j)f_j, \end{aligned}$$

where we use the property, $\hat{q}_i f_i = q_i \times f_i = -\hat{f}_i q_i$. Substituting these into (3.184), we obtain

$$\begin{bmatrix} \frac{2M_{11}I_{3\times 3}}{1+f_1\cdot f_1} & -\frac{2M_{12}\hat{q}_1(\hat{q}_2+q_2 f_2^T)}{1+f_2\cdot f_2} & \cdots & -\frac{2M_{1n}\hat{q}_1(\hat{q}_n+q_n f_n^T)}{1+f_n\cdot f_n} \\ -\frac{2M_{21}\hat{q}_2(\hat{q}_1+q_1 f_1^T)}{1+f_1\cdot f_1} & \frac{2M_{22}I_{3\times 3}}{1+f_2\cdot f_2} & \cdots & -\frac{2M_{2n}\hat{q}_2(\hat{q}_n+q_n f_n^T)}{1+f_n\cdot f_n} \\ \vdots & \vdots & & \vdots \\ -\frac{2M_{n1}\hat{q}_n(\hat{q}_1+q_1 f_1^T)}{1+f_1\cdot f_1} & -\frac{2M_{n2}\hat{q}_n(\hat{q}_2+q_2 f_2^T)}{1+f_2\cdot f_2} & \cdots & \frac{2M_{nn}I_{3\times 3}}{1+f_n\cdot f_n} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}, \tag{3.186}$$

which is an equation equivalent to (3.184), written in terms of local coordinates for $F_i$ using the Cayley transformation. Any standard numerical method to solve nonlinear equations can be applied to find $f_i$. Then, $F_i q_i$ is computed by using (3.185). In particular, (3.186) is written in a form for which a fixed point iteration method can be readily applied (see Kelley 1995).

If there are no coupling terms in the kinetic energy, we can obtain an explicit solution of (3.184). When $M_{ij} = 0$ for $i \neq j$, (3.186) 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 (3.185) 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-time equations of motion given by (3.49)–(3.52)

in an explicit form as in Corollary 3.5.

### 3.4.8 Summary of Computational Properties

We have derived discrete-time Euler-Lagrange equations for several dynamic systems evolving on a product of two-spheres, and some computational results are presented. Here, we summarize the computational properties of the Lie homogeneous variational integrator compared to other numerical integrators.

*Computational Accuracy.* Similar to the Lie group variational integrator, the distinct feature of the Lie homogeneous variational integrators is that they preserve the symplecticity as well as the structure of the two-spheres. As shown at Figure 3.9(b), 3.11(b), 3.12(b), 3.13(b), and 3.14(b), the computed total energy of the Lie homogeneous variational integrator oscillates around the initial value, but there is no drift on long time scales. The computed total energy of the explicit Runge-Kutta methods increases (or decreases) linearly with respect to the simulation time. As a result, the total energy deviation for the Runge-Kutta method is larger than the Lie homogeneous variational integrator by several times as seen in Figure 3.10(b). We find a similar property for the momentum map at Figure 3.9(c). Figure 3.9(d), 3.11(c), 3.12(c), and 3.13(c) show that the unit length error of the Lie homogeneous variational integrator is at the level of $10^{-14}$–$10^{-15}$, but the unit length error of the explicit Runge-Kutta method is almost $10^{10}$ times larger.

In summary, the Lie homogeneous variational integrators are geometrically accurate in the sense that they preserve all of the geometric properties of the dynamics, such as symplecticity, momentum map, total energy, and the structure of the two-spheres.

*Computational Efficiency.* We have shown that if the inertia matrix for the kinetic energy is diagonal, the Lie homogeneous variational integrator becomes explicit. In this case, the variational integrator given in Corollary 3.5 is faster than the explicit Runge-Kutta method with the same second order accuracy, as it requires one-time evaluation of the potential derivatives per time step.

The RATTLE algorithm is a symplectic numerical integrator for a constrained Hamiltonian system (see Leimkuhler and Reich 2004). By considering the unit norm structure of the two-sphere as a nonlinear constraint, one can construct a symplectic numerical integrator evolving on two-spheres. It turns out that the variational integrator given at Corollary 3.4 is equivalent to the RATTLE algorithm if the computational implementational details are ignored. But, rather than imposing the constraint on the unit length at each time step, we present a computational algorithm to update the element of the two-sphere using the group action of $\mathsf{SO}(3)$.

The purpose of this section is to provide a systematic method to derive a geometric numerical integrator evolving on a homogeneous space. The specific form of the discrete Lagrangian given in (3.33) yields an equivalent expression to the RATTLE algorithm. But, it can be extended in various ways, such as a configuration dependent inertia or an abstract homogeneous manifold, by following the approach presented in this section. The presented computational approach is generally

slower than the RATTLE algorithm, but it represents a specific numerical method that updates an element of a homogeneous space using the corresponding Lie group action explicitly. In particular, a subgroup that guarantees the transitivity property is found for overall efficiency.

In summary, the presented form of the Lie homogeneous variational integrator is more numerically efficient if the inertia matrix is diagonal. It may be slightly slower than the RATTLE algorithm using the presented computational algorithm. But, the computational approach provides an explicit method to generalize the current results to mechanical systems evolving on an abstract homogeneous manifold.

## 3.5 Conclusions

In this chapter, geometric numerical integrators, referred to as Lie group variational integrators, are developed for dynamic systems with a Lie group configuration manifold. The presented method represents the first time that the Lie group approach is explicitly adopted in the context of variational integrators for an arbitrary Lie group. They provide a systematic way to develop a class of geometric numerical integrators that preserve the geometric properties of the dynamics as well as the Lie group structure.

Numerical simulations in Section 3.3 show that it is critical to preserve both the symplectic property of dynamics and the structure of the Lie group. The Lie group variational integrators have substantial computational advantages compared to other geometric integrators that preserve either none or one of these properties. They are more efficient than considering the Lie group structure as a nonlinear algebraic constraint to be satisfied at each time step.

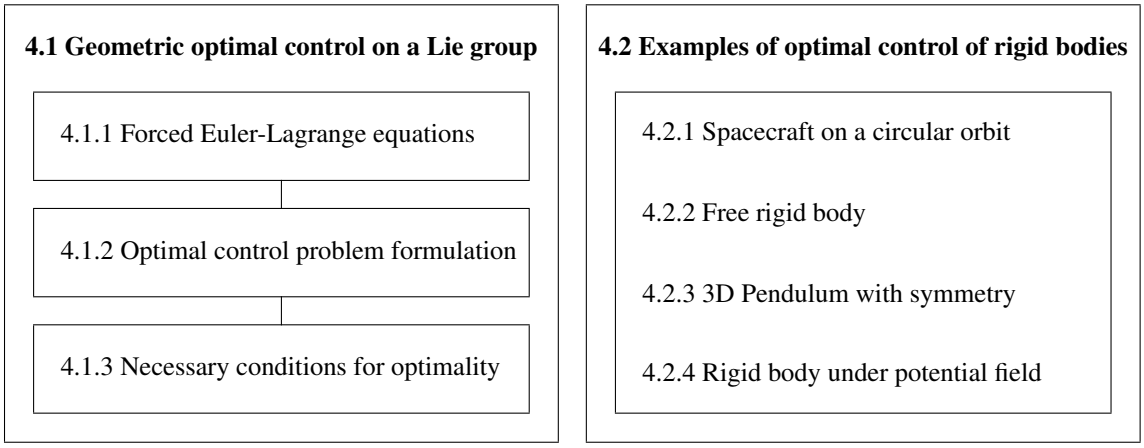
Compared with other geometric integrators for a rigid body, as in the work of Hulbert (1992); Krysl (2005); Lewis and Simo (1994); Simo and Wong (1991), the Lie group variational integrator provides a systematic method to obtain a class of numerical integrators that preserve all of the geometric features, rather than developing a specific numerical integrator that preserves only a few geometric characteristics. Compared with discrete-time mechanics on a Lie group developed by Bobenko and Suris (1999); Marsden et al. (1999); Moser and Veselov (1991), the Lie group variational integrator can be applied to a wide class of rigid body dynamics acting under a potential field as shown in Section 3.3.

These results are extended to mechanical systems on a product of two-spheres, to obtain Lie homogeneous variational integrators in Section 3.2. They preserve the geometric feature of dynamics as well as the structure of the two-spheres. Other geometric numerical integrators on $\mathsf{S}^2$ are developed for a kinematics equations, and they do not necessarily preserve the symplectic property or momentum map (see Lewis and Nigam 2003; Lewis and Olver 2001; Munthe-Kaas and Zanna 1997). The development in Section 3.2 also provides an overall framework to construct variational integrators for dynamic systems on an abstract homogeneous manifold.

CHAPTER 4

# GEOMETRIC OPTIMAL CONTROL OF RIGID BODIES ON A LIE GROUP

In this chapter, we formulate optimal control problems for mechanical systems that evolve on a Lie group, and we derive necessary conditions for optimality. These results are illustrated by several rigid body systems introduced in the previous chapters. This chapter is focused on developing an intrinsic form of optimality conditions in continuous-time. Computational approaches to find optimal control inputs are discussed in the next chapter.




This chapter is organized as follows. In Section 4.1, we develop geometric optimal control theory for dynamics on an arbitrary Lie group; the Euler-Lagrange equations derived in Chapter 2 are extended to include the effect of external control inputs, a general form of an optimal control problem is formulated, and necessary conditions for optimality are developed. In Section 4.2, these results are applied to several optimal control problems for rigid body dynamics.

## 4.1 Geometric Optimal Control on a Lie Group

Optimal control problems deal with finding trajectories, such that a certain optimality condition is satisfied under prescribed constraints (see, for example, Bryson and Ho 1975; Kirk 1970; Sussmann and Willems 1997). For example, a minimum time optimal control problem is studied for a spacecraft to change its attitude in a desired way while minimizing the maneuver time, and subject to bounded control moments.

Geometric optimal control on a Lie group has been studied by Baillieul (1978); Bloch and

Crouch (1996); Brockett (1972); Jurdjevic (1998a,b, 1997) (see also references therein). But, these approaches are based on kinematics equations on a Lie group, and assume that group elements are directly controlled by control input elements in the Lie algebra. For example, an optimal attitude control problem of a rigid body is considered in Jurdjevic (1997) by viewing the angular velocity to be a control input.

This chapter deals with optimal control problems of *dynamic systems* with a Lie group configuration manifold. More precisely, it may be considered as an optimal control problem on a tangent bundle of a Lie group $\mathsf{G}$, identified with $\mathsf{G} \times \mathfrak{g}$ by left trivialization. For example, in Section 4.2.2, we consider an optimal attitude control problem of a rigid body controlled by an external moment.

Another distinct feature of the presented geometric optimal control theory is that it is a coordinate-free approach. Most of the prior work related to optimal control problems of a rigid body is based on local coordinates of $\mathsf{SO}(3)$, such as Euler angles, or on quaternions (see, for example, Bilimoria and Wie 1993; Byers and Vadali 1993; Scrivener and Thompson 1994; Seywald and Kumar 1993). Here, we develop an intrinsic form of optimal control problems and necessary conditions. Optimality conditions and the resulting optimal control input are independent of the choice of representation for the rigid body configuration. They are globally represented by group elements without any singularity and ambiguity, and the optimality conditions are more compact than expressions written in terms of local parameterizations.

In summary, we develop a geometric optimal control theory to treat optimal control problems for dynamic systems on a Lie group, expressed in a coordinate-free form. The geometric optimal control approach is based on the Lagrangian mechanics on a Lie group presented in Chapter 2; necessary conditions for optimality are derived using variational arguments. Here, we first extend the Euler-Lagrange equations to include the effect of control inputs, and we formulate optimal control problems for controlled Euler-Lagrange systems on a Lie group.

### 4.1.1 Forced Euler-Lagrange Equations on a Lie Group

Consider a mechanical system evolving on a Lie group $\mathsf{G}$. As discussed in Chapter 2, we assume that a Lagrangian of the system is expressed as $L(g, \xi) : \mathsf{G} \times \mathfrak{g} \to \mathbb{R}$ by left-trivialization, $\xi = g^{-1}\dot{g}$.

Suppose that there exists a generalized force $u(t) : [t_0, t_f] \to \mathfrak{g}^*$ acting on the system. Forced Euler-Lagrange equations are obtained according to the Lagrange-d'Alembert principle:

$$\delta \int_{t_0}^{t_f} L(g, \xi)\, dt + \int_{t_0}^{t_f} u(t) \cdot \eta\, dt = 0 \tag{4.1}$$

for any $\eta = g^{-1}\delta g \in \mathfrak{g}$ vanishing at the endpoints. This is equivalent to Hamilton's principle with the additional forcing term. From (2.8), this can be written as

$$\int_{t_0}^{t_f} \left\langle \mathsf{T}_e^*\mathsf{L}_g \cdot \mathbf{D}_g L(g, \xi) + \mathrm{ad}_\xi^* \cdot \mathbf{D}_\xi L(g, \xi),\ \eta \right\rangle - \left\langle \frac{d}{dt}\mathbf{D}_\xi L(g, \xi),\ \eta \right\rangle + \langle u,\ \eta \rangle\ dt = 0,$$

which yields the forced Euler-Lagrange equations on $\mathsf{G}$.

**Proposition 4.1** *Consider a mechanical system evolving on a Lie group* $\mathsf{G}$*. We identify the tangent bundle* $\mathsf{TG}$ *with* $\mathsf{G} \times \mathfrak{g}$ *by left-trivialization. Suppose that the Lagrangian is defined as* $L(g, \xi) : \mathsf{G} \times \mathfrak{g} \to \mathbb{R}$*, and there exists a generalized force* $u : [t_0, t_f] \to \mathfrak{g}^*$ *acting on the system. The corresponding forced Euler-Lagrange equations are given by*

$$\frac{d}{dt}\mathbf{D}_\xi L(g,\xi) - \mathrm{ad}^*_\xi \cdot \mathbf{D}_\xi L(g,\xi) - \mathsf{T}^*_e\mathsf{L}_g \cdot \mathbf{D}_g L(g,\xi) = u, \tag{4.2}$$

$$\dot{g} = g\xi. \tag{4.3}$$

*Special Form of Forced Euler-Lagrange Equations.* Proposition 4.1 gives the forced Euler-Lagrange equations for mechanical systems evolving on a Lie group $\mathsf{G}$ for a general form of the Lagrangian. While it is possible to formulate an optimal control problem and derive necessary optimality conditions for these systems, we consider mechanical systems with a structured form of the Lagrangian. This allows us to obtain a more compact form of necessary conditions in Section 4.1.3.

Let $\mathbf{J} : \mathfrak{g} \to \mathfrak{g}^*$ be the inertia operator, which is linear, positive definite, and symmetric. More explicitly, it satisfies

$$\langle \mathbf{J}(\xi),\ \xi\rangle > 0, \tag{4.4}$$

$$\mathbf{J}(c_1\xi_1 + c_2\xi_2) = c_1\mathbf{J}(\xi_1) + c_2\mathbf{J}(\xi_2), \tag{4.5}$$

$$\langle \mathbf{J}(\xi_1),\ \xi_2\rangle = \langle \mathbf{J}(\xi_2),\ \xi_1\rangle, \tag{4.6}$$

for any $c_1, c_2 \in \mathbb{R}$, $\xi \neq 0, \xi_1, \xi_2 \in \mathfrak{g}$.

We assume that the Lagrangian is the difference between a kinetic energy, expressed in term of the inertia operator, and a configuration dependent potential $U : \mathsf{G} \to \mathbb{R}$, given by

$$L(g,\xi) = \frac{1}{2}\langle \mathbf{J}(\xi),\ \xi\rangle - U(g). \tag{4.7}$$

We apply Proposition 4.1 to this Lagrangian. The derivatives of the Lagrangian are given by

$$\mathbf{D}_\xi L(g,\xi)\cdot\delta\xi = \frac{1}{2}\langle \mathbf{J}(\delta\xi),\ \xi\rangle + \frac{1}{2}\langle \mathbf{J}(\xi),\ \delta\xi\rangle = \langle \mathbf{J}(\xi),\ \delta\xi\rangle,$$

$$\mathsf{T}^*_e\mathsf{L}_g \cdot \mathbf{D}_g L(g,\xi) = -\mathsf{T}^*_e\mathsf{L}_g \cdot \mathbf{D}_g U(g) \equiv M(g),$$

where the force due to the potential is denoted by $M : \mathsf{G} \to \mathfrak{g}^*$. Then, the forced Euler-Lagrange equations reduce to the following.

**Corollary 4.1** *Consider a mechanical system evolving on a Lie group* $\mathsf{G}$*. We identify the tangent bundle* $\mathsf{TG}$ *with* $\mathsf{G} \times \mathfrak{g}$ *by left-trivialization. Suppose that the Lagrangian is given by (4.7) for a positive definite, symmetric, and linear inertia operator* $\mathbf{J} : \mathfrak{g} \to \mathfrak{g}^*$ *and the configuration dependent potential* $U : \mathsf{G} \to \mathbb{R}$*. There exists a force* $u : [t_0, t_f] \to \mathfrak{g}^*$ *acting on the system. Then, the*

*corresponding forced Euler-Lagrange equations are given by*

$$\frac{d}{dt}\mathbf{J}(\xi) - \mathrm{ad}^*_\xi \cdot \mathbf{J}(\xi) - M(g) = u, \tag{4.8}$$

$$\dot{g} = g\xi, \tag{4.9}$$

*where the force due to the potential* $M : \mathsf{G} \to \mathfrak{g}^*$ *is given by* $M(g) = -\mathsf{T}^*_e\mathsf{L}_g \cdot \mathbf{D}_g U(g)$.

### 4.1.2 Optimal Control Problem Formulation

We consider an optimal control problem for forced Euler-Lagrange systems described by Corollary 4.1. The optimal control problem is to find the control input that minimizes the following cost functional:

$$\mathcal{J} = \int_{t_0}^{t_f} \phi(g(t), \xi(t), u(t))\, dt,$$

where $\phi : \mathsf{G} \times \mathfrak{g} \times \mathfrak{g}^* \to \mathbb{R}$ is given. Several types of boundary conditions can be considered: free terminal time, fixed terminal time, free terminal conditions, fixed terminal conditions, and terminal states lying on a given surface. Additionally, equality constraints or inequality constraints may be imposed on the trajectory of the mechanical system and the control input. For simplicity, here we consider optimal control problems with a free terminal time and fixed terminal conditions. The optimal control problem with a fixed final time and fixed terminal conditions is considered as a special case. The subsequent development can be easily extended to other optimal control problems.

The optimal control problem is summarized as follows.

$$\text{For given } t_0, g(t_0), \xi(t_0), g^f, \xi^f$$

$$\min_{u(t), t_f} \left\{ \mathcal{J} = \int_{t_0}^{t_f} \phi(g(t), \xi(t), u(t))\, dt \right\},$$

$$\text{such that } g(t_f) = g^f,\ \xi(t_f) = \xi^f,$$

$$\text{subject to the Euler-Lagrange equations (4.8), (4.9).}$$

### 4.1.3 Necessary Conditions for Optimality

We derive necessary conditions for optimality using the calculus of variations: the Euler-Lagrange equations are constrained by using Lagrange multipliers, and the variation of the corresponding augmented cost functional is set to zero. The resulting necessary conditions are expressed as a two point boundary value problem.

More explicitly, define the augmented cost functional:

$$\mathcal{J}_a = \int_{t_0}^{t_f} \phi(g, \xi, u) + \left\langle \frac{d}{dt}\mathbf{J}(\xi) - \mathrm{ad}^*_\xi \mathbf{J}(\xi) - M(g) - u,\ \lambda^1 \right\rangle + \left\langle \lambda^2,\ g^{-1}\dot{g} - \xi \right\rangle\, dt, \tag{4.10}$$

where $\lambda^1 \in \mathfrak{g}$, $\lambda^2 \in \mathfrak{g}^*$ are Lagrange multipliers. The variation of the augmented cost functional

is composed of two terms: the term due to the the terminal time variation and the term due to the variation of trajectories for a fixed terminal time:

$$\delta \mathcal{J}_a = \delta \mathcal{J}_a^1 + \delta \mathcal{J}_a^2.$$

The variation of the augmented cost functional due to the terminal time change is given by

$$\delta \mathcal{J}_a^1 = \left[ \phi(g, \xi, u) + \left\langle \frac{d}{dt}\mathbf{J}(\xi) - \mathrm{ad}_\xi^* \mathbf{J}(\xi) - M(g) - u,\ \lambda^1 \right\rangle + \left\langle \lambda^2,\ g^{-1}\dot{g} - \xi \right\rangle \right] \Bigg|_{t=t_f} \delta t_f. \tag{4.11}$$

The variation of the augmented cost functional, for a fixed terminal time, is given by

$$\begin{aligned} \delta \mathcal{J}_a^2 = &\int_{t_0}^{t_f} \mathbf{D}_g \phi(g, \xi, u) \cdot \delta g + \mathbf{D}_\xi \phi(g, \xi, u) \cdot \delta \xi + \mathbf{D}_u \phi(g, \xi, u) \cdot \delta u \, dt \\ &+ \int_{t_0}^{t_f} \left\langle \delta \left( \frac{d}{dt}\mathbf{J}(\xi) \right) - \mathrm{ad}_{\delta\xi}^* \mathbf{J}(\xi) - \mathrm{ad}_\xi^* \mathbf{J}(\delta\xi) - \mathsf{T}_g M(g) \cdot \delta g - \delta u,\ \lambda^1 \right\rangle \\ &\qquad + \left\langle \lambda^2,\ \dot{\eta} + \mathrm{ad}_\xi \eta - \delta\xi \right\rangle dt. \end{aligned} \tag{4.12}$$

We apply several properties and identities on the variations and the $\mathrm{ad}$ operator. Since the variation and the derivative commute, we obtain $\delta\left(\frac{d}{dt}\mathbf{J}(\xi)\right) = \frac{d}{dt}\mathbf{J}(\delta\xi)$ (See Appendix A.4 for details). Therefore, using the symmetry of the inertial operator, we obtain $\left\langle \delta\left(\frac{d}{dt}\mathbf{J}(\xi)\right),\ \lambda^1 \right\rangle = \left\langle \mathbf{J}(\delta\dot{\xi}),\ \lambda^1 \right\rangle = \left\langle \mathbf{J}(\lambda^1),\ \delta\dot{\xi} \right\rangle$. From the definition of the $\mathrm{ad}^*$ operator and the skew-symmetry of the $\mathrm{ad}$ operator, $\left\langle \mathrm{ad}_{\delta\xi}^* \mathbf{J}(\xi),\ \lambda^1 \right\rangle = \left\langle \mathbf{J}(\xi),\ \mathrm{ad}_{\delta\xi}\lambda^1 \right\rangle = -\left\langle \mathbf{J}(\xi),\ \mathrm{ad}_{\lambda^1}\delta\xi \right\rangle = -\left\langle \mathrm{ad}_{\lambda^1}^* \mathbf{J}(\xi),\ \delta\xi \right\rangle$. Using the symmetry of the inertia operator, $\left\langle \mathrm{ad}_\xi^* \mathbf{J}(\delta\xi),\ \lambda^1 \right\rangle = \left\langle \mathbf{J}(\delta\xi),\ \mathrm{ad}_\xi \lambda^1 \right\rangle = \left\langle \mathbf{J}(\mathrm{ad}_\xi \lambda^1),\ \delta\xi \right\rangle$. Define $\mathcal{M}(g, \lambda^1) : \mathsf{G} \times \mathfrak{g} \to \mathfrak{g}^*$ such that

$$\left\langle \mathsf{T}_g M(g) \cdot \delta g,\ \lambda^1 \right\rangle = \left\langle \mathsf{T}_g M(g) \cdot (\mathsf{T}_e \mathsf{L}_g \cdot \eta),\ \lambda^1 \right\rangle = \left\langle \mathcal{M}(g, \lambda^1),\ \eta \right\rangle. \tag{4.13}$$

Note that this map is linear with respect to $\lambda^1$. Using these properties and the definition of the map $\mathcal{M}(g, \lambda^1)$, (4.12) can be written as

$$\begin{aligned} \delta \mathcal{J}_a^2 = &\int_{t_0}^{t_f} \left\langle \mathbf{J}(\lambda^1), \delta\dot{\xi} \right\rangle + \left\langle \lambda^2,\ \dot{\eta} \right\rangle dt \\ &+ \int_{t_0}^{t_f} \left\langle \delta u,\ \mathbf{D}_u \phi - \lambda^1 \right\rangle + \left\langle \mathsf{T}_e^* \mathsf{L}_g \cdot \mathbf{D}_g \phi - \mathcal{M}(g, \lambda^1) + \mathrm{ad}_\xi^* \lambda^2,\ \eta \right\rangle \\ &\qquad + \left\langle \mathbf{D}_\xi \phi + \mathrm{ad}_{\lambda^1}^* \mathbf{J}(\xi) - \mathbf{J}(\mathrm{ad}_\xi \lambda^1) - \lambda^2,\ \delta\xi \right\rangle dt. \end{aligned}$$

Using integration by parts for the first two terms, we obtain

$$\begin{aligned}\delta\mathcal{J}_a^2 = &\left[\left\langle \mathbf{J}(\lambda^1), \delta\xi\right\rangle + \left\langle \lambda^2,\, \eta\right\rangle\right]\Big|_{t=t_f} \\ &+ \int_{t_0}^{t_f} \left\langle \delta u,\, \mathbf{D}_u\phi - \lambda^1\right\rangle + \left\langle -\dot{\lambda}^2 + \mathsf{T}_e^*\mathsf{L}_g\cdot\mathbf{D}_g\phi - \mathcal{M}(g,\lambda^1) + \mathrm{ad}_\xi^*\lambda^2,\, \eta\right\rangle \\ &\qquad + \left\langle -\mathbf{J}(\dot{\lambda}^1) + \mathbf{D}_\xi\phi + \mathrm{ad}_{\lambda^1}^*\mathbf{J}(\xi) - \mathbf{J}(\mathrm{ad}_\xi\lambda^1) - \lambda^2,\, \delta\xi\right\rangle dt, \end{aligned} \tag{4.14}$$

where we use the fact that the initial conditions are fixed, i.e. $\delta\xi(t_0) = \eta(t_0) = 0$. Since the desired terminal conditions are given, $\delta t_f$ and $\xi(t_f), \eta(t_f)$ are related as

$$\dot{g}(t_f)\delta t_f + g(t_f)\eta(t_f) = 0, \tag{4.15}$$

$$\dot{\xi}(t_f)\delta t_f + \delta\xi(t_f) = 0. \tag{4.16}$$

Substituting these into the first two terms of (4.14), we obtain

$$\begin{aligned}\delta\mathcal{J}_a^2 = &-\left[\left\langle \mathbf{J}(\lambda^1), \dot{\xi}\right\rangle + \left\langle \lambda^2,\, g^{-1}\dot{g}\right\rangle\right]\Big|_{t=t_f}\delta t_f \\ &+ \int_{t_0}^{t_f} \left\langle \delta u,\, \mathbf{D}_u\phi - \lambda^1\right\rangle + \left\langle -\dot{\lambda}^2 + \mathsf{T}_e^*\mathsf{L}_g\cdot\mathbf{D}_g\phi - \mathcal{M}(g,\lambda^1) + \mathrm{ad}_\xi^*\lambda^2,\, \eta\right\rangle \\ &\qquad + \left\langle -\mathbf{J}(\dot{\lambda}^1) + \mathbf{D}_\xi\phi + \mathrm{ad}_{\lambda^1}^*\mathbf{J}(\xi) - \mathbf{J}(\mathrm{ad}_\xi\lambda^1) - \lambda^2,\, \delta\xi\right\rangle dt, \end{aligned} \tag{4.17}$$

Therefore, the variation of the augmented cost functional is given by the sum of (4.11) and (4.17) as follows.

$$\begin{aligned}\delta\mathcal{J}_a = &\left[\phi(g,\xi,u) - \left\langle \mathrm{ad}_\xi^*\mathbf{J}(\xi) + M(g) + u,\ \lambda^1\right\rangle - \left\langle \lambda^2,\ \xi\right\rangle\right]\Big|_{t=t_f}\delta t_f \\ &+ \int_{t_0}^{t_f} \left\langle \delta u,\, \mathbf{D}_u\phi - \lambda^1\right\rangle + \left\langle -\dot{\lambda}^2 + \mathsf{T}_e^*\mathsf{L}_g\cdot\mathbf{D}_g\phi - \mathcal{M}(g,\lambda^1) + \mathrm{ad}_\xi^*\lambda^2,\, \eta\right\rangle \\ &\qquad + \left\langle -\mathbf{J}(\dot{\lambda}^1) + \mathbf{D}_\xi\phi + \mathrm{ad}_{\lambda^1}^*\mathbf{J}(\xi) - \mathbf{J}(\mathrm{ad}_\xi\lambda^1) - \lambda^2,\, \delta\xi\right\rangle dt. \end{aligned} \tag{4.18}$$

This is zero for any variation about the optimal trajectory. Thus, we obtain necessary conditions for optimality as follows.

**Proposition 4.2** *Consider a forced Euler-Lagrange system evolving on a Lie group, whose Lagrangian is given by (4.7). Necessary conditions for the optimal control problem presented in Section 4.1.2 are as follows.*

- *Optimality condition*

$$\mathbf{D}_u\phi(g,\xi,u) - \lambda^1 = 0, \tag{4.19}$$

- *Multiplier equations*

$$\mathbf{J}(\dot{\lambda}^1) - \mathrm{ad}^*_{\lambda^1}\mathbf{J}(\xi) + \mathbf{J}(\mathrm{ad}_\xi \lambda^1) + \lambda^2 - \mathbf{D}_\xi \phi(g, \xi, u) = 0, \tag{4.20}$$

$$\dot{\lambda}^2 - \mathrm{ad}^*_\xi \lambda^2 - \mathsf{T}^*_e \mathsf{L}_g \cdot \mathbf{D}_g \phi(g, \xi, u) + \mathcal{M}(g, \lambda^1) = 0, \tag{4.21}$$

  *where* $\mathcal{M} : \mathsf{G} \times \mathfrak{g} \to \mathfrak{g}^*$ *is given by (4.13).*

- *Boundary conditions*

$$\left[\phi(g, \xi, u) - \left\langle \mathrm{ad}^*_\xi \mathbf{J}(\xi) + M(g) + u,\ \lambda^1 \right\rangle - \left\langle \lambda^2,\ \xi \right\rangle\right] \Big|_{t=t_f} = 0, \tag{4.22}$$

$$g(t_f) = g^f, \quad \xi(t_f) = \xi^f. \tag{4.23}$$

**Remark 4.1** These necessary conditions are expressed in a coordinate-free form. Therefore, the optimality conditions and the resulting optimal control are independent of the choice of any coordinates. They are globally represented by group elements without any singularity and ambiguity, and the optimality conditions are more compact than expressions written in terms of local parameterizations.

**Remark 4.2** The necessary conditions have the form of a two point boundary value problem: find the control input $u$, state $(g, \xi)$, multipliers $(\lambda^1, \lambda^2)$, and terminal time $t_f$ that satisfy the optimality condition (4.19), the Euler-Lagrange equations (4.2), (4.3), the multiplier equations (4.20), (4.21), and the given boundary conditions. A computational approach to solve this two point boundary value problem is discussed in Chapter 5.

**Remark 4.3** The presented necessary conditions are for optimal control problems with fixed terminal conditions and a free terminal time. They can be extended to other optimal control problems. For example, if the terminal time is fixed, then the necessary conditions for the corresponding optimal control problem are the same as presented in Proposition 4.2 without (4.22). If there is a constraint on the control input, for example, the control input lies in a given submanifold $\mathcal{U} \subset \mathfrak{g}^*$, then only the optimality condition given by (4.19) is changed to

$$\int_{t_0}^{t_f} \left\langle \delta u,\ \mathbf{D}_u \phi(g, \xi, u) - \lambda^1 \right\rangle \, dt \geq 0 \tag{4.24}$$

for all admissible variations of the control input $\delta u$ and for all $t \in [t_0, t_f]$.

## 4.2 Examples of Optimal Control of Rigid Bodies

Based on the optimal control problem formulation and the necessary conditions developed in the previous section, we study the following optimal control problems for rigid body dynamics on a Lie group. We formulate the specific optimal control problems and we develop necessary conditions for optimality. Computational approaches for solving the corresponding two point boundary value problems are discussed in Section 5.2.

| Section | Optimal Control Problem | G |
|---|---|---|
| 4.2.1 | Fuel optimal attitude control of a spacecraft on a circular orbit | $\mathsf{SO(3)}$ |
| 4.2.2 | Time optimal attitude control of a free rigid body | $\mathsf{SO(3)}$ |
| 4.2.3 | Fuel optimal attitude control of a 3D pendulum with symmetry | $\mathsf{SO(3)}$ |
| 4.2.4 | Fuel optimal control of a rigid body | $\mathsf{SE(3)}$ |

The fuel optimal attitude control of a spacecraft in Section 4.2.1, and the fuel optimal control of a rigid body in Section 4.2.4 are direct applications of Proposition 4.2 on the Lie groups $\mathsf{SO(3)}$ and $\mathsf{SE(3)}$, respectively. In the time optimal attitude control problem presented in Section 4.2.2, we consider a bounded control input, where the optimality condition is as discussed in Remark 4.3. In Section 4.2.3, we consider an optimal control problem for an underactuated control input that guarantees satisfaction of a symmetry property.

### 4.2.1 Fuel Optimal Attitude Control of a Spacecraft on a Circular Orbit

Consider the attitude dynamics of a rigid spacecraft on a circular orbit about a massive central body, including gravity gradient effects (see Hughes 1986; Wie 1998). The configuration manifold is $\mathsf{SO(3)}$. We study a minimum fuel optimal control problem that rotates the spacecraft to a desired terminal attitude and angular velocity during a fixed maneuver time.

In this section, forced Euler-Lagrange equations are derived according to Corollary 4.1, and a mathematical formulation of the optimal control problem is presented. Necessary conditions for optimality are developed from Proposition 4.2.

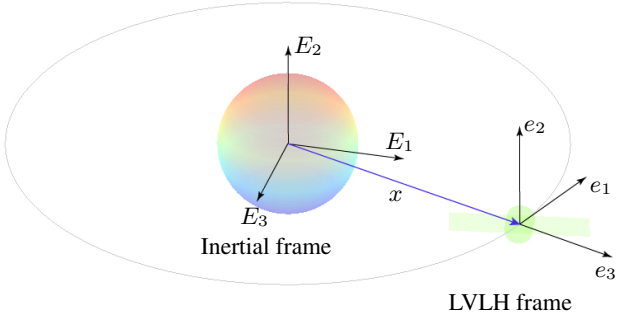



Figure 4.1: Spacecraft on a circular orbit

**Forced Euler-Lagrange Equations**

We assume that the rigid spacecraft is on a circular orbit with a constant orbital angular velocity $\omega_0 \in \mathbb{R}$. We three frames; a reference frame, a body fixed frame, and a local vertical and local horizontal (LVLH) frame. The first axis of the LVLH frame is tangential to the orbit, and the second axis is perpendicular to the orbital plane. The spacecraft model with the LVLH frame is shown in Figure 4.1.

Since the orbital angular velocity is fixed, the linear transformation from the LVLH frame to the reference frame is given by $\exp(\omega_0 t\hat{e}_2)$. Let $R \in \mathsf{SO}(3)$ be the linear transformation from the reference frame to the body fixed frame. Therefore, the attitude of the spacecraft with respect to the LVLH frame is given by $\exp(-\omega_0 t\hat{e}_2)R$. The gravitational potential can be written as

$$U(R) = -\frac{GM}{r_0} - \frac{1}{2}\omega_0^2(\text{tr}[J] - 3e_3^T \exp(-\omega_0 t\hat{e}_2)RJR^T \exp(\omega_0 t\hat{e}_2)e_3), \tag{4.25}$$

where constants $G$ and $M$ are the gravitational constant and the mass of the central body, respectively, and the orbital radius is given by $r_0 = (GM/\omega_0^2)^{1/3}$ (see Wie 1998). The second term describes the gravity gradient potential that arises from the gravity variation over the spacecraft.

Let $\Omega \in \mathbb{R}^3$ be the angular velocity of the spacecraft with respect to the reference frame, represented in the body fixed frame. The Lagrangian of the spacecraft on a circular orbit can be written as (4.7), with the inertia operator $\mathbf{J} : \mathfrak{so}(3) \to \mathfrak{so}(3)^*$ given by

$$\left\langle \mathbf{J}(\hat{\Omega}_1), \hat{\Omega}_2 \right\rangle = \left\langle \widehat{J\Omega_1}, \hat{\Omega}_2 \right\rangle = (J\Omega_1)^T\Omega_2 \tag{4.26}$$

for $\Omega_1, \Omega_2 \in \mathbb{R}^3$.

Now we find the expression for the generalized force $u \in \mathfrak{so}(3)^*$. We assume that an external control moment $\tau \in \mathbb{R}^3$ is applied to the rigid body. This represents the control moment expressed in the body fixed frame. Let $\rho \in \mathbb{R}^3$ be the vector from the mass center to a mass element of the rigid body. Then, the virtual displacement of the element due to the variation of $R$ is given by $\delta R\rho = R\hat{\eta}\rho$ in the reference frame. Thus, the virtual displacement of the mass element is represented by $\hat{\eta}\rho$ in the body fixed frame. Let $dF(\rho)$ be the force acting on the mass element expressed in the body fixed frame. Note that $\int_{\mathcal{B}} dF(\rho) = 0$ as the inter-particle forces cancel out, and $\int_{\mathcal{B}} \rho \times dF(\rho) = \tau$ as the external moment $\tau$ is applied to the rigid body. Then, the virtual work done by the control moment on the rigid body is given by

$$\int_{\mathcal{B}} \langle dF(\rho), \hat{\eta}\rho\rangle = \int_{\mathcal{B}} \eta \cdot (\rho \times dF(\rho)) = \eta \cdot \tau.$$

Therefore, the generalized force acting on the rigid body is equal to the external control moment applied to the rigid body, i.e. $u = \tau \in \mathfrak{so}(3)^* \simeq (\mathbb{R}^3)^*$. From now on, we denote the external moment by $u$ for simplicity.

Recall that the $\text{ad}$ operator on $\mathsf{SO}(3)$ is given by $\text{ad}_{\Omega'}\Omega = \Omega' \times \Omega$ $\text{ad}^*_{\Omega'}\Omega = -\Omega' \times \Omega$ for $\Omega, \Omega' \in \mathfrak{so}(3) \simeq \mathbb{R}^3$. According to Corollary 4.1, the forced Euler-Lagrange equations for the

attitude dynamics of the spacecraft on a circular orbit are given by

$$J\dot{\Omega} + \Omega \times J\Omega = M + u, \tag{4.27}$$

$$\dot{R} = R\hat{\Omega}, \tag{4.28}$$

$$M = 3\omega_0^2(R^T \exp(-\omega_0 t\hat{e}_2)e_3) \times (JR^T \exp(-\omega_0 t\hat{e}_2)e_3). \tag{4.29}$$

**Optimal Control Problem**

The objective of the optimal control problem is to rotate the spacecraft from the initial attitude and angular velocity $(R(t_0), \Omega(t_0))$ to the desired terminal attitude and angular velocity $(R^f, \Omega^f)$ for a fixed terminal time $t_f$, while minimizing the control effort.

$$\begin{gathered}
\text{For given: } t_0, (R(t_0), \Omega(t_0)), t_f, (R^f, \Omega^f) \\
\min_u \left\{ \mathcal{J} = \int_{t_0}^{t_f} \frac{1}{2} u^T u \, dt \right\}, \\
\text{such that } R(t_f) = R^f,\ \Omega(t_f) = \Omega^f, \\
\text{subject to (4.27), (4.28), (4.29).}
\end{gathered}$$

**Necessary Conditions for Optimality**

This is a special case of the general optimal control problem introduced in Section 4.1.2 with $\phi(R, \Omega, u) = \frac{1}{2} u^T u$. Therefore,

$$\mathbf{D}_R \phi = 0, \quad \mathbf{D}_\Omega \phi = 0, \quad \mathbf{D}_u \phi = u. \tag{4.30}$$

We find the expression for $\mathcal{M}(R, \lambda^1) : \mathsf{SO}(3) \times \mathfrak{so}(3) \to \mathfrak{so}(3)^*$. The derivative of the gravitational moment is given by

$$\begin{aligned}
\mathsf{T}_g M(R) \cdot R\hat{\eta} &= -3\omega_0^2(JR^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \, (-\hat{\eta} R^T \exp(\omega_0 t\hat{e}_2)e_3) \\
&\quad - 3\omega_0^2(R^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \, (J\hat{\eta} R^T \exp(\omega_0 t\hat{e}_2)e_3) \\
&= -3\omega_0^2(JR^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \, (R^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \, \eta \\
&\quad + 3\omega_0^2(R^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \, J \, (R^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \, \eta.
\end{aligned}$$

According to (4.13), the expression for $\mathcal{M}(R, \lambda^1)$ is obtained as follows.

$$\begin{aligned}
\mathcal{M}(R, \lambda^1) = 3\omega^2 \Big[ &- (JR^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \, (R^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \\
&+ (R^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \, J \, (R^T \exp(\omega_0 t\hat{e}_2)e_3)^\wedge \Big]^T \lambda^1.
\end{aligned} \tag{4.31}$$

Substituting (4.26), (4.30), and (4.31) into (4.19)–(4.23), we obtain the following necessary conditions for optimality according to Proposition 4.2:

- Optimality condition

$$u = \lambda^1. \tag{4.32}$$

- Multiplier equations

$$J\dot{\lambda}^1 + \lambda^1 \times J\Omega + J(\Omega \times \lambda^1) + \lambda^2 = 0, \tag{4.33}$$

$$\dot{\lambda}^2 + \Omega \times \lambda^2 + \mathcal{M}(R, \lambda^1) = 0. \tag{4.34}$$

- Boundary conditions

$$R(t_f) = R^f, \quad \Omega(t_f) = \Omega^f, \tag{4.35}$$

where the expression for $\mathcal{M}(R, \lambda^1)$ is given by (4.31).

This is an intrinsic form of necessary conditions for an optimal attitude control problem on $\mathsf{SO}(3)$. Using the fact that $\mathfrak{so}(3)$ is isomorphic to $\mathbb{R}^3$, the multiplier equations are expressed as compact vector equations on $\mathbb{R}^6$ in (4.33), (4.34). If the spacecraft is inertially symmetric, i.e. $J = I_{3\times 3}$, the multiplier equation (4.33) reduces to $\dot{\lambda}^1 + \lambda^2 = 0$. These are much more compact than necessary conditions expressed in terms of Euler angles or quaternions, and there is no singularity in representing the attitude of the spacecraft. Therefore, these can be used to study large angle spacecraft rotational maneuvers. A computational approach to solve these necessary conditions to obtain optimal attitude maneuvers of spacecraft is presented in Section 5.2.1.

### 4.2.2 Time Optimal Attitude Control of a Free Rigid Body

We study a time optimal control problem for the attitude dynamics of a free rigid body, that is a rigid body acted on by a control moment but no other external moment. This rigid body model corresponds to the spacecraft on a circular orbit presented in the previous section, where the gravity variation over the spacecraft body is ignored. Time optimal attitude maneuvers have been extensively studied in the literature (see, for example, a survey paper by Scrivener and Thompson 1994, and references therein). For example, the time optimal control of spacecraft has received consistent interest as rapid attitude maneuvers are critical to various space missions such as military surveillance and satellite communications.

The time optimal maneuver for a single degree of freedom rigid body model, where the attitude maneuver is constrained to an eigen-axis rotation, is derived in Etter (1989). Later, it is shown that the eigen-axis rotation is not generally time optimal by Bilimoria and Wie (1993); Seywald and Kumar (1993). The attitude dynamics are often simplified in optimality analyses, e.g., by assuming an inertially symmetric rigid body model (see Bilimoria and Wie 1993; Modgalya and Bhat 2006; Seywald and Kumar 1993), linearization of the dynamics (see Byers and Vadali 1993) and constant magnitude angular velocity (see Modgalya and Bhat 2006). Here, we present necessary conditions for the time optimal attitude control problem on $\mathsf{SO}(3)$, without any simplifying assumptions on the

rigid body (see Lee et al. 2008a).

In this section, forced Euler-Lagrange equations are derived according to Corollary 4.1, and a mathematical formulation of the optimal control problem is presented. Necessary conditions for optimality are developed from Proposition 4.2; as discussed in Remark 4.3, the optimality condition includes the effect of bounded control inputs.

### Forced Euler-Lagrange Equations

Define two frame; a reference frame and a body fixed frame. Let $R \in \mathsf{SO}(3)$ be the linear transformation from the body fixed frame to the reference frame, and $\Omega \in \mathbb{R}^3$ be the angular velocity of the rigid body represented in the body fixed frame.

In this optimal control problem, we assume that there is no configuration dependent potential, i.e. $U(R) \equiv 0$. According to Corollary 4.1, the forced Euler-Lagrange equations for the attitude dynamics of the free rigid body are given by

$$J\dot{\Omega} + \Omega \times J\Omega = u, \tag{4.36}$$

$$\dot{R} = R\hat{\Omega}. \tag{4.37}$$

### Optimal Control Problem

We assume that the magnitude of the control moment is uniformly bounded by a constant $\overline{u} \in \mathbb{R}$, i.e. $\|u(t)\|_2 \leq \overline{u}$ for any $t \in [t_0, t_f]$. The objective of the time optimal attitude control problem is to transfer the rigid body with a given initial attitude and angular velocity $(R(t_0), \Omega(t_0))$ to the desired terminal condition $(R^f, \Omega^f)$, with the constrained control moment in a minimal maneuver time $t_f$.

$$\text{For given: } t_0, (R(t_0), \Omega(t_0)), (R^f, \Omega^f), \bar{u}$$

$$\min_{u, t_f} \left\{ \mathcal{J} = \int_{t_0}^{t_f} 1 \, dt \right\},$$

$$\text{such that } R(t_f) = R^f,\ \Omega(t_f) = \Omega^f,$$

$$\text{subject to } \|u(t)\| \leq \bar{u} \ \ \forall t \in [t_0, t_f] \text{ and } (4.36), (4.37).$$

### Necessary Conditions for Optimality

This is a special case of the general optimal control problem introduced in Section 4.1.2 with $\phi(R, \Omega, u) = 1$. Therefore, the derivatives of $\phi$ are zero. As discussed in Remark 4.3, the optimality condition is given by

$$\int_{t_0}^{t_f} \left\langle \delta u, \, -\lambda^1 \right\rangle \, dt \geq 0$$

for $\lambda^1 \in \mathfrak{so}(3) \simeq \mathbb{R}^3$ and all admissible variation $\delta u$. For the given constraint, the optimal control input is

$$u = \overline{u}\frac{\lambda^1}{\|\lambda^1\|}.$$

According to Proposition 4.2, the necessary conditions for optimality are given by

- Optimality condition

$$u = \overline{u}\frac{\lambda^1}{\|\lambda^1\|}. \tag{4.38}$$

- Multiplier equations

$$J\dot{\lambda}^1 + \lambda^1 \times J\Omega + J(\Omega \times \lambda^1) + \lambda^2 = 0, \tag{4.39}$$

$$\dot{\lambda}^2 + \Omega \times \lambda^2 = 0. \tag{4.40}$$

- Boundary conditions

$$\left[1 - \lambda^1 \cdot (-\Omega \times J\Omega + u) - \lambda^2 \cdot \Omega\right]\Big|_{t=t_f} = 0, \tag{4.41}$$

$$R(t_f) = R^f, \quad \Omega(t_f) = \Omega^f. \tag{4.42}$$

In (4.38), the optimal control input is not well-defined if $\lambda^1$ is equal to zero for a finite period of time. This is referred to as a singular arc (see Bell and Jacobson 1975). By following the approach presented in Lee et al. (2008a), it can be shown that there is no singular arc in this optimal control problem. Suppose that there is a singular arc, i.e. $\lambda^1 = \dot{\lambda}^1 = 0$ for a finite time period. Then, $\lambda^2 = 0$ from (4.39), and $\dot{\lambda}^2 = 0$ from (4.40). Due to the linear structure of the multiplier equations, it follows that $\lambda^1 = \lambda^2 = 0$ for $t \in [t_0, t_f]$. Then, it is clear that the boundary condition (4.41) cannot be satisfied. Therefore, there is no singular arc along the optimal solution.

Here we do not impose any simplifying assumptions on the rigid body, such as appear in the current literature (see, for example, Bilimoria and Wie 1993; Byers and Vadali 1993; Modgalya and Bhat 2006; Seywald and Kumar 1993). But the presented necessary conditions are compact, and they have no singularity. A computational approach to solve these necessary conditions to obtain time optimal attitude maneuvers is presented in Section 5.2.3.

### 4.2.3 Fuel Optimal Attitude Control of a 3D Pendulum with Symmetry

Consider the 3D pendulum model presented in Section 2.3.2. The 3D pendulum is a rigid body supported by a frictionless pivot point acting under uniform gravitational potential. We have shown that the 3D pendulum has a symmetry represented by a group action of $\mathsf{SO}(2) \simeq \mathsf{S}^1$, and consequently, the angular momentum about the gravity direction is conserved, and the configuration manifold $\mathsf{SO}(3)$ can be reduced to a quotient space $\mathsf{SO}(3)/\mathsf{S}^1 \simeq \mathsf{S}^2$.

We study an optimal attitude control of the 3D pendulum (see Lee et al. 2007f). The external control moment does not have any component about the gravity direction, and therefore, the angular momentum 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.

In particular, we consider the case where the angular momentum about the gravity direction is zero, and we choose the desired maneuver as a rest-to-rest rotation about the gravity direction. This is interesting since the under-actuated control moment cannot generate any direct effect on the rotation about the gravity direction. The desired maneuver depends on the geometric phase effect discussed in Appendix A.3.3.

In this section, forced Euler-Lagrange equations are derived according to Corollary 4.1, and a mathematical formulation of the optimal control problem is presented. Necessary conditions for optimality are developed. Since the control input has a special structure, the necessary conditions presented in Proposition 4.2 are appropriately modified.

**Forced Euler-Lagrange Equations**

We assume that the external control moment is expressed in the body fixed frame as

$$u = R^T e_3 \times u_p$$

for a control parameter $u_p \in \mathbb{R}^3$. Since the vector $R^T e_3$ represents the gravity direction in the body fixed frame, the external control moment has no component along the gravity direction. Therefore, the angular momentum about the gravity direction is preserved in the controlled dynamics.

According to Corollary 4.1, the forced Euler-Lagrange equations are given by

$$J\dot{\Omega} + \Omega \times J\Omega = mg\rho_c \times R^T e_3 + R^T e_3 \times u_p, \tag{4.43}$$

$$\dot{R} = R\hat{\Omega}. \tag{4.44}$$

**Optimal Control Problem**

The objective of the optimal control problem is to rotate the 3D pendulum from the initial attitude $R(t_0) = I$ to the terminal attitude $R(t_f) = \exp(\theta\hat{e}_3)$ for a fixed terminal time $t_f$ and a rotation angle $\theta \in \mathsf{S}^1$, while minimizing the control effort. The initial angular velocity and the terminal angular velocity are set to zero.

$$\text{For given: } t_0, (R(t_0) = I, \Omega(t_0) = 0), t_f, \theta$$

$$\min_u \left\{ \mathcal{J} = \int_{t_0}^{t_f} u_p^T u_p \, dt \right\},$$

$$\text{such that } R(t_f) = \exp(\theta\hat{e}_3),\ \Omega(t_f) = 0,$$

$$\text{subject to (4.43), (4.44).}$$

**Necessary Conditions for Optimality**

This optimal control problem is not a special case of the general optimal control problem presented in Section 4.1.2, where the external control inputs are dependent on time. Here, the generalized force depends on the rotation matrix as well as the control parameter $u_p$. Therefore, the variations of the generalized force include the effect of the variation of the rotation matrix. But, necessary condition for optimality can be developed by following an approach that is similar to the one discussed in Section 4.1.3.

Define the augmented cost functional:

$$\mathcal{J}_a = \int_{t_0}^{t_f} \frac{1}{2} u_p^T u_p + \left\langle \frac{d}{dt}\mathbf{J}(\Omega) + \Omega \times J\Omega - M(R) - R^T e_3 \times u_p,\ \lambda^1 \right\rangle + \left\langle \lambda^2,\ (R^T\dot{R})^\vee - \Omega \right\rangle dt, \tag{4.45}$$

where the gravitational moment of the 3D pendulum is given by $M(R) = mg\rho_c \times R^T e_3$. This is equal to (4.10), except that two terms $\phi$ and $u$ in (4.10) are replaced by $\frac{1}{2}u_p^T u_p$ and $R^T e_3 \times u_p$ in (4.45). Therefore, from (4.14), we obtain the variation of the augmented cost functional as follows.

$$\begin{aligned}\delta\mathcal{J}_a = \int_{t_0}^{t_f} \langle \delta u_p,\ u_p \rangle + \left\langle -\dot{\lambda}^2 - \mathcal{M}(R,\lambda^1) - \Omega \times \lambda^2,\ \eta \right\rangle \\ + \left\langle -J\dot{\lambda}^1 - \lambda^1 \times J\Omega - J(\Omega \times \lambda^1) - \lambda^2,\ \delta\xi \right\rangle - \left\langle \delta(R^T e_3 \times u_p),\ \lambda^1 \right\rangle dt.\end{aligned} \tag{4.46}$$

We find the expression for $\mathcal{M}(R,\lambda^1)$. From (4.13), we obtain

$$\left\langle -mg\hat{\rho}_c\hat{\eta}R^T e_3, \lambda^1 \right\rangle = \left\langle mg\hat{\rho}_c\widehat{R^T e_3}\eta, \lambda^1 \right\rangle = \left\langle mg\widehat{R^T e_3}\hat{\rho}_c\lambda^1, \eta \right\rangle = \left\langle \mathcal{M}(R,\lambda^1),\ \eta \right\rangle. \tag{4.47}$$

Thus, $\mathcal{M} : \mathsf{SO}(3) \times \mathfrak{so}(3) \to \mathfrak{so}(3)^*$ is given by $\mathcal{M}(R,\lambda^1) = mg\widehat{R^T e_3}\hat{\rho}_c\lambda^1$. Similarly, the last term of (4.46) is given by

$$\begin{aligned}\left\langle \delta(R^T e_3 \times u_p),\ \lambda^1 \right\rangle &= \left\langle -(\hat{\eta}R^T e_3) \times u_p + R^T e_3 \times \delta u_p,\ \lambda^1 \right\rangle \\ &= \left\langle -\hat{u}_p\widehat{R^T e_3}\,\eta + R^T e_3 \times \delta u_p,\ \lambda^1 \right\rangle \\ &= \left\langle -\widehat{R^T e_3}\hat{u}_p\lambda^1,\ \eta \right\rangle + \left\langle \delta u_p,\ \lambda^1 \times R^T e_3 \right\rangle.\end{aligned} \tag{4.48}$$

Substituting (4.47) and (4.48) into (4.46), we obtain necessary conditions for optimality as follows.

- Optimality condition

$$u_p = -R^T e_3 \times \lambda^1. \tag{4.49}$$

- Multiplier equations

$$J\dot{\lambda}^1 + \lambda^1 \times J\Omega + J(\Omega \times \lambda^1) + \lambda^2 = 0, \tag{4.50}$$

$$\dot{\lambda}^2 + \Omega \times \lambda^2 + mg\widehat{R^T e_3}\hat{\rho}_c\lambda^1 - \widehat{R^T e_3}\hat{u}_p\lambda^1 = 0. \tag{4.51}$$

- Boundary conditions

$$R(t_f) = \exp(\theta\hat{e}_3), \quad \Omega(t_f) = 0. \tag{4.52}$$

It is interesting to observe that the quantity $e_3^T R\lambda^2$ is preserved along the optimal solution since

$$\frac{d}{dt}e_3^T R\lambda^2 = e_3^T R\hat{\Omega}\lambda^2 + e_3^T R\dot{\lambda}^2 = e_3^T R\hat{\Omega}\lambda^2 - e_3^T R(\hat{\Omega}\lambda^2) = 0.$$

Together with the conservation of the angular momentum about the gravity direction, this yields numerical ill-conditioning in the optimal control problem. A simple computational approach to overcome this difficulty is presented in Lee et al. (2007f), and it is discussed in Section 5.2.3.

### 4.2.4 Fuel Optimal Control of a Rigid Body

Consider a general maneuver of a rigid body, acting under a potential that is dependent on the attitude and the position of the body, presented in Section 2.3.5. We consider an optimal control problem to transfer the rigid body to a desired position and attitude using minimal effort during a fixed maneuver time. In many optimal control problems, the optimal attitude control problem and the optimal position control problem are independent and can be solved separately. Here, we explicitly consider coupling effects of the rotational and the translational dynamics of the rigid body.

In this section, forced Euler-Lagrange equations are derived according to Corollary 4.1, and a mathematical formulation of the optimal control problem is presented. Necessary conditions for optimality are developed from Proposition 4.2.

#### Forced Euler-Lagrange Equations

The configuration manifold of the general motion of a rigid body in $\mathbb{R}^3$ is the special Euclidean group $\mathsf{SE(3)}$. The configuration of the rigid body is represented by $g = (R, x) \in \mathsf{SE(3)}$ for a rotation matrix $R \in \mathsf{SO(3)}$ representing the attitude and a vector $x$ representing the location of the mass center in a reference frame. The Lagrangian is given by (2.95) for a configuration dependent potential $U(R, x) : \mathsf{SE(3)} \to \mathbb{R}$. Recall that the inertia operator $\mathbf{J} : \mathfrak{se}(3) \to \mathfrak{se}(3)^*$ and the $\mathrm{ad}$ operator on $\mathsf{SE(3)}$ are given by

$$\mathbf{J}((\Omega, V)) = \begin{bmatrix} \widehat{J\Omega} & mV \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} J_d\hat{\Omega} + \hat{\Omega}J_d & mV \\ 0 & 0 \end{bmatrix}, \tag{4.53}$$

$$\mathrm{ad}_{(\Omega,V)} = \begin{bmatrix} \hat{\Omega} & 0 \\ \hat{V} & \hat{\Omega} \end{bmatrix}, \quad \mathrm{ad}^*_{(\Omega,V)} = \begin{bmatrix} -\hat{\Omega} & -\hat{V} \\ 0 & -\hat{\Omega} \end{bmatrix}. \tag{4.54}$$

We assume that the external control input is given by $u = (u^f, u^m) \in \mathfrak{se}(3)^*$, where the control force $u^f \in (\mathbb{R}^3)^*$ and the control moment $u^m \in (\mathbb{R}^3)^*$, are expressed in the body fixed frame. According to Corollary 4.1, the forced Euler-Lagrange equations are given by

$$J\dot{\Omega} + \Omega \times J\Omega = M + u^f, \tag{4.55}$$

$$m\dot{V} + m\Omega \times V = -R^T \frac{\partial U}{\partial x} + u^m, \tag{4.56}$$

$$\dot{R} = R\hat{\Omega}, \tag{4.57}$$

$$\dot{x} = RV. \tag{4.58}$$

The force due to the potential, in the body fixed frame, is $-R^T \frac{\partial U}{\partial x}$, and the moment due to potential is determined by $\hat{M} = \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R}$ as shown in (2.101).

**Optimal Control Problem**

The objective of the optimal control problem is to transfer the rigid body with a given initial condition $(R(t_0), x(t_0), \Omega(t_0), V(t_0))$ to a desired terminal condition $(R^f, x^f, \Omega^f, V^f)$ at a fixed terminal time $t_f$, while using minimal control effort.

$$\begin{gathered}
\text{For given: } t_0, (R(t_0), x(t_0), \Omega(t_0), V(t_0)), t_f, (R^f, x^f, \Omega^f, V^f) \\
\min_u \left\{ \mathcal{J} = \int_{t_0}^{t_f} \frac{1}{2} \langle u,\, W(u) \rangle \, dt \right\}, \\
\text{such that } R(t_f) = R^f,\, x(t_f) = x^f,\, \Omega(t_f) = \Omega^f,\, V(t_f) = V^f, \\
\text{subject to (4.55)–(4.58),}
\end{gathered}$$

where $W : \mathfrak{se}(3)^* \to \mathfrak{se}(3)$ is a weighting function given by $W(u) = (W^m u^m, W^f u^f)$ for symmetric positive definite matrices $W^f, W^m \in \mathbb{R}^{3\times 3}$.

**Necessary Conditions for Optimality**

This is a special case of the general optimal control problem introduced in Section 4.1.2 with $\phi(g, \xi, u) = \langle u,\, Wu \rangle$. Since the terminal time is fixed, the boundary condition (4.22) is omitted.

Let $\lambda^1 = (\lambda^1_R, \lambda^1_x) \in \mathfrak{se}(3)$ and $\lambda^2 = (\lambda^2_R, \lambda^2_x) \in \mathfrak{se}(3)^*$ be Lagrange multipliers with $\lambda^1_R, \lambda^1_x \in (\mathbb{R}^3)^*$ and $\lambda^2_R, \lambda^2_x \in \mathbb{R}^3$. According to (4.19), the optimality conditions are easily obtained. The multiplier equations can be obtained by directly substituting the expressions for the inertia operator (4.53) and $\mathrm{ad}, \mathrm{ad}^*$ operators (4.54) on $\mathsf{SE}(3)$ into the multiplier equations given by (4.20) and (4.21). The expression for $\mathcal{M}(R, x, \lambda^1_R, \lambda^1_x) : \mathsf{SE}(3) \times \mathfrak{se}(3) \to \mathfrak{se}(3)^*$ is determined by (4.13) as follows.

$$\begin{aligned}
&\left\langle \mathsf{T}_{(R,x)} \left( \left( \frac{\partial U}{\partial R}^T R - R^T \frac{\partial U}{\partial R} \right)^\vee, -R^T \frac{\partial U}{\partial x} \right) \cdot (R\hat{\eta}, \delta x),\ (\lambda^1_R, \lambda^1_x) \right\rangle \\
&\qquad = \left\langle \mathcal{M}(R, x, \lambda^1_R, \lambda^1_x),\ (\eta, \delta x) \right\rangle. \qquad (4.59)
\end{aligned}$$

In summary, the necessary conditions for optimality are given by

- Optimality condition

$$u^m = (W^m)^{-1}\lambda_R^1, \quad u^f = (W^f)^{-1}\lambda_x^1. \tag{4.60}$$

- Multiplier equations

$$\begin{bmatrix} J\dot{\lambda}_R^1 \\ m\dot{\lambda}_x^1 \end{bmatrix} - \begin{bmatrix} -\hat{\lambda}_R^1 & -\hat{\lambda}_x^1 \\ 0 & -\hat{\lambda}_R^1 \end{bmatrix} \begin{bmatrix} J\Omega \\ mV \end{bmatrix} + \begin{bmatrix} J & 0 \\ 0 & mI_{3\times 3} \end{bmatrix} \begin{bmatrix} \hat{\Omega} & 0 \\ \hat{V} & \hat{\Omega} \end{bmatrix} \begin{bmatrix} \lambda_R^1 \\ \lambda_x^1 \end{bmatrix} + \begin{bmatrix} \lambda_R^2 \\ \lambda_x^2 \end{bmatrix} = 0, \tag{4.61}$$

$$\begin{bmatrix} \dot{\lambda}_R^2 \\ \dot{\lambda}_x^2 \end{bmatrix} - \begin{bmatrix} -\hat{\Omega} & -\hat{V} \\ 0 & -\hat{\Omega} \end{bmatrix} \begin{bmatrix} \lambda_R^2 \\ \lambda_x^2 \end{bmatrix} + \mathcal{M}(R, x, \lambda_R^1, \lambda_x^1) = 0, \tag{4.62}$$

  where the expression for $\mathcal{M}(R, x, \lambda_R^1, \lambda_x^1) : \mathsf{SE}(3) \times \mathfrak{se}(3) \to \mathfrak{se}(3)^*$ is determined by (4.59).

- Boundary conditions

$$R(t_f) = R^f,\ x(t_f) = x^f,\ \Omega(t_f) = \Omega^f,\ V(t_f) = V^f. \tag{4.63}$$

These necessary conditions can be used to study optimal rigid body maneuvers with explicit consideration for the coupling effects of rotational dynamics and translational dynamics. A computational approach to solve these necessary conditions for optimal transfer maneuvers of spacecraft is presented in Section 5.2.4.

## 4.3 Conclusions

In Section 4.1, we have formulated an optimal control problem for dynamic systems that evolve on a Lie group, and we have derived necessary conditions for optimality by the calculus of variations. These results are illustrated by optimal control problems for several rigid body systems in Section 4.2.

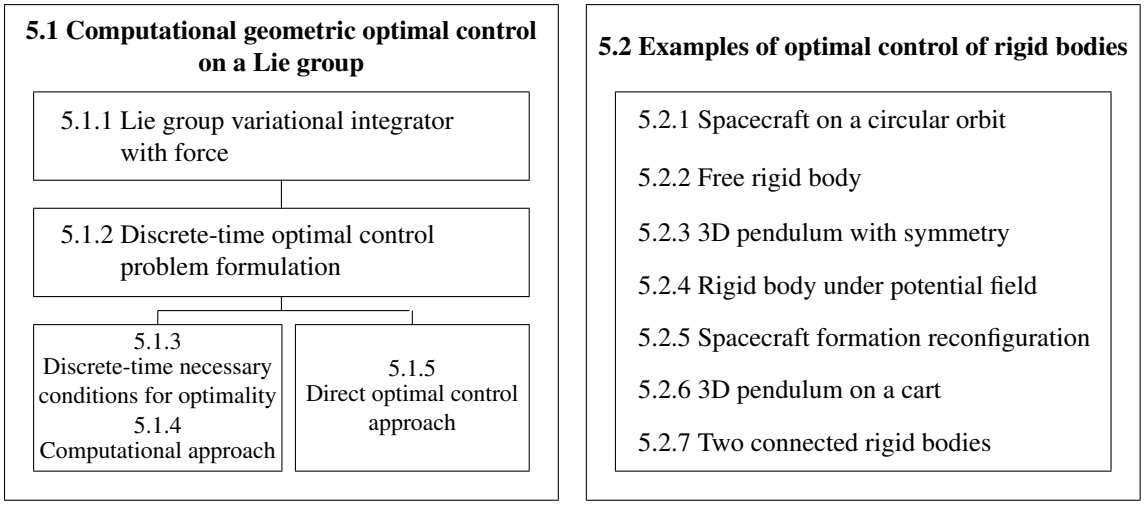
The intrinsic form of necessary conditions for optimality developed in this chapter can be applied to a wide class of optimal control problems on a Lie group. In particular, this chapter dealt with optimal control problems for dynamic systems with a Lie group configuration manifold. This is distinguished from the existing optimal control theories developed for kinematics equations on a Lie group.

As shown in Section 4.2, this approach provides a coordinate-free form of necessary conditions. Therefore, the optimality condition and the resulting optimal control input are independent of the choice of representations for rigid body configurations. They are globally represented by group elements without any singularity and ambiguity, and the optimality conditions are more compact than expressions written in terms of local parameterizations. These results can be applied to optimal control problems for complex multibody systems.

## CHAPTER 5

# COMPUTATIONAL GEOMETRIC OPTIMAL CONTROL OF RIGID BODIES ON A LIE GROUP

In the previous chapter, we developed necessary conditions for optimality for an optimal control problem for mechanical systems evolving on a Lie group. While the presented necessary conditions are expressed in a global and relatively compact form that is suitable for analytical study, the procedure to find optimal control inputs generally relies on numerical computations. This chapter discusses computational geometric approaches for finding the optimal control input numerically while preserving geometric properties of the dynamics. These approaches are based on the formulation of computational geometric mechanics discussed in Chapter 3.




This chapter is organized as follows. In Section 5.1, we develop computational geometric optimal control theory for dynamics on an arbitrary Lie group; the Lie group variational integrator derived in Chapter 3 are extended to include effect of external control inputs, and a general form of a discrete-time optimal control problem is formulated. Two optimal control approaches are presented; in Section 5.1.3 and Section 5.1.4, discrete-time necessary conditions for optimality and a computational approach to solve the corresponding two point boundary value problem are developed, and in Section 5.1.5, a direct optimal control approach based on a parameter optimization technique is presented. These approaches are applied to several optimal control problems for rigid body dynamics in Section 5.2

## 5.1 Computational Geometric Optimal Control on a Lie Group

Computational geometric optimal control is concerned with developing numerical methods for optimal control problems, that conserve geometric properties of the dynamics and of the optimality conditions. The essential idea is the same as that of computational geometric mechanics presented in Chapter 3; computational methods are developed according to a discrete-time analogue of the underlying fundamental principles. We developed computational geometric mechanics by discretizing Hamilton's principle, and the resulting Lie group variational integrators have desirable geometric properties. In computational geometric optimal control, we discretize an optimal control problem at the problem formulation stage using a structure-preserving geometric integrator, and we develop discrete-time optimality conditions using the calculus of variations. This is in contrast to the usual optimal control approaches, where discretization appears at the last stage when solving the necessary conditions numerically.

Computational geometric optimal control has substantial computational advantages. As discussed in Chapter 3, discrete-time equations derived from computational geometric mechanics are more faithful to the continuous equations of motion, and consequently a more accurate solution to the optimal control problem is obtained. Although external control inputs often break the Lagrangian and Hamiltonian system structure, the computational superiority of numerical solutions obtained from discrete-time geometric mechanics still holds for controlled systems. It has been shown that they compute the energy dissipation rate of controlled systems more accurately (see Marsden and West 2001). For example, this feature is extremely important in computing accurate optimal trajectories for long time maneuvers using low energy control inputs.

Computational geometric optimal control not only provides an accurate optimal control input, but it also enables us to find it efficiently. The optimal solutions are usually sensitive to small variations in the multipliers. This causes difficulties, such as numerical ill-conditioning, when solving the necessary conditions for optimality in the form of a two point boundary value problem. Computed sensitivity derivatives are not influenced by the numerical dissipation introduced by conventional numerical integration schemes. Therefore, they are numerically more robust, and necessary conditions can be solved in a computationally efficient way.

In summary, computational geometric optimal control is a discrete-time analogue of geometric optimal control discussed in Chapter 4. It deals with optimal control problems for dynamic systems on a Lie group, expressed in a coordinate-free form. The computational geometric optimal control approach develops efficient numerical algorithms for optimal control problems, that preserve the geometrical features. This approach is based on computation geometric mechanics presented in Chapter 3. Here, we first extend Lie group variational integrators to include the effect of control inputs, and we formulate discrete-time optimal control problems. Discrete-time necessary conditions for optimality are developed according to a discrete analogue of the calculus of variations, and a direct optimal control approach is discussed.

### 5.1.1 Lie Group Variational Integrator with Generalized Forces

Consider a mechanical system evolving on a Lie group $\mathsf{G}$. Suppose that there exists a generalized force $u(t) : [t_0, t_f] \to \mathfrak{g}^*$ acting on the system. The discrete generalized forces $u_{d_k}^+, u_{d_k}^- \in \mathfrak{g}^*$ are chosen such that they approximate the virtual work in the Lagrange-d'Alembert principle given by (4.1):

$$\int_{t_k}^{t_{k+1}} u(t) \cdot \eta \, dt \approx u_{d_k}^- \cdot \eta_k + u_{d_k}^+ \cdot \eta_{k+1}.$$

The discrete Lagrange-d'Alembert principle states that

$$\delta \sum_{k=0}^{N-1} L_d(g_k, f_k) + \sum_{k=0}^{N-1} u_{d_k}^- \cdot \eta_k + u_{d_k}^+ \cdot \eta_{k+1} = 0. \tag{5.1}$$

As the discrete Hamilton's principle given by (3.3) approximates Hamilton's principle given by (2.2), this discrete Lagrange-d'Alembert principle approximates the Lagrange-d'Alembert principle given by (4.1). This is equivalent to discrete Hamilton's principle with the additional forcing term included. From (3.9), and using the fact that $\eta_0 = \eta_N = 0$, this can be written as

$$\begin{aligned}\sum_{k=1}^{N-1} \Big\langle \mathsf{T}_e^*\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_{d_k} - \mathrm{Ad}_{f_k^{-1}}^* \cdot (\mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_{d_k}) + \mathsf{T}_e^*\mathsf{L}_{f_{k-1}} \cdot \mathbf{D}_{f_{k-1}} \mathsf{L}_{d_{k-1}}, \ \eta_k \Big\rangle \\ + \sum_{k=1}^{N-1} \Big\langle u_{d_k}^- + u_{d_{k-1}}^+, \ \eta_k \Big\rangle = 0,\end{aligned}$$

which yields the forced discrete Euler-Lagrange equations on $\mathsf{G}$.

**Proposition 5.1** *Consider a mechanical system evolving on a Lie group $\mathsf{G}$. The discrete-time kinematics equation is given by (3.11), where the group element $g_{k+1}$ is updated by the right group action of $f_k \in \mathsf{G}$ on $g_k$. Suppose that the discrete Lagrangian is defined as $L_d(g_k, f_k) : \mathsf{G} \times \mathsf{G} \to \mathbb{R}$, and there exist discrete generalized forces $u_k^-, u_k^+ \in \mathfrak{g}^*$. The corresponding forced discrete-time Euler-Lagrange equations are given by*

$$\begin{aligned}\mathsf{T}_e^*\mathsf{L}_{f_{k-1}} \cdot \mathbf{D}_{f_{k-1}} L_d(g_{k-1}, f_{k-1}) - \mathrm{Ad}_{f_k^{-1}}^* \cdot (\mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_d(g_k, f_k)) \\ + \mathsf{T}_e^*\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_d(g_k, f_k) + u_{d_k}^- + u_{d_{k-1}}^+ = 0,\end{aligned} \tag{5.2}$$

$$g_{k+1} = g_k f_k. \tag{5.3}$$

*Using the discrete Legendre transformation, we obtain the equivalent discrete-time Hamilton's*

*equations are follows.*

$$\mu_k = \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}^*_e\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_{d_k}) - \mathsf{T}^*_e\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_{d_k} - u^-_{d_k}, \tag{5.4}$$

$$g_{k+1} = g_k f_k, \tag{5.5}$$

$$\mu_{k+1} = \mathrm{Ad}^*_{f_k} \cdot (\mu_k + \mathsf{T}^*_e\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} L_{d_k} + u^-_{d_k}) + u^+_{d_k}. \tag{5.6}$$

*Special Form of Discrete-time Forced Euler-Lagrange Equations.* Proposition 5.1 provides forced discrete-time Euler-Lagrange equations for mechanical systems evolving on a Lie group $\mathsf{G}$ with the general form of the discrete Lagrangian. While it is possible to formulate an optimal control problem and derive necessary optimality conditions for these systems, we consider mechanical systems with a discrete Lagrangian as in Section 4.1. This allow us to illustrate the essential idea of the computational geometric optimal control approach more explicitly, and discrete-time necessary conditions for optimality in Section 5.1.3 have a more compact form.

We assume that the discrete Lagrangian has the following form.

$$L_d(g_k, f_k) = T_d(f_k) - (1-c)hU(g_k) - chU(g_k f_k), \tag{5.7}$$

for a discrete kinetic energy $T_d : \mathsf{G} \to \mathbb{R}$ and a configuration dependent potential $U : \mathsf{G} \to \mathbb{R}$. The constant $c \in [0, 1]$ is a free parameter. If it is chosen as $c = \frac{1}{2}$, the discrete Lagrangian represents the trapezoidal rule for the Lagrangian given by (4.7), and it has second-order accuracy. The control inputs are parameterized by their values at each discrete time step, and the discrete generalized forces are chosen as

$$u^-_{d_k} = (1-c)hu_k, \quad u^+_{d_k} = chu_{k+1}. \tag{5.8}$$

From Proposition 5.1, the corresponding forced discrete Euler-Lagrange equations are given as follows.

**Corollary 5.1** *Consider a mechanical system evolving on a Lie group $\mathsf{G}$. The discrete-time kinematics equation is defined in (3.11), where the group element $g_{k+1}$ is updated by the right group action of $f_k \in \mathsf{G}$ on $g_k$. Suppose that the discrete Lagrangian is given by (5.7), and there exist discrete generalized forces $u^-_k, u^+_k \in \mathfrak{g}^*$ given by (5.8) for a free parameter $c \in [0, 1]$. The corresponding forced discrete Euler-Lagrange equations are given by*

$$\mathsf{T}^*_e\mathsf{L}_{f_{k-1}} \cdot \mathbf{D}_{f_{k-1}} T_d(f_{k-1}) - \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}^*_e\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} T_d(f_k)) + hM(g_k) + hu_k = 0, \tag{5.9}$$

$$g_{k+1} = g_k f_k, \tag{5.10}$$

*where $M(g_k) = -\mathsf{T}^*_e\mathsf{L}_g \cdot \mathbf{D}U(g_k)$. Using the discrete Legendre transformation, we obtain the equivalent discrete-time Hamilton's equations as follows:*

$$\mu_k = \mathrm{Ad}^*_{f_k^{-1}} \cdot (\mathsf{T}^*_e \mathsf{L}_{f_k} \cdot \mathbf{D}T_d(f_k)) - (1-c)hM(g_k) - (1-c)hu_k, \tag{5.11}$$

$$\mu_{k+1} = \mathrm{Ad}^*_{f_k} \cdot (\mu_k + (1-c)hM(g_k) + (1-c)hu_k) + chM(g_{k+1}) + chu_{k+1}, \tag{5.12}$$

$$g_{k+1} = g_k f_k. \tag{5.13}$$

These equations are obtained by substituting (5.7) and (5.8) into (5.2). The detailed proof is given in Appendix A.5.

### 5.1.2 Discrete-time Optimal Control Problem Formulation

We consider an optimal control problem for the mechanical system described by Corollary 5.1. The optimal control problem is to find the control input sequence that minimizes a cost function given by

$$\mathcal{J}_d = \sum_{k=0}^{N-1} \phi_d(g_k, f_k, u_k),$$

where $\phi_d : \mathsf{G} \times \mathsf{G} \times \mathfrak{g}^* \to \mathbb{R}$ is given. Several types of optimal control problems can be considered. For simplicity, we study an optimal control problem with a fixed terminal time and fixed terminal conditions. The subsequent development can be easily extended to other types of optimal control problems.

The discrete-time optimal control problem is summarized as follows:

$$\text{For given } t_0, (g_0, \mu_0), (g^f, \mu^f), N$$

$$\min_{u_k} \left\{ \mathcal{J}_d = \sum_{k=0}^{N-1} \phi_d(g_k, f_k, u_k) \right\},$$

$$\text{such that } g_N = g^f,\ \mu_N = \mu^f,$$

subject to the discrete Euler-Lagrange equations (5.11), (5.12), (5.13).

### 5.1.3 Discrete-time Necessary Conditions for Optimality

We derive discrete-time necessary conditions for optimality. Instead of discretizing the continuous-time necessary conditions derived in Section 4.1.3, we derive optimality conditions using a discrete-time version of the calculus of variations: discrete-time Hamilton's equations are enforced by using Lagrange multipliers, and the variation of the corresponding augmented cost function is set to zero. The resulting necessary conditions are expressed as a two point boundary value problem in a discrete-time setting.

More explicitly, we define the augmented cost functional as,

$$\begin{aligned}\mathcal{J}_{a_d} = \sum_{k=0}^{N-1} \phi_d(g_k, f_k, u_k) + \left\langle \mu_k - \mathrm{Ad}^*_{f_k^{-1}} \cdot T'_d(f_k) + (1-c)hM(g_k) + (1-c)hu_k,\ \lambda^0_k \right\rangle \\ + \left\langle \mu_{k+1} - \mathrm{Ad}^*_{f_k} \cdot (\mu_k + (1-c)hM(g_k) + (1-c)hu_k) - chM(g_{k+1}) - chu_{k+1},\ \lambda^1_k \right\rangle \\ + \left\langle \lambda^2_k,\ \log(g_k^{-1} g_{k+1}) - \log f_k \right\rangle, \end{aligned} \tag{5.14}$$

where $\lambda^0_k, \lambda^1_k \in \mathfrak{g}$, $\lambda^2_k \in \mathfrak{g}^*$ are sequences of Lagrange multipliers, and $T'_d(f_k) \in \mathfrak{g}^*$ is defined as $T'_d(f_k) = \mathsf{T}^*_e \mathsf{L}_{f_k} \cdot \mathbf{D} T_d(f_k)$ for $f_k \in \mathsf{G}$. The logarithm for $f_k, g_k^{-1} g_{k+1} \in \mathsf{G}$ is well-defined since they are close to the identity element if we choose the time step size $h$ to be sufficiently small.

The variation of the augmented cost functional yields discrete-time necessary conditions for optimality. By the Lagrange multiplier theorem, $\mu_k$, $g_k$ and $f_k$ are considered as independent variables. The infinitesimal variation of $g_k$ is defined as $\delta g_k = g_k \eta_k$ for $\eta_k \in \mathfrak{g}$ in (3.5). Similarly, the infinitesimal variation of $f_k$ is defined as

$$\delta f_k = f_k \chi_k \tag{5.15}$$

for $\chi_k \in \mathfrak{g}$.

We find expressions for the variation of each term of (5.14). The variation of the first term of (5.14) is given by

$$\begin{aligned}\delta J^1_{a_d} &= \mathbf{D}_{g_k} \phi_d(g_k, f_k, u_k) \cdot \delta g_k + \mathbf{D}_{f_k} \phi_d(g_k, f_k, u_k) \cdot \delta f_k + \mathbf{D}_{u_k} \phi_d(g_k, f_k, u_k) \cdot \delta u_k \\ &= \langle \mathsf{T}^*_e \mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k} \phi_{d_k},\ \eta_k \rangle + \langle \mathsf{T}^*_e \mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} \phi_{d_k},\ \chi_k \rangle + \langle \delta u_k,\ \mathbf{D}_{u_k} \phi_d(g_k, f_k, u_k) \rangle . \end{aligned} \tag{5.16}$$

The variation of the second term of (5.14) is given by

$$\begin{aligned}\delta J^2_{a_d} = \left\langle \delta\mu_k - \mathrm{Ad}^*_{f_k^{-1}} \cdot \mathsf{T}_{f_k} T'_d(f_k) \cdot \delta f_k + \mathrm{Ad}^*_{f_k^{-1}} (\mathrm{ad}^*_{\chi_k} T'_d(f_k)),\ \lambda^0_k \right\rangle \\ + \left\langle (1-c)h \mathsf{T}_{g_k} M(g_k) \cdot \delta g_k + (1-c)h\delta u_k,\ \lambda^0_k \right\rangle, \end{aligned}$$

where we use the formula for the derivative of the $\mathrm{Ad}^*$ operator given by (A.53). In (4.13), we have defined $\mathcal{M}(g, \lambda^1) : \mathsf{G} \times \mathfrak{g} \to \mathfrak{g}^*$ such that

$$\left\langle \mathsf{T}_g M(g) \cdot \delta g,\ \lambda^1 \right\rangle = \left\langle \mathsf{T}_g M(g) \cdot (\mathsf{T}_e \mathsf{L}_g \cdot \eta),\ \lambda^1 \right\rangle = \left\langle \mathcal{M}(g, \lambda^1),\ \eta \right\rangle \tag{5.17}$$

for any $\eta, \lambda^1 \in \mathfrak{g}$. Similarly, we define $\mathcal{T}(f, \lambda^1) : \mathsf{G} \times \mathfrak{g} \to \mathfrak{g}^*$ such that

$$\left\langle \mathsf{T}_f T'_d(f) \cdot \delta f,\ \lambda^0 \right\rangle = \left\langle \mathsf{T}_f T'_d(f) \cdot (\mathsf{T}_e \mathsf{L}_f \chi),\ \lambda^0 \right\rangle = \left\langle \mathcal{T}(f, \lambda^0),\ \chi \right\rangle \tag{5.18}$$

for any $\chi, \lambda^0 \in \mathfrak{g}$. Thus, the second term of $\delta \mathcal{J}^2_{a_d}$ is given by

$$\left\langle \mathrm{Ad}^*_{f_k^{-1}} \cdot \mathsf{T}_{f_k} T'_d(f_k) \cdot \delta f_k,\ \lambda^0_k \right\rangle = \left\langle \mathsf{T}_{f_k} T'_d(f_k) \cdot \delta f_k,\ \mathrm{Ad}_{f_k^{-1}} \lambda^0_k \right\rangle = \left\langle \mathcal{T}(f_k, \mathrm{Ad}_{f_k^{-1}} \lambda^0_k),\ \chi_k \right\rangle .$$

Using the property of the $\mathrm{Ad}^*$ operator, the third term of $\delta\mathcal{J}^2_{a_d}$ is given by

$$\begin{aligned}\left\langle \mathrm{Ad}^*_{f_k^{-1}}(\mathrm{ad}^*_{\chi_k} T'_d(f_k)),\ \lambda^0_k \right\rangle &= \left\langle T'_d(f_k),\ \mathrm{ad}_{\chi_k}(\mathrm{Ad}_{f_k^{-1}}\lambda^0_k) \right\rangle \\ &= \left\langle T'_d(f_k),\ -\mathrm{ad}_{\mathrm{Ad}_{f_k^{-1}}\lambda^0_k}(\chi_k) \right\rangle = \left\langle -\mathrm{ad}^*_{\mathrm{Ad}_{f_k^{-1}}\lambda^0_k}(T'_d(f_k)),\ \chi_k \right\rangle.\end{aligned}$$

Using these expressions, the variation of the second term of (5.14), $\delta\mathcal{J}^2_{a_d}$ is given by

$$\begin{aligned}\delta\mathcal{J}^2_{a_d} = \left\langle \delta\mu_k,\ \lambda^0_k \right\rangle &+ \left\langle -\mathcal{T}(f_k, \mathrm{Ad}_{f_k^{-1}}\lambda^0_k) - \mathrm{ad}^*_{\mathrm{Ad}_{f_k^{-1}}\lambda^0_k}(T'_d(f_k)),\ \chi_k \right\rangle \\ &+ \left\langle (1-c)h\mathcal{M}(g_k, \lambda^0_k),\ \eta_k \right\rangle + \left\langle (1-c)h\delta u_k,\ \lambda^0_k \right\rangle. \end{aligned} \tag{5.19}$$

The variation of the third term of (5.14) is given by

$$\begin{aligned}\delta\mathcal{J}^3_{a_d} = &\left\langle \delta\mu_{k+1} - \mathrm{Ad}^*_{f_k}\cdot(\delta\mu_k + (1-c)h\mathsf{T}_{g_k}M(g_k)\cdot\delta g_k + (1-c)h\delta u_k),\ \lambda^1_k \right\rangle \\ &+ \left\langle -\mathrm{Ad}^*_{f_k}(\mathrm{ad}^*_{\mathrm{Ad}_{f_k}\chi_k}(\mu_k + (1-c)hM(g_k) + (1-c)hu_k)),\ \lambda^1_k \right\rangle \\ &+ \left\langle -ch\mathsf{T}_{g_{k+1}}M(g_{k+1})\cdot\delta g_{k+1} - ch\delta u_{k+1},\ \lambda^1_k \right\rangle,\end{aligned}$$

where we use the formula for the derivative of the $\mathrm{Ad}^*$ operator given by (A.52). Using the property of the $\mathrm{Ad}^*$ map, it can be shown that this is equal to

$$\begin{aligned}\delta\mathcal{J}^3_{a_d} = &\left\langle \delta\mu_{k+1},\ \lambda^1_k \right\rangle - \left\langle \delta\mu_k + (1-c)h\delta u_k,\ \mathrm{Ad}_{f_k}\lambda^1_k \right\rangle - \left\langle (1-c)h\mathcal{M}(g_k, \mathrm{Ad}_{f_k}\lambda^1_k),\ \eta_k \right\rangle \\ &+ \left\langle \mathrm{Ad}^*_{f_k}(\mathrm{ad}^*_{\mathrm{Ad}_{f_k}\lambda^1_k}(\mu_k + (1-c)hM(g_k) + (1-c)hu_k)),\ \chi_k \right\rangle \\ &- \left\langle ch\mathcal{M}(g_{k+1}, \lambda^1_k),\ \eta_{k+1} \right\rangle - \left\langle ch\delta u_{k+1},\ \lambda^1_k \right\rangle.\end{aligned} \tag{5.20}$$

Instead of taking a variation of $\log$ in the fourth term of (5.14), we take a variation of $f_k$ and $g_k^{-1}g_{k+1}$. Using (3.6), we obtain

$$\delta\mathcal{J}^4_{a_d} = \left\langle \lambda^2_k,\ -\chi_k + \eta_{k+1} - \mathrm{Ad}_{f_k^{-1}}\eta_k \right\rangle = \left\langle \lambda^2_k,\ -\chi_k + \eta_{k+1} \right\rangle - \left\langle \mathrm{Ad}^*_{f_k^{-1}}\lambda^2_k,\ \eta_k \right\rangle. \tag{5.21}$$

From (5.16), (5.19), (5.20), and (5.21), the variation of the cost functional is given by

$$\begin{aligned}\delta J_{a_d} = &\sum_{k=0}^{N-1} \left\langle \delta\mu_{k+1},\ \lambda^1_k \right\rangle + \left\langle \lambda^2_k - ch\mathcal{M}(g_{k+1}, \lambda^1_k),\ \eta_{k+1} \right\rangle + \left\langle \delta u_{k+1},\ -ch\lambda^1_k \right\rangle \\ &+ \left\langle \mathsf{T}^*_e\mathsf{L}_{f_k}\cdot\mathbf{D}_{f_k}\phi_{d_k} - \mathcal{T}(f_k, \mathrm{Ad}_{f_k^{-1}}\lambda^0_k) - \mathrm{ad}^*_{\mathrm{Ad}_{f_k^{-1}}\lambda^0_k}(T'_d(f_k)),\ \chi_k \right\rangle \\ &+ \left\langle \mathrm{Ad}^*_{f_k}(\mathrm{ad}^*_{\mathrm{Ad}_{f_k}\lambda^1_k}(\mu_k + (1-c)hM(g_k) + (1-c)hu_k)) - \lambda^2_k,\ \chi_k \right\rangle \\ &+ \left\langle \delta\mu_k,\ \lambda^0_k - \mathrm{Ad}_{f_k}\lambda^1_k \right\rangle \\ &+ \left\langle \mathsf{T}^*_e\mathsf{L}_{g_k}\cdot\mathbf{D}_{g_k}\phi_{d_k} + (1-c)h\mathcal{M}(g_k, \lambda^0_k) - (1-c)h\mathcal{M}(g_k, \mathrm{Ad}_{f_k}\lambda^1_k) - \mathrm{Ad}^*_{f_k^{-1}}\lambda^2_k,\ \eta_k \right\rangle \\ &+ \left\langle \delta u_k,\ \mathbf{D}_{u_k}\phi_{d_k} + (1-c)h\lambda^0_k - (1-c)h\mathrm{Ad}_{f_k}\lambda^1_k \right\rangle.\end{aligned}$$

Using the fact that the variations vanish at the end points, the summation can be reindexed as follows:

$$\begin{aligned}\delta J_{a_d} = \sum_{k=0}^{N-1} &\left\langle \mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k}\phi_{d_k} - \mathcal{T}(f_k, \mathrm{Ad}_{f_k^{-1}}\lambda_k^0) - \mathrm{ad}^*_{\mathrm{Ad}_{f_k^{-1}}\lambda_k^0}(T_d'(f_k)),\ \chi_k \right\rangle \\ &+ \left\langle \mathrm{Ad}^*_{f_k}(\mathrm{ad}^*_{\mathrm{Ad}_{f_k}\lambda_k^1}(\mu_k + (1-c)hM(g_k) + (1-c)hu_k)) - \lambda_k^2,\ \chi_k \right\rangle \\ &+ \left\langle \delta\mu_k,\ \lambda_{k-1}^1 + \lambda_k^0 - \mathrm{Ad}_{f_k}\lambda_k^1 \right\rangle + \left\langle \lambda_{k-1}^2 - ch\mathcal{M}(g_k, \lambda_{k-1}^1),\ \eta_k \right\rangle \\ &+ \left\langle \mathsf{T}_e^*\mathsf{L}_{g_k} \cdot \mathbf{D}_{g_k}\phi_{d_k} + (1-c)h\mathcal{M}(g_k, \lambda_k^0) - (1-c)h\mathcal{M}(g_k, \mathrm{Ad}_{f_k}\lambda_k^1) - \mathrm{Ad}^*_{f_k^{-1}}\lambda_k^2,\ \eta_k \right\rangle \\ &+ \left\langle \delta u_k,\ \mathbf{D}_{u_k}\phi_{d_k} + (1-c)h\lambda_k^0 - (1-c)h\mathrm{Ad}_{f_k}\lambda_k^1 - ch\lambda_{k-1}^1 \right\rangle. \end{aligned} \tag{5.22}$$

This is equal to zero for all possible variations along the optimal solution. Thus, we obtain discrete-time necessary conditions for optimality as follows.

**Proposition 5.2** *Consider a mechanical system evolving on a Lie group that is expressed by Corollary 5.1 for a free parameter $c \in [0,1]$. Discrete-time necessary conditions for optimality for the discrete-time optimal control problem presented in Section 5.1.2 are as follows.*

- *Optimality condition*

$$\mathbf{D}_{u_k}\phi_{d_k} + (1-c)h\lambda_k^0 - (1-c)h\mathrm{Ad}_{f_k}\lambda_k^1 - ch\lambda_{k-1}^1 = 0, \tag{5.23}$$

- *Multiplier equations*

$$\begin{aligned} &\mathsf{T}_e^*\mathsf{L}_g \cdot \mathbf{D}_{f_k}\phi_{d_k} - \mathcal{T}(f_k, \mathrm{Ad}_{f_k^{-1}}\lambda_k^0) - \mathrm{ad}^*_{\mathrm{Ad}_{f_k^{-1}}\lambda_k^0}(T_d'(f_k)) \\ &\qquad + \mathrm{Ad}^*_{f_k}(\mathrm{ad}^*_{\mathrm{Ad}_{f_k}\lambda_k^1}(\mu_k + (1-c)hM(g_k) + (1-c)hu_k)) - \lambda_k^2 = 0, \end{aligned} \tag{5.24}$$

$$\lambda_{k-1}^1 + \lambda_k^0 - \mathrm{Ad}_{f_k}\lambda_k^1 = 0, \tag{5.25}$$

$$\begin{aligned} \lambda_{k-1}^2 - \mathrm{Ad}^*_{f_k^{-1}}\lambda_k^2 + \mathsf{T}_e^*\mathsf{L}_g \cdot \mathbf{D}_{g_k}\phi_{d_k} - ch\mathcal{M}(g_k, \lambda_{k-1}^1) \\ + (1-c)h\mathcal{M}(g_k, \lambda_k^0) - (1-c)h\mathcal{M}(g_k, \mathrm{Ad}_{f_k}\lambda_k^1) = 0, \end{aligned} \tag{5.26}$$

  *where $T_d'(f_k) = \mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}T_d(f_k)$, and $\mathcal{M}, \mathcal{T} : \mathsf{G} \times \mathfrak{g} \to \mathfrak{g}^*$ are given by (5.17) and (5.18), respectively.*

- *Boundary conditions*

$$g_N = g^f, \quad \mu_N = \mu^f. \tag{5.27}$$

**Remark 5.1** These necessary conditions are expressed in a coordinate-free form. Therefore, the optimality condition and the resulting optimal control input are independent of the choice of representations. They are globally represented by group elements without any singularity and ambiguity.

**Remark 5.2** The necessary conditions for optimality are represented by a discrete-time two point boundary value problem. It is to find the control input sequence $u_k$, state $(g_k, \mu_k)$, and multipliers $(\lambda_k^0, \lambda_k^1, \lambda_k^2)$ that satisfy the optimality condition (5.23), the equations of motion (5.11), (5.12), (5.13), and the multiplier equation (5.24), (5.25), (5.26), under the given boundary conditions. A computational approach to solve this discrete-time two point boundary value problem is discussed in Section 5.1.4.

**Remark 5.3** Here, the multiplier equations evolve on a $3n$-dimensional linear space for a $2n$- dimensional tangent bundle $\mathsf{G} \times \mathfrak{g}$. The $n$-dimensional redundancy is resolved as follows. The multiplier equation (5.24) is linear with respect to the multiplier $\lambda_k^0$. Thus, it is possible to express the multiplier $\lambda_k^0$ in terms of other variables at the same time step. The resulting expression can be substituted into (5.25) and (5.26) so that the multiplier equations evolves on a $2n$-dimensional space. The detailed procedure is described in Section 5.2.1. Alternatively, (5.11) is not explicitly constrained by using the Lagrange multiplier $\lambda_k^0$, and we find the expression for the constrained variation of $f_k$. This procedure is described in Section 5.2.2.

**Remark 5.4** The presented necessary conditions are for the optimal control problem with fixed terminal conditions and a fixed terminal time. But, they can be extended to other types of optimal control problems. For example, in Section 5.2.2, a time optimal control problem with a free terminal time and a control input constraint is studied.

**Remark 5.5** The free parameter $c \in [0, 1]$ determines the accuracy of discrete-time equations of motion and necessary conditions for optimality. They are second-order accurate if and only if $c = \frac{1}{2}$. Otherwise, they are first-order accurate. But, the equations are simplified if the parameter is chosen as either $c = 0$ or $c = 1$. In the optimal control problems presented in Section 5.2, we consider two cases when $c = \frac{1}{2}$ and $c = 1$.

### 5.1.4 Computational Approach for Discrete-time Necessary Conditions

We discuss a computational approach to solve the presented discrete-time necessary conditions for optimality. It is based on a neighboring extremal method (see Bryson and Ho 1975). The essential procedure is as follows. We first guess the unspecified initial multipliers. Then, the corresponding nominal trajectory, satisfying all of the necessary conditions except the boundary conditions, is determined by the multiplier equations and the equations of motion. The initial multipliers are 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 the control input history or multiplier trajectories are iterated, so the number of optimization parameters is proportional to the number of discrete time steps.

One 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 may result in numerical ill-conditioning. Therefore, it is important to compute the sensitivities accurately in order to apply the neighboring extremal method.

We compute the sensitivity of an optimal solution as follows. We first find an expression for the multiplier $\lambda_k^0$ by solving (5.24), and we substitute it into the multiplier equations (5.25), (5.26). We also substitute the optimality condition (5.23) into the equations of motion and the multiplier equations. For the given fixed initial conditions and a guess of the initial multipliers, we can obtain the trajectories for $(g_k, \mu_k)$ and $(\lambda_k^1, \lambda_k^2)$, and consequently the terminal condition $(g_N, \mu_N)$. Then, we obtain a solution of the optimal control problem that satisfies all of the necessary conditions except the terminal boundary condition.

The sensitivities of the specified terminal boundary conditions with respect to the unspecified initial multipliers is obtained by a linear analysis. By following the variational analysis that we used to derive the necessary conditions, we can develop linearized equations of the equations of motion and of the optimality condition. The corresponding solution of the linearized equations can be expressed as follows.

$$\begin{bmatrix} (\eta_N, \delta\mu_N) \\ (\delta\lambda_N^1, \delta\lambda_N^2) \end{bmatrix} = \begin{bmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{bmatrix} \begin{bmatrix} (\eta_0, \delta\mu_0) \\ (\delta\lambda_0^1, \delta\lambda_0^2) \end{bmatrix}, \tag{5.28}$$

where $\Psi^{ij}$ for $i, j \in \{1, 2\}$ represents a computable linear operator from $\mathfrak{g} \times \mathfrak{g}^*$ to $\mathfrak{g} \times \mathfrak{g}^*$.

For the given two point boundary value problem, $(\eta_0, \delta\mu_0) = (0, 0)$ since the initial condition is fixed. The terminal multipliers are free. Thus, we obtain

$$(\eta_N, \delta\mu_N) = \Psi^{12}(\delta\lambda_0^1, \delta\lambda_0^2). \tag{5.29}$$

The linear operator $\Psi^{12}$ represents the sensitivity of the specified terminal boundary conditions with respect to the unspecified initial multipliers. 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 described in Kelley (1995).

### 5.1.5 Direct Optimal Control Approach

We have developed discrete-time necessary conditions for optimality, and a computational approach to solve the necessary condition has been discussed. This method is referred to as an indirect optimal control approach.

Alternatively, a numerical approach for a constrained parameter optimization, such as a sequential quadratic programming, can be directly applied to the optimal control problem without deriving necessary conditions. The discrete-time equations of motion are considered as nonlinear equality constraints, and the parameterized control inputs are optimized to minimize the cost functional. This method is referred to as a direct optimal control approach.

Since the discrete-time Euler-Lagrange equations are faithful to the continuous equations of motion, more accurate solutions to the optimal control problems are obtained. They are more ef-

ficient, since optimal control inputs can be obtained by using larger time steps which requires less computational effort. This approach is preferable if the necessary conditions for optimality become too complicated. We apply a direct optimal control approach to a fuel optimal control of a 3D pendulum on a cart in Section 5.2.6, and to a fuel optimal control of two connected rigid bodies in Section 5.2.7.

## 5.2 Examples of Optimal Control of Rigid Bodies

We consider the following discrete-time optimal control problems for rigid body dynamics on a Lie group. We mathematically formulate an optimal control problem and we derive discrete-time necessary conditions for optimality. The direct optimal control approach is applied in Section 5.2.7.

| Section | Optimal Control Problem | G |
|---|---|---|
| 5.2.1 | Fuel optimal attitude control of a spacecraft on a circular orbit | $\mathsf{SO}(3)$ |
| 5.2.2 | Time optimal attitude control of a free rigid body | $\mathsf{SO}(3)$ |
| 5.2.3 | Fuel optimal attitude control of a 3D pendulum with symmetry | $\mathsf{SO}(3)$ |
| 5.2.4 | Fuel optimal control of a rigid body | $\mathsf{SE}(3)$ |
| 5.2.5 | Combinatorial optimal control of formation reconfiguration | $\mathsf{SE}(3)^n$ |
| 5.2.6 | Fuel optimal control of a 3D Pendulum on a cart | $\mathsf{SO}(3) \times \mathbb{R}^2$ |
| 5.2.7 | Fuel optimal control of two rigid bodies connected by a ball joint | $\mathsf{SO}(3) \times \mathsf{SO}(3)$ |

The fuel optimal attitude control of a spacecraft in Section 5.2.1, and the fuel optimal control of a rigid body in Section 5.2.4 are direct applications of Proposition 5.2 on $\mathsf{SO}(3)$ and $\mathsf{SE}(3)$, respectively. Generalized optimal control problems are considered in Section 5.2.2 for bounded control inputs, and in Section 5.2.3 for structured control inputs. The fuel optimal control problem of a rigid body is extended to a combinatorial optimal control problem for formation reconfiguration in Section 5.2.5. A direct optimal control approach is applied to a 3D pendulum on a cart in Section 5.2.6, and two rigid bodies connected by a ball joint in Section 5.2.7. These optimal control problems have been studied in Lee et al. (2005a, 2006a, 2007d,e,f,g, 2008a).

### 5.2.1 Fuel Optimal Attitude Control of a Spacecraft on a Circular Orbit

We study a fuel optimal control problem for the attitude dynamics of a spacecraft on a circular orbit, presented in Section 4.2.1. A rigid spacecraft lies on a circular orbit about a massive central body with an orbital angular velocity $\omega_0$. The configuration manifold is $\mathsf{SO}(3)$. The objective is to rotate the spacecraft from a given initial condition to a desired terminal condition using minimal control input during a fixed maneuver time.

In this section, forced discrete-time Hamilton's equations are derived according to Corollary 5.1, and a mathematical formulation of the optimal control problem is presented. Discrete-time necessary conditions for optimality are derived and computational results are presented.

#### Forced Hamilton's Equations

The discrete Lagrangian of the spacecraft on a circular orbit is chosen as

$$L_d(R_k, F_k) = \frac{1}{h}\mathrm{tr}[(I - F_k)J_d] - \frac{h}{2}U(R_k) - \frac{h}{2}U(R_kF_k). \tag{5.30}$$

This corresponds to the discrete Lagrangian given by (5.7) with $c = \frac{1}{2}$. The expression for the gravitational potential is given by (4.25). We assume the external control input $\hat{u}_k \in \mathfrak{so}(3)^*$ is

applied to the rigid body.

We derive forced discrete-time Hamilton's equations. Recall that the $\mathrm{Ad}$ operator for $\mathsf{SO}(3)$ is given as follows,

$$\mathrm{Ad}_F \hat{\eta} = F\hat{\eta}F^T = \widehat{F\eta}, \quad \mathrm{Ad}^*_F \hat{\eta} = F^T\hat{\eta}F = \widehat{F^T\eta}, \tag{5.31}$$

are

$$T_d'(F_k) = \mathsf{T}^*_e\mathsf{L}_{F_k} \cdot \mathbf{D}_{F_k} T_d(F_k) = \frac{1}{h}(J_d F_k - F_k^T J_d), \tag{5.32}$$

$$\mathrm{Ad}^*_{F_k^T}(T_d'(F_k)) = \frac{1}{h}(F_k J_d - J_d F_k^T). \tag{5.33}$$

According to Corollary 5.1, forced Hamilton's equations are given by

$$\hat{\Pi}_k + \frac{h}{2}\hat{u}_k + \frac{h}{2}M_k = \frac{1}{h}(F_k J_d - J_d F_k^T), \tag{5.34}$$

$$\Pi_{k+1} = F_k^T \Pi_k + \frac{h}{2}F_k^T(u_k + M_k) + \frac{h}{2}u_{k+1} + \frac{h}{2}M_{k+1}, \tag{5.35}$$

$$R_{k+1} = R_k F_k, \tag{5.36}$$

where the expression for the gravitational moment $M_k$ is given by (4.29).

$$M = 3\omega_0^2(R^T \exp(\omega_0 t\hat{e}_2)e_3) \times (JR^T \exp(\omega_0 t\hat{e}_2)e_3).$$

**Optimal Control Problem**

The objective of this optimal control problem is to transfer the spacecraft with a given initial attitude and angular velocity $(R_0, \Omega_0)$ to a desired terminal condition $(R^f, \Omega^f)$ during a fixed maneuver time $Nh$ while minimizing the control effort.

$$\text{For given: } (R_0, \Omega_0), (R^f, \Omega^f), N$$

$$\min_{u_k} \left\{ \mathcal{J}_d = \sum_{k=0}^{N-1} \frac{h}{2} u_k^T u_k \right\},$$

$$\text{such that } R_N = R^f,\ \Omega_N = \Omega^f,$$

$$\text{subject to (5.34), (5.35), (5.36).}$$

**Discrete-time Necessary Conditions for Optimality**

We derive discrete-time necessary conditions for optimality from Proposition 5.2. We first solve (5.24) to obtain an explicit expression for the multiplier $\lambda_k^0$, and we substitute it into the remaining necessary conditions.

*Expression for $\lambda_k^0$.* Using (5.18), we find the expression for $\mathcal{T}(F,\lambda) : \mathsf{SO}(3) \times \mathfrak{so}(3) \to \mathfrak{so}(3)^*$ as follows.

$$\begin{aligned}\left\langle \mathsf{T}_{F_k} T_d'(F_k)\cdot \delta F_k,\ \hat{\lambda}_k \right\rangle &= \frac{1}{h}\left\langle J_d F_k \hat{\chi}_k + \hat{\chi}_k F_k^T J_d,\ \hat{\lambda}_k \right\rangle = \frac{1}{2h}\mathrm{tr}\Big[(J_d F_k \hat{\chi}_k + \hat{\chi}_k F_k^T J_d)^T \hat{\lambda}_k\Big] \\ &= \frac{1}{2h}\mathrm{tr}\Big[(\hat{\lambda}_k J_d F_k + F_k^T J_d \hat{\lambda}_k^1)^T \hat{\chi}_k\Big] = \langle \mathcal{T}(F_k,\lambda_k), \chi_k\rangle\,.\end{aligned}$$

Thus, we obtain

$$\mathcal{T}(F_k,\lambda_k) = \frac{1}{h}(\hat{\lambda}_k J_d F_k + F_k^T J_d \hat{\lambda}_k), \tag{5.37}$$

which yields

$$\mathcal{T}(F_k, \mathrm{Ad}_{F_k^T}\lambda_k^0) = \frac{1}{h}(F_k^T \hat{\lambda}_k^0 F_k J_d F_k + F_k^T J_d F_k^T \hat{\lambda}_k^0 F_k). \tag{5.38}$$

We find expressions for the remaining terms in (5.24). Since the $\mathrm{ad}^*$ operator for a matrix Lie group is given by $\mathrm{ad}^*_\eta \alpha = \eta^T\alpha - \alpha\eta^T$ for $\eta \in \mathfrak{g}$, $\alpha \in \mathfrak{g}^*$, we obtain

$$\begin{aligned}\mathrm{ad}^*_{\mathrm{Ad}_{F_k^T}\lambda_k^0}(T_d'(F_k)) &= \frac{1}{h}(F_k^T\hat{\lambda}_k^0 F_k)^T(J_d F_k - F_k^T J_d) - \frac{1}{h}(J_d F_k - F_k^T J_d)(F_k^T\hat{\lambda}_k^0 F_k)^T \\ &= \frac{1}{h}(F_k^T\hat{\lambda}_k^0 J_d + J_d\hat{\lambda}_k^0 F_k) - \frac{1}{h}(F_k^T\hat{\lambda}_k^0 F_k J_d F_k + F_k^T J_d F_k^T\hat{\lambda}_k^0 F_k).\end{aligned} \tag{5.39}$$

Similarly,

$$\begin{aligned}&\mathrm{Ad}^*_{F_k}(\mathrm{ad}^*_{\mathrm{Ad}_{F_k}\lambda_k^1}(\Pi_k + \frac{h}{2}M_k + \frac{h}{2}u_k)) \\ &\quad= F_k^T\left((F_k\hat{\lambda}_k^1 F_k^T)^T(\hat{\Pi}_k + \frac{h}{2}\hat{M}_k + \frac{h}{2}\hat{u}_k) - (\hat{\Pi}_k + \frac{h}{2}\hat{M}_k + \frac{h}{2}\hat{u}_k)(F_k\hat{\lambda}_k^1 F_k^T)^T\right)F_k \\ &\quad= -\hat{\lambda}_k^1 F_k^T(\hat{\Pi}_k + \frac{h}{2}\hat{M}_k + \frac{h}{2}\hat{u}_k)F_k + F_k^T(\hat{\Pi}_k + \frac{h}{2}\hat{M}_k + \frac{h}{2}\hat{u}_k)F_k\hat{\lambda}_k^1 \\ &\quad= \left((F_k^T(\Pi_k + \frac{h}{2}M_k + \frac{h}{2}u_k))^\wedge\lambda_k^1\right)^\wedge.\end{aligned} \tag{5.40}$$

Substituting these into (5.24), the multiplier equation for $\lambda_k^0$ is given by

$$-\frac{1}{h}(F_k^T\hat{\lambda}_k^0 J_d + J_d\hat{\lambda}_k^0 F_k) + \left((F_k^T(\Pi_k + \frac{h}{2}M_k + \frac{h}{2}u_k))^\wedge\lambda_k^1\right)^\wedge - \hat{\lambda}_k^2 = 0.$$

Since $F_k^T\hat{\lambda}_k^0 = \widehat{F_k^T\lambda_k^0}F_k^T$, and using (A.9), given by $\hat{x}A + A^T\hat{x} = (\{\mathrm{tr}[A]\,I - A\}x)^\wedge$ for any $x \in \mathbb{R}^3$, $A \in \mathbb{R}^{3\times 3}$, this is equivalent to

$$\frac{1}{h}(\mathrm{tr}\big[F_k^T J_d\big]\,I - F_k^T J_d)F_k^T\lambda_k^0 = (F_k^T(\Pi_k + \frac{h}{2}M_k + \frac{h}{2}u_k))^\wedge\lambda_k^1 - \lambda_k^2, \tag{5.41}$$

which yields an explicit expression for the $\lambda_k^0$.

*Necessary conditions.* According to Proposition 5.2, discrete-time necessary conditions for optimality are given as follows.

- Optimality condition

$$hu_k + \frac{h}{2}\lambda_k^0 - \frac{h}{2}F_k\lambda_k^1 - \frac{h}{2}\lambda_{k-1}^1 = 0, \tag{5.42}$$

- Multiplier equations

$$\lambda_k^0 = hF_k(\mathrm{tr}\big[F_k^T J_d\big]\, I - F_k^T J_d)^{-1}\left\{(F_k^T(\Pi_k + \frac{h}{2}M_k + \frac{h}{2}u_k))^\wedge \lambda_k^1 - \lambda_k^2\right\}, \tag{5.43}$$

$$\lambda_{k-1}^1 + \lambda_k^0 - F_k\lambda_k^1 = 0, \tag{5.44}$$

$$\lambda_{k-1}^2 - F_k\lambda_k^2 + \frac{h}{2}\mathcal{M}(R_k, \lambda_k^0) - \frac{h}{2}\mathcal{M}(R_k, \lambda_{k-1}^1) - \frac{h}{2}\mathcal{M}(R_k, F_k\lambda_k^1) = 0, \tag{5.45}$$

- Boundary conditions

$$R_N = R^f, \quad \Omega_N = J\Omega^f. \tag{5.46}$$

where the expression for $\mathcal{M}(R_k, \lambda_k)$ is given by (4.31).

$$\begin{aligned}\mathcal{M}(R, \lambda) = 3\omega^2\Big[ &- (JR^T \exp(-\omega_0 t\hat{e}_2)e_3)^\wedge \, (R^T \exp(-\omega_0 t\hat{e}_2)e_3)^\wedge \\ &+ (R^T \exp(-\omega_0 t\hat{e}_2)e_3)^\wedge \, J \, (R^T \exp(-\omega_0 t\hat{e}_2)e_3)^\wedge\Big]^T \lambda.\end{aligned}$$

**Numerical Results**

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 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.

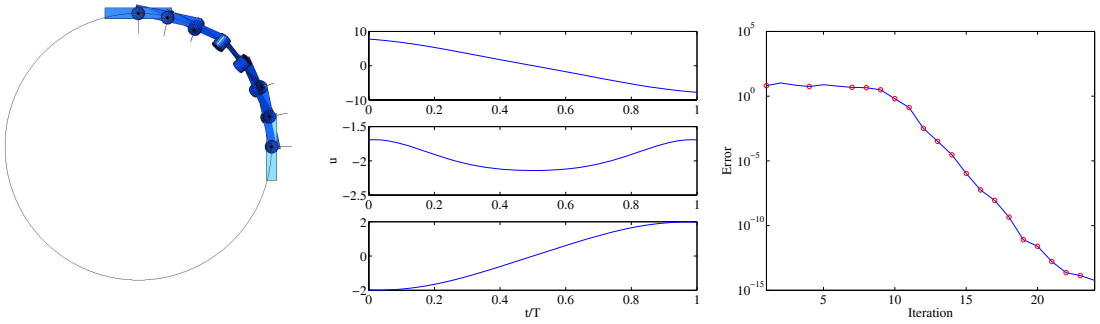
(i) Rotation maneuver about the LVLH axis $e_1$:

$$R_0 = I,\ \Omega_0 = \omega_0 R_0^T e_2, \quad R^f = \mathrm{diag}\,[1, -1, -1]\,,\ \Omega^f = \omega_0 R_N^{d,T} e_2.$$

(ii) Rotation maneuver about the LVLH axes $e_1$ and $e_2$:

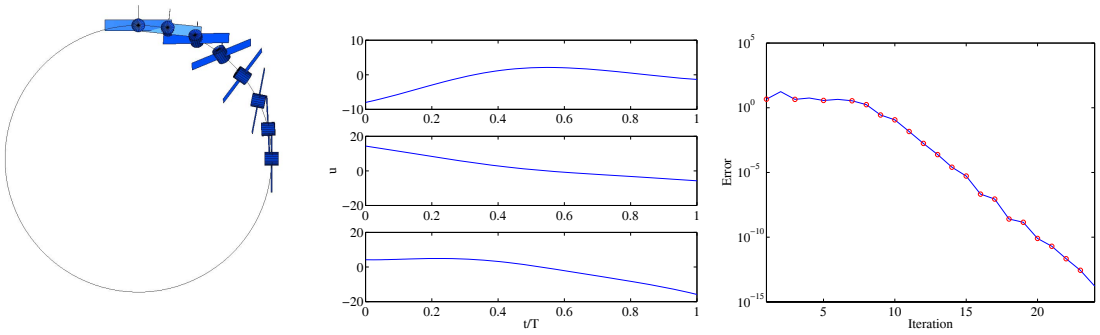
$$R_0 = \mathrm{diag}\,[1, -1, -1]\,,\ \Omega_0 = \omega_0 R_0^T e_2, \quad R^f = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix},\ \Omega^f = \omega_0 R_N^{d,T} e_2.$$

The discrete-time necessary conditions for optimality are solved using the shooting method described in Section 5.1.4. The optimized costs and the violation of the constraints are 23.35, $2.90 \times 10^{-15}$, and 70.74, $7.31 \times 10^{-15}$, respectively for each case. Figures 5.1 and 5.2 show




(a) Attitude maneuver (b) Control input $u$ (c) Convergence rate

Figure 5.1: Optimal attitude control of a spacecraft: rotation about the LVLH tangential axis




(a) Attitude maneuver (b) Control input $u$ (c) Convergence rate

Figure 5.2: Optimal attitude control of a spacecraft: rotation about about the LVLH tangential and normal axes

the attitude maneuver of the spacecraft (clockwise direction), control inputs, and violation of the terminal boundary condition as a function of the number of iterations.

As shown by Figure 5.1(c) and 5.2(c), the computational geometric optimal control approach exhibits excellent numerical convergence properties. The circles denote outer iterations to compute the sensitivity derivatives in the Newton-Armijo iteration, and 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 control inputs are close to zero throughout the maneuver time. The error in satisfaction of the terminal boundary conditions converges to machine precision within 25 iterations. These convergence rates demonstrate that the sensitivity derivatives are being computed accurately. This is because the proposed computational algorithms are geometrically exact and numerically accurate. There is no numerical dissipation caused by the algorithm, and therefore, we obtain more accurate sensitivities along the optimal solution.

### 5.2.2 Time Optimal Attitude Control of a Free Rigid Body

We study a time optimal control problem for the attitude dynamics of a free rigid body, presented in Section 4.2.2. The objective is to rotate a rigid body with a given initial condition to a desired terminal condition during a minimal maneuver time using a bounded control input. The discrete-time necessary conditions described in Proposition 5.2 cannot be directly applied to this optimal control problem, since the maneuver time is not fixed, and control inputs are bounded.

In this section, forced discrete-time Hamilton's equations are derived according to Corollary 5.1, and a mathematical formulation of the optimal control problem is presented. Modified discrete-time necessary conditions for optimality are developed by following a similar approach to that presented in Section 5.1.3, and a numerical example is shown.

#### Forced Hamilton's Equations

The discrete Lagrangian of the attitude dynamics of a free rigid body is chosen as

$$L_d(R_k, F_k) = \frac{1}{h}\mathrm{tr}[(I - F_k)J_d]\,. \tag{5.47}$$

The discrete generalized forces are $u_{d_k}^- = 0$, $u_{d_k}^+ = hu_{k+1}$. This simplifies the subsequent development for discrete-time necessary conditions.

From Corollary 5.1, the forced discrete-time Hamilton's equations for the attitude dynamics of a free rigid body are given by

$$\hat{\Pi}_k = \frac{1}{h}(F_k J_d - J_d F_k^T), \tag{5.48}$$

$$\Pi_{k+1} = F_k^T \Pi_k + hu_{k+1}, \tag{5.49}$$

$$R_{k+1} = R_k F_k. \tag{5.50}$$

#### Optimal Control Problem

We assume that the magnitude of the control moment is bounded by a constant $\overline{u} \in \mathbb{R}$, i.e. $\|u_k\|_2 \leq \overline{u}$ for any $k \in \{0, \ldots, N\}$. The objective of the time optimal attitude control problem is to transfer the rigid body with a given initial attitude and angular velocity $(R_0, \Omega_0)$ to a desired terminal condition $(R^f, \Omega^f)$ within a minimal maneuver time $Nh$, where the 2-norm of the control moment is constrained by the given limit $\overline{u}$.

$$\text{For given: } (R_0, \Omega_0), (R^f, \Omega^f), \bar{u}$$

$$\min_{u_{k+1}, N} \left\{ \mathcal{J}_d = \sum_{k=0}^{N-1} 1\, dt \right\},$$

$$\text{such that } R_N = R^f,\ \Omega_N = \Omega^f,$$

$$\text{subject to } \|u_k\| \leq \bar{u}\ \ \forall k \in \{0, \ldots, N\} \text{ and } (5.48), (5.49), (5.50).$$

**Discrete-time Necessary Conditions for Optimality**

This optimal control problem is not a special case of Proposition 5.2, since the terminal maneuver time $N$ is not fixed. But, the presented results can be easily extended to handle the time optimal control problem.

We define the augmented cost functional

$$\mathcal{J}_{a_d} = \sum_{k=0}^{N-1} 1 + \left\langle -\Pi_{k+1} + F_k^T \Pi_k + h u_{k+1},\ \lambda_k^1 \right\rangle + \left\langle \lambda_k^2,\ \frac{1}{2}(F_k - F_k^T)^\vee - (R_k^T R_{k+1} - R_{k+1}^T R_k)^\vee \right\rangle \tag{5.51}$$

for Lagrange multipliers $\lambda_k^1 \in \mathfrak{so}(3)$, $\lambda_k^2 \in \mathfrak{so}(3)^*$. Comparing this with (5.14), the discrete equation (5.48) is not explicitly constrained by using a multiplier. Here we develop an expression for the constrained variation of $F_k$ as discussed in Remark 5.3. In addition, the matrix logarithms for $F_k$, $R_k^T R_{k+1}$ are replaced by $(F_k - F_k^T)^\vee$, and $(R_k^T R_{k+1} - R_{k+1}^T R_k)^\vee$. This is based on the fact that if the step size is chosen sufficiently small so that the relative attitude rotation between integration steps is less than $\pi/2$, then the rotation matrix $F_k$ is equal to $R_k^T R_{k+1}$ if and only if their skew parts are identical. This is a property of the matrix logarithm on $\mathsf{SO}(3)$, and it is easier to handle the skew part of a rotation matrix than a matrix logarithm of it.

We find a constrained variation of $F_k$ from (5.48): the variation of $F_k$ is expressed as $\delta F_k = F_k \hat{\chi}_k$ for $\chi_k \in \mathbb{R}^3$, and $\chi_k$ is written in terms of $F_k$ and $\delta\Pi_k$. Taking a variation of (5.48), we obtain

$$h\delta\hat{\Pi}_k = F_k \hat{\chi}_k J_d + J_d \chi_k F_k^T = \widehat{F_k \chi_k} F_k J_d + J_d F_k^T \widehat{F_k \chi_k} = ((\mathrm{tr}[F_k J_d]\, I - F_k J_d) F_k \chi_k)^\wedge,$$

where we use the property of the hat map given by (A.9). Thus, $\chi_k$ is given by

$$\chi_k = h F_k^T (\mathrm{tr}[F_k J_d]\, I - F_k J_d)^{-1} \delta\Pi_k \equiv \mathcal{B}_k \delta\Pi_k, \tag{5.52}$$

where $\mathcal{B}_k = h F_k^T (\mathrm{tr}[F_k J_d]\, I - F_k J_d)^{-1}$.

Now we consider the terminal time variation. Since the terminal conditions are fixed, the variations of the terminal attitude and angular velocity are related to the variation of the maneuver time as follows,

$$\delta\Omega_N + (\Omega_N - \Omega_{N-1})\delta N = 0, \tag{5.53}$$

$$R_N \hat{\eta}_N + \frac{1}{2} R_N (F_{N-1} - F_{N-1}^T)\delta N = 0, \tag{5.54}$$

where the terminal angular velocity is approximated by $\hat{\Omega}_N = R_N^T \dot{R}_N \approx \frac{1}{2h}(R_N^T + R_{N-1}^T)(R_N - R_{N-1}) = \frac{1}{2h}(F_{N-1} - F_{N-1}^T)$.

The next step is taking a variation of the augmented cost functional (5.51), and substituting the constrained variation (5.52) and the variations of the terminal states (5.53), (5.54). Then the

variation of the augmented cost functional is given by

$$\delta\mathcal{J}_{a_d} = \sum_{k=0}^{N-1} \left\langle h\delta u_{k+1},\ \lambda_k^1 \right\rangle + \sum_{k=1}^{N-1} \left\langle \delta\Pi_k,\ -\lambda_{k-1}^1 + (F_k - \mathcal{B}_k^T \widehat{F_k^T \Pi_k})\lambda_k^1 + \frac{1}{2}\mathcal{B}_k^T(\text{tr}[F_k]\, I - F_k)\lambda_k^2 \right\rangle$$
$$+ \sum_{k=1}^{N-1} \left\langle \frac{1}{2}(\text{tr}[F_{k-1}]\, I - F_{k-1})\lambda_{k-1}^2 - \frac{1}{2}F_k(\text{tr}[F_k]\, I - F_k)\lambda_k^2,\ \eta_k \right\rangle$$
$$+ \left\{ 1 + \lambda_{N-1}^1 \cdot \left\{ -\Pi_{N-1} + F_{N-1}^T \Pi_{N-1} + h u_N \right\} + \lambda_{N-1}^R \cdot \frac{1}{4}\left( (F_{N-1})^2 - (F_{N-1}^T)^2 \right)^\vee \right\} \delta N.$$

The multiplier equations are chosen such that the expressions in the pairing are equal to zero, and the transversality condition is chosen such that the expression in the last braces is equal to zero. Then, the variation of the cost functional reduces to

$$\delta\mathcal{J}_{a_d} = \sum_{k=0}^{N-1} \left\langle h\delta u_{k+1},\ \lambda_k^1 \right\rangle.$$

The optimality condition is chosen such that $\delta\mathcal{J}_{a_d} \geq 0$ for all admissible variations of the control input.

In summary, discrete-time necessary conditions for optimality are given by

- Optimality condition

$$u_{k+1} = -\bar{u}\,(\lambda_k^1 / \left\|\lambda_k^1\right\|), \tag{5.55}$$

- Multiplier equations

$$-\lambda_{k-1}^1 + (F_k - \mathcal{B}_k^T \widehat{F_k^T \Pi_k})\lambda_k^1 + \frac{1}{2}\mathcal{B}_k^T(\text{tr}[F_k]\, I - F_k)\lambda_k^2 = 0, \tag{5.56}$$

$$(\text{tr}[F_{k-1}]\, I - F_{k-1})\lambda_{k-1}^2 - F_k(\text{tr}[F_k]\, I - F_k)\lambda_k^2, \tag{5.57}$$

$$\mathcal{B}_k = F_k^T(\text{tr}[F_k J_d]\, I - F_k J_d)^{-1}, \tag{5.58}$$

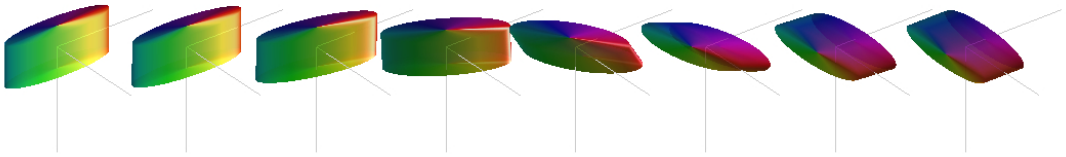
- Boundary conditions

$$1 + \lambda_{N-1}^1 \cdot \left\{ -\Pi_{N-1} + F_{N-1}^T \Pi_{N-1} + h u_N \right\} + \lambda_{N-1}^2 \cdot \frac{1}{4}\left( (F_{N-1})^2 - (F_{N-1}^T)^2 \right)^\vee = 0, \tag{5.59}$$

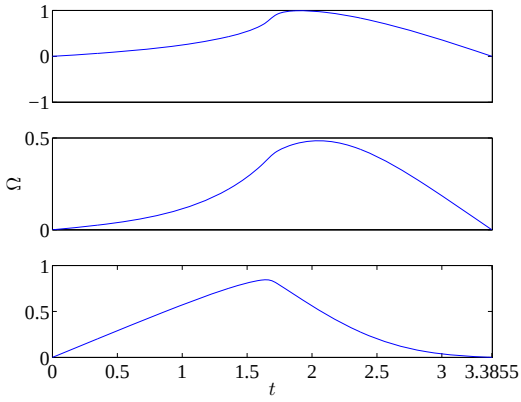
$$R_N = R^f, \quad \Omega_N = \Omega^f. \tag{5.60}$$

**Numerical Results**

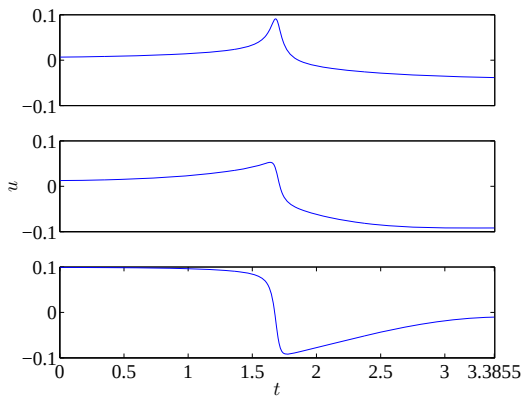
We choose an elliptic cylinder for a rigid body model with semi-major axis $0.8\,\text{m}$, semi-minor axis $0.2\,\text{m}$, height $0.6\,\text{m}$, and mass $1\,\text{kg}$. The inertia matrix is $J = \text{diag}[0.04,\ 0.19,\ 0.17]\,\text{kgm}^2$, and the maximum control limit is chosen as $\overline{u} = 0.1\,\text{Nm}$. The desired attitude maneuver is a rest-to-rest large angle rotation described by $(R_\circ, \Omega_\circ) = (I_{3\times3}, 0)$, $(R_f, \Omega_f) = (\exp\theta\hat{v}, 0)$, where $v = \frac{1}{\sqrt{3}}[1,\ 1,\ 1] \in \mathbb{R}^3$, and $\theta = 120^\circ$.



(a) Attitude maneuver




(b) Angular velocity $\Omega$




(c) Control input $u$

Figure 5.3: Time optimal attitude control of a free rigid body

The discrete-time necessary conditions for optimality are solved using the shooting method described in Section 5.1.4. In the numerical computations, we fix the number of steps as $N = 1000$ in this particular numerical example, and we vary the step size $h$. In essence, we find the seven parameters, initial multipliers and the time step $h$, satisfying the seven-dimensional terminal boundary conditions under the discrete-time equations of motion, the multiplier equation, and the optimality condition. The optimal solutions are found in $94$ 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 histories are presented in Figure 5.3. The optimized maneuver times are $3.39$ sec, and the control input is saturated during the entire maneuver time.

### 5.2.3 Fuel Optimal Attitude Control of a 3D Pendulum with Symmetry

We study a fuel optimal attitude control of a 3D pendulum introduced in Section 4.2.3 (see Lee et al. 2007f). The 3D pendulum is a rigid body supported by a frictionless pivot point acting under the gravitational potential. We have shown that the 3D pendulum has a symmetry represented by a group action of $\mathsf{SO}(2) \simeq \mathsf{S}^1$, and consequently, the angular momentum about the gravity direction is conserved, and the configuration manifold $\mathsf{SO}(3)$ can be reduced to a quotient space $\mathsf{SO}(3)/\mathsf{S}^1 \simeq \mathsf{S}^2$.

The external control moment does not have any component about the gravity direction, and therefore, 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.

In this section, forced discrete-time Hamiltons equations are derived according to Corollary 5.1, and a mathematical formulation of the optimal control problem is presented. This optimal control problem cannot be considered as a special case of Proposition 5.2, since the control input has a special structure. Here, modified discrete-time necessary conditions for optimality are developed by following the approach presented in Section 5.1.3. The necessary conditions have numerical ill-conditioning due to the conservation property. We also present a computational approach to avoid this numerical ill-conditioning. The key idea is to decompose the sensitivity derivatives into symmetric parts and asymmetric parts. The resulting numerical results are studied according to geometric phase effects.

#### Forced Hamilton Equations

The discrete Lagrangian of the 3D pendulum is chosen as

$$L_d(R_k, F_k) = \frac{1}{h}\mathrm{tr}[(I - F_k)J_d] - hU(R_kF_k). \tag{5.61}$$

The gravitational potential is $U(R) = mg\rho_c R^T e_3$. As discussed in Section 4.2.3, control moment is expressed in the body fixed frame as

$$u_k = R_k^T e_3 \times u_{p_k}$$

for a control parameter $u_{p_k} \in \mathbb{R}^3$. Since the vector $R^T e_3$ represents the gravity direction in the body fixed frame, the external control moment has no component along the gravity direction. We choose the discrete generalized forces as $u_{d_k}^- = 0$, $u_{d_k}^+ = hR_{k+1}^T e_3 \times u_{p_{k+1}}$.

From Corollary 5.1, forced discrete-time Hamilton's equations are given by

$$\hat{\Pi}_k = \frac{1}{h}(F_k J_d - J_d F_k^T), \tag{5.62}$$

$$\Pi_{k+1} = F_k^T \Pi_k + hM_{k+1} + hR_{k+1}^T e_3 \times u_{p_{k+1}}, \tag{5.63}$$

$$R_{k+1} = R_k F_k, \tag{5.64}$$

where the moment due to the potential is given by $M = mg\rho_c \times R^T e_3$.

**Optimal Control Problem**

The objective of the optimal control problem is to rotate the 3D pendulum from an initial attitude $R_0 = I$ to a desired terminal attitude $R^f = \exp(\theta \hat{e}_3)$ for a fixed terminal time $Nh$ and a rotation angle $\theta \in \mathsf{S}^1$, while minimizing the control effort. The initial angular velocity and the terminal angular velocity are zero.

$$\begin{gathered}
\text{For given: } (R_0 = I, \Omega_0 = 0), N, \theta \\
\min_{u_{p_{k+1}}} \left\{ \mathcal{J}_d = \sum_{k=0}^{N-1} \frac{h}{2} u_{p_{k+1}}^T u_{p_{k+1}} \right\}, \\
\text{such that } R_N = \exp(\theta \hat{e}_3),\ \Omega_N = 0, \\
\text{subject to (5.62), (5.63), (5.64).}
\end{gathered}$$

**Discrete-time Necessary Conditions for Optimality**

This optimal control problem cannot be considered as a special case of the Proposition 5.2, as the generalized force is dependent on the rotation matrix $R_k$. But, discrete-time necessary conditions can be easily obtained by generalizing Proposition 5.2 to include the variation of the control moment as shown in (4.48).

Define the augmented cost functional:

$$\begin{aligned}
\mathcal{J}_{a_d} = \sum_{k=0}^{N-1} \frac{h}{2} u_{p_{k+1}}^T u_{p_{k+1}} + \left\langle \hat{\Pi}_k - \frac{1}{h}(F_k J_d - J_d F_k^T),\ \lambda_k^0 \right\rangle \\
+ \left\langle \Pi_{k+1} - F_k^T \Pi_k - h M_{k+1} - h R_{k+1}^T e_3 \times u_{p_{k+1}},\ \lambda_k^1 \right\rangle + \left\langle \lambda_k^2,\ \log(R_k^T R_{k+1}) - \log F_k \right\rangle.
\end{aligned}$$

This is equal to (5.14) when $c = 1$, except that two terms $\phi$ and $u_{k+1}$ in (5.14) are replaced by $\frac{h}{2} u_{p_{k+1}}^T u_{p_{k+1}}$ and $R_{k+1}^T e_3 \times u_{p_{k+1}}$. Therefore, we can obtain the variation of the augmented cost functional from (5.22) with $c = 1$ as follows.

$$\begin{aligned}
\delta J_{a_d} = \sum_{k=0}^{N-1} \Bigg\langle -\mathcal{T}(F_k, \mathrm{Ad}_{F_k^T} \lambda_k^0) - \mathrm{ad}^*_{\mathrm{Ad}_{F_k^T}\lambda_k^0}(T_d'(f_k)) + \mathrm{Ad}^*_{F_k}(\mathrm{ad}^*_{\mathrm{Ad}_{F_k}\lambda_k^1} \Pi_k) - \lambda_k^2,\ \chi_k \Bigg\rangle \\
+ \left\langle \delta\Pi_k,\ \lambda_{k-1}^1 + \lambda_k^0 - \mathrm{Ad}_{F_k} \lambda_k^1 \right\rangle + \left\langle \lambda_{k-1}^2 - h\mathcal{M}(R_k, \lambda_{k-1}^1) - \mathrm{Ad}^*_{F_k^T} \lambda_k^2,\ \eta_k \right\rangle \\
+ \left\langle \delta u_{p_{k+1}},\ h u_{p_{k+1}} \right\rangle - \left\langle \delta(h R_{k+1}^T e_3 \times u_{p_{k+1}}),\ \lambda_k^1 \right\rangle. \qquad (5.65)
\end{aligned}$$

The last term of (5.65) is given by

$$\begin{aligned}
&\left\langle \delta(h R_{k+1}^T e_3 \times u_{p_{k+1}}),\ \lambda_k^1 \right\rangle \\
&\quad = \left\langle h(-\hat{\chi}_k R_{k+1}^T - F_k^T \hat{\eta}_k R_k^T) e_3 \times u_{p_{k+1}},\ \lambda_k^1 \right\rangle + \left\langle h R_{k+1}^T e_3 \times \delta u_{p_{k+1}},\ \lambda_k^1 \right\rangle \\
&\quad = -\left\langle h \hat{u}_{p_{k+1}} \widehat{R_k^T e_3} \chi_k + h \hat{u}_{p_{k+1}} F_k^T \widehat{R_k^T e_3} \eta_k,\ \lambda_k^1 \right\rangle + \left\langle h \widehat{R_{k+1}^T e_3} \delta u_{p_{k+1}},\ \lambda_k^1 \right\rangle
\end{aligned}$$

$$= -\left\langle h\widehat{R_k^T e_3}\hat{u}_{p_{k+1}}\lambda_k^1,\ \chi_k\right\rangle - \left\langle h\widehat{R_k^T e_3}F_k\hat{u}_{p_{k+1}}\lambda_k^1,\ \eta_k\right\rangle - \left\langle \delta u_{p_{k+1}},\ h\widehat{R_{k+1}^T e_3}\lambda_k^1\right\rangle. \quad (5.66)$$

From (4.47), the expression for the derivative of the gravity moment is given by $M(R,\lambda^1) = mg\widehat{R^T e_3}\hat{\rho}_c\lambda^1$. Now, we find an expression for the first term of (5.65). From (5.38), (5.39), and (5.40), we obtain

$$\mathcal{T}(F_k, \mathrm{Ad}_{F_k^T}\lambda_k^0) = \frac{1}{h}(F_k^T\hat{\lambda}_k^0 F_k J_d F_k + F_k^T J_d F_k^T\hat{\lambda}_k^0 F_k),$$
$$\mathrm{ad}^*_{\mathrm{Ad}_{F_k^T}\lambda_k^0}(T_d'(F_k)) = \frac{1}{h}(F_k^T\hat{\lambda}_k^0 J_d + J_d\hat{\lambda}_k^0 F_k) - \frac{1}{h}(F_k^T\hat{\lambda}_k^0 F_k J_d F_k + F_k^T J_d F_k^T\hat{\lambda}_k^0 F_k),$$
$$\mathrm{Ad}^*_{F_k}(\mathrm{ad}^*_{\mathrm{Ad}_{F_k}\lambda_k^1}\Pi_k) = \left(\widehat{F_k^T\Pi_k}\lambda_k^1\right)^\wedge.$$

Using these, the first term of (5.65) can be written as

$$\left\langle -\mathcal{T}(F_k, \mathrm{Ad}_{F_k^{-1}}\lambda_k^0) - \mathrm{ad}^*_{\mathrm{Ad}_{F_k^{-1}}\lambda_k^0}(T_d'(f_k)) + \mathrm{Ad}^*_{F_k}(\mathrm{ad}^*_{\mathrm{Ad}_{F_k}\lambda_k^1}\Pi_k) - \lambda_k^2,\ \chi_k\right\rangle$$
$$= \left\langle -\frac{1}{h}(F_k^T\hat{\lambda}_k^0 J_d + J_d\hat{\lambda}_k^0 F_k)^\vee + \widehat{F_k^T\Pi_k}\lambda_k^1 - \lambda_k^2,\ \chi_k\right\rangle. \quad (5.67)$$

Substituting (5.66), (5.67) into (5.65), we obtain the variation of the augmented cost functional as follows.

$$\delta J_{a_d} = \sum_{k=0}^{N-1}\left\langle -\frac{1}{h}(F_k^T\hat{\lambda}_k^0 J_d + J_d\hat{\lambda}_k^0 F_k)^\vee + \widehat{F_k^T\Pi_k}\lambda_k^1 - \lambda_k^2 + h\widehat{R_k^T e_3}\hat{u}_{p_{k+1}}\lambda_k^1,\ \chi_k\right\rangle$$
$$+ \left\langle \delta\Pi_k,\ \lambda_{k-1}^1 + \lambda_k^0 - F_k\lambda_k^1\right\rangle + \left\langle \delta u_{p_{k+1}},\ hu_{p_{k+1}} + h\widehat{R_{k+1}^T e_3}\lambda_k^1\right\rangle$$
$$+ \left\langle \lambda_{k-1}^2 - hmg\widehat{R_k^T e_3}\hat{\rho}_c\lambda_{k-1}^1 - F_k\lambda_k^2 + h\widehat{R_k^T e_3}F_k\hat{u}_{p_{k+1}}\lambda_k^1,\ \eta_k\right\rangle.$$

This is equal to zero for all admissible variations. Following the procedure used to obtain (5.41), we can find an explicit expression for $\lambda_k^0$.

In summary, discrete-time necessary conditions for optimality are as follows.

- Optimality condition

$$u_{p_{k+1}} = -R_{k+1}^T e_3 \times \lambda_k^1, \quad (5.68)$$

- Multiplier equations

$$\lambda_k^0 = hF_k(\mathrm{tr}\left[F_k^T J_d\right] I - F_k^T J_d)^{-1}\left\{(\widehat{F_k^T\Pi_k} + h\widehat{R_k^T e_3}\hat{u}_{p_{k+1}})\lambda_k^1 - \lambda_k^2\right\}, \quad (5.69)$$
$$\lambda_{k-1}^1 + \lambda_k^0 - F_k\lambda_k^1 = 0, \quad (5.70)$$
$$\lambda_{k-1}^2 - F_k\lambda_k^2 - hmg\widehat{R_k^T e_3}\hat{\rho}_c\lambda_{k-1}^1 + h\widehat{R_k^T e_3}F_k\hat{u}_{p_{k+1}}\lambda_k^1 = 0, \quad (5.71)$$

- Boundary conditions

$$R_N = \exp(\theta \hat{e}_3), \quad \Omega_N = 0. \tag{5.72}$$

**Modified Computational Approach**

We solve these discrete-time necessary conditions using the computational approach discussed in Section 5.1.4. We first substitute the expression for $\lambda_k^0$ into (5.69) and the optimality condition (5.68) into (5.70) and (5.71). Then, the multiplier equations are expressed in terms of multipliers $\lambda_k^1, \lambda_k^2$, attitudes, and angular velocities $R_k, \Omega_k$. According to (5.29), we obtain

$$\begin{bmatrix} \eta_N \\ \delta\Pi_N \end{bmatrix} = \begin{bmatrix} \Psi_1^{12} & \Psi_2^{12} \\ \Psi_3^{12} & \Psi_4^{12} \end{bmatrix} \begin{bmatrix} \delta\lambda_0^1 \\ \delta\lambda_0^2 \end{bmatrix}, \tag{5.73}$$

where the linear operator $\Psi^{12}$ from $\mathfrak{so}(3) \times \mathfrak{so}(3)^*$ to $\mathfrak{so}(3) \times \mathfrak{so}(3)^*$ is represented by four submatrices $\Psi_i^{12} \in \mathbb{R}^{3\times 3}$ for $i \in \{1, 2, 3, 4\}$. At each iteration, we require the inverse of the sensitivity derivative represented by the matrix $\Psi^{12}$ to update the initial multipliers that satisfy the terminal boundary condition.

By Noether's theorem, the symmetry of the 3D pendulum yields a conserved quantity, which causes a fundamental singularity in the sensitivity derivatives for the two-point boundary value problem. The sensitivity matrix $\Psi^{12}$ 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, (5.73) is numerically ill-conditioned.

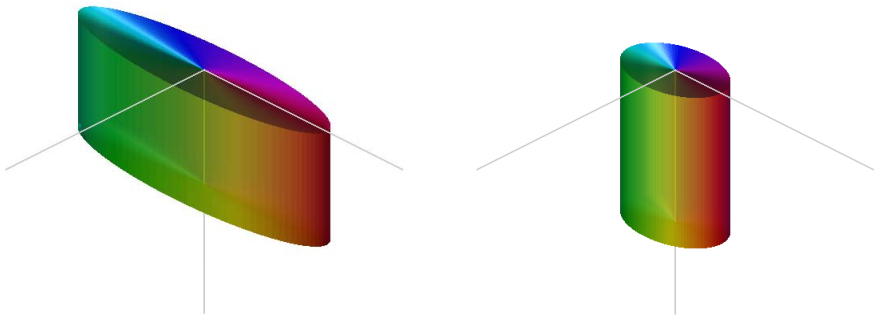
Here, we present a simple numerical scheme to avoid the numerical ill-conditioning caused by the symmetry. The essential idea is to decompose the sensitivity derivative into a symmetric part and an asymmetric part. Since the angular moment of the 3D pendulum is expressed in the reference frame as $\pi_k = R_k \Pi_k$, its infinitesimal variation is given by

$$\begin{aligned} \delta\pi_N &= \delta(R_N\Pi_N) = \delta R_N \Pi_N + R_N \delta\Pi_N \\ &= -R_N \hat{\Pi}_N \eta_N + R_N \delta\Pi_N \\ &= -R_N \hat{\Pi}_N (\Psi_1^{12}\delta\lambda_0^1 + \Psi_2^{12}\delta\lambda_0^2) + R_N(\Psi_3^{12}\delta\lambda_0^1 + \Psi_4^{12}\delta\lambda_0^2). \end{aligned}$$

Now, the sensitivity derivative equation (5.73) can be rewritten in terms of the inertial angular momentum variation as

$$\begin{bmatrix} \eta_N \\ \delta\pi_N \end{bmatrix} = \begin{bmatrix} \Psi_1^{12} & \Psi_2^{12} \\ R_N(\Psi_3^{12} - \hat{\Pi}_N\Psi_1^{12}) & R_N(\Psi_4^{12} - \hat{\Pi}_N\Psi_2^{12}) \end{bmatrix} \begin{bmatrix} \delta\lambda_0^1 \\ \delta\lambda_0^2 \end{bmatrix}. \tag{5.74}$$

From the symmetry, the third component of the inertial angular momentum variation is zero; thus $\delta(\pi_N)_3 = 0$. That is, 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 based on the pseudo-inverse of the $5 \times 6$ matrix composed of the first five rows of the transformed sensitivity derivative in (5.74).



(a) Body (A) (b) Body (B)

Figure 5.4: Two types of the 3D pendulum body for the optimal control problem with symmetry

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.

### Numerical Results

Numerical optimization results for the 3D pendulum are given. Two elliptical cylinders, shown in Figure 5.4, are used as rigid pendulum models. Properties of the 3D pendulum models are chosen as

$$\text{Body (A): } m = 1,\ J = \text{diag}[0.13, 0.28, 0.17], \rho = 0.3e_3.$$
$$\text{Body (B): } m = 1,\ J = \text{diag}[0.22, 0.23, 0.03], \rho = 0.4e_3.$$

Each maneuver involves a transfer from an initial hanging equilibrium to another hanging equilibrium corresponding a rotation about the vertical axis. The rotation angle is chosen as $180^\circ$. Thus, $R_0 = I$, $R^f = \exp(\pi\hat{e}_3) = \text{diag}[-1, -1, 1]$. Since the vertical component of the angular momentum is exactly zero, the rotation is purely caused by the geometric phase effect given in (A.37). 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.

The optimized cost functions and the violations of the terminal boundary conditions are 7.32, $4.80 \times 10^{-15}$, and 3.37, $3.06 \times 10^{-14}$, respectively. It takes 2.72 minutes and 5.05 minutes with an Intel Pentium M 740 1.73GHz processor on MATLAB. Figures 5.5 and 5.6 show snapshots of the attitude maneuvers, the direction of the gravity in the body fixed frame $\Upsilon = R^T e_3$ displayed on a sphere, control input histories, and convergence rate.

The convergence rate figures show the violation of the terminal boundary conditions as it depends on the number of iterations. Red circles denote outer iterations in the 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 50th iteration. These convergence results are consistent with

the quadratic convergence rates expected of Newton methods with accurately computed gradients. If the sensitivity derivative is not decomposed as in (5.73), then the condition number of the sensitivity matrix is at the level of $10^{19}$, and the numerical iterations fail to converge.

*Geometric Phase of the 3D Pendulum.* We interpret the optimization results using the geometric phase formula given by (A.37).

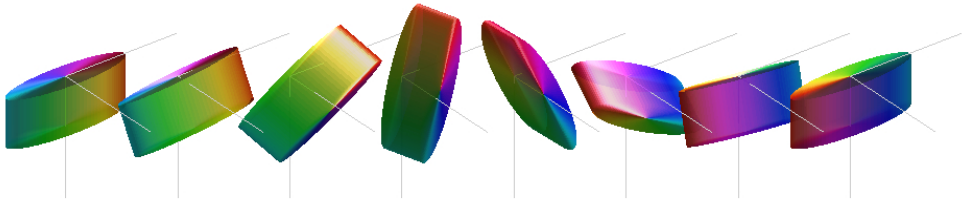
$$\theta_{\text{geo}}(T) = \int_{\mathcal{S}} \frac{2\left\|J\Upsilon\right\|^2 - \text{tr}[J]\left(\Upsilon \cdot J\Upsilon\right)}{\left(\Upsilon \cdot J\Upsilon\right)^2}\, dA, \tag{5.75}$$

where $\mathcal{S}$ is the region in $\mathsf{S}^2$ that is enclosed by the curve $\Upsilon(t)$ for $t \in [t_0, t_f]$. For the 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 $\mathsf{S}^2$ whose boundary is the reduced trajectory $\Upsilon$, it is more efficient for the reduced trajectory to enclose the area at which the absolute value of the integrand of (5.75) is maximized.

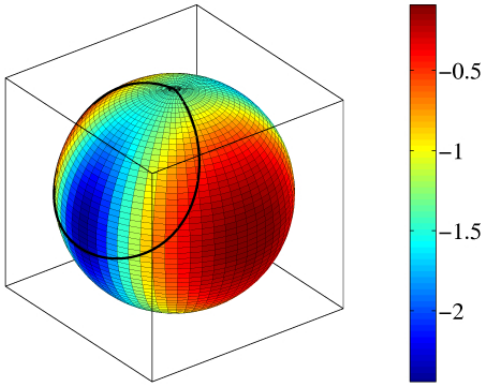
In Figure 5.5(b) and Figure 5.6(b), the infinitesimal geometric phase per unit area is shown by color shading. The reduced trajectory, which represents the gravity direction in the body fixed frame, is shown by a solid line. The north pole of the sphere corresponds to the hanging equilibrium manifold, and the reduced trajectory starts and ends at the same north pole for the given boundary conditions.

Comparing Figure 5.5(b) with Figure 5.6(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.

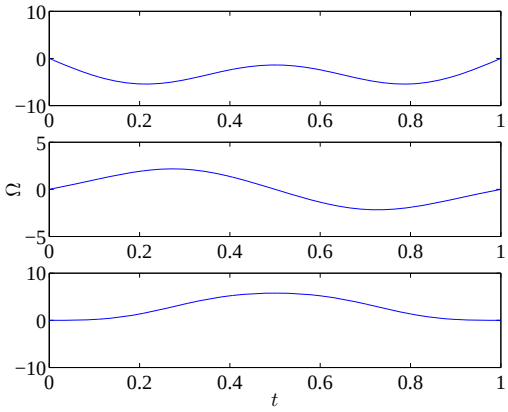
As a result, the optimized attitude maneuver of Body (A) is distinguished from that of Body (B). The attitude maneuver of Body (A) is relatively more aggressive than that of Body (B) since the reduced trajectory passes near the equator corresponding to a horizontal position. Body (B) does not have to move far away from the hanging equilibrium since the infinitesimal geometric phase is maximized at that point. The resulting attitude maneuver is relatively benign.



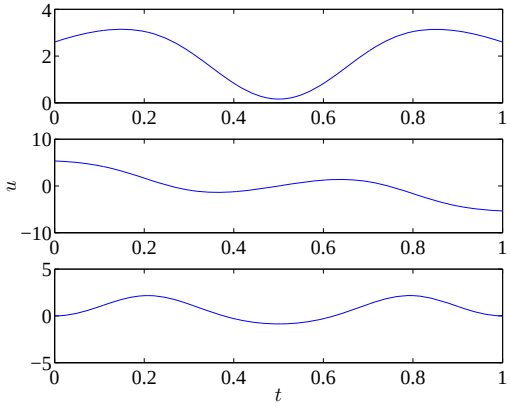
(a) Attitude maneuver



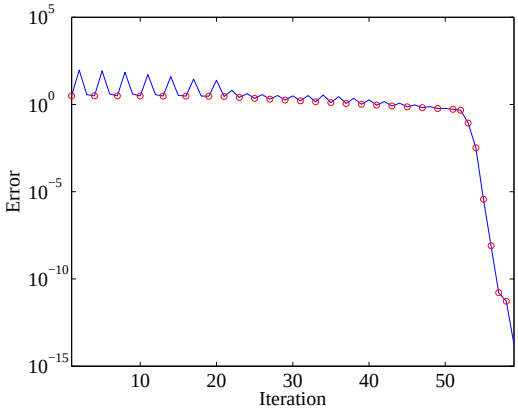

(b) Geometric phase




(c) Angular velocity

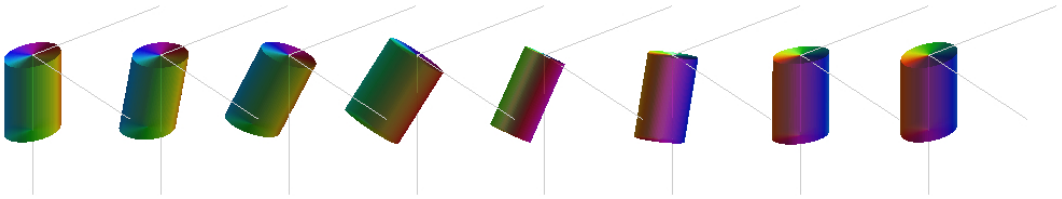



(d) Control input

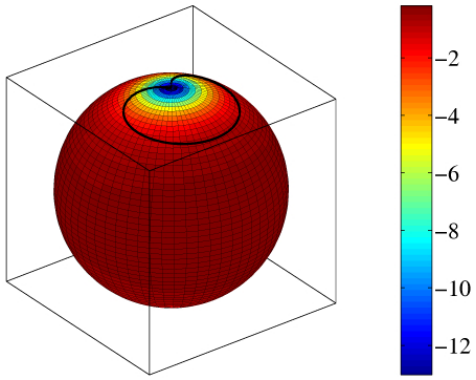



(e) Convergence rate

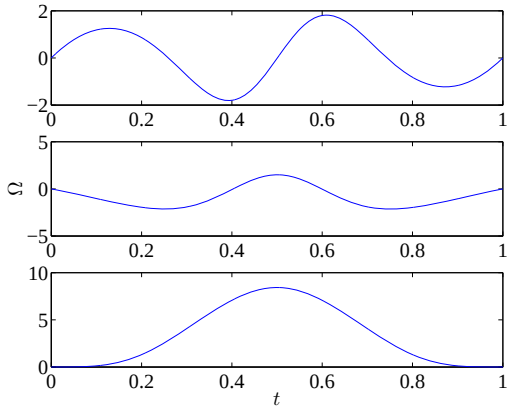
Figure 5.5: Optimal control of a 3D pendulum: body A, 180° rotation about the gravity direction



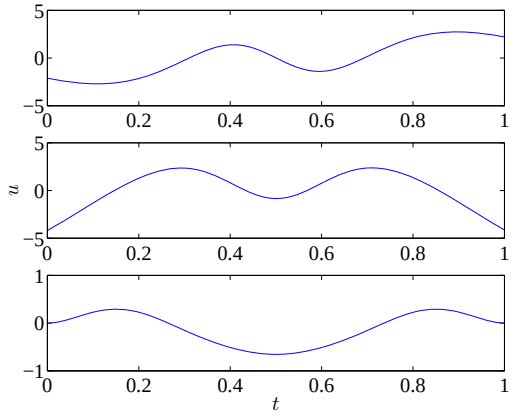
(a) Attitude maneuver



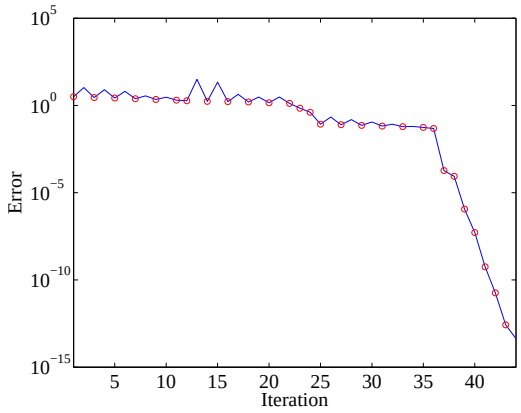

(b) Geometric phase




(c) Angular velocity




(d) Control input




(e) Convergence rate

Figure 5.6: Optimal control of a 3D pendulum: body B, 180° rotation about the gravity direction

### 5.2.4 Fuel Optimal Control of a Rigid Body

We study a fuel optimal control problem of a rigid body acting under a potential field, introduced in Section 4.2.4. The objective is to change the position, the attitude, the linear velocity, and the angular velocity of a rigid body during a fixed maneuver time using a minimal control force and control moment. Here, we explicitly consider the coupling effects of the rotational attitude maneuver and the translational maneuver of a rigid body.

In this section, forced discrete-time Hamilton's equations are derived according to Corollary 5.1, and a mathematical formulation of the optimal control problem is presented. Discrete-time necessary conditions for optimality are derived from Proposition 5.2 and computational results are presented.

#### Forced Hamilton's Equations

Consider a rigid body that is acting under the configuration dependent potential $U(R,x) : \mathsf{SE}(3) \to \mathbb{R}$. The configuration manifold is $\mathsf{SE}(3)$. This rigid body model is presented in Section 2.3.5, and the corresponding Lie group variational integrator is developed in Section 3.3.5. The discrete Lagrangian is chosen as

$$L(g_k, f_k) = \frac{1}{2h} m \Delta x_k^T \Delta x_k - \frac{1}{h} \mathrm{tr}[(I - F_k) J_d] + hU(R_k F_k, x_k + \Delta x_k).$$

This corresponds to the discrete Lagrangian given by (5.7) with $c = 1$. We assume that external control force $u_k^f \in (\mathbb{R}^3)^*$ and external control moment $u_k^m \in \mathfrak{so}(3)^*$ are applied to the rigid body, and they are expressed in the body fixed frame. The discrete generalized forces are given by $u_{d_k}^- = 0$, $u_{d_k}^+ = (hu_{k+1}^m, hu_{k+1}^f)$.

From Corollary 5.1, the forced discrete-time Hamilton's equations are given as follows.

$$h\hat{\Pi}_k = F_k J_d - J_d F_k^T, \tag{5.76}$$

$$R_{k+1} = R_k F_k, \tag{5.77}$$

$$x_{k+1} = x_k + \frac{h}{m}\gamma_k, \tag{5.78}$$

$$\gamma_{k+1} = \gamma_k + hf_{k+1} + hu_{k+1}^f, \tag{5.79}$$

$$\Pi_{k+1} = F_k^T \Pi_k + hM_{k+1} + hu_{k+1}^m, \tag{5.80}$$

where $f_k \in \mathbb{R}^3$ is the force due to the potential, and $M_k \in \mathbb{R}^3$ is the moment due to the potential. They are determined by the following expressions

$$f_k = -\frac{\partial U_k}{\partial x_k}, \tag{5.81}$$

$$\hat{M}_k = \frac{\partial U_k}{\partial R_k}^T R_k - R_k^T \frac{\partial U_k}{\partial R_k}. \tag{5.82}$$

**Optimal Control Problem**

The objective is to transfer the rigid body from a given initial condition $(R_0, x_0, \Omega_0, \dot{x}_0)$ to a desired terminal condition $(R^f, x^f, \Omega^f, \dot{x}^f)$ during a fixed terminal time $Nh$ using minimal control effort.

$$\begin{gathered}\text{For given: } (R_0, x_0, \Omega_0, \dot{x}_0), N, (R^f, x^f, \Omega^f, \dot{x}^f) \\ \min_u \left\{ \mathcal{J}_d = \sum_{k=0}^{N-1} \frac{h}{2} \langle W(u_k),\, u_k \rangle \right\}, \\ \text{such that } R_N = R^f,\, x_N = x^f,\, \Omega_N = \Omega^f,\, \dot{x}_N = \dot{x}^f, \\ \text{subject to (5.76)–(5.82),}\end{gathered}$$

where $W : \mathfrak{se}(3) \to \mathfrak{se}(3)^*$ is a weighting function given by $W(u) = (W^m u^m, W^f u^f)$ for symmetric positive definite matrices $W^f, W^m \in \mathbb{R}^{3\times 3}$.

**Discrete-time Necessary Conditions for Optimality**

This optimal control problem is a special case of Proposition 5.2, applied to the Lie group $\mathsf{SE}(3)$, when the free parameter of the discrete Lagrangian is $c = 1$ and the cost is $\phi_d = \frac{h}{2} \langle W(u_k),\, u_k \rangle$. From (5.23)–(5.27), necessary conditions for optimality can be obtained as follows.

- Optimality condition

$$u^m_{k+1} = (W^m)^{-1} \lambda^1_{R_k}, \quad u^f_{k+1} = (W^f)^{-1} \lambda^1_{x_k}. \tag{5.83}$$

- Multiplier equations

$$\begin{bmatrix} \lambda^1_{R_{k-1}} \\ \lambda^1_{x_{k-1}} \\ \lambda^2_{R_{k-1}} \\ \lambda^2_{x_{k-1}} \end{bmatrix} = \begin{bmatrix} A^{11}_k & A^{12}_k & A^{13}_k & A^{14}_k \\ A^{21}_k & A^{22}_k & A^{23}_k & A^{24}_k \\ 0 & 0 & A^{33}_k & A^{34}_k \\ A^{41}_k & A^{42}_k & 0 & 0 \end{bmatrix}^T \begin{bmatrix} \lambda^1_{R_k} \\ \lambda^1_{x_k} \\ \lambda^2_{R_k} \\ \lambda^2_{x_k} \end{bmatrix}, \tag{5.84}$$

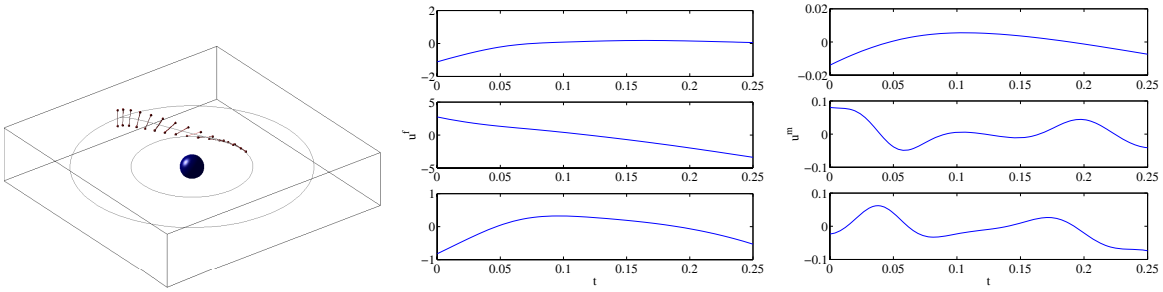
where matrices $A^{ij}_k \in \mathbb{R}^{3\times 3}$ for $i, j \in \{1, \ldots, 4\}$ are given by

$$\begin{aligned} A^{11}_k &= h\mathcal{M}_{x_{k+1}}, & A^{12}_k &= \frac{h^2}{m}\mathcal{M}_{x_{k+1}}, & A^{13}_k &= h\mathcal{M}_{R_{k+1}} A^{33}_k, \\ A^{14}_k &= F^T_k + h\mathcal{M}_{R_{k+1}} A^{34}_k + \widehat{F^T_k \Pi_k} A^{34}_k, & A^{21}_k &= h\mathcal{F}_{x_{k+1}}, & A^{22}_k &= I_{3\times 3} + \frac{h^2}{m}\mathcal{F}_{x_{k+1}}, \\ A^{23}_k &= h\mathcal{F}_{R_{k+1}} A^{33}_k, & A^{24}_k &= h\mathcal{F}_{R_{k+1}} A^{34}_k, & A^{33}_k &= F^T_k, \\ A^{34}_k &= hF^T_k \left\{ \text{tr}[F_k J_d]\, I_{3\times 3} - F_k J_d \right\}^{-1}, & A^{41} &= I_{3\times 3}, & A^{42} &= \frac{h}{m} I_{3\times 3}, \end{aligned}$$

The matrices $\mathcal{F}_{x_k}, \mathcal{F}_{R_k}, \mathcal{M}_{x_k}, \mathcal{M}_{R_k} \in \mathbb{R}^{3\times 3}$ are determined by the following equations.

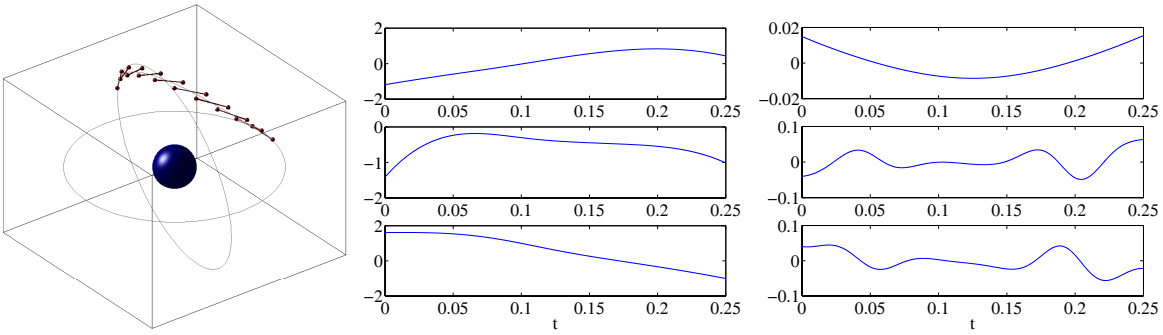
$$\delta f_k = \mathcal{F}_{x_k}\, \delta x_k + \mathcal{F}_{R_k}\, \eta_k, \tag{5.85}$$

$$\delta M_k = \mathcal{M}_{x_k} \delta x_k + \mathcal{M}_{R_k} \eta_k \tag{5.86}$$



(a) Spacecraft maneuver (b) Control force $u^f$ (c) Control moment $u^m$

Figure 5.7: Optimal orbit transfer of a dumbbell spacecraft (Orbital radius change)



(a) Spacecraft maneuver (b) Control force $u^f$ (c) Control moment $u^m$

Figure 5.8: Optimal orbit transfer of a dumbbell spacecraft (Orbital inclination change)

- Boundary conditions

$$R_N = R^f,\ x_N = x^f,\ \Omega_N = \Omega^f,\ V_N = V^f. \tag{5.87}$$

### Numerical Results

We study optimal maneuvers of a dumbbell spacecraft model presented in Section 3.3.6. The dumbbell spacecraft is composed of two spheres connected by a massless rod. We assume that it is acting under a central gravity field, and the mass of the spacecraft is negligible compared to the mass of a central body. The resulting model is referred to as a Restricted Full Two Body Problem (RF2BP).

Initially, the spacecraft is on a circular orbit. Two maneuvers are considered. The first maneuver is to change the orbital radius to twice the initial orbital radius, and the second maneuver is to increase the orbital inclination by $60\,\mathrm{deg}$. The maneuver time is chosen to be a quarter of the orbital period of the initial circular orbit. The boundary conditions are follows. Here 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, respectively.

(i) Orbital radius change

$$
\begin{aligned}
x_0 &= [1,0,0], & x^f &= [0,2,0],\\
R_0 &= \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & R^f &= \begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix},\\
\dot{x}_0 &= [0, 0.9835, 0], & \dot{x}^f &= [-0.7041, 0, 0],\\
\Omega_0 &= [0,0,0.9835], & \Omega^f &= [-0.3521, 0, 0].
\end{aligned}
$$

(ii) Orbital inclination change

$$
\begin{aligned}
x_0 &= [1,0,0], & x^f &= [-0.3536, 0.3536, 0.8660],\\
R_0 &= \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & R^f &= \begin{bmatrix} -0.7071 & 0.3535 & 0.6123 \\ -0.7071 & -0.3535 & -0.6123 \\ 0 & -0.8660 & 0.5 \end{bmatrix},\\
\dot{x}_0 &= [0, 0.9835, 0], & \dot{x}^f &= [-0.6954, -0.6954, 0],\\
\Omega_0 &= [0,0,0.9835], & \Omega^f &= [0,0,0.9835].
\end{aligned}
$$

For all cases, an 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 25th iterations. These convergence results are consistent with the quadratic convergence rates expected of Newton methods with accurately computed gradients. The optimal costs and the violation of the terminal boundary conditions are 22.23, $1.85 \times 10^{-14}$, and 13.03, $9.32 \times 10^{-15}$ respectively. The optimal maneuver and control input histories are shown at Figure 5.7 and 5.8.

### 5.2.5 Combinatorial Optimal Control of Spacecraft Formation Reconfiguration

The objective of spacecraft formation control is to use multiple spacecraft for cooperative missions such as long base-line interferometers. Formation reconfiguration can be classified into two types: (i) each spacecraft is assigned a specified location in the desired reconfigured formation; (ii) a specified location in the desired formation can be occupied by any single spacecraft (see Wang and Hadaegh 1999). In general, a formation is composed of identical spacecraft or groups of spacecraft of the same type, and the total fuel consumption depends on the permutational assignment of positions in the formation configuration as well as the maneuver of each spacecraft.

In this section, we study an optimal spacecraft formation problem integrated with an integer/combinatorial optimization approach for the assignment. Usually in combinatorial optimization problems for multiple agents, the dynamics of each agent is either ignored or simplified (see Savla et al. 2006). Here, we use the rigid body model shown in Section 5.2.4, including both translational dynamics and rotational attitude dynamics under a central gravitational potential. Thus, finding optimal control inputs for spacecraft assigned to a fixed desired location is demanding even if the

combinatorial assignment optimization problem is not considered. This combined problem is interesting and challenging since it requires combining a combinatorial optimization approach and an optimal control method for the non-trivial dynamics of multiple spacecraft evolving on $\mathsf{SE}(3)$.

**Combinatorial Optimal Control Problem**

We study an optimal formation control problem for $n$ identical rigid bodies where the dynamics of each body is described by (5.76)–(5.80). The objective is to find the optimal control forces and moments such that the group moves from a given initial configuration $(R_0^i, x_0^i, \Omega_0^i, \dot{x}_0^i)$ for $i \in \{1, 2, \ldots, n\}$ to a desired target $\mathfrak{T}$ during a given maneuver time $Nh$, where the superscript $i$ denotes the $i$-th rigid body.

More precisely, the desired target formation is defined as follows. We assume that the $n$ desired positions $\left\{x^{f,i}(\theta)\right\}_{i=1}^n$, at which all rigid bodies are located at the terminal maneuver time, are given as functions of parameters $\theta \in \mathbb{R}^l$. The desired attitude, angular velocity, and linear velocity at the terminal time, $(R^f, \Omega^f, \dot{x}^f)$, are assumed to be fixed and to be the same for all rigid bodies. This type of formation appears in spacecraft interferometric imaging applications, where spacecraft should be aligned in a single imagining plane while pointing at an object. In this case, the image quality performance of the spacecraft is invariant under a rotation of the spacecraft formation about the axis perpendicular to the imagining plane, passing through the object. The parameters $\theta$ allow us to account for this symmetry in combinatorial optimal formation control.

Since all rigid bodies are identical, there are $n!$ possible combinatorial assignments for $n$ rigid bodies to these $n$ desired locations. Let $\{a_{ij}\}$ be a $n \times n$ matrix composed of binary elements $\{0, 1\}$, referred to as an assignment or a permutation matrix. Each element of the assignment matrix $a_{ij}$ represent a possible assignment of the $i$-th rigid body to the $j$-th desired terminal position $x_d^j$. If $a_{ij} = 1$, the $i$-th rigid body is assigned to the $j$-th node, and if $a_{ij} = 0$, the $i$-th rigid body is not assigned to the $j$-th node. An assignment is valid when each rigid body is assigned to a single node, to which other rigid bodies are not assigned. Therefore, an assignment matrix is valid if $\sum_{j=1}^n a_{ij} = \sum_{i=1}^n a_{ij} = 1$ for all $i, j$, i.e. there is exactly a single 1 entry for every row and every column of the matrix $A$.

For a given assignment $A$, let $A_i$ be the assignment for the $i$-th rigid body. In other words, the $(i, A_i)$-th element of the assignment matrix $A$ is equal to 1. The $i$-th rigid body is assigned to the $A_i$-th desired location, $x^{f,A_i}$. Alternatively, an assignment can be expressed as a set of pairs $(i, A_i)$ for $1 \leq i \leq n$.

The target is defined in terms of a parameter $\theta$ and an assignment $A$ as follows.

$$\mathfrak{T}(\theta, A) = \left\{x^{f,A_i}(\theta)\right\}_{i=1}^n \in \mathbb{R}^{3n}.$$

Thus, for a given parameter $\theta \in \mathbb{R}^l$ and a given assignment $A$, the terminal boundary conditions for all rigid bodies are completely determined.

The optimal control problem for a formation reconfiguration of $n$ rigid bodies is formulated as

$$\text{given: } \left\{(R_0^i, x_0^i, \Omega_0^i, \dot{x}^i)\right\}_{i=1}^n, \, (R^f, \left\{x^{f,i}(\theta)\right\}_{i=1}^n, \Omega^f, \dot{x}^f), \, N,$$

$$\min_{u,\theta,A} \left\{ \mathcal{J} = \sum_{i=1}^{n} \sum_{k=0}^{N-1} \frac{h}{2} (u_{k+1}^{f,i})^T W_f u_{k+1}^{f,i} + \frac{h}{2} (u_{k+1}^{m,i})^T W_m u_{k+1}^{m,i} \right\},$$

$$\text{such that } \left\{(R_N^i, x_N^i, \Omega_N^i, \dot{x}_N^i) = (R^f, x^{f,A_i}(\theta), \Omega^f, \dot{x}^f)\right\}_{i=1}^n,$$

$$a_{ij} \in \{0,1\}, \quad \sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{n} a_{ij} = 1 \quad \text{for any } 1 \leq i,j \leq n,$$

$$\text{subject to (5.76)–(5.80).}$$

Since we have neglected the interaction between rigid bodies, the dynamics of the rigid bodies are only coupled through the terminal boundary conditions. If the parameter $\theta$ and the assignment $A$ are prescribed, then the optimal control problems for $n$ rigid bodies can be solved independently using the computational approach presented in Section 5.2.4. The formation cost is the sum of the resulting costs for each rigid body. Therefore, the optimal formation control problem for multiple rigid bodies consists in finding the optimal value of the parameter and the optimal assignment of rigid bodies among the $n!$ possible assignments. This is similar to the optimal formation reconfiguration problem presented in Junge et al. (2006) except that we include the combinatorial assignment problem explicitly.

**Hierarchical Optimization Approach**

We solve this combinatorial optimal control problem using a hierarchical optimization approach. We consider an optimal control problem for a fixed assignment, and we consider an assignment optimization problem for a fixed target parameter. These two optimization problems define the combinatorial optimal control problem.

*Optimal Control of $n$ Rigid Bodies.* We first solve the optimal formation control problem assuming that an assignment $A$ is pre-determined and fixed. Since the parameter $\theta$ completely defines the terminal configuration for the fixed assignment $A$, it also determines the corresponding cost by summing the cost of the optimal trajectories for each rigid body. Thus, the optimization problem can be decomposed into an outer optimization problem to find the optimal value of $\theta$ that minimizes the total cost, and an inner optimization problem to find the optimal control inputs for the given value of $\theta$. This is a consequence of the fact that

$$\min_{u,\theta} \mathcal{J}(u,\theta) = \min_{\theta'} \left\{ \min_{u} \left\{ \mathcal{J}(u,\theta) | \theta = \theta' \right\} \right\}. \tag{5.88}$$

The inner optimization problem is solved by using the computational approach described in Section 5.2.4. The optimal value of $\theta$ is found by using a parameter optimization method. From the optimality condition given by (5.83), the cost is dependent on the multipliers. Since the sen-

sitivity of the multipliers with respect to the terminal boundary condition is available through the computational approach of the optimal control problem given by (5.28), it is possible to obtain the sensitivity of the cost with respect to the target parameter by applying the chain rule properly. Then, a gradient-based parameter optimization technique can be applied. We apply the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method presented in Kelley (1995).

*Assignment Optimization.* Now, we solve an assignment optimization problem assuming that the target parameter $\theta$ is fixed. For the given value of $\theta$, the $n$ desired points $\left\{x^{f,i}\right\}_{i=1}^{n}$, at which all rigid bodies are located at the terminal time, are completely defined. Thus, there are $n!$ possible combinatorial assignments.

Let $\left\{c^{ij}\right\}$ be a $n \times n$ matrix, referred to as a cost matrix. Each element $c^{ij}$ represents the optimal cost of the $i$-th rigid body transferred to the $j$-th desired location. For an assignment $\{a_{ij}\}$, the total cost is given by $\mathcal{J} = \sum_{i,j=1}^{n} c^{ij} a_{ij}$. The optimal assignment problem is given by

$$\min_{a_{ij}} \sum_{i,j=1}^{n} c^{ij} a_{ij},$$

subject to $\sum_{j=1}^{n} a_{ij} = 1$, $\sum_{i=1}^{n} a_{ij} = 1$, and $a_{ij} \in \{0, 1\}$ for all $i, j \in \{1, 2, \ldots, n\}$.

Since we assume that there is no interaction between rigid bodies, the cost matrix is independent of the assignment. For the given value of the target parameter $\theta$, we must solve at most $n^2$ optimal control problems to obtain the cost matrix. Once we have the complete cost matrix, the optimal assignment can be obtained by comparing costs for all possible assignments or by using the Hungarian method (see Murty 1985).

It is often expensive to obtain the cost matrix. Each element of the cost matrix is a solution of an optimal control problem presented in Section 5.2.4. For the formation optimization problem, we need to find the cost matrix with varying values of the target parameter $\theta$. Thus, the objective of this subsection is to find the optimal assignment without solving all $n^2$ optimal control problems. We start with an initial single spacecraft optimal trajectory computation, and use its optimal cost and sensitivities to populate the remaining entries of the cost matrix. We construct a combinatorial assignment method using these approximations.

(i) Guess an initial assignment and solve the corresponding optimal control problems for this assignment.

(ii) Estimate the cost matrix using linear approximations.

(iii) Find a new assignment using the estimated cost matrix, and solve the corresponding optimal control problems for this assignment.

(iv) Find the best assignment using all of the solutions of the optimal control problems obtained so far.

(v) Construct a second-order approximation based on the best assignment, and estimate the cost matrix.

(vi) Repeat (iii)-(v) until the best assignment is repeated for a pre-determined multiple times in a row at (v).

*Combinatorial Optimal Control.* We have presented two optimization approaches; finding the optimal value of the target parameter for a given assignment, and finding the optimal assignment for a given value of the target parameter. We integrate both methods using a hierarchical optimization approach.

The original optimization problem is stated as finding the optimal control inputs, target parameter, and assignment that minimizes the total cost:

$$\min_{u,\theta,A} \mathcal{J}(u,\theta,A).$$

Equivalently, this can be stated as finding the optimal assignment over the optimal control inputs and the target parameter:

$$\min_{A'} \left\{ \min_{u,\theta} \left\{ \mathcal{J}(u,\theta,A) | A = A' \right\} \right\}.$$

In the inner stage, we optimize the target parameter and the control inputs using a continuous optimization approach, and in the outer stage, we find the optimal assignment using the combinatorial optimization approach. The optimization process is terminated when the iterations yield a solution that is optimal for both the inner and outer optimization stages.

**Numerical Results**

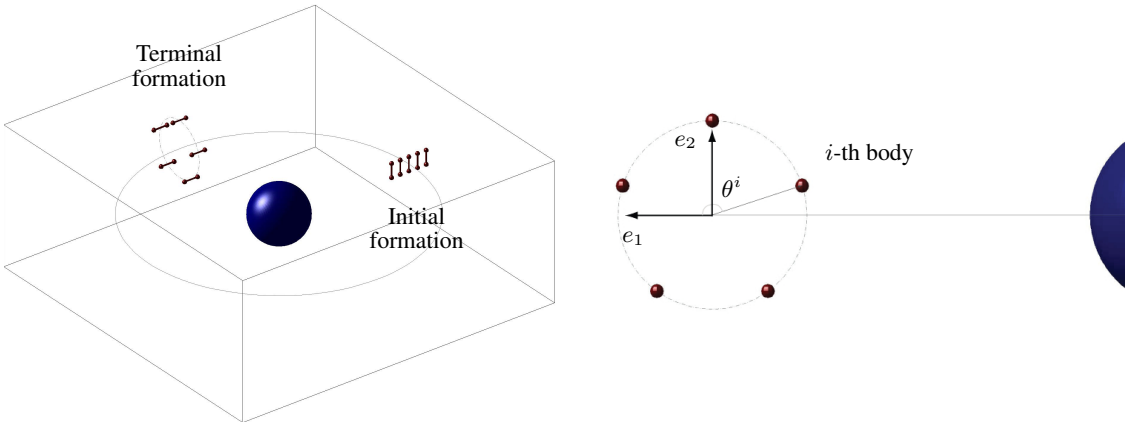
We study a maneuver involving 5 identical rigid spacecraft orbiting in a central gravity field. Each spacecraft is modeled as a dumbbell, which consists of two equal spheres and a massless rod, as presented in Section 5.2.4.

The spacecraft are initially aligned along a radial direction as shown in Figure 5.9(a). At the terminal time, we require that the spacecraft are equally distributed on a target circle described by the location of its center $x_\circ \in \mathbb{R}^3$, the radius $r_\circ \in \mathbb{R}$, and the unit normal vector $n_\circ \in \mathsf{S}^2$. Let $\theta^i \in \mathsf{S}^1$ be the angle of the $i$-th spacecraft on the target circle as shown in Figure 5.9(b). We choose the target parameter as the angle of the first rigid body. The target $\mathfrak{T}$ is given by

$$\mathfrak{T}(\theta^1, A) = \left\{ x_\circ + r_\circ \cos\theta^i e_1 + r_\circ \sin\theta^i e_2 \right\}_{i=1}^n,$$

where $e_1 = \frac{x_\circ}{\|x_\circ\|}$, $e_2 = e_1 \times n_\circ$ are unit vectors in the target plane, and the angle $\theta^i$ is chosen to distribute the spacecraft uniformly on the circle, i.e. $\theta^i = \theta^1 + 2\pi/5(A_i - i)$.

Since the target parameter $\theta^1$ determines the terminal position of the first spacecraft completely, we require that the first spacecraft be assigned to the first desired location, i.e. $A_1 = 1$. There




(a) The initial formation and the terminal formation

(b) The terminal formation on a target circle

Figure 5.9: The initial formation and the desired terminal formation of spacecraft

remain $4!$ assignments for the other four spacecraft. Thus, the optimization parameters are the angle of the first spacecraft on the target circle, the $4!$ assignments for the remaining spacecraft, and the control inputs.

The iteration procedure for a particular numerical implementation of the optimization are shown as follows with computation time on an Intel Pentium M 1.73GHz processor using MATLAB.

i) The initial guess of the assignment is given by $A = \{(1,1),(2,4),(3,2),(4,3),(5,5)\}$.

ii.a) For the given assignment, the optimal value of $\theta^1 = 2.4520$ is obtained in $48.32$ minutes with $\mathcal{J} = 8.6984$.

ii.b) For the given value of $\theta^1$, the optimal assignment of $A = \{(1,1),(2,5),(3,2),(4,3),(5,4)\}$ is obtained in $3.04$ minutes with cost $\mathcal{J} = 8.6905$.

ii.c) For the given $\theta^1$ and the given assignment, we check $\frac{\partial \mathcal{J}}{\partial \theta^1} = 2.42 \times 10^{-2}$. Repeat iteration.

iii.a) For the given assignment, the optimal value of $\theta^1 = 2.5084$ is obtained in $12.69$ minutes with $\mathcal{J} = 8.6898$.

iii.b) For the given value of $\theta^1$, the same assignment of $A = \{(1,1),(2,5),(3,2),(4,3),(5,4)\}$ is obtained in $2.98$ minutes with cost $\mathcal{J} = 8.6898$.

iii.c) For the given $\theta^1$ and the corresponding optimal assignment, we check $\frac{\partial \mathcal{J}}{\partial \theta^1} = 9.99 \times 10^{-5}$.

iv) The optimization is terminated in $67.03$ minutes with $\mathcal{J} = 8.6898$ for $\theta^1 = 2.5084$ and $A = \{(1,1),(2,5),(3,2),(4,3),(5,4)\}$.

The corresponding maneuvers for the spacecraft are shown in Figure 5.10. At each iteration, we use the optimization data accumulated in the previous iterations in order to initialize the initial multiplier values for the optimal control problems. This reduces the computation time as the iterations proceed.

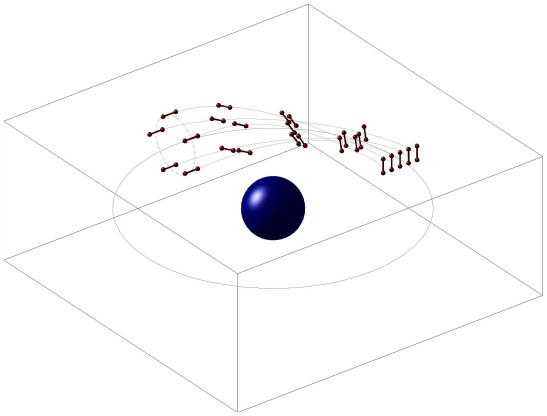
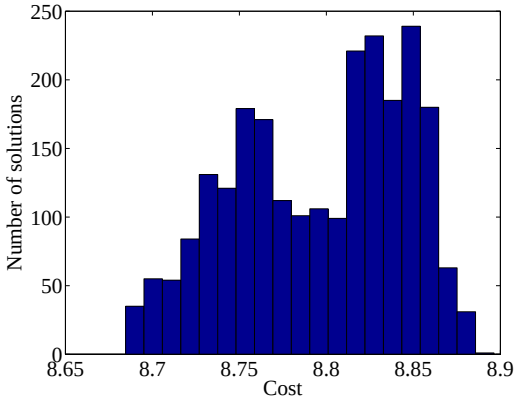
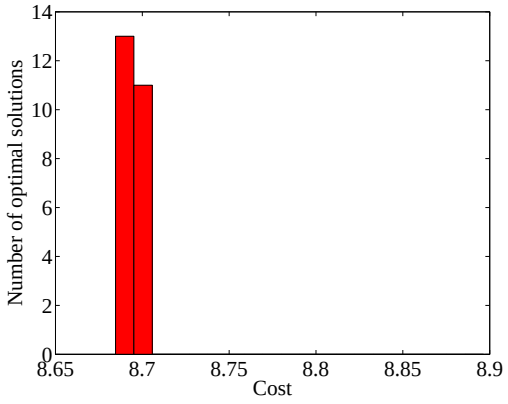


Figure 5.10: Optimal spacecraft formation reconfiguration maneuver




(a) Histogram of total costs for 2400 solutions with varying target parameters and assignments




(b) Histogram of total cost for 24 optimized solutions with varying initial guesses of assignment

Figure 5.11: Distribution of the total costs before and after optimization

In order to estimate the distribution of all possible solutions, we uniformly discretize the interval $[0, 2\pi)$ using 100 points for the target parameters and we find the total costs of $4!$ assignments for each value of the target parameter. The histogram for total costs of the corresponding $100 \times 4! = 2400$ solutions is shown in Figure 5.11(a).

Numerical simulations show that the optimized solution obtained depends more strongly on the initial guess of the assignment than the initial guesses for the target parameter and the initial multiplier values. We repeat the numerical optimization for all possible $4!$ initial guesses of the assignments. Figure 5.11(b) shows the histogram of the optimized total costs for varying initial assignment. Six initial assignments converge to the optimal solution with the minimal cost $\mathcal{J} = 8.6898$ in the pre-computed $100 \times 4! = 2400$ solutions; this is assumed to be close to the global optimal solution. Seven initial assignments converged to a local optimal solution with $\mathcal{J} = 8.6985$.

### 5.2.6 Fuel Optimal Control of a 3D Pendulum on a Cart

Consider a 3D pendulum whose pivot is fixed to a cart moving on a horizontal plane, introduced in Section 2.3.4. In this section, we study an optimal control problem of the 3D pendulum on a cart model. We assume that an external control force is applied to the cart, and the 3D pendulum is not actuated. Therefore, the 3D pendulum on a cart model is underactuated, and the motion of the pendulum is achieved through the coupling between the cart dynamics and the pendulum dynamics. We solve this optimal control problem using the direct optimal control approach discussed in Section 5.1.5.

The 3D pendulum on a cart has symmetry represented by a rotation about the vertical axis, and the total angular momentum about the vertical axis is preserved. In this optimal control problem, the external control force acting on the cart breaks this symmetry, and the total angular momentum is not conserved 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 5.2.3, and from the optimal control of connected rigid bodies, discussed in Section 5.2.7, where the control inputs respect the symmetry of free dynamics, and the momentum map is preserved in the controlled dynamics.

We develop forced discrete-time Hamilton's equations from the Lie group variational integrator presented in Section 3.3.4, and a mathematical formulation of the optimal control problem is presented. Computational results obtained by the direct optimal control approach are shown.

#### Force Hamilton's Equations

The discrete Lagrangian of the 3D pendulum on a cart is chosen as

$$L_d(R_k, x_k, y_k, F_k, \Delta x_k, \Delta y_k) = \frac{1}{2h}(M+m)((\Delta x_k)^2 + (\Delta y_k)^2) + \frac{1}{h}\mathrm{tr}[(I - F_k)J_d] + \frac{m}{h}\Delta x_k e_1^T R_k(F_k - I)\rho_c + \frac{m}{h}\Delta y_k e_2^T F_k(R_k - I)\rho_c + \frac{h}{2}mge_3^T R_k\rho_c + \frac{h}{2}mge_3^T R_k F_k\rho_c.$$

We assume that an external control force $u = (u_x, u_y) \in \mathbb{R}^{2*}$ is applied to the cart along the horizontal plane. The discrete generalized forces are chosen as $u_{d_k}^- = (0, \frac{h}{2}u_k)$, $u_{d_k}^+ = (0, \frac{h}{2}u_{k+1}) \in \mathfrak{so}(3)^* \times \mathbb{R}^{2*}$. From Proposition 5.1, the forced Hamilton's equations for the 3D pendulum on a cart are given as follows.

$$p_{x_k} = \frac{1}{h}(M+m)(x_{k+1} - x_k) + \frac{m}{h}e_1(R_{k+1} - R_k)\rho_c + \frac{h}{2}u_{x_k}, \tag{5.89}$$

$$p_{y_k} = \frac{1}{h}(M+m)(y_{k+1} - y_k) + \frac{m}{h}e_2(R_{k+1} - R_k)\rho_c + \frac{h}{2}u_{y_k}, \tag{5.90}$$

$$\hat{p}_{\Omega_k} = \frac{1}{h}(F_k J_d - J_d F_k^T) + \left\{\frac{m}{h}(x_{k+1} - x_k)\hat{\rho}_c R_k^T e_1 + \frac{m}{h}(y_{k+1} - y_k)\hat{\rho}_c R_k^T e_2 - \frac{h}{2}mg\hat{\rho}_c R_k^T e_3\right\}^\wedge, \tag{5.91}$$

$$R_{k+1} = R_k F_k, \tag{5.92}$$

$$p_{x_{k+1}} = p_{x_k} + \frac{h}{2} u_{x_{k+1}}, \tag{5.93}$$

$$p_{y_{k+1}} = p_{y_k} + \frac{h}{2} u_{y_{k+1}}, \tag{5.94}$$

$$\begin{aligned} \hat{p}_{\Omega_{k+1}} = {} & \frac{1}{h}(J_d F_k - F_k^T J_d) \\ & + \left\{ \frac{m}{h}(x_{k+1} - x_k)\hat{\rho}_c R_{k+1}^T e_1 + \frac{m}{h}(y_{k+1} - y_k)\hat{\rho}_c R_{k+1}^T e_2 + \frac{h}{2} mg \hat{\rho}_c R_{k+1}^T e_3 \right\}^\wedge . \end{aligned} \tag{5.95}$$

**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 control effort.

$$\begin{gathered} \text{For given: } (R_0, x_0, y_0, \Omega_0, \dot{x}_0, \dot{y}_0), N, (R^f, x^f, y^f, \Omega^f, \dot{x}^f, \dot{y}^f) \\ \min_{u_k} \left\{ \mathcal{J}_d = \sum_{k=0}^{N} \frac{h}{2} u_k^T W u_k \right\}, \\ \text{such that } (R_N, x_N, y_N, \Omega_N, \dot{x}_N, \dot{y}_N) = (R^f, x^f, y^f, \Omega^f, \dot{x}^f, \dot{y}^f), \\ \text{subject to (5.89)–(5.95)}, \end{gathered}$$

where $W \in \mathbb{R}^{2\times 2}$ is a symmetric positive-definite matrix.

**Numerical Results**

Properties of the 3D pendulum and the cart are chosen as

$$M = m = 1\,\text{kg}, \quad \rho_c = [0.25, 0.25, 1]\,\text{m}, \quad J = \begin{bmatrix} 1.09 & -0.06 & -0.25 \\ -0.05 & 1.10 & -0.25 \\ -0.25 & -0.25 & 0.15 \end{bmatrix} \text{kgm}^2.$$

The desired maneuver is a rest-to-rest 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

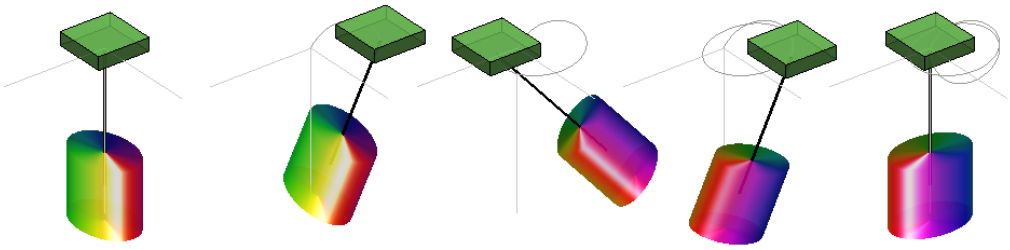
$$\begin{gathered} R_0 = I, \quad \Omega_0 = 0, \quad x_0 = y_0 = 0, \quad \dot{x}_0 = \dot{y}_0 = 0, \\ R^f = \exp(\theta \hat{e}_3), \quad \Omega^f = 0, \quad x^f = y^f = 0, \quad \dot{x}^f = \dot{y}^f = 0, \end{gathered}$$

where the rotation angle $\theta \in \mathsf{S}^1$ varies as $\theta = \frac{\pi}{2}$ and $\theta = \pi$. The maneuver time is $t_f = 2$ 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 dynamics and the pendulum dynamics.

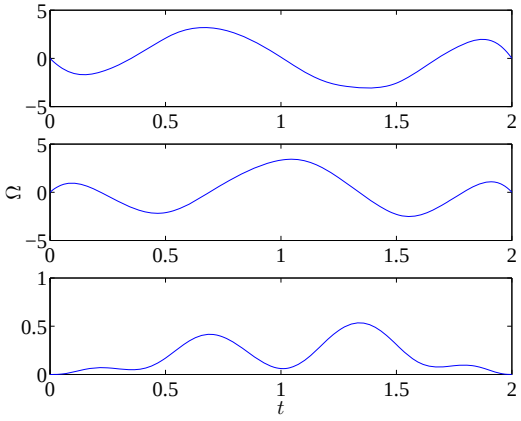
We apply the direct optimal control approach discussed in Section 5.1.5. Each component of the control inputs $u = (u_x, u_y)$ is parameterized by 7 points, and the control inputs are obtained by cubic spline interpolation. The resulting 14 control input parameters are optimized using a sequential quadratic programming method to satisfy the terminal boundary conditions while minimizing the cost function.

The optimized maneuver, the angular velocity of the pendulum, the velocity of the cart, and the trajectory of the cart in the horizontal plane, for $\theta = \frac{\pi}{2}$, and $\theta = \pi$, are shown in Figure 5.12 and Figure 5.13, respectively. Blue circles denote the optimized control input parameters. The optimal cost and the violation of the terminal boundary conditions are $\mathcal{J}_d = 171.60$, $7.65 \times 10^{-6}$, and $\mathcal{J}_d = 297.43$, $1.83 \times 10^{-8}$, respectively for each case. The optimal motion of the cart on the horizontal plane consists of loops, and the optimal optimal maneuver of the 3D pendulum consists of large angle rotations.

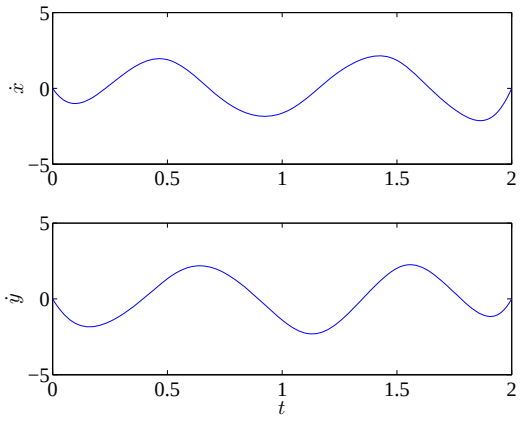
This demonstrates the advantage of the computational geometric optimal control approach: it is difficult to study this kind of aggressive maneuver of a multibody system using local coordinates, due to the singularity and the complexity. The presented computational geometric optimal control approach completely utilizes the nonlinear coupling between the cart and the pendulum dynamics to obtain a nontrivial optimal maneuver of the 3D pendulum on a cart.



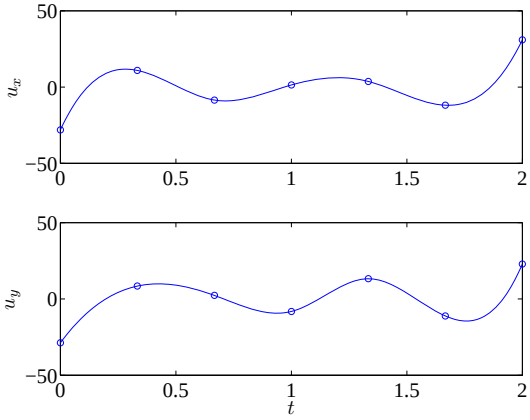
(a) Optimal maneuver



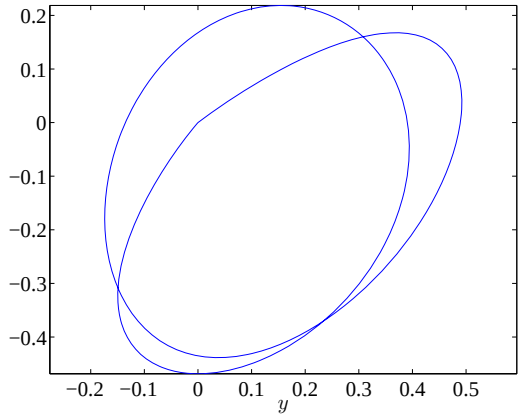

(b) Angular velocity $\Omega$




(c) Velocity of cart $\dot{x}, \dot{y}$

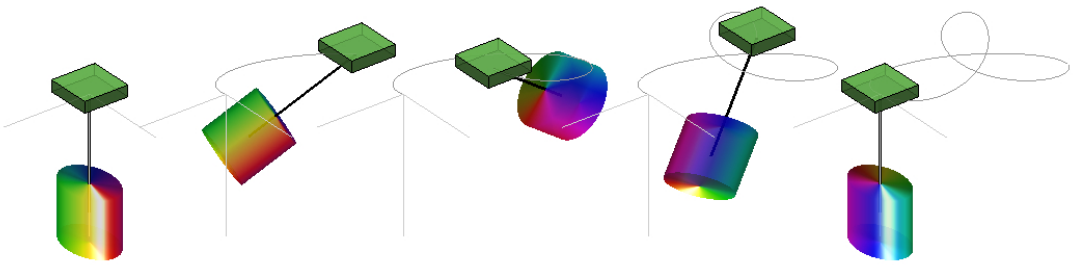



(d) Control input $u$

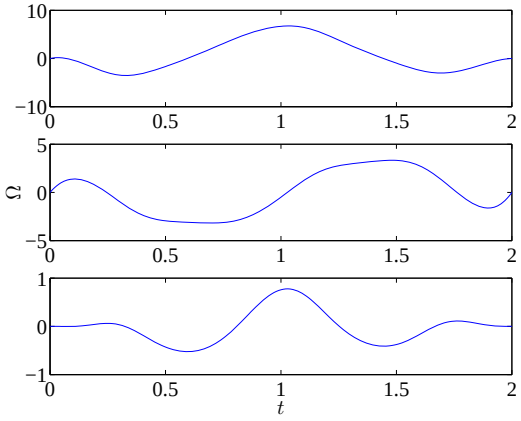



(e) Trajectory of cart in horizontal plane

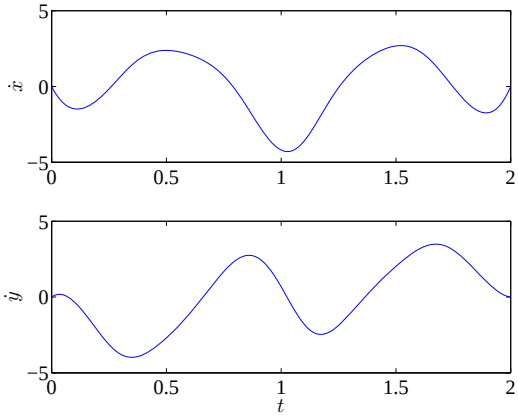
Figure 5.12: Optimal control of a 3D pendulum on a cart ($\theta = \pi/2$)



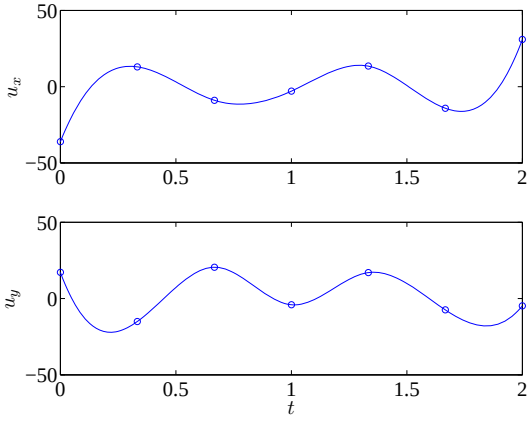
(a) Optimal maneuver



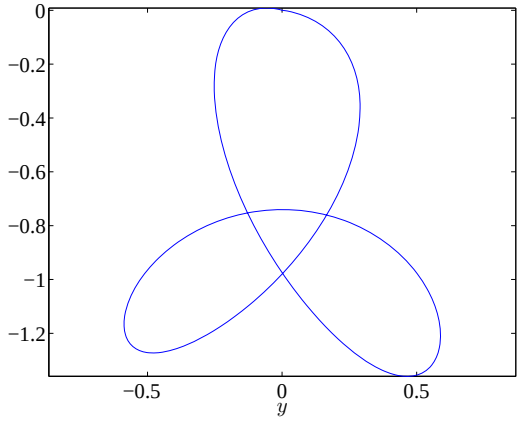

(b) Angular velocity $\Omega$




(c) Velocity of cart $\dot{x}, \dot{y}$




(d) Control input $u$




(e) Trajectory of cart in horizontal plane

Figure 5.13: Optimal control of a 3D pendulum on a cart ($\theta = \pi$)

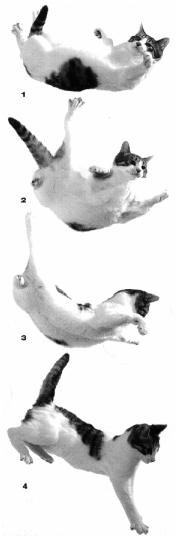


Figure 5.14: Falling cat problem

### 5.2.7 Fuel Optimal Control of Two Rigid Bodies Connected by a Ball Joint with Symmetry

Consider two rigid bodies connected by a ball joint, described in Section 2.3.7. In this section, we study an optimal control problem for the connected rigid body model. Similar to the optimal control of the 3D pendulum with symmetry discussed in Section 5.2.3, we construct a control input that respects a symmetry of the system, and the optimal maneuver is indirectly achieved by geometric phase effect. We solve this optimal control problem using a direct optimal control approach discussed in Section 5.1.5.

We develop forced discrete-time Hamilton's equations from the Lie group variational integrator presented in Section 3.3.7, and a mathematical formulation of the optimal control problem is presented. Computational results obtained by the direct optimal control approach are shown.

#### Forced Hamilton's Equations

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 $\mathsf{SO}(3)$. Due to these symmetries, the corresponding momentum maps are preserved, and the configuration manifold can be reduced to a quotient space.

In this optimal control problem, we reduce the configuration manifold to $\mathsf{SO}(3) \times \mathsf{SO}(3)$ using the symmetry of the translational action of $\mathbb{R}^3$. The corresponding value of the total linear momentum is zero. The resulting connected rigid bodies with a fixed mass center represents a freely rotating system of coupled rigid bodies; this is closely related to the falling cat problem (see Enos 1993). Interestingly, a cat, when dropped back-first from rest, is able to reorient itself and land on its feet. A proper change of shape of the body yields a rotation of the cat according to the geometric phase effect (see Montgomery 1991).

Similar to the falling cat problem, we assume that an internal control moment is applied at the joint, that controls the relative attitude between the two rigid bodies. 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.

The reduced discrete equations of motion on $\mathsf{SO}(3) \times \mathsf{SO}(3)$ can be derived by applying the discrete Routh reduction procedure discussed in Section 3.1.4. Alternatively, they can be directly derived from (3.156)–(3.161) by using the fact that the total linear momentum, denoted by $p_{3_k}$, is conserved in the absence of a potential: we find an expression for $(x_{k+1} - x_k)$ from (3.159), and substitute it into (3.156) and (3.158). The resulting forced Hamilton's equations are given as follows.

$$\begin{aligned}\hat{p}_{1_k} = \frac{1}{h}\left\{F_{1_k}(J_{d_1} - \alpha m_1 d_1 d_1^T) - (J_{d_1} - \alpha m_1 d_1 d_1^T)F_{1_k}^T\right\} \\ - \beta\frac{m_1}{h}(R_{1_k}^T R_{2_k} F_{2_k} d_2 d_1^T - d_1 d_2^T F_{2_k}^T R_{2_k}^T R_{1_k}) + \beta\frac{m_1}{h}(R_{1_k}^T R_{2_k} d_2 d_1^T - d_1 d_2^T R_{2_k}^T R_{1_k}),\end{aligned} \tag{5.96}$$

$$\begin{aligned}\hat{p}_{2_k} = \frac{1}{h}\left\{F_{2_k}(J_{d_2} - \beta m_2 d_2 d_2^T) - (J_{d_2} - \beta m_2 d_2 d_2^T)F_{2_k}^T\right\} \\ - \alpha\frac{m_2}{h}(R_{2_k}^T R_{1_k} F_{1_k} d_1 d_2^T - d_2 d_1^T F_{1_k}^T R_{1_k}^T R_{2_k}) + \alpha\frac{m_2}{h}(R_{2_k}^T R_{1_k} d_1 d_2^T - d_2 d_1^T R_{1_k}^T R_{2_k}),\end{aligned} \tag{5.97}$$

$$R_{i_{k+1}} = R_{i_k} F_{i_k}, \tag{5.98}$$

$$p_{1_{k+1}} = F_{1_k}^T(p_{1_k} - (B_{1_k} - B_{1_k}^T)^\vee) + hR_{1_{k+1}}^T u_{k+1}, \tag{5.99}$$

$$p_{2_{k+1}} = F_{2_k}^T(p_{2_k} - (B_{2_k} - B_{2_k}^T)^\vee) - hR_{2_{k+1}}^T u_{k+1}, \tag{5.100}$$

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

$$B_{i_k} = \frac{m_i}{h}(F_{i_k} - I)d_i\left\{-\alpha R_{1_k}(F_{1_k} - I)d_1 - \beta R_{2_k}(F_{2_k} - I)d_2\right\}^T R_{i_k}. \tag{5.101}$$

**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 control effort.

$$\begin{gathered}\text{For given: } (R_{1_0}, R_{2_0}, \Omega_{1_0}, \Omega_{2_0}), N, (R_1^f, R_2^f, \Omega_1^f, \Omega_2^f) \\ \min_{u_k}\left\{\mathcal{J}_d = \sum_{k=0}^{N}\frac{h}{2}u_k^T W u_k\right\}, \\ \text{such that } (R_{1_N}, R_{2_N}, \Omega_{1_N}, \Omega_{2_N}) = (R_1^f, R_2^f, \Omega_1^f, \Omega_2^f), \\ \text{subject to (5.96)–(5.101),}\end{gathered}$$

where $W \in \mathbb{R}^{3\times3}$ is a symmetric positive-definite matrix. In particular, we choose attitude maneuvers that can be described by rest-to-rest rotations of the entire system while the relative attitude at the terminal time is the same as that at the initial time.

**Numerical Results**

The properties of rigid bodies are chosen as

$$m_1 = 1.5\text{kg}, \quad J_1 = \begin{bmatrix} 0.18 & 0.32 & 0.32 \\ 0.32 & 1.88 & -0.06 \\ 0.32 & -0.06 & 1.86 \end{bmatrix} \text{kg}\cdot\text{m}^2, \quad d_1 = [-1.08, 0.20, 0.20]\text{m},$$

$$m_2 = 1\text{kg}, \quad J_2 = \begin{bmatrix} 0.11 & -0.18 & -0.18 \\ -0.18 & 0.89 & -0.04 \\ -0.18 & -0.04 & 0.88 \end{bmatrix} \text{kg}\cdot\text{m}^2, \quad d_2 = [0.9, 0.2, 0.2]\text{m}.$$

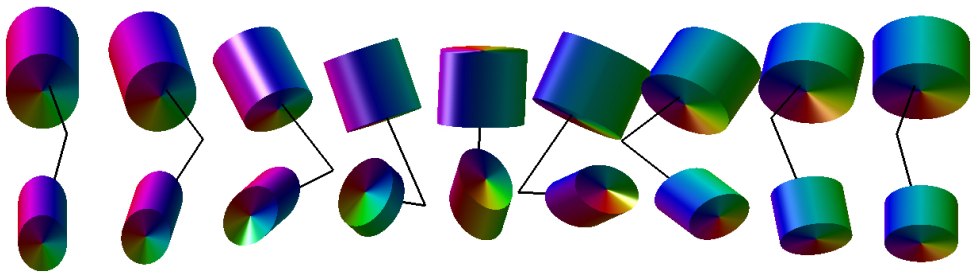
The desired maneuver is a rest-to-rest rotation about the $x$ axis.

$$R_{1_0} = I, \quad \Omega_{1_0} = 0, \quad R_{2_0} = I, \quad \Omega_{2_0} = 0,$$
$$R_1^f = \exp(\theta\hat{e}_1), \quad \Omega_1^f = 0, \quad R_2^f = \exp(\theta\hat{e}_1), \quad \Omega_2^f = 0,$$

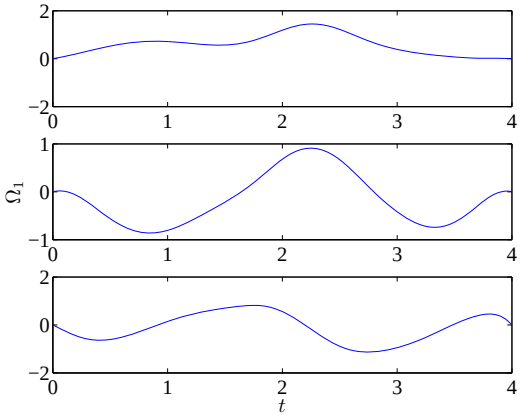
where the rotation angle is denoted by $\theta$. Two cases are considered, when $\theta = \frac{\pi}{2}$, and $\theta = \pi$. The maneuver time is $t_f = 4$ sec, and the step size is $h = 0.01$.

We apply the direct optimal control approach discussed in Section 5.1.5. 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 using a sequential quadratic programming method to satisfy the terminal boundary conditions while minimizing the cost function. The terminal angular velocity constraint for the second body is not imposed since it is satisfied automatically by the other three constraints due to the total angular momentum preservation property. By formulating the optimization process this way, we eliminate a source of numerical ill-conditioning. This is similar to the modified computational approach for the indirect optimal control discussed in Section 5.2.3 for the optimal control of the 3D pendulum with symmetry.

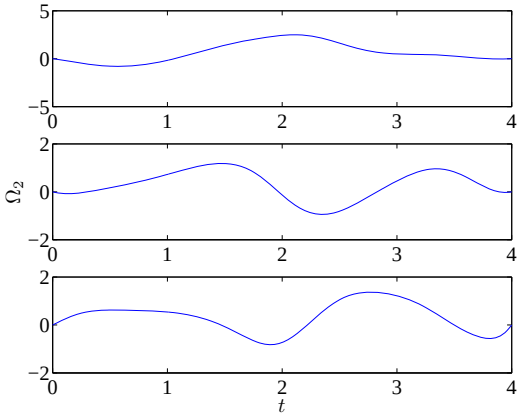
The optimized maneuver, the angular velocity, the control input trajectories, and the total angular momentum in the reference frame are shown in Figure 5.15 and Figure 5.16. The optimal cost and the violation of the terminal boundary conditions are $\mathcal{J}_d = 0.154$, $1.19\times10^{-8}$, and $\mathcal{J}_d = 0.574$, $2.48 \times 10^{-8}$, respectively for $\theta = \frac{\pi}{2}$, and $\theta = \pi$. Throughout this complicated maneuver, the total angular momentum is zero, and the rotation about the $e_1$ axis is purely caused by 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.



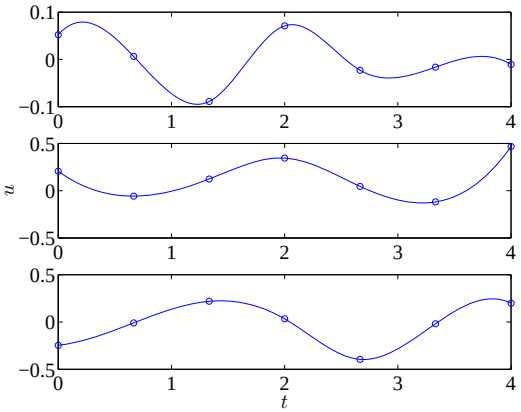
(a) Optimal maneuver



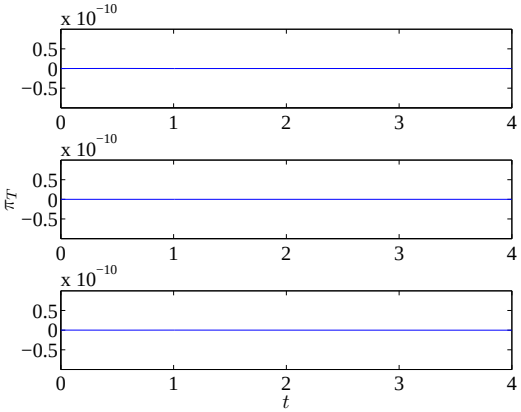

(b) Angular velocity $\Omega_1$




(c) Angular velocity $\Omega_2$

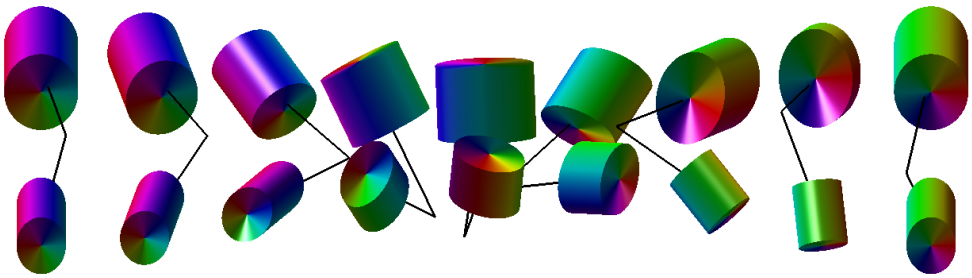



(d) Control input $u$

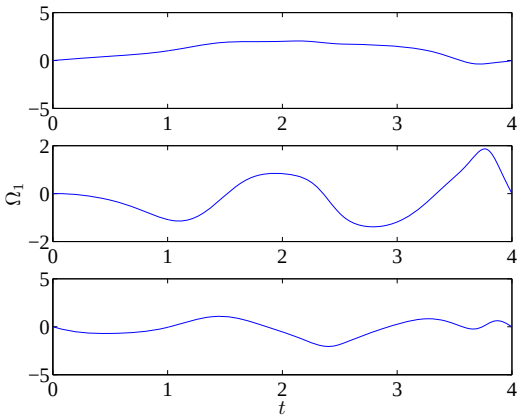



(e) Total angular momentum

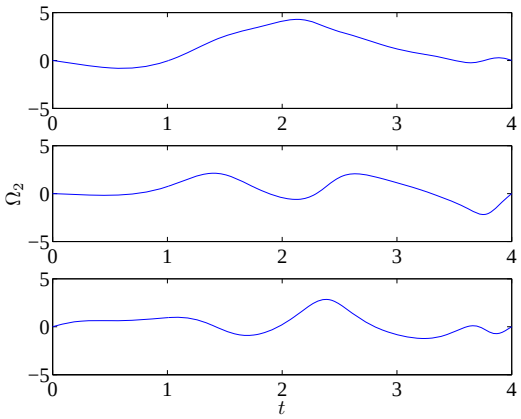
Figure 5.15: Optimal control of two connected rigid bodies ($\theta = \pi/2$)



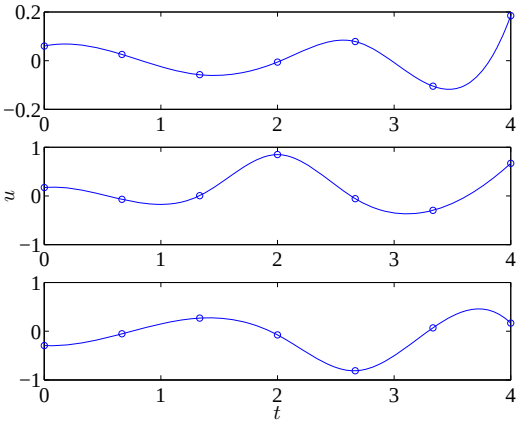
(a) Optimal maneuver



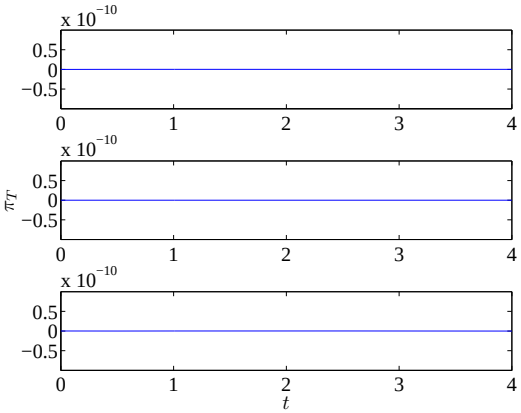

(b) Angular velocity $\Omega_1$




(c) Angular velocity $\Omega_2$




(d) Control input $u$




(e) Total angular momentum

Figure 5.16: Optimal control of two connected rigid bodies ($\theta = \pi$)

## 5.3 Conclusions

In this chapter, a computational geometric approach for an optimal control problem of rigid bodies has been discussed. The essential idea is formulating a discrete-time optimal control problem using a Lie group variational integrator, and applying standard optimal control approaches, such as an indirect optimal control approach or a direct optimal control approach to the discrete-time equations of motion. This method is in contrast to the usual optimal control approach, where the discretization appears in the last stage to find the optimal control inputs numerically.

The computational geometric optimal control approach has substantial advantages in terms of preserving the geometric properties of the optimality conditions. As discussed in Chapter 3, a discrete flow of a Lie group variational integrator has desirable geometric properties, and it is more reliable over long time periods. The computational geometric optimal control approach inherits the desirable properties of the Lie group variational integrator. In necessary conditions for optimality, multiplier equations can be considered as a dual system of linearized equations of motion. Since the linearized flow of a Lagrangian/Hamiltonian system is symplectic, the multiplier equations also have geometric properties. The discrete-time necessary conditions, presented in Section 5.1.3, preserve the geometric properties of the optimality conditions, as they are derived from a symplectic discrete flow.

The computational geometric optimal control approach allows us to find the optimal control inputs efficiently. In indirect optimal control, the shooting method may be prone to numerical ill-conditioning, since a small change in the initial multipliers can cause highly nonlinear behavior of the terminal conditions. It is difficult to compute the Jacobian matrix for Newton iterations accurately, and consequently, the numerical error may not converge to machine precision. However, as shown in Figure 5.1(c), 5.2(c), 5.5(e), and 5.6(e), 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 caused by the numerical algorithm, and therefore, we obtain more accurate sensitivities along the optimal solution.

Another advantage of 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. For example, we can study large-angle optimal attitude maneuvers as presented in Section 5.2.1, Section 5.2.6, and Section 5.2.7, and we can consider nontrivial coupling effects between the translational maneuver and the rotational maneuver of a rigid body, as presented in Section 5.2.4.

CHAPTER 6

# CONCLUSIONS

## 6.1 Conclusions

In this dissertation, computational geometric mechanics and optimal control have been developed for dynamic systems evolving on a Lie group, with applications to rigid body dynamics. Theoretical results are presented both for continuous-time dynamic systems and for discrete-time dynamic systems in parallel, and they are applied to several nontrivial rigid body dynamics.

| **Geometric Mechanics for Rigid Bodies on a Lie Group** | **Computational Geometric Mechanics for Rigid Bodies on a Lie Group** |
|---|---|
| Generalized Euler-Poincaré equation on $\mathsf{G}$<br>Euler-Lagrange equations on $(\mathsf{S}^2)^n$ | Lie group variational integrator on $\mathsf{G}$<br>Homogeneous variational integrator on $(\mathsf{S}^2)^n$ |
| Planar pendulum, 3D pendulum, 3D pendulum with an internal degree of freedom, 3D pendulum on a cart,<br>Single rigid body, Full body problem, Two rigid bodies connected by a ball joint<br>Double Spherical Pendulum, $n$-body problem on a sphere, Pure bending of elastic rod,<br>Spatial array of magnetic dipoles, Molecular dynamics on a sphere | |

| **Geometric Optimal Control for Rigid Bodies on a Lie Group** | **Computational Geometric Optimal Control for Rigid Bodies on a Lie Group** |
|---|---|
| Necessary conditions for optimality on $\mathsf{G}$ | Discrete-time necessary conditions for optimality on $\mathsf{G}$<br>Direct optimal control approach |
| Fuel optimal attitude control of a spacecraft, Time optimal attitude control of a free rigid body,<br>Fuel optimal attitude control of a 3D pendulum with symmetry, Fuel optimal control of a rigid body,<br>Combinatorial optimal control of spacecraft formation reconfiguration,<br>Fuel optimal control of a 3D pendulum on a cart,<br>Fuel optimal control of two rigid bodies connected by a ball joint | |

The continuous-time equations of motion for dynamic system on a Lie group are presented in Chapter 2. They can be considered as either a generalized form of Euler-Poincaré equations or a left-trivialized form of Euler-Lagrange equations on a manifold. The tangent bundle of a Lie group $\mathsf{TG}$ is identified with $\mathsf{G} \times \mathfrak{g}$ by the left trivialization, and the equations of motion are expressed in terms of Lie group elements and Lie algebra elements. Even if the Lagrangian is not left invariant, this approach is still desirable: since a Lie algebra is a linear vector space at the fixed identity

element, there is no need to deal with covariant derivatives or Christoffel symbols. The resulting equations of motion for multiple rigid body systems are more compact and concise.

Geometric numerical integrators, referred to as Lie group variational integrators, are developed for dynamic systems with a Lie group configuration manifold. The method presented in this dissertation represents the first time that the Lie group approach has been explicitly adopted in the context of variational integrators for an arbitrary Lie group. They provide a systematic way to develop a class of geometric numerical integrators that preserve the geometric properties of the dynamics as well as the Lie group structure. Numerical simulations show that it is critical to preserve both the symplectic property of the dynamics and the structure of the Lie group. The Lie group variational integrators have substantial computational advantages compared to other geometric integrators that preserve either none or one of these properties. They are more efficient than considering the Lie group structure as a nonlinear algebraic constraint to be satisfied at each time step. The Lie group variational integrator is extended to mechanical systems evolving on a product of two-spheres, which provides an overall framework to develop variational integrators on a homogeneous manifold.

Geometric optimal control, as presented in this dissertation, is used to treat optimal control problems for dynamic systems on a Lie group. This is distinguished from the existing optimal control theories developed for kinematics equations on a Lie group. In Chapter 4, an intrinsic form of necessary conditions for optimality is developed. They are applied to a wide class of dynamic systems on a Lie group, and they are more compact than optimality conditions expressed in terms of local coordinates.

Computational geometric optimal control formulates a discrete-time optimal control problem based on geometric numerical integrators. This is in contrast to other optimal control approaches where a discretization appears at the terminal stage in solving optimality conditions numerically. The computational geometric optimal control preserves the geometric structure of the optimal control problem as well as the geometric properties of the dynamics. It turns out that the approach also allow us to find the optimal trajectory more efficiently since there is no numerical dissipation in the discrete-time numerical flow. This is applied to several optimal control problems for rigid bodies.

In summary, this dissertation develops computational geometric mechanics and computational geometric optimal control approaches for dynamic systems on a Lie group. The essential idea is to derive a computational algorithm from a discrete analogue of the underlying fundamental principles so that the physical properties are preserved naturally. They are applied to several nontrivial rigid body dynamics.

## 6.2 Future Work

This dissertation has a broad scope; it combines system and control theory in engineering and differential geometry in applied mathematics. The resulting computational geometric methods have broad potential impact on numerous scientific and engineering problems varying from formation re-

configuration in aerospace engineering to molecular dynamics simulations in chemistry. Therefore, the results in this dissertation can be extended in various directions.

*Generalized Lie Group Variational Integrator.* The variational integrators have been extended in various ways. For example, a variational integrator for a Lagrangian system with a degenerate Lagrangian has been developed by Rowley and Marsden (2002) with application to point vortices, and multisymplectic variational integrators have been developed for Hamiltonian continuum mechanics (see Marsden et al. 2001). Asynchronous variational integrators in the work by Lew et al. (2003) consider variational integrators with varying step size. Generalized variational integrators are developed for nonsmooth systems by Fetecau et al. (2003), and for nonholonomic systems by Cortés and Martínez (2001).

The presented Lie group variational integrators can be extended in similar ways to obtain generalized Lie group variational integrators for various types of mechanical systems. In particular, it would be interesting to develop a multiple time-scale Lie group variational integrator for highly oscillatory systems. This is motivated by scientific and engineering problems from molecular dynamics, astrophysics, structural dynamics, and nonlinear wave equations.

*Discrete-time Geometric Control Systems on a Lie Group.* Geometric mechanics provides fundamental insights into mechanical systems and yields new approaches in control system design. For example, the method of controlled Lagrangian is a constructive technique for stabilizing mechanical systems, where control inputs are obtained from a Lagrangian system with a modified Lagrangian. This approach originated by Bloch et al. (1992) has been extended in various ways (see, for example, Bloch et al. 2000, 2001; Zenkov et al. 2000, 2002, and references therein).

In particular, a discrete-time theory of controlled Lagrangian systems was developed for variational integrators by Bloch et al. (2005, 2006), and applied to the feedback stabilization of the unstable inverted equilibrium of a planar pendulum on a cart. It would be natural to combine the techniques of the controlled Lagrangian method and the Lie group variational integrators developed in this dissertation to obtain a discrete-time geometric control approach for mechanical systems on a Lie group.

*Uncertainty Propagation and Estimation on a Lie Group.* A mathematical model of a dynamic system may not capture all of the dynamic characteristics of the system exactly. It always includes approximations, simplifications, or system identification errors. Therefore, uncertainty propagation and estimation provide important information for many scientific and engineering problems. However, it is challenging in the sense that uncertainty propagation for a general nonlinear system is expressed as a partial differential equation. The standard linearization approach is only applicable to a system with frequent measurements.

A deterministic attitude estimation scheme has been developed by Lee et al. (2006b, 2007h); Sanyal et al. (2008) using the Lie group variational integrator; limitations of the linearization approach for uncertainty propagation of attitude dynamics have been presented in Lee et al. (2007a).

Recently, a global uncertainty propagation scheme for the attitude dynamics of a rigid body is developed by Lee et al. (2008b). This approach is based on the symplectic property of the attitude dynamics and noncommutative harmonic analysis on a Lie group: the symplectic property of Hamiltonian systems imposes a fundamental limit on the uncertainty propagation (see Hsiao and Scheeres. 2007; Scheeres et al. 2005), which yields a particle based uncertainty propagation method. The propagated probability density function is expressed in term of noncommutative harmonic analysis (see, for example Biedenharn and Louck 1981; Chirikjian and Kyatkin 2001; Sugiura 1990).

*Application to Large Scale Full $n$-body Problems.* The full $n$ body problem studies the motion of $n$ interacting bodies modeled as rigid bodies with arbitrary shapes. Since the motion of each rigid body is described by its translation and rotation with respect to a given reference frame, the configuration space for the full $n$ body problem is $\mathsf{SE}(3)^n$. This is particularly important for various physical systems where the mutual interaction depends on the relative attitude as well as the relative position. For example, the mutual gravitational forces and moments in binary asteroids vary with the rotation of each body, and the electrostatic forces between charged molecules depend on their electric pole directions.

The computational accuracy and efficiency of Lie group variational integrators has been illustrated by a full two body problem. Since the computational superiority of the Lie group variational integrators increases as the complexity of the system or the simulation time increases, they can be applied to challenging full $n$ body problems such as simulations of the asteroid belt or molecular dynamics, using powerful parallel computing resources.

*Applications to Multibody Systems.* Multibody systems appear in advanced mechanical systems in the area of automobiles, aerospace, robotics, and power plants. For example, solar panel deployment in satellites, bipedal robots, cooperative multiple vehicles, or flexible bodies are often analyzed as multibody systems.

Lie group variational integrators can be applied to any multibody system whose configuration manifold is expressed as a product of Lie groups, Euclidean spaces or two-spheres, thereby generating structure-preserving geometric numerical integrators for the multibody system. The continuous-time counterpart of this development provides a remarkably compact form for global equations of motion. These give both an efficient computational approach and a powerful analysis tool for multibody systems.

In particular, it would be be interesting to study optimal maneuvers of a large space tether, which controls its orbit by changing its shape. This requires significantly less expensive propulsion than spacecraft using rocket engines. The elastic rod model presented in Section 3.4.4, and the underactuated optimal control approach discussed in Section 5.2.3 and Section 5.2.7 might serve as a basis for this research.

# APPENDIX

**APPENDIX A**

# PROPERTIES AND PROOFS

## A.1 Properties of the *hat* Map

The hat map $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ is defined as

$$\hat{x} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \tag{A.1}$$

for $x = [x_1; x_2; x_3] \in \mathbb{R}^3$. This identifies the Lie algebra $\mathfrak{so}(3)$ with $\mathbb{R}^3$ using the vector cross product in $\mathbb{R}^3$.

Several properties of the hat map are summarized as follows.

$$\hat{x}y = x \times y = -y \times x = -\hat{y}x, \tag{A.2}$$

$$\hat{x}^T\hat{x} = (x^T x)I - xx^T, \tag{A.3}$$

$$\hat{x}\hat{y}\hat{x} = -(y^T x)\hat{x}, \tag{A.4}$$

$$-\frac{1}{2}\text{tr}[\hat{x}\hat{y}] = x^T y, \tag{A.5}$$

$$\widehat{x \times y} = \hat{x}\hat{y} - \hat{y}\hat{x} = yx^T - xy^T, \tag{A.6}$$

$$\text{tr}[\hat{x}A] = \frac{1}{2}\text{tr}\left[\hat{x}(A - A^T)\right], \tag{A.7}$$

$$\widehat{Ax} = \hat{x}\left(\frac{1}{2}\text{tr}[A]\, I - A\right) + \left(\frac{1}{2}\text{tr}[A]\, I - A\right)^T \hat{x}, \tag{A.8}$$

$$\hat{x}A + A^T\hat{x} = \left(\{\text{tr}[A]\, I_{3\times 3} - A\}\, x\right)^\wedge. \tag{A.9}$$

for any $x, y \in \mathbb{R}^3$, $A \in \mathbb{R}^{3\times 3}$.

## A.2 Inertia Matrix of a Rigid Body

Consider a rigid body $\mathcal{B}$. We define a body fixed frame. Let $\rho \in \mathbb{R}^3$ be the vector from the origin of the body fixed frame to a mass element of the rigid body.

The inertia matrix $J \in \mathbb{R}^{3\times 3}$ of the rigid body is given by

$$J = \int_{\mathcal{B}} \hat{\rho}^T \hat{\rho}\, dm(\rho) = \int_{\mathcal{B}} \begin{bmatrix} y^2+z^2 & -xy & -zx \\ -xy & z^2+x^2 & -yz \\ -zx & -yz & x^2+y^2 \end{bmatrix} dm(\rho), \tag{A.10}$$

where $x, y, z \in \mathbb{R}$ are the coordinates of the vector $\rho$, i.e. $\rho = [x; y; z]$. We define a nonstandard inertia matrix $J_d \in \mathbb{R}^{3\times 3}$ as

$$J_d = \int_{\mathcal{B}} \rho\rho^T\, dm(\rho) = \int_{\mathcal{B}} \begin{bmatrix} x^2 & xy & zx \\ xy & y^2 & yz \\ zx & yz & z^2 \end{bmatrix} dm(\rho). \tag{A.11}$$

From the coordinate expressions, it is clear that the standard inertia matrix $J$ and the nonstandard inertia matrix $J_d$ represent the second order mass distributions of the rigid body. Using the property, $\hat{\rho}^T\hat{\rho} = (\rho^T\rho)I_{3\times 3} - \rho\rho^T$, it can be shown that

$$J_d = \frac{1}{2}\text{tr}[J]\, I_{3\times 3} - J, \quad J = \text{tr}[J_d]\, I_{3\times 3} - J_d. \tag{A.12}$$

In Section 2.3.2, we show that the rotational kinetic energy of the rigid body, when it rotates about the origin with an angular velocity $\Omega \in \mathbb{R}^3$, can be written as

$$T = \frac{1}{2}\Omega^T J \Omega = \frac{1}{2}\text{tr}\left[\hat{\Omega} J_d \hat{\Omega}^T\right]. \tag{A.13}$$

Furthermore, the following equation is satisfied for any $\Omega \in \mathbb{R}^3$.

$$\widehat{J\Omega} = \hat{\Omega} J_d + J_d \hat{\Omega}. \tag{A.14}$$

A proof for (A.14) is as follows. Let $\Omega = [\Omega_1, \Omega_2, \Omega_3]$ for any $\Omega_1, \Omega_2, \Omega_3 \in \mathbb{R}$. By substitution and rearrangement, it is straightforward to show that both the left hand side expression and the right hand side expression of (A.14) are equal to

$$\begin{bmatrix} (J_{yy}+J_{zz})\Omega_1 - J_{xy}\Omega_2 - J_{zx}\Omega_3 \\ -J_{xy}\Omega_1 + (J_{zz}+J_{xx})\Omega_2 - J_{yz}\Omega_3 \\ -J_{zx}\Omega_1 - J_{yz}\Omega_2 + (J_{xx}+J_{yy})\Omega_3 \end{bmatrix}^{\wedge},$$

where $J_{xy} = \int_{\mathcal{B}} xy\, dm(\rho) \in \mathbb{R}$, and other terms are defined similarly.

## A.3 Lagrangian-Routh Reduction of the 3D Pendulum

The 3D pendulum is a rigid body supported by a frictionless pivot point acting under a gravitational potential (see Section 2.3.2). The configuration manifold is $\mathsf{SO}(3)$. The Lagrangian of the 3D pendulum has a symmetry: it is invariant under an action of $\mathsf{H} = \mathsf{SO}(2) \simeq \mathsf{S}^1$ given by $\Phi : \mathsf{S}^1 \times \mathsf{SO}(3) \to \mathsf{SO}(3)$

$$\Phi(\theta, R) = \exp_{\mathsf{SO}(3)}(\theta \hat{e}_3) R, \tag{A.15}$$

which represents the rotation of the 3D pendulum about the gravity direction $e_3$. As a result, the configuration manifold can be reduced to the quotient space $\mathsf{SO}(3)/\mathsf{SO}(2) \simeq \mathsf{S}^2$.

Here, we summarize the Lagrange-Routh reduction and reconstruction procedure to obtain the reduced equations of motion of the 3D pendulum on $\mathsf{S}^2$ (see Marsden et al. 2000). This reduction procedure is interesting and challenging, since the projection $\mathrm{proj} : \mathsf{SO}(3) \to \mathsf{S}^2$ given by $\mathrm{proj}(R) = R^T e_3 = \Upsilon$ 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.

### A.3.1 Reduction

The procedure for Lagrange-Routh reduction is as follows. We find expressions for the mechanical connection, from which a Routhian is defined. The Routhian satisfies the Lagrange-d'Alembert principle with a magnetic two-form, which yields the reduced Euler-Lagrange equations.

*Routhian* We identify the Lie algebra of $\mathsf{S}^1$ with $\mathbb{R}$. At (2.71), the momentum map of the 3D pendulum $J_L : \mathsf{SO}(3) \times \mathfrak{so}(3) \to \mathbb{R}^*$ is given by

$$J_L(R, \hat{\Omega}) = e_3^T R J \Omega. \tag{A.16}$$

The locked inertia tensor $\mathbb{I}(R) : \mathbb{R} \to \mathbb{R}^*$ is defined as

$$\langle \mathbb{I}(R)\varrho,\, \zeta \rangle = \langle\!\langle \varrho_{\mathsf{SO}(3)},\, \zeta_{\mathsf{SO}(3)} \rangle\!\rangle$$

for $\varrho, \zeta \in \mathbb{R}$. Substituting the expression for the infinitesimal generator $\zeta_{\mathsf{SO}(3)} = \zeta \widehat{R^T e_3}$, $\varrho_{\mathsf{SO}(3)} = \varrho \widehat{R^T e_3}$ given by (3.84), we obtain

$$\varrho\zeta \mathbb{I}(R) = \varrho\zeta e_3^T R J R^T e_3.$$

Thus, the locked inertia tensor is given by $\mathbb{I}(R) = e_3^T R J R^T e_3$. The mechanical connection $\mathcal{A} : \mathsf{SO}(3) \times \mathfrak{so}(3) \to \mathbb{R}$ is given by

$$\mathcal{A}(R, \hat{\Omega}) = \mathbb{I}^{-1}(R) J_L(R, \hat{\Omega}) = \frac{e_3^T R J \Omega}{e_3^T R J R^T e_3}. \tag{A.17}$$

Using the mechanical connection, we decompose the tangent space into a vertical space and a

horizontal space as follows. For a given $(R, \hat{\Omega}) \in \mathsf{TSO(3)}$, the horizontal part and the vertical part are obtained by

$$\begin{aligned}
\mathrm{ver}(R, \hat{\Omega}) &= (\mathcal{A}(R, \hat{\Omega}))_{\mathsf{SO(3)}}(R) = \frac{e_3^T RJ\Omega}{e_3^T RJR^T e_3} \widehat{R^T e_3} = \frac{\Upsilon \cdot J\Omega}{\Upsilon \cdot J\Upsilon} \hat{\Upsilon}, \\
\mathrm{hor}(R, \hat{\Omega}) &= \hat{\Omega} - \mathrm{ver}(R, \hat{\Omega}) = \left( \Omega - \frac{R^T e_3 e_3^T RJ}{e_3^T RJR^T e_3} \Omega \right)^\wedge = \left( \Omega - \frac{\Upsilon \Upsilon^T J}{\Upsilon \cdot J\Upsilon} \Omega \right)^\wedge .
\end{aligned}$$

Since $\dot{\Upsilon} = \dot{R}^T e_3 = -\hat{\Omega} R^T e_3 = \Upsilon \times \Omega$, we obtain

$$\dot{\Upsilon} \times \Upsilon = (\Upsilon \times \Omega) \times \Upsilon = (\Upsilon^T \Upsilon)\Omega - (\Upsilon^T \Omega)\Upsilon = \Omega - \Upsilon \Upsilon^T \Omega.$$

Using this expression, the horizontal part of $\Omega$ can be written in terms of $\Upsilon$ as

$$\Omega_{\mathrm{hor}} = \dot{\Upsilon} \times \Upsilon - \frac{J\Upsilon \cdot (\dot{\Upsilon} \times \Upsilon)}{\Upsilon \cdot J\Upsilon} \Upsilon = \dot{\Upsilon} \times \Upsilon - b\Upsilon$$

for $b = \frac{J\Upsilon \cdot (\dot{\Upsilon} \times \Upsilon)}{\Upsilon \cdot J\Upsilon}$. For the given value of the momentum map $\nu \in \mathbb{R}^*$, the Routhian of the 3D pendulum is given by

$$\begin{aligned}
R^\nu(\Upsilon, \dot{\Upsilon}) &= \frac{1}{2} \Omega_{\mathrm{hor}} \cdot J\Omega_{\mathrm{hor}} - mg\Upsilon \cdot \rho - \frac{1}{2} \frac{\nu^2}{\Upsilon \cdot J\Upsilon} \\
&= \frac{1}{2} (\dot{\Upsilon} \times \Upsilon) \cdot J(\dot{\Upsilon} \times \Upsilon) - \frac{1}{2} (b^2 + \gamma^2)(\Upsilon \cdot J\Upsilon) + mg\Upsilon \cdot \rho, \qquad \text{(A.18)}
\end{aligned}$$

where $\gamma = \frac{\nu}{\Upsilon \cdot J\Upsilon}$.

*Variation of the Routhian* The Routhian satisfies the variational Lagrange-d'Alembert principle. The infinitesimal variation for $\Upsilon \in \mathsf{S}^2$ is chosen as:

$$\delta\Upsilon = \Upsilon \times \eta, \qquad \text{(A.19)}$$

$$\delta\dot{\Upsilon} = \dot{\Upsilon} \times \eta + \Upsilon \times \dot{\eta}. \qquad \text{(A.20)}$$

Here we assume that $\eta \cdot \Upsilon = 0$, since the component of $\eta$ parallel to $\Upsilon$ has no effect on $\delta\Upsilon$. These expressions are essential for developing the reduced equations of motion.

Using (A.19), (A.20), and the properties $\Upsilon \cdot \dot{\Upsilon} = 0$, $\Upsilon \cdot \eta = 0$, the variation of the Routhian is given by

$$\begin{aligned}
\delta R^\nu = \dot{\eta} \cdot J(\dot{\Upsilon} \times \Upsilon - b\Upsilon) \\
- \eta \cdot \Upsilon \times \left[ -\dot{\Upsilon} \times J(\dot{\Upsilon} \times \Upsilon) + (b^2 + \lambda^2) J\Upsilon - bJ(\dot{\Upsilon} \times \Upsilon) + b(\dot{\Upsilon} \times J\Upsilon) + mg\rho \right].
\end{aligned} \qquad \text{(A.21)}$$

*Magnetic two-form* From the given mechanical connection $\mathcal{A}$ and a value of the momentum map $\nu \in \mathbb{R}^*$, define a one-form $\mathcal{A}_\nu$ on $\mathsf{SO}(3) \times \mathfrak{so}(3)$ by

$$\mathcal{A}_\nu(R) \cdot (R,\, \hat{\Omega}) = \left\langle \nu, \mathcal{A}(R, \hat{\Omega}) \right\rangle = \nu \frac{e_3^T R J \Omega}{e_3^T R J R^T e_3}.$$

The magnetic two-form $\beta_\nu$ is the exterior derivative of $\mathcal{A}_\nu$, which can be obtained by using the identity

$$\mathbf{d}\mathcal{A}_\nu(X,Y) = X[\mathcal{A}_\nu(Y)] - Y[\mathcal{A}_\nu(X)] - \mathcal{A}_\nu([X,Y])$$

for $X = R\hat{\eta}, Y = R\hat{\xi} \in \mathsf{T}_R\mathsf{SO}(3)$, where $X[\mathcal{A}_\nu(Y)]$ denotes the Lie derivative of $\mathcal{A}_\nu(Y)$ along $X$ (see Marsden and Ratiu 1999). The first term of the above equation is obtained as

$$\begin{aligned} X[\mathcal{A}_\nu(Y)] &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \mathcal{A}_\nu(Y)_{R=R\exp\epsilon\hat{\eta}} \\ &= \frac{\nu}{(e_3^T R J R^T e_3)^2}\left[e_3^T R\hat{\eta} J\xi (e_3^T R J R^T e_3) + 2 e_3^T R J \xi (e_3^T R J \hat{\eta} R^T e_3)\right] \\ &= \frac{\nu}{(e_3^T R J R^T e_3)^2}\left[-J\xi\cdot(\eta \times R^T e_3)(e_3^T R J R^T e_3) + 2(\xi \cdot J R^T e_3)(J R^T e_3 \cdot (\eta \times R^T e_3))\right]. \end{aligned}$$

Similarly, we find expressions for $Y[\mathcal{A}_\nu(X)]$. The Lie bracket is given by $[X,Y] = R\widehat{\eta \times \xi}$. Therefore, we obtain

$$\begin{aligned} \mathbf{d}\mathcal{A}_\nu(X,Y) &= -\frac{\nu}{(e_3^T R J R^T e_3)}\left[J\xi\cdot(\eta\times R^T e_3) - J\eta\cdot(\xi\times R^T e_3) + (\eta\times\xi)\cdot J R^T e_3\right] \\ &- \frac{\nu}{(e_3^T R J R^T e_3)^2}\left[-2(\xi\cdot J R^T e_3)(J R^T e_3\cdot(\eta\times R^T e_3)) + 2(\eta\cdot J R^T e_3)(J R^T e_3\cdot(\xi\times R^T e_3))\right]. \end{aligned}$$

Substituting $\Upsilon = R^T e_3$ into this, the magnetic two form is given by

$$\begin{aligned} \beta_\nu(\Upsilon\times\eta, \Upsilon\times\xi) = &-\frac{\nu}{(\Upsilon\cdot J\Upsilon)}\left[J\xi\cdot(\eta\times\Upsilon) - J\eta\cdot(\xi\times\Upsilon) + (\eta\times\xi)\cdot J\Upsilon\right] \\ &- \frac{\nu}{(\Upsilon\cdot J\Upsilon)^2}\left[-2(\xi\cdot J\Upsilon)(J\Upsilon\cdot(\eta\times\Upsilon)) + 2(\eta\cdot J\Upsilon)(J\Upsilon\cdot(\xi\times\Upsilon))\right]. \end{aligned}$$

Since $(a\cdot x)(b\cdot y) - (a\cdot y)(b\cdot x) = (a\times b)\cdot(x\times y)$ for any $a,b,x,y\in\mathbb{R}^3$, the last two terms of the above equation are reduced to

$$\begin{aligned} -2(\xi\cdot J\Upsilon)(J\Upsilon\cdot(\eta\times\Upsilon)) + 2(\eta\cdot J\Upsilon)(J\Upsilon\cdot(\xi\times\Upsilon)) &\\ = -2(\xi\cdot J\Upsilon)(\eta\cdot(\Upsilon\times J\Upsilon)) + 2(\eta\cdot J\Upsilon)(\xi\cdot(\Upsilon\times J\Upsilon)) &\\ = 2(J\Upsilon\times(\Upsilon\times J\Upsilon))\cdot(\eta\times\xi) &\\ = 2\left\{\|J\Upsilon\|^2\,\Upsilon - (\Upsilon\cdot J\Upsilon)J\Upsilon\right\}\cdot(\eta\times\xi). & \end{aligned}$$

Thus, we obtain

$$\begin{aligned}&\beta_\nu(\Upsilon\times\eta,\Upsilon\times\xi)\\&=-\frac{\nu}{(\Upsilon\cdot J\Upsilon)^2}\left[(\Upsilon\cdot J\Upsilon)\left\{\Upsilon\cdot(J\xi\times\eta)-\Upsilon\cdot(J\eta\times\xi)-J\Upsilon\cdot(\eta\times\xi)\right\}+2\left\|J\Upsilon\right\|^2\Upsilon\cdot(\eta\times\xi)\right].\end{aligned}$$

Using the identity, $\Upsilon\cdot(J\xi\times\eta)+\Upsilon\cdot(\xi\times J\eta)+J\Upsilon\cdot(\xi\times\eta)=\text{tr}[J]\,\Upsilon\cdot(\xi\times\eta)$, the magnetic two form $\beta_\nu$ is given by

$$\beta_\nu(\Upsilon\times\eta,\Upsilon\times\xi)=-\frac{\nu}{(\Upsilon\cdot J\Upsilon)^2}\left[-(\Upsilon\cdot J\Upsilon)\text{tr}[J]+2\left\|J\Upsilon\right\|^2\right]\Upsilon\cdot(\eta\times\xi). \tag{A.22}$$

Since $\dot\Upsilon=\Upsilon\times\Omega$, and $\Upsilon\cdot(\omega\times\eta)=\eta\cdot(\Upsilon\times\omega)=\eta\cdot\dot\Upsilon$, the interior product of the magnetic two-form is given by

$$\mathbf{i}_{\dot\Upsilon}\beta_\nu(\delta\Upsilon)=\beta_\nu(\Upsilon\times\Omega,\Upsilon\times\eta)=\lambda\left\{\text{tr}[J]-2\frac{\left\|J\Upsilon\right\|^2}{\Upsilon\cdot J\Upsilon}\right\}\dot\Upsilon\cdot\eta, \tag{A.23}$$

where $\lambda=\frac{\nu}{\Upsilon\cdot J\Upsilon}$.

*Lagrange-d'Alembert Principle* The Routhian satisfies the Lagrange-d'Alembert principle with the magnetic term, given by

$$\delta\int_0^T R^\nu(\Upsilon,\dot\Upsilon)dt=\int_0^T\mathbf{i}_{\dot\Upsilon}\beta_\nu(\delta\Upsilon)dt. \tag{A.24}$$

Substituting (A.21) and (A.23) into (A.24), and integrating by parts, we obtain

$$-\int_0^T\eta\cdot\left[J(\ddot\Upsilon\times\Upsilon-b\dot\Upsilon-\dot b\Upsilon)+\Upsilon\times\mathcal{X}+c\dot\Upsilon\right]dt=0, \tag{A.25}$$

where

$$\mathcal{X}=-\dot\Upsilon\times J(\dot\Upsilon\times\Upsilon)+(b^2+\lambda^2)J\Upsilon-bJ(\dot\Upsilon\times\Upsilon)+b(\dot\Upsilon\times J\Upsilon)+mg\rho, \tag{A.26}$$

and $c$ is given by (2.74):

$$c=\Upsilon\left\{\text{tr}[J]-2\frac{\left\|J\Upsilon\right\|^2}{\Upsilon\cdot J\Upsilon}\right\}.$$

Since (A.25) is satisfied for all $\eta$ with $\Upsilon\cdot\eta=0$, we obtain

$$J(\ddot\Upsilon\times\Upsilon-b\dot\Upsilon-\dot b\Upsilon)+\Upsilon\times\mathcal{X}+c\dot\Upsilon=a\Upsilon, \tag{A.27}$$

for a constant $a\in\mathbb{R}$. This is the reduced equation of motion. However, this equation has an ambiguity since the value of the constant $a$ is unknown; this equation is implicit for $\ddot\Upsilon$ since the term $\dot b$ is expressed in terms of $\ddot\Upsilon$. The next step is to determine expressions for $a$ and $\dot b$ using the definition of $b$ and several vector identities.

We first find an expression for the constant $a$ in terms of $\Upsilon, \dot{\Upsilon}$. Taking the dot product of (A.27) with $\Upsilon$, we obtain

$$\Upsilon \cdot J(\ddot{\Upsilon} \times \Upsilon - b\dot{\Upsilon} - \dot{b}\Upsilon) = a. \tag{A.28}$$

From the definition of $b$, we can show the following identity: $\Upsilon \cdot J(\dot{\Upsilon} \times \Upsilon - b\Upsilon) = 0$. Differentiating this with time, and substituting into (A.28), we find an expression for the constant $a$ in terms of $\Upsilon, \dot{\Upsilon}$ as

$$a = -\dot{\Upsilon} \cdot J(\dot{\Upsilon} \times \Upsilon - b\Upsilon). \tag{A.29}$$

Substituting (A.29) into (A.27), and taking the dot product of the result with $\Upsilon$, we obtain an expression for $\dot{b}$ in terms of $\Upsilon, \dot{\Upsilon}$ as

$$\dot{b} = \Upsilon \cdot J^{-1}\left\{\Upsilon \times \mathcal{X} + c\dot{\Upsilon} + (\dot{\Upsilon} \cdot J(\dot{\Upsilon} \times \Upsilon - b\Upsilon))\Upsilon\right\}. \tag{A.30}$$

Substituting (A.30) into (A.27), and using the vector identity $Y - (\Upsilon \cdot Y)\Upsilon = (\Upsilon \cdot \Upsilon)Y - (\Upsilon \cdot Y)\Upsilon = -\Upsilon \times (\Upsilon \times Y)$ for any $Y \in \mathbb{R}^3$, we obtain the following form for the reduced equation of motion

$$\ddot{\Upsilon} \times \Upsilon - b\dot{\Upsilon} - \Upsilon \times \left[\Upsilon \times J^{-1}\left\{\Upsilon \times \mathcal{X} + c\dot{\Upsilon} + (\dot{\Upsilon} \cdot J(\dot{\Upsilon} \times \Upsilon - b\Upsilon))\Upsilon\right\}\right] = 0.$$

Now, we simplify this equation. The above expression is equivalent to the following equation

$$\Upsilon \times \left[\ddot{\Upsilon} \times \Upsilon - b\dot{\Upsilon} - \Upsilon \times \left[\Upsilon \times J^{-1}\left\{\Upsilon \times \mathcal{X} + c\dot{\Upsilon} + (\dot{\Upsilon} \cdot J(\dot{\Upsilon} \times \Upsilon - b\Upsilon))\Upsilon\right\}\right]\right] = 0. \tag{A.31}$$

Since $\Upsilon \cdot \ddot{\Upsilon} = -\|\dot{\Upsilon}\|^2$, the first term is given by

$$\Upsilon \times (\ddot{\Upsilon} \times \Upsilon) = (\Upsilon \cdot \Upsilon)\ddot{\Upsilon} - (\Upsilon \cdot \ddot{\Upsilon})\Upsilon = \ddot{\Upsilon} + \|\dot{\Upsilon}\|^2\Upsilon.$$

Using the property $\Upsilon \times (\Upsilon \times (\Upsilon \times Y)) = -(\Upsilon \cdot \Upsilon)\Upsilon \times Y = -\Upsilon \times Y$ for $Y \in \mathbb{R}^3$, the third term of (A.31) can be simplified.

*Reduced Euler-Lagrange Equations* Substituting (A.26) and rearranging, the reduced Euler-Lagrange equations for the 3D pendulum on $\mathsf{S}^2$ are given by

$$\ddot{\Upsilon} = -\|\dot{\Upsilon}\|^2\Upsilon + \Upsilon \times \Sigma, \tag{A.32}$$

where $\Sigma = b\dot{\Upsilon} + J^{-1}\left[(J(\dot{\Upsilon} \times \Upsilon) - bJ\Upsilon) \times ((\dot{\Upsilon} \times \Upsilon) - b\Upsilon) + \lambda^2 J\Upsilon \times \Upsilon - mg\Upsilon \times \rho - c\dot{\Upsilon}\right]$, which recovers (2.72)–(2.74).

### A.3.2 Reconstruction

For a given integral curve of the reduced equation $\Upsilon(t) : [0, T] \to \mathsf{S}^2$, the reconstruction is to find the corresponding curve $R(t) : [0, T] \to \mathsf{SO}(3)$ that is projected to the given curve and satisfies the Euler-Lagrange equations.

We first choose a curve $\tilde{R}(t) \in \mathsf{SO}(3)$ that is projected into the reduced curve, i.e. $\mathrm{proj}(\tilde{R}(t)) = \Upsilon(t)$. Then, the reconstructed curve can be written as $R(t) = \Phi_{\theta(t)}(\tilde{R}(t))$ for some $\theta(t) \in \mathsf{S}^1$. Conservation of the momentum map yields the following reconstruction equation for $\theta(t)$ (see Marsden et al. 2000).

$$\theta(t)^{-1}\dot{\theta}(t) = \mathbb{I}^{-1}(\tilde{R}(t))\nu - \mathcal{A}(\dot{\tilde{R}}(t)).$$

If we choose $\tilde{R}(t)$ as the horizontal lift, the last term vanishes as the horizontal space is annihilated by the mechanical connection. Since the symmetry group is abelian, the solution reduces to a quadrature as

$$\theta(t) = \theta(0)\exp_{\mathsf{SO}(2)}\left[\int_0^t \mathbb{I}^{-1}(\tilde{R}(s))\nu\, ds\right].$$

In summary, the reconstruction procedure is as follows.

(i) Horizontally lift $\Upsilon(t)$ to obtain $R_{\mathrm{hor}}(t)$ by integrating the following equation with $R_{\mathrm{hor}}(0) = R(0)$

$$\dot{R}_{\mathrm{hor}}(t) = R_{\mathrm{hor}}(t)\hat{\Omega}_{\mathrm{hor}}(t), \tag{A.33}$$

where

$$\Omega_{\mathrm{hor}}(t) = \dot{\Upsilon}(t) \times \Upsilon(t) - \frac{J\Upsilon(t)\cdot(\dot{\Upsilon}(t)\times\Upsilon(t))}{\Upsilon(t)\cdot J\Upsilon(t)}\Upsilon(t). \tag{A.34}$$

(ii) Determine $\theta_{\mathrm{dyn}}(t) \in \mathsf{S}$ by the following equation

$$\theta_{\mathrm{dyn}}(t) = \int_0^t \frac{\nu}{\Upsilon(s)\cdot J\Upsilon(s)}\, ds. \tag{A.35}$$

(iii) Reconstruct the curve in $\mathsf{SO}(3)$ as

$$R(t) = \theta_{\mathrm{dyn}}(t)\cdot R_{\mathrm{hor}}(t) = \exp[\theta_{\mathrm{dyn}}(t)\hat{e}_3]R_{\mathrm{hor}}(t). \tag{A.36}$$

### A.3.3 Geometric Phase

Suppose that the integral curve in $\mathsf{S}^2$ is a closed curve, i.e. $\Upsilon(0) = \Upsilon(T)$. Since $R_{\mathrm{hor}}(0) = R(0)$ and $R_{\mathrm{hor}}(T)$ are in the same fiber, $R_{\mathrm{hor}}(T)$ can be written as

$$R_{\mathrm{hor}}(T) = \theta_{\mathrm{geo}}(T)\cdot R(0),$$

for $\theta_{\mathrm{geo}}(T) \in \mathsf{S}^1$. From the reconstruction equation (A.36), we obtain

$$R(T) = \theta_{\mathrm{dyn}}(T)\cdot\theta_{\mathrm{geo}}(T)\cdot R(0),$$

where $\theta_{\mathrm{dyn}}(T) \in \mathsf{S}^1$ and $\theta_{\mathrm{geo}}(T) \in \mathsf{S}^1$ are referred to as the dynamic phase and the geometric phase, respectively. If the value of the momentum map is zero, then the dynamic phase is also zero from (A.35), and the attitude is changed only by the geometric phase effect.

It can be shown that the geometric phase is the negative of the integral of the magnetic two-form on the area enclosed by the integral curve $\Upsilon(t)$

$$\theta_{\mathrm{geo}}(T) = \int_{\mathcal{S}} \frac{2\left\|J\Upsilon(t)\right\|^2 - \mathrm{tr}[J]\left(\Upsilon(t)\cdot J\Upsilon(t)\right)}{\left(\Upsilon(t)\cdot J\Upsilon(t)\right)^2}\, dA, \tag{A.37}$$

where $\mathcal{S}$ is an area in $\mathsf{S}^2$ with $\Upsilon(t)$ as a boundary.

## A.4 Commutativity of the Variation and the Time Derivative

Here, we show the commutativity of the variation operation and the time derivative for $\xi = g^{-1}\dot{g} \in \mathfrak{g}$, i.e, $\frac{d}{dt}(\delta\xi) = \delta(\frac{d}{dt}\xi)$. We give a proof for a matrix Lie group; an extension to general Lie groups can be developed using the results presented in Bloch et al. (1996).

Recall that for a curve $g$ in G, we define $\xi = g^{-1}\dot{g}$. The variation is chosen as $g^\epsilon = g\exp\epsilon\eta$ for $\eta \in \mathfrak{g}$ so that the infinitesimal variation is given by $\delta g = g\eta$. Thus, $\eta = g^{-1}\delta g$. We first find the expression for $\delta\xi$ in (2.5) (see Marsden and Ratiu 1999)

$$\delta\xi = \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\left((g^\epsilon)^{-1}\frac{dg^\epsilon}{dt}\right) = -(g^{-1}\delta g g^{-1})\dot{g} + g^{-1}\frac{d^2g^\epsilon}{d\epsilon dt}\bigg|_{\epsilon=0} = -\eta\xi + g^{-1}\frac{d^2g^\epsilon}{d\epsilon dt}\bigg|_{\epsilon=0}. \quad \text{(A.38)}$$

But, $\frac{d}{dt}\eta$ is given by

$$\dot{\eta} = \frac{d}{dt}\left(g^{-1}\frac{dg}{d\epsilon}\right)\bigg|_{\epsilon=0} = -(g^{-1}\dot{g}g^{-1})\delta g + g^{-1}\frac{d^2g^\epsilon}{dtd\epsilon}\bigg|_{\epsilon=0} = -\xi\eta + g^{-1}\frac{d^2g^\epsilon}{dtd\epsilon}\bigg|_{\epsilon=0}. \quad \text{(A.39)}$$

Since the partial derivatives commute for smooth maps, the difference is given by $\delta\xi - \dot{\eta} = \xi\eta - \eta\xi = [\xi,\,\eta]$, which yields

$$\delta\xi = \dot{\eta} + [\xi,\,\eta]. \quad \text{(A.40)}$$

This gives (2.5).

Now we show that $\frac{d}{dt}(\delta\xi) = \delta(\frac{d}{dt}\xi)$. The time derivative of $\delta\xi$ is given by

$$\frac{d}{dt}(\delta\xi) = \ddot{\eta} + [\dot{\xi},\,\eta] + [\xi,\,\dot{\eta}]. \quad \text{(A.41)}$$

Now, we find an expression for the variation of $\dot{\xi}$.

$$\begin{aligned}
\delta(\frac{d}{dt}\xi) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\left\{\frac{d}{dt}\left((g^\epsilon)^{-1}\frac{dg^\epsilon}{dt}\right)\right\} \\
&= \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\left\{-\left((g^\epsilon)^{-1}\frac{dg^\epsilon}{dt}(g^\epsilon)^{-1}\right)\frac{dg^\epsilon}{dt} + (g^\epsilon)^{-1}\frac{d^2g^\epsilon}{dt^2}\right\} \\
&= (g^{-1}\delta g g^{-1})\dot{g}g^{-1}\dot{g} - g^{-1}\frac{d^2g^\epsilon}{d\epsilon dt}\bigg|_{\epsilon=0} g^{-1}\dot{g} + g^{-1}\dot{g}(g^{-1}\delta g g^{-1})\dot{g} - (g^{-1}\dot{g}g^{-1})\frac{d^2g^\epsilon}{d\epsilon dt}\bigg|_{\epsilon=0} \\
&\quad - (g^{-1}\delta g g^{-1})\ddot{g} + g^{-1}\frac{d^3g^\epsilon}{d\epsilon dt^2}\bigg|_{\epsilon=0}
\end{aligned}$$

Since $g^{-1}\delta g = \eta$ and $g^{-1}\dot{g} = \xi$, this reduces to

$$\delta(\frac{d}{dt}\xi) = \eta\xi\xi - g^{-1}\frac{d^2g^\epsilon}{d\epsilon dt}\bigg|_{\epsilon=0}\xi + \xi\eta\xi - \xi g^{-1}\frac{d^2g^\epsilon}{d\epsilon dt}\bigg|_{\epsilon=0} - \eta g^{-1}\ddot{g} + g^{-1}\frac{d^3g^\epsilon}{d\epsilon dt^2}\bigg|_{\epsilon=0}. \quad \text{(A.42)}$$

From (A.39), we have $g^{-1}\frac{d^2g^\epsilon}{d\epsilon dt}\big|_{\epsilon=0} = \dot{\eta} + \xi\eta$. Therefore, we obtain

$$g^{-1}\frac{d^3g^\epsilon}{d\epsilon dt^2}\bigg|_{\epsilon=0} = g^{-1}\frac{d}{dt}(g\dot{\eta} + g\xi\eta) = (\xi\dot{\eta} + \ddot{\eta} + \xi\xi\eta + \dot{\xi}\eta + \xi\dot{\eta}).$$

Since $\dot{g} = g\xi$, we have $g^{-1}\ddot{g} = \xi\xi + \dot{\xi}$. Substituting these into (A.42), we obtain

$$\begin{aligned}\delta(\frac{d}{dt}\xi) &= \eta\xi\xi - (\dot{\eta} + \xi\eta)\xi + \xi\eta\xi - \xi(\dot{\eta} + \xi\eta) - \eta(\xi\xi + \dot{\xi}) + (\xi\dot{\eta} + \ddot{\eta} + \xi\xi\eta + \dot{\xi}\eta + \xi\dot{\eta}) \\ &= \ddot{\eta} + (\dot{\xi}\eta - \eta\dot{\xi}) + (\xi\dot{\eta} - \dot{\eta}\xi),\end{aligned}$$

which is equal to (A.41). Therefore, $\frac{d}{dt}(\delta\xi) = \delta(\frac{d}{dt}\xi)$.

## A.5 Proof of Corollary 5.1

Corollary 5.1 shows the forced discrete-time Euler-Lagrange equations and the discrete-time Hamilton's equations for a special form of the discrete Lagrangian and the generalized forces. Suppose that the discrete Lagrangian and the generalized forces are given (5.7) and (5.8):

$$\begin{aligned} L_d(g_k, f_k) &= T_d(f_k) - (1-c)hU(g_k) - chU(g_k f_k), \\ u_{d_k}^- &= (1-c)hu_k, \quad u_{d_k}^+ = chu_{k+1}, \end{aligned}$$

where the constant $c$ is a free parameter lies in the interval $[0, 1]$. We develop the corresponding discrete-time equations of motion from Proposition 5.1.

*Derivatives of the Discrete Lagrangian* We first find expressions for the derivatives of the discrete Lagrangian. The derivative of the discrete Lagrangian with respect to $f_k$ is given by

$$\begin{aligned} \mathbf{D}_{f_k} L_{d_k} \cdot \delta f_k &= \mathbf{D}_{f_k} T_{d_k} \cdot \delta f_k - ch\mathbf{D}U(g_k f_k) \cdot \mathsf{T}_{f_k}\mathsf{L}_{g_k}\delta f_k \\ &= \mathbf{D}_{f_k} T_{d_k} \cdot (\mathsf{T}_e\mathsf{L}_{f_k} \circ \mathsf{T}_{f_k}\mathsf{L}_{f_k^{-1}}\delta f_k) - ch\mathbf{D}U(g_k f_k) \cdot (\mathsf{T}_{f_k}\mathsf{L}_{g_k} \circ \mathsf{T}_e\mathsf{L}_{f_k} \circ \mathsf{T}_{f_k}\mathsf{L}_{f_k^{-1}}\delta f_k) \\ &= \left\langle \mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} T_{d_k} - ch\mathsf{T}_e^*\mathsf{L}_{g_k f_k} \cdot \mathbf{D}U(g_k f_k),\ \mathsf{T}_{f_k}\mathsf{L}_{f_k^{-1}}\delta f_k \right\rangle. \end{aligned}$$

Since the force due to the potential is defined as $M(g) = -\mathsf{T}_e^*\mathsf{L}_g \cdot \mathbf{D}U(g)$, we obtain

$$\mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_{d_k} = \mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} T_{d_k} + chM(g_{k+1}). \tag{A.43}$$

From the definition of the $\mathrm{Ad}$ operator, for any $\alpha \in \mathfrak{g}^*$ and $\eta \in \mathfrak{g}$, we have

$$\begin{aligned} \left\langle \alpha,\ \mathrm{Ad}_{f^{-1}}\eta \right\rangle &= \left\langle \alpha,\ \mathsf{T}_f\mathsf{L}_{f^{-1}} \circ \mathsf{T}_e\mathsf{R}_f\eta \right\rangle = \left\langle \alpha,\ \mathsf{T}_{f^{-1}}\mathsf{R}_f \circ \mathsf{T}_e\mathsf{L}_{f^{-1}}\eta \right\rangle \\ &= \left\langle \mathsf{T}_e^*\mathsf{R}_f \circ \mathsf{T}_f^*\mathsf{L}_{f^{-1}}\alpha,\ \eta \right\rangle = \left\langle \mathsf{T}_e^*\mathsf{L}_{f^{-1}} \circ \mathsf{T}_{f^{-1}}^*\mathsf{R}_f\alpha,\ \eta \right\rangle = \left\langle \mathrm{Ad}_{f^{-1}}^*\alpha,\ \eta \right\rangle \end{aligned}$$

Thus, $\mathrm{Ad}_{f^{-1}}^*\alpha = \mathsf{T}_e^*\mathsf{L}_{f^{-1}} \circ \mathsf{T}_{f^{-1}}^*\mathsf{R}_f = \mathsf{T}_e^*\mathsf{R}_f \circ \mathsf{T}_f^*\mathsf{L}_{f^{-1}}$. Using this, we obtain

$$\begin{aligned} \mathrm{Ad}_{f_k^{-1}}^* \cdot (\mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} L_{d_k}) &= \mathrm{Ad}_{f_k^{-1}}^* \cdot (\mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} T_{d_k}) \\ &\quad - ch\mathsf{T}_e^*\mathsf{R}_{f_k} \circ \mathsf{T}_{f_k}^*\mathsf{L}_{f_k^{-1}}(\mathsf{T}_e^*\mathsf{L}_{g_k f_k} \cdot \mathbf{D}U(g_k f_k)) \\ &= \mathrm{Ad}_{f_k^{-1}}^* \cdot (\mathsf{T}_e^*\mathsf{L}_{f_k} \cdot \mathbf{D}_{f_k} T_{d_k}) - ch\mathsf{T}_e^*\mathsf{R}_{f_k} \circ \mathsf{T}_{f_k}^*\mathsf{L}_{g_k} \cdot \mathbf{D}U(g_k f_k)). \end{aligned} \tag{A.44}$$

The derivative of the discrete Lagrangian with respect to $g_k$ is given by

$$\begin{aligned} \mathbf{D}_{g_k} L_{d_k} \cdot \delta g_k &= -ch\mathbf{D}U(g_k) \cdot \delta g_k - (1-c)h\mathbf{D}U(g_k f_k) \cdot \mathsf{TR}_{f_k}\delta g_k \\ &= -ch\mathbf{D}U(g_k) \cdot (\mathsf{T}_e\mathsf{L}_{g_k} \circ \mathsf{T}_{g_k}\mathsf{L}_{g_k^{-1}}\delta g_k) \\ &\quad - (1-c)h\mathbf{D}U(g_k f_k) \cdot (\mathsf{T}_{g_k}\mathsf{R}_{f_k} \circ \mathsf{T}_e\mathsf{L}_{g_k} \circ \mathsf{T}_{g_k}\mathsf{L}_{g_k^{-1}}\delta g_k) \end{aligned}$$

$$= -\left\langle ch\mathsf{T}_e^*\mathsf{L}_{g_k}\cdot\mathbf{D}U(g_k) + (1-c)h\mathsf{T}_e^*\mathsf{L}_{g_k}\circ\mathsf{T}_{g_k}^*\mathsf{R}_{f_k}\cdot\mathbf{D}U(g_kf_k),\ \mathsf{T}_{g_k}\mathsf{L}_{g_k^{-1}}\delta g_k\right\rangle$$

Thus, we obtain

$$\mathsf{T}_e^*L_{g_k}\mathbf{D}_{g_k}L_{d_k} = (1-c)hM(g_k) - ch\mathsf{T}_e^*\mathsf{L}_{g_k}\circ\mathsf{T}_{g_k}^*\mathsf{R}_{f_k}\cdot\mathbf{D}U(g_kf_k). \tag{A.45}$$

*Forced Discrete-time Euler-Lagrange Equations* Substituting (A.43), (A.44), and (A.45) into (5.2), we obtain

$$\begin{aligned}&\mathsf{T}_e^*\mathsf{L}_{f_{k-1}}\cdot\mathbf{D}_{f_{k-1}}T_{d_{k-1}} + chM(g_k) - \mathrm{Ad}_{f_k^{-1}}^*\cdot(\mathsf{T}_e^*\mathsf{L}_{f_k}\cdot\mathbf{D}_{f_k}T_{d_k})\\ &\quad + ch\mathsf{T}_e^*\mathsf{R}_{f_k}\circ\mathsf{T}_{f_k}^*\mathsf{L}_{g_k}\cdot\mathbf{D}U_{k+1} + (1-c)hM(g_k) - ch\mathsf{T}_e^*\mathsf{L}_{g_k}\circ\mathsf{T}_{g_k}^*\mathsf{R}_{f_k}\cdot\mathbf{D}U_{k+1} + hu_k = 0.\end{aligned}$$

Since $\mathsf{T}^*\mathsf{L}_g\circ\mathsf{T}^*\mathsf{R}_f = \mathsf{T}^*\mathsf{R}_f\circ\mathsf{T}^*\mathsf{L}_g$, this reduces to

$$\mathsf{T}_e^*\mathsf{L}_{f_{k-1}}\cdot\mathbf{D}_{f_{k-1}}T_{d_{k-1}} - \mathrm{Ad}_{f_k^{-1}}^*\cdot(\mathsf{T}_e^*\mathsf{L}_{f_k}\cdot\mathbf{D}_{f_k}T_d(f_k)) + hM(g_k) + hu_k = 0, \tag{A.46}$$

which yields (5.9).

*Forced Discrete-time Hamilton's Equations* The forced discrete Legendre transformation is given by

$$\mu_k = -\mathsf{T}_e^*\mathsf{L}_{g_k}\cdot\mathbf{D}_{g_k}L_{d_k} + \mathrm{Ad}_{f_k^{-1}}^*\cdot(\mathsf{T}_e^*\mathsf{L}_{f_k}\cdot\mathbf{D}_{f_k}L_{d_k}) - u_{d_k}^-.$$

Substituting (A.44) and (A.45) into this, we obtain

$$\mu_k = \mathrm{Ad}_{f_k^{-1}}^*\cdot(\mathsf{T}_e^*\mathsf{L}_{f_k}\cdot\mathbf{D}_{f_k}T_{d_k}) - (1-c)hM(g_k) - (1-c)hu_k,$$

which yields (5.11). From this, we obtain

$$\mathsf{T}_e^*\mathsf{L}_{f_k}\cdot\mathbf{D}_{f_k}T_{d_k} = \mathrm{Ad}_{f_k}^*\cdot(\mu_k + (1-c)hM(g_k) + (1-c)hu_k).$$

Therefore $\mu_{k+1}$ can be written as

$$\mu_{k+1} = \mathrm{Ad}_{f_{k+1}^{-1}}^*\cdot(\mathsf{T}_e^*\mathsf{L}_{f_{k+1}}\cdot\mathbf{D}_{f_{k+1}}T_{d_{k+1}}) - (1-c)hM(g_{k+1}) - (1-c)hu_{k+1}.$$

We shift the time index of (A.46) by one step, and we substitute the above two equations to obtain

$$\mathrm{Ad}_{f_k}^*\cdot(\mu_k + (1-c)hM(g_k) + (1-c)hu_k) - \mu_{k+1} + chM(g_{k+1}) + chu_{k+1} = 0, \tag{A.47}$$

which is equivalent to (5.12).

## A.6 Derivatives of the Adjoint Operator

We derive expressions for the derivatives of the $\mathrm{Ad}$ operator. For $g \in \mathsf{G}$, the adjoint operator $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is the tangential map of the inner automorphism

$$\mathrm{Ad}_g \xi = \mathsf{T}_{g^{-1}} \mathsf{L}_g \cdot \mathsf{T}_e \mathsf{R}_{g^{-1}} \cdot \xi \tag{A.48}$$

where $\xi \in \mathfrak{g}$. The derivative of $\mathrm{Ad}_g \xi$ with respect to $g$ at $e$ in the direction $\eta$ corresponds to the ad operator $\mathrm{ad}_\eta \xi = [\eta,\ \xi]$ (see Marsden and Ratiu 1999).

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{\exp \epsilon\eta}\, \xi = [\eta,\ \xi]. \tag{A.49}$$

**Proposition A.1** *The derivatives of the* $\mathrm{Ad}$ *operator are given as follows.*

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{g \exp \epsilon\eta}\, \xi = \mathrm{Ad}_g [\eta,\ \xi] = [\mathrm{Ad}_g \eta,\ \mathrm{Ad}_g \xi], \tag{A.50}$$

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{(g \exp \epsilon\eta)^{-1}}\, \xi = [\mathrm{Ad}_{g^{-1}} \xi,\ \eta], \tag{A.51}$$

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}^*_{g \exp \epsilon\eta}\, \alpha = \mathrm{Ad}^*_g (\mathrm{ad}^*_{\mathrm{Ad}_g \eta} \alpha), \tag{A.52}$$

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}^*_{(g \exp \epsilon\eta)^{-1}}\, \alpha = -\mathrm{Ad}^*_{g^{-1}} (\mathrm{ad}^*_\eta \alpha). \tag{A.53}$$

**Proof.** We find the expression for the derivative of $\mathrm{Ad}_g \xi$ with respect to $g$ at $g$ in the direction $\eta$. Since $\mathrm{Ad}_{gf} = \mathrm{Ad}_g \circ \mathrm{Ad}_f$ for any $g, f \in \mathsf{G}$, we obtain

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{g \exp \epsilon\eta}\, \xi = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_g \circ \mathrm{Ad}_{\exp \epsilon\eta}\, \xi$$

Since $\mathrm{Ad}_g$ is a linear map, this is equal to

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{g \exp \epsilon\eta}\, \xi = \mathrm{Ad}_g \circ \left( \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{\exp \epsilon\eta}\, \xi \right) = \mathrm{Ad}_g [\eta,\ \xi],$$

which shows the first equality of (A.50). The second equality follows from the general property $\mathrm{Ad}_g [\eta,\ \xi] = [\mathrm{Ad}_g \eta,\ \mathrm{Ad}_g \xi]$ (see Marsden and Ratiu 1999). Similarly, we have

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{(g \exp \epsilon\eta)^{-1}}\, \xi = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{\exp(-\epsilon\eta)} \circ (\mathrm{Ad}_{g^{-1}}\, \xi) = -[\eta,\ \mathrm{Ad}_{g^{-1}}\, \xi],$$

which is equal to (A.51).

We find the expressions for the derivatives of the co-adjoint operator. For $\alpha \in \mathfrak{g}^*$, we obtain

$$\begin{aligned}\left\langle \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}^*_{g\exp\epsilon\eta}\,\alpha,\ \xi \right\rangle &= \left\langle \alpha,\ \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{g\exp\epsilon\eta}\xi \right\rangle = \langle \alpha,\ [\mathrm{Ad}_g\eta,\ \mathrm{Ad}_g\xi]\rangle \\ &= \left\langle \alpha,\ \mathrm{ad}_{\mathrm{Ad}_g\eta}(\mathrm{Ad}_g\xi)\right\rangle = \left\langle \mathrm{ad}^*_{\mathrm{Ad}_g\eta}\alpha,\ \mathrm{Ad}_g\xi \right\rangle \\ &= \left\langle \mathrm{Ad}^*_g(\mathrm{ad}^*_{\mathrm{Ad}_g\eta}\alpha),\ \xi \right\rangle,\end{aligned}$$

which yields (A.52). Similarly, we have

$$\begin{aligned}\left\langle \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}^*_{(g\exp\epsilon\eta)^{-1}}\,\alpha,\ \xi \right\rangle &= \left\langle \alpha,\ \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Ad}_{(g\exp\epsilon\eta)^{-1}}\xi \right\rangle = \left\langle \alpha,\ [\mathrm{Ad}_{g^{-1}}\xi,\ \eta]\right\rangle \\ &= \left\langle \alpha,\ -\mathrm{ad}_\eta(\mathrm{Ad}_{g^{-1}}\xi)\right\rangle = \left\langle -\mathrm{ad}^*_\eta\alpha,\ \mathrm{Ad}_{g^{-1}}\xi \right\rangle \\ &= \left\langle -\mathrm{Ad}^*_{g^{-1}}(\mathrm{ad}^*_\eta\alpha),\ \xi \right\rangle,\end{aligned}$$

which yields (A.53). ■

# BIBLIOGRAPHY

R. Abraham and J. Marsden. *Foundations of Mechanics*. Benjamin/Cummings Publishing Company, second edition, 1978.

M. Allen and D. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, 1987.

V. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 1989.

J. Baillieul. Geometric methods for nonlinear optimal control problems. *Journal of Optimization Theory and Applications*, 25:519–548, 1978.

D. Bell and D. Jacobson. *Singular Optimal Control Problems*. Academic Press, 1975.

S. Bendersky and B. Sandler. Investigation of spatial double pendulum: an engineering approach. *Discrete Dynamics in Nature and Society*, 2006:1–12, 2006.

G. Benettin and A. Giorgilli. On the Hamiltonian interpolation of near-to-the identity symplectic mappings with application to symplectic integration algorithms. *Journal of Statistical Physics*, 74(5-6):1117–1143, 1994.

L. Biedenharn and J. Louck. *Angular Momentum in Quantum Physics*. Addison-Wesley, 1981.

K. Bilimoria and B. Wie. Time-optimal three-axis reorientation of a rigid spacecraft. *Journal of Guidance, Control, and Dynamics*, 16(3):446–452, 1993.

A. Bloch. *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, 2003a.

A. Bloch. *Nonholonomic Mechanics and Control: Internet Supplement*, volume 24 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, 2003b. URL `http://www.cds.caltech.edu/mechanics_and_control/`.

A. Bloch and P. Crouch. Nonholonomic and vakonomic control systems on Riemannian manifolds. In M. Enos, editor, *Dynamics and Control of mechanical systems: the falling cat and related problems*, Fields Institute Communications, pages 25–52. American Mathematical Society, 1993.

A. Bloch and P. Crouch. Optimal control, optimization, and analytical mechanics. In J. Baillieul and J. Willems, editors, *Mathematical Control Theory*, chapter 8, pages 268–321. Springer, 1998.

A. Bloch and P. Crouch. Nonholonomic control systems on Riemannian manifolds. *SIAM Journal of Control and Optimization*, 33(1):126–148, 1995.

A. Bloch and P. Crouch. Optimal control and geodesic flows. *System and Control Letters*, 28:65–72, 1996.

A. Bloch, P. Krishnaprasad, J. Marsden, and G. Sánchez de Alvarez. Stabilization of rigid body dynamics by internal and external torques. *Automatica*, 28:745–756, 1992.

A. Bloch, P. Krishnaprasad, J. Marsden, and T. Ratiu. The Euler-Poincaré equations and double bracket dissipation. *Communications in Mathematical Physics*, 175:1–42, 1996.

A. Bloch, P. Crouch, J. Marsden, and T. Ratiu. Discrete rigid body dynamics and optimal control. In *IEEE Conference on Decision and Control*, pages 2249–2254, 1998.

A. Bloch, N. Leonard, and J. Marsden. Controlled Lagrangians and the stabilization of mechanical systems I: The first matching theorem. *IEEE Transactions on Systems and Control*, 45:2253–2270, 2000.

A. Bloch, D. Chang, N. Leonard, and J. Marsden. Controlled Lagrangians and the stabilization of mechanical systems II: Potential shaping. *IEEE Transactions on Automatic Control*, 46:1556–1571, 2001.

A. Bloch, P. Crouch, J. Marsden, and T. Ratiu. The symmetric representation of the rigid body equations and their discretization. *Nonlinearity*, 15(4):1309–1341, 2002.

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*, pages 6579–6584, 2005.

A. Bloch, M. Leok, J. Marsden, and D. 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.

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.

A. Bobenko and Y. Suris. Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top. *Communications in Mathematical Physics*, 204:147–188, 1999.

R. Brockett. Lie theory and control systems defined on spheres. *SIAM Journal of Applied Mathematics*, 25:213–225, 1973.

R. Brockett. System theory on group manifolds and coset spaces. *SIAM Journal on Control*, 10: 265–284, 1972.

A. Bryson and Y. Ho. *Applied Optimal Control*. Hemisphere Publishing Corporation, 1975.

C. Budd and A. Iserles. Geometric integration: Numerical solution of differential equations on manifolds. *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences*, 357:945–956, 1999.

F. Bullo and A. Lewis. *Geometric control of mechanical systems*, volume 49 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 2005. Modeling, analysis, and design for simple mechanical control systems.

R. Byers and S. Vadali. Quasi-closed form solution to the time-optimal rigid spacecraft reorientation problem. *Journal of Guidance, Control, and Dynamics*, 16(3):453–461, 1993.

N. Chaturvedi, T. Lee, M. Leok, and N. H. McClamroch. Nonlinear dynamics of the 3D pendulum, 2007. URL `http://arxiv.org/abs/0707.1196`.

X. Cheng, M. Jalil, and H. Lee. Time-quantified monte carlo algorithm for interacting spin array micromagnetic dynamics. *Physical Review B*, 73:224438, 2006.

G. Chirikjian and A. Kyatkin. *Engineering applications of noncommutative harmonic analysis*. CRC Press, Boca Raton, FL, 2001.

J. Cortés and S. Martínez. Nonholonomic integrators. *Nonlinearity*, 14:1365–1392, 2001.

P. Crouch and R. Grossman. Numerical integration of ordinary differential equations on manifolds. *Journal of Nonlinear Science*, 3:1–33, 1993.

L. Dieci, R. Russell, and E. van Vleck. Unitary integrators and applications to continuous orthonormalization techniques. *SIAM Journal on Numerical Analysis*, 31(1):261–281, 1994.

M. Enos, editor. *Dynamics and Control of Mechanical Systems: the falling cat and related problems*. Fields Institute Communications. American Mathematical Society, 1993.

J. Etter. A solutions of the time optimal Euler rotation problem. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, pages 1441–1449, 1989.

E. Fahnestock, T. Lee, M. Leok, N. H. McClamroch, and D. Scheeres. Polyhedral potential and variational integrator computation of the full two body problem. In *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 2006. URL `http://arxiv.org/abs/math.OC/0601424`. AIAA 2006-6289.

R. Fetecau, J. Marsden, M. Ortiz, and M. West. Nonsmooth Lagrangian mechanics and variational collision integrators. *SIAM Journal of Applied Dynamical Systems*, 2:381–416, 2003.

H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics*. Addison Wesley, 2001.

O. Gonzalez and J. Simo. On the stability of symplectic and energy-momentum algorithms for non-linear Hamiltonian systems with symmetry. *Computer Methods in Applied Mechanics and Engineering*, 134:197–222, 1996.

O. Gonzalez, D. Higham, and A. Stuart. Qualitative properties of modified equations. *IMA Journal of Numerical Analysis*, 19:169–190, 1990.

D. Greenspan. *Computer Oriented Mathematical Physics*. Pergamon Press, 1981.

D. Greenspan. New forms of discrete mechanics. *Kybernetes*, 1:87–101, 1972.

E. Hairer. Backward analysis of numerical integrators and symplectic methods. *Annals of Numerical Mathematics*, 1(1-4):107–132, 1994. Scientific computation and differential equations (Auckland, 1993).

E. Hairer and C. Lubich. Long-time energy conservation of numerical methods for oscillatory differential equations. *SIAM Journal of Numerical Analysis*, 38:414–441, 2000.

E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential Algebraic Problems*. Springer Series in Computational Mechanics 14. Springer, 1996.

E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*. Springer Series in Computational Mechanics 31. Springer, 2000.

E. Hairer, C. Lubich, and G. Wanner. Geometric numerical integration illustrated by the Störmer-Verlet method. In *Acta Numerica*, volume 12, pages 399–450. Cambridge University Press, 2003.

F. Hsiao and D. Scheeres. Fundamental constraints on uncertainty evolution in Hamiltonian systems. *IEEE Transactions on Automatic Control*, 52(4):686–691, 2007.

P. Hughes. *Spacecraft attitude dynamcis*. John Wiley & Sons, 1986.

G. Hulbert. Explicit momentum conserving algorithms for rigid body dynamics. *Computers & Structures*, 44(6):1291–1303, 1992.

I. Hussein. *Motion Planning for Multi-Spacecraft Interferometric Imaging Systems*. PhD thesis, University of Michigan, 2005.

I. Hussein and A. Bloch. Optimal control on Riemannian manifolds with potential fields. In *Proceedings of IEEE Conference on Decision and Control*, pages 1982–1987, 2004a.

I. Hussein and A. Bloch. Dynamic interpolat ion on Riemannian manifolds: An application to interferometric imaging. In *Proceeding of the American Control Conference*, pages 685–690, 2004b.

I. Hussein and A. Bloch. Optimal control of under-actuated systems with application to Lie groups. In *Proceedings of the American Control Conference*, pages 1472–1477, 2005a.

I. Hussein and A. Bloch. Optimal control of underactuated nonholonomic mechanical systems. In *Proceedings of the American Control Conference*, page 55905595, 2006.

I. Hussein and A. Bloch. Constrained optimal trajectory tracking on the group of rigid body motions. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2152–2157, 2005b.

I. Hussein, M. Leok, A. Sanyal, and A. Bloch. A discrete variational integrator for optimal control problems on SO(3). In *Proceedings of the IEEE Conference on Decision and Control*, pages 6636–6641, 2006. URL `http://arxiv.org/math.NA/0509536`.

A. Iserles, H. Munthe-Kaas, S. Nørsett, and A. Zanna. Lie-group methods. In *Acta Numerica*, volume 9, pages 215–365. Cambridge University Press, 2000.

S. Jalnapurkar, M. Leok, J. Marsden, and M. West. Discrete Routh reduction. *Journal of Physics A: Mathematical and General*, 39:5521–5544, 2006.

O. Junge, J. Marsden, and S. Ober-Blöbaum. Discrete mechanics and optimal control. In *IFAC Congress*, Praha, 2005.

O. Junge, J. Marsden, and S. Ober-Blöbaum. Optimal reconfiguration of formation flying spacecraft a decentralized approach. In *Proceedings of the IEEE Conference on Decision and Control*, pages 5210–5215, 2006.

V. Jurdjevic. Optimal control problems on Lie groups: Crossroads between geometry and mechanics. In B. Jackubczyk and W. Respondek, editors, *Geometry of Feedback and Optimal Control*, Pure and Applied Mathematics: A series of monographs and textbooks, chapter 7, pages 257–304. CRC, 1998a.

V. Jurdjevic. Optimal control, geometry, and mechanics. In J. Baillieul and J. C. Willems, editors, *Mathematical Control Theory*, chapter 7, pages 227–267. Springer, 1998b.

V. Jurdjevic. *Geometric Control Theory*. Cambridge University, 1997.

C. Kane, J. Marsden, and M. Ortiz. Symplectic-energy-momentum preserving variational integrators. *Journal of Mathematical Physics*, 40(7):3353–3371, 1999.

C. Kane, J. Marsden, M. Ortiz, and M. West. Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *International Journal for Numerical Methods in Engineering*, 49(10):1295–1325, 2000.

C. Kelley. *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.

H. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.

D. Kirk. *Optimal Control Theory: an introduction*. Prentice-Hall, 1970.

V. Kozlov and A. Harin. Kepler's problem in constant curvature spaces. *Celestial Mechanics and Dynamical Astronomy*, 54:393–399, 1992.

P. Krysl. Explicit momentum-conserving integrator for dynamics of rigid bodies approximating the midpoint Lie algorithm. *International Journal for Numerical Methods in Engineering*, 63(15): 2171–2193, 2005.

R. LaBudde and D. Greenspan. Discrete mechancis-a general treatment. *Journal of Computational Physics*, 15:134–167, 1974.

R. 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*, 25(4):323–346, 1976.

F. Lasagni. Canonial Runge-Kutta methods. *ZAMP*, 39:952–953, 1988.

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, Minneapolis, Minnesota, Jun 2005a. URL `http://arxiv.org/abs/math.NA/0509299`.

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*, pages 962–967, Toronto, Canada, Aug 2005b.

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 2170–2175, San Diego, California, Dec 2006a. URL `http://arxiv.org/abs/math.OC/0602588`.

T. Lee, A. Sanyal, M. Leok, and N. H. McClamroch. Deterministic global attitude estimation. In *Proceedings of the IEEE Conference on Decision and Control*, pages 3174–3179, San Diego, California, Dec 2006b. URL `http://arxiv.org/abs/math.OC/0602589`.

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*, pages 2689–2694, New Orleans, Louisiana, Dec 2007a. URL `http://arxiv.org/abs/math.DS/0702737`.

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 2007b. doi: 10.1007/s10569-007-9073-x.

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 2007c. doi: 10.1016/j.cma.2007.01.017.

T. Lee, M. Leok, and N. H. McClamroch. Computational geometric optimal control of rigid bodies. *Communications in Information and Systems, special issue dedicated to R. W. Brockett*, 2007d. submitted.

T. Lee, M. Leok, and N. H. McClamroch. Optimal attitude control of a rigid body using geometrically exact computations on SO(3). *Journal of Dynamical and Control Systems*, 2007e. URL `http://arxiv.org/abs/math.OC/0601424`. accepted.

T. Lee, M. Leok, and N. H. McClamroch. Optimal attitude control for a rigid body with symmetry. In *Proceedings of the American Control Conference*, pages 1073–1078, New York, July 2007f. URL `http://arxiv.org/abs/math.OC/06009482`.

T. Lee, M. Leok, and N. H. McClamroch. A combinatorial optimal control problem for spacecraft formation reconfiguration. In *Proceedings of the IEEE Conference on Decision and Control*, pages 5370–5375, New Orleans, Louisiana, Dec 2007g. URL `http://arxiv.org/abs/math.OC/0702738`.

T. Lee, M. Leok, N. H. McClamroch, and A. Sanyal. Global attitude estimation using single direction measurements. In *Proceedings of the American Control Conference*, pages 3659–3664, New York, July 2007h. URL `http://arxiv.org/abs/math.OC/06009481`.

T. Lee, M. Leok, and N. H. McClamroch. Time optimal attitude control for a rigid body. In *Proceedings of the American Control Conference*, 2008a. accepted.

T. Lee, M. Leok, and N. H. McClamroch. Global symplectic uncertainty propagation on SO(3). In *Proceedings of the IEEE Conference on Decision and Control*, 2008b. URL `http://arxiv.org/abs/0803.1515`. submitted.

B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 2004.

J. Lennard-Jones. Cohesion. *The Proceedings of the Physical Society*, 43:461–482, 1931.

M. Leok. *Foundations of computational geometric mechanics*. PhD thesis, California Instittute of Technology, 2004.

A. Lew, J. Marsden, M. Ortiz, and M. West. Asynchronous variational integrators. *Archive for Rational Mechanics and Analysis*, 167:85–146, 2003.

D. Lewis and N. Nigam. Geometric integration on spheres and some interesting applications. *Journal of Computational and Applied Mathematics*, 151(1):141–170, 2003. doi: 10.1016/S0377-0427(02)00743-4.

D. Lewis and P. Olver. Geometric integration algorithms on homogeneous manifolds. *Foundations of Computational Mathematics*, 2(4):363–392, 2001. doi: 10.1007/s102080010028.

D. Lewis and J. Simo. Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups. *Journal of Nonlinear Science*, 4:253–299, 1994.

J. Marsden. *Lectures on Mechanics*. London Mathematical Society Lecture Note Series 174. Cambridge University Press, 1992.

J. Marsden and T. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, second edition, 1999.

J. Marsden and M. West. Discrete mechanics and variational integrators. In *Acta Numerica*, volume 10, pages 317–514. Cambridge University Press, 2001.

J. Marsden, R. Montgomery, and T. Ratiu. *Reduction, Symmetry and Phases in Mechanics*. American Mathematical Society, 1990.

J. Marsden, J. Scheurle, and J. Wendlandt. Visualization of orbits and pattern evocation for the double spherical pendulum. *ZAMP*, 44(1):17–43, 1993.

J. Marsden, S. Pekarsky, and S. Shkoller. Discrete Euler–Poincaré and Lie–Poisson equations. *Nonlinearity*, 12:1647–1662, 1999.

J. Marsden, T. Ratiu, and J. Scheurle. Reduction theory and the Lagrange-Routh equations. *Journal of Mathematical Physics*, 41(6):3379–3429, 2000.

J. Marsden, S. Pekarsky, S. Shkoller, and M. West. Variational methods, multisymplectic geometry and continuum mechanics. *Journal of Geometry and Physics*, 38:253–284, 2001.

R. McLachlan and R. Quispel. *Six Lectures on The Geometric Integration of ODEs, in Foundations of Computational Mathematics*, pages 155–210. London Mathematical Society Lecture Note, 284. Cambridge University Press, 2001.

L. Meirovitch. *Methods of Analytical Dynamics*. Dover Publications, 2004.

M. Modgalya and S. Bhat. Time-optimal attitude reorientation at constant angular velocity magnitude with bounded angular acceleration. In *Proceedings of IEEE Conference on Decision and Control*, pages 223–228, 2006.

R. Montgomery. Guage theory of the falling cat. In M. Enos, editor, *Dynamics and control of mechanical systems: the falling cat and related problems*, Fields Institute Communications, pages 193–218. American Mathematical Society, 1991.

J. Moser and A. Veselov. Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Communications in Mathematical Physics*, 139:217–243, 1991.

H. Munthe-Kaas. Runge-Kutta methods on Lie groups. *BIT*, 35:572–587, 1995.

H. Munthe-Kaas and A. Zanna. *Numerical Integration of Differential Equations on Homogeneous Manifolds, in Foundations of Computational Mathematics*, pages 305–315. Springer, 1997.

R. Murray, Z. Li, and S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1993.

K. Murty. *Linear and combinatorial programming*. Wiley, 1985.

H. Nijmeijer and A. van der Schaft. *Nonlinear Dynamical Systems*. Springer, 1990.

F. Potra and W. Rheinbold. On the numerical solution of Euler-Lagrange equations. *Mechanics Based Design of Structures and Machines*, 19(1):1–18, 1991.

C. Rowley and J. Marsden. Variational integrators for degenerate lagrangians, with application to point vortices. In *Proceedings of IEEE Conference on Decision and Control*, pages 1521–1527, 2002.

A. Sanyal, T. Lee, M. Leok, and N. H. McClamroch. Global optimal attitude estimation using uncertainty ellipsoids. *Systems and Control Letters*, 57(3):236–245, 2008. doi: 10.1016/j.sysconle.2007.08.014.

J. Sanz-Serna. Symplectic integrators for Hamiltonian problems: an overview. In *Acta Numerica*, volume 1, pages 243–286. Cambridge University Press, 1992.

J. Sanz-Serna. Runge-kutta schemes for Hamiltonian systems. *BIT*, 28(4):877–883, 1988. doi: http://dx.doi.org/10.1007/BF01954907.

K. Savla, F. Bullo, and E. Frazzoli. On traveling salesperson problems for a double integrator. In *Proceedings of the IEEE Conference on Decision and Control*, pages 5305–5310, 2006.

D. Scheeres, F. Hsiao, R. Park, B. Villac, and J. Maruskin. Fundamental limits on spacecraft orbit uncertainty and distribution propagation. In *Proceedings of the AAS Malcolm D. Shuster Astronautics Symposium*, 2005. AAS 05-471.

D. Scheeres, E. Fahnestock, S. Ostro, J. Margot, L. Benner, S. Broschart, J. Bellerose, J. Giorgini, M. Nolan, C. Magri, P. Pravec, P. Scheirich, R. Rose, R. Jurgens, E. D. Jong, and S. Suzuki. Dynamical configuration of binary near-Earth asteroid (66391) 1999 KW4. *Science*, 314:1280–1283, 2006.

S. Scrivener and R. Thompson. Survey of time-optimal attitude maneuvers. *Journal of Guidance, Control, and Dynamics*, 17(2):225–233, 1994.

H. Seywald and R. Kumar. Singular control in minimum time spacecraft reorientation. *Journal of Guidance, Control, and Dynamics*, 16(4):686–694, 1993.

J. Shen, A. Sanyal, N. Chaturvedi, D. Bernstein, and N. H. McClamroch. Dynamics and control of a 3D pendulum. In *Proceedings of IEEE Conference on Decision and Control*, pages 323–328, Dec. 2004.

M. Shuster. Survey of attitude representations. *Journal of the Astronautical Sciences*, 41:439–517, 1993.

J. Simo and K. Wong. Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum. *International Journal for Numerical Methods in Engineering*, 31(1):19–52, 1991.

J. Simo, T. Posbergh, and J. Marsden. Stability of coupled rigid body and geometrically exact rods: block diagonalization and the energy-momentum method. *Physics Reports*, 193(6):279–360, 1990.

J. Simo, N. Tarnow, and K. Wong. Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics. *Computer Methods in Applied Mechanics and Engineering*, 100:63–116, 1992.

J. Stuelpnagel. On the parametrization of the three-dimensional rotation group. *SIAM Review*, 6(4): 422–430, 1964.

M. Sugiura. *Unitary Representations and Harmonic Analysis*. Kodansha, 1990.

H. Sussmann. An introduction to the coordinate-free maximum principle. In B. Jackubczyk and W. Respondek, editors, *Geometry of Feedback and Optimal Control*, Pure and Applied Mathematics: A series of monographs and textbooks, chapter 12, pages 463–557. CRC, 1998a.

H. Sussmann. Geometry and optimal control. In J. Baillieul and J. C. Willems, editors, *Mathematical Control Theory*, chapter 5, pages 140–198. Springer, 1998b.

H. Sussmann and J. Willems. 300 years of optimal control: from the Brachystrochrone to the maximum principle. *IEEE Control Systems Magazine*, 17:32–44, 1997.

V. Varadarajan. *Lie Groups, Lie Algebras, and Their Representation*. Springer, 1984.

A. Veselov. Integrable discrete-time systems and difference operators. *Journal Functional Analysis and Its Applications*, 22:83–93, 1988.

L. Wang. *Geometry, dynamics and control of coupled systems*. PhD thesis, University of Maryland, 1990.

P. Wang and F. Hadaegh. Minimum-fuel formation reconfiguration of multiple free-flying spacecraft. *Journal of the Astronautical Sciences*, 47(1-2):77–102, 1999.

J. Wendlandt and J. Marsden. Mechanical integrators derived from a discrete variational principle. *Physica D*, 106(3-4):223–246, 1997.

B. Wie. *Space Vehicle Dynamics and Control*. AIAA, 1998.

J. Yen. Constrained equations of motion in multibody dynamics as ODEs on manifolds. *SIAM Journal on Numerical Analysis*, 30(2):553–568, 1993.

D. Zenkov, A. Bloch, N. Leonard, and J. Marsden. Matching and stabilization of low-dimensional nonholonomic systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 1289–1295, 2000.

D. Zenkov, A. Bloch, N. Leonard, and J. Marsden. Flat nonholonomic matching. In *Proceedings of the American Control Conference*, pages 2812–2817, 2002.