AN OVERVIEW OF LIE GROUP VARIATIONAL INTEGRATORS AND THEIR APPLICATIONS TO OPTIMAL CONTROL

MELVIN LEOK

ABSTRACT. We introduce a general framework for the construction of variational integrators of arbitrarily high-order that incorporate Lie group techniques to automatically remain on a Lie group, while retaining the geometric structure-preserving properties characteristic of variational integrators, including symplecticity, momentum-preservation, and good long-time energy behavior. This is achieved by constructing G-invariant discrete Lagrangians in the context of Lie group methods through the use of natural charts and interpolation at the level of the Lie algebra. In the presence of symmetry, the reduction of these G-invariant Lagrangians yield a higher-order analogue of discrete Euler–Poincaré reduction.

As an illustrative example, we consider the full body problem from orbital mechanics, which is concerned with the dynamics of rigid bodies in space interacting under their mutual gravitational potential. The importance of simultaneously preserving the symplectic and Lie group properties of the full body dynamics is demonstrated in numerical simulations comparing Lie group variational integrators with integrators that are not symplectic or do not preserve the Lie group structure.

Lastly, we demonstrate the application of Lie group variational integrators to the construction of optimal control algorithms on Lie groups, and describe a modified scheme that improves the numerical efficiency of the computation, while maintaining the accuracy of the computed solutions.

CONTENTS

1. INTRODUCTION

Many problems involving long-time integration in science and engineering, such as solar system dynamics (see, for example, Sussman and Wisdom, 1992) and molecular dynamics (see, for example, Skeel et al., 1997), involve systems that are highly nonlinear, and sensitive to small perturbations. Consequently, accurately computing particular trajectories for long-time integration is typically prohibitively expensive, and it is instead desirable to construct simulations that correctly reflect qualitative properties of the system.

Geometric integrators are numerical methods that preserve the geometric structure of a continuous dynamical system (see, for example, Hairer et al., 2006; Leimkuhler and Reich, 2004, and references therein). In the problems we consider, the underlying geometric structure affects the qualitative behavior of solutions, and as such, numerical methods that preserve the geometry of a problem typically yield more qualitatively accurate simulations. This qualitative property of geometric integrators can be better understood by adopting the viewpoint that a numerical method is a discrete dynamical system that approximates the flow of the continuous system (see, for example, Benettin and Giorgilli, 1994; Tang, 1994), as opposed to the traditional

view that a numerical method approximates individual trajectories. In particular, this viewpoint allows questions about long-time stability to be addressed, which would otherwise be difficult to answer.

We consider geometric integrators based on discretizing Hamilton's principle in the context of discrete mechanics (see, for example, Marsden and West (2001)), that yield variational integrators which are automatically symplectic and momentum-preserving, as well as exhibiting good energy behavior. By incorporating essential ideas from Lie group methods (see, for example, Iserles et al. (2000)), we obtain a general framework for constructing Lie group variational integrators of arbitrarily high-order that retain the structure-preserving properties of variational integrators, while automatically evolving on Lie groups without the use of reprojection, constraints, and local coordinates. As we will demonstrate, in addition to exhibiting excellent geometric conservation properties, Lie group variational integrators are exceptionally efficient numerically. Furthermore, we will show how Lie group variational integrators serve as the basis of an efficient, geometrically exact algorithm for solving optimal control problems on Lie groups.

Outline. After recalling the construction of variational integrators, in §2 we will develop a general framework for constructing Lie group variational integrators of arbitrarily high order, and the discrete Euler–Poincaré reduction of such schemes. In §3, we explicitly construct a Lie group variational integrator for the full body problem from celestial mechanics, and perform an in depth numerical comparison with non-symplectic, symplectic, and Lie group methods of the same order of accuracy. Lastly, in §4, we demonstrate how Lie group variational integrators can be adapted to optimal control problems on Lie groups.

Related Literature. This paper is intended to be a self-contained survey of the ongoing research on Lie group variational integrators, and their applications to celestial and astrodynamics simulations, and geometric and optimal control. It is based on a body of related literature that has been performed by the author and his collaborators.

The general theory of Lie group variational integrators was developed in Leok (2004), and adapted to rigid body dynamics applications in Lee et al. (2005, 2006a, 2007a,b). Numerical comparisons for astrodynamics applications were performed in Fahnestock et al. (2006), and Lie group variational integrators were applied to actual simulations of binary near-Earth asteroids in Scheeres et al. (2006).

The abstract formulation of discrete optimal control problems on Lie groups is introduced in Bloch et al. (2006); Hussein et al. (2006), and applications to the satellite control are considered in Lee et al. (2006b, 2007c,e,f). These techniques are also applied to the problem of attitude estimation based on the use of ellipsoidal bounds on uncertainty in Lee et al. (2006c, 2007d); Sanyal et al. (2006).

2. General Theory of Lie Group Variational Integrators

We will review some of the previous work on discrete mechanics (see, for example, Marsden and West (2001)), and the construction of high-order variational integrators, before constructing Lie group analogues of variational integrators. Furthermore, we will consider the discrete Euler–Poincaré reduction (see, for example, Marsden et al. (1999)) of these Lie group variational integrators.

2.1. **Standard Formulation of Discrete Mechanics.** The standard formulation of discrete variational mechanics (see, for example, Marsden and West (2001)) is to consider the ***discrete Hamilton's principle***,

$$\delta \mathbb{S}_d = 0,$$

where the ***discrete action sum***, $\mathbb{S}_d : Q^{n+1} \to \mathbb{R}$, is given by

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

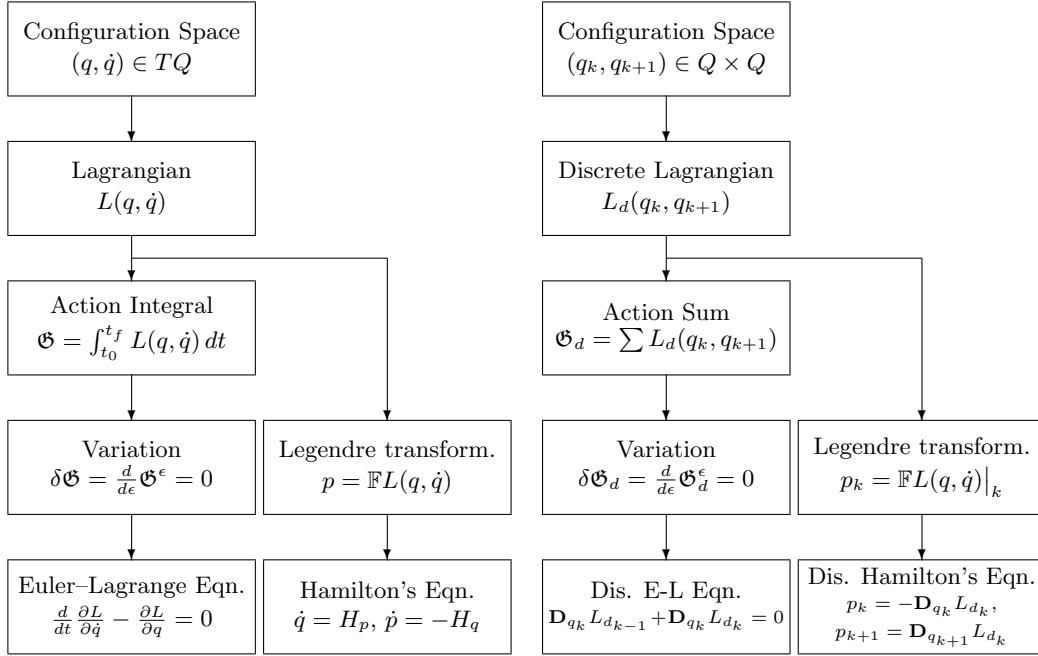| Configuration Space $(q, \dot{q}) \in TQ$		Configuration Space $(q_k, q_{k+1}) \in Q \times Q$	
Lagrangian $L(q, \dot{q})$		Discrete Lagrangian $L_d(q_k, q_{k+1})$	
Action Integral $\mathfrak{G} = \int_{t_0}^{t_f} L(q, \dot{q})\, dt$		Action Sum $\mathfrak{G}_d = \sum L_d(q_k, q_{k+1})$	
Variation $\delta\mathfrak{G} = \frac{d}{d\epsilon}\mathfrak{G}^\epsilon = 0$	Legendre transform. $p = \mathbb{F}L(q, \dot{q})$	Variation $\delta\mathfrak{G}_d = \frac{d}{d\epsilon}\mathfrak{G}_d^\epsilon = 0$	Legendre transform. $p_k = \mathbb{F}L(q, \dot{q})\big|_k$
Euler–Lagrange Eqn. $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$	Hamilton's Eqn. $\dot{q} = H_p,\ \dot{p} = -H_q$	Dis. E-L Eqn. $\mathbf{D}_{q_k} L_{d_{k-1}} + \mathbf{D}_{q_k} L_{d_k} = 0$	Dis. Hamilton's Eqn. $p_k = -\mathbf{D}_{q_k} L_{d_k}$, $p_{k+1} = \mathbf{D}_{q_{k+1}} L_{d_k}$

Figure 1: Procedures to derive continuous and discrete equations of motion

The **discrete Lagrangian**, $L_d : Q \times Q \to \mathbb{R}$, is a generating function of the symplectic flow, and is an approximation to the **exact discrete Lagrangian**,

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

where $q_{01}(0) = q_0$, $q_{01}(h) = q_1$, and q_{01} satisfies the Euler–Lagrange equation in the time interval $(0, h)$. The exact discrete Lagrangian is related to the Jacobi solution of the Hamilton–Jacobi equation. The discrete variational principle then yields the **discrete Euler–Lagrange (DEL)** equation,

$$D_2 L_d(q_0, q_1) + D_1 L_d(q_1, q_2) = 0,$$

which yields an implicit update map $(q_0, q_1) \mapsto (q_1, q_2)$ that is valid for initial conditions (q_0, q_1) that as sufficiently close to the diagonal of $Q \times Q$. The relationship between continuous and discrete variational mechanics is summarized in Figure 1.

2.2. **Lie Group Variational Integrators.** Here, we will introduce higher-order Lie group variational integrators. The basic idea behind all Lie group techniques is to express the update map of the numerical scheme in terms of the exponential map,

$$g_1 = g_0 \exp(\xi_{01}),$$

and thereby reduce the problem to finding an appropriate Lie algebra element $\xi_{01} \in \mathfrak{g}$, such that the update scheme has the desired order of accuracy. This is a desirable reduction, as the Lie algebra is a vector space, and as such the interpolation of elements can be easily defined. In our construction, the interpolatory method we use on the Lie group relies on interpolation at the level of the Lie algebra.

For a more in depth review of Lie group methods, please refer to Iserles et al. (2000). In the case of variational Lie group methods, we will express the variational problem in terms of finding Lie algebra elements, such that the discrete action is stationary.

As we will consider the reduction of these higher-order Lie group integrators, we will chose a construction that yields a G-invariant discrete Lagrangian whenever the continuous Lagrangian is G-invariant. This is achieved through the use of G-equivariant interpolatory functions, and in particular, natural charts on G.

2.2.1. *Galerkin Variational Integrators.* We first recall the construction of higher-order Galerkin variational integrators, as originally described in Marsden and West (2001). Given a Lie group G, the associated **state space** is given by the tangent bundle TG. In addition, the dynamics on G is described by a **Lagrangian**, $L : TG \to \mathbb{R}$. Given a time interval $[0, h]$, the **path space** is defined to be

$$\mathcal{C}(G) = \mathcal{C}([0, h], G) = \{g : [0, h] \to G \mid g \text{ is a } C^2 \text{ curve}\},$$

and the **action map**, $\mathfrak{S} : \mathcal{C}(G) \to \mathbb{R}$, is given by

$$\mathfrak{S}(g) \equiv \int_0^h L(g(t), \dot{g}(t))dt.$$

We approximate the action map, by numerical quadrature, to yield $\mathfrak{S}^s : \mathcal{C}([0, h], G) \to \mathbb{R}$,

$$\mathfrak{S}^s(g) \equiv h \sum_{i=1}^s b_i L(g(c_i h), \dot{g}(c_i h)),$$

where $c_i \in [0, 1]$, $i = 1, \ldots, s$ are the quadrature points, and b_i are the quadrature weights.

Recall that the discrete Lagrangian should be an approximation of the form

$$L_d(g_0, g_1, h) \approx \underset{g \in \mathcal{C}([0,h],G), g(0)=g_0, g(h)=g_1}{\text{ext}} \mathfrak{S}(g).$$

If we restrict the extremization procedure to the subspace spanned by the interpolatory function that is parameterized by $s+1$ internal points, $\varphi : G^{s+1} \to \mathcal{C}([0, h], G)$, we obtain the following discrete Lagrangian,

$$L_d(g_0, g_1) = \underset{g^\nu \in G; g^0=g_0; g^s=g_1}{\text{ext}} \mathfrak{S}(T\varphi(g^\nu; \cdot))$$

$$= \underset{g^\nu \in G; g^0=g_0; g^s=g_1}{\text{ext}} h \sum_{i=1}^s b_i L(T\varphi(g^\nu; c_i h)).$$

The interpolatory function is G-equivariant if

$$\varphi(gg^\nu; t) = g\varphi(g^\nu; t).$$

Lemma 2.1. *If the interpolatory function $\varphi(g^\nu; t)$ is G-equivariant, and the Lagrangian, $L : TG \to \mathbb{R}$, is G-invariant, then the discrete Lagrangian, $L_d : G \times G \to \mathbb{R}$, given by*

$$L_d(g_0, g_1) = \underset{g^\nu \in G; g^0=g_0; g^s=g_1}{\text{ext}} h \sum_{i=1}^s b_i L(T\varphi(g^\nu; c_i h)),$$

is G-invariant.

Proof.

$$L_d(gg_0, gg_1) = \underset{\tilde{g}^\nu \in G; \tilde{g}^0=gg_0; \tilde{g}^s=gg_1}{\text{ext}} h \sum_{i=1}^s b_i L(T\varphi(\tilde{g}^\nu; c_i h)),$$

$$= \underset{g^\nu \in g^{-1}G; g^0=g_0; g^s=g_1}{\text{ext}} h \sum_{i=1}^s b_i L(T\varphi(gg^\nu; c_i h)),$$

$$= \underset{g^\nu \in G; g^0=g_0; g^s=g_1}{\text{ext}} h \sum_{i=1}^s b_i L(TL_g \cdot T\varphi(g^\nu; c_i h)),$$

$$= \operatorname*{ext}_{g^\nu \in G; g^0 = g_0; g^s = g_1} h \sum_{i=1}^{s} b_i L(T\varphi(g^\nu; c_i h)),$$

$$= L_d(g_0, g_1),$$

where we used the G-equivariance of the interpolatory function in the third equality, and the G-invariance of the Lagrangian in the forth equality. $\qquad\square$

Remark 2.1. *While G-equivariant interpolatory functions provide a computationally efficient method of constructing G-invariant discrete Lagrangians, we can construct a G-invariant discrete Lagrangian (when G is compact) by averaging an arbitrary discrete Lagrangian. In particular, given a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$, the averaged discrete Lagrangian, given by*

$$\bar{L}_d(q_0, q_1) = \frac{1}{|G|} \int_{g \in G} L_d(gq_0, gq_1) dg$$

is G-equivariant. Therefore, in the case of compact symmetry groups, a G-invariant discrete Lagrangian always exists.

2.2.2. *Natural Charts.* Following the construction in Marsden et al. (1999), we use the group exponential map at the identity, $\exp_e : \mathfrak{g} \to G$, to construct a G-equivariant interpolatory function, and a higher-order discrete Lagrangian. As shown in Lemma 2.1, this construction yields a G-invariant discrete Lagrangian if the Lagrangian itself is G-invariant.

In a finite-dimensional Lie group G, $\exp_e$ is a local diffeomorphism, and thus there is an open neighborhood $U \subset G$ of e such that $\exp_e^{-1} : U \to \mathfrak{u} \subset \mathfrak{g}$. When the group acts on the left, we obtain a chart $\psi_g : L_g U \to \mathfrak{u}$ at $g \in G$ by

$$\psi_g = \exp_e^{-1} \circ L_{g^{-1}}.$$

We would like to construct an interpolatory function that is described by a set of control points $\{g^\nu\}_{\nu=0}^s$ in the group G at control times $0 = d_0 < d_1 < d_2 < \ldots < d_{s-1} < d_s = 1$. Our natural chart based at g^0 induces a set of control points $\xi^\nu = \psi_{g^0}^{-1}(g^\nu)$ in the Lie algebra $\mathfrak{g}$ at the same control times. Let $\tilde{l}_{\nu,s}(t)$ denote the Lagrange polynomials associated with the control times d_ν, which yields an interpolating polynomial at the level of the Lie algebra,

$$\xi_d(\xi^\nu; \tau h) = \sum_{\nu=0}^{s} \xi^\nu \tilde{l}_{\nu,s}(\tau).$$

Applying $\psi_{g^0}^{-1}$ yields an interpolating curve in G of the form,

$$\varphi(g^\nu; \tau h) = \psi_{g^0}^{-1}\Big(\sum_{\nu=0}^{s} \psi_{g^0}(g^\nu)\tilde{l}_{\nu,s}(\tau)\Big),$$

where $\varphi(d_\nu h) = g^\nu$ for $\nu = 0, \ldots, s$. Furthermore, this interpolant is G-equivariant, as shown in the following Lemma.

Lemma 2.2. *The interpolatory function given by*

$$\varphi(g^\nu; \tau h) = \psi_{g^0}^{-1}\Big(\sum_{\nu=0}^{s} \psi_{g^0}(g^\nu)\tilde{l}_{\nu,s}(\tau)\Big),$$

is G-equivariant.

Proof.

$$\varphi(gg^\nu; \tau h) = \psi_{(gg^0)}^{-1}\Big(\sum_{\nu=0}^{s} \psi_{gg^0}(gg^\nu)\tilde{l}_{\nu,s}(\tau)\Big)$$

$$= L_{gg^0} \exp_e \Big(\sum_{\nu=0}^{s} \exp_e^{-1}((gg^0)^{-1}(gg^\nu))\tilde{l}_{\nu,s}(\tau)\Big)$$

$$= L_g L_{g^0} \exp_e \Big(\sum_{\nu=0}^{s} \exp_e^{-1}((g^0)^{-1}g^{-1}gg^\nu)\tilde{l}_{\nu,s}(\tau)\Big)$$

$$= L_g \psi_{g^0}^{-1} \left(\sum_{\nu=0}^{s} \exp_e^{-1} \circ L_{(g^0)^{-1}}(g^\nu) \tilde{l}_{\nu,s}(\tau) \right)$$

$$= L_g \psi_{g^0}^{-1} \left(\sum_{\nu=0}^{s} \psi_{g^0}(g^\nu) \tilde{l}_{\nu,s}(\tau) \right)$$

$$= L_g \varphi(g^\nu; \tau h). \qquad \qquad \square$$

Remark 2.2. *In the proof that φ is G-equivariant, it was important that the base point for the chart should transform in the same way as the internal points g^ν. As such, the interpolatory function will be G-equivariant for a chart that it based at any one of the internal points g^ν that parameterize the function, but will not be G-equivariant if the chart is based at a fixed $g \in G$. Without loss of generality, we will consider the case when the chart is based at the first point g^0.*

We will now consider a discrete Lagrangian based on the use of interpolation in a natural chart, which is given by

$$L_d(g_0, g_1) = \underset{g^\nu \in G; g^0 = g_0; g^s = g_0^{-1} g_1}{\text{ext}} h \sum_{i=1}^{s} b_i L(T\varphi(\{g^\nu\}_{\nu=0}^{s}; c_i h)).$$

To further simplify the expression, we will express the extremal in terms of the Lie algebra elements ξ^ν associated with the ν-th control point. This relation is given by

$$\xi^\nu = \psi_{g_0}(g^\nu),$$

and the interpolated curve in the algebra is given by

$$\xi(\xi^\nu; \tau h) = \sum_{\kappa=0}^{s} \xi^\kappa \tilde{l}_{\kappa,s}(\tau),$$

which is related to the curve in the group,

$$g(g^\nu; \tau h) = g_0 \exp(\xi(\psi_{g_0}(g^\nu); \tau h)).$$

The velocity $\dot{\xi} = g^{-1}\dot{g}$ is given by

$$\dot{\xi}(\tau h) = g^{-1}\dot{g}(\tau h) = \frac{1}{h} \sum_{\kappa=0}^{s} \xi^\kappa \dot{\tilde{l}}_{\kappa,s}(\tau).$$

Using the standard formula for the derivative of the exponential,

$$T_\xi \exp = T_e L_{\exp(\xi)} \cdot \text{dexp}_{\text{ad}_\xi},$$

where

$$\text{dexp}_w = \sum_{n=0}^{\infty} \frac{w^n}{(n+1)!},$$

we obtain the following expression for discrete Lagrangian,

$$L_d(g_0, g_1) = \underset{\xi^\nu \in \mathfrak{g}; \xi^0 = 0; \xi^s = \psi_{g_0}(g_1)}{\text{ext}} h \sum_{i=1}^{s} b_i L\Big(L_{g_0} \exp(\xi(c_i h)),$$

$$T_{\exp(\xi(c_i h))} L_{g_0} \cdot T_e L_{\exp(\xi(c_i h))} \cdot \text{dexp}_{\text{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h)) \Big).$$

More explicitly, we can compute the conditions on the Lie algebra elements for the expression above to be extremal. This implies that

$$L_d(g_0, g_1) = h \sum_{i=1}^{s} b_i L\Big(L_{g_0} \exp(\xi(c_i h)), T_{\exp(\xi(c_i h))} L_{g_0} \cdot T_e L_{\exp(\xi(c_i h))} \cdot \text{dexp}_{\text{ad}_{\xi(c_i h)}}(\dot{\xi}(c_i h)) \Big)$$

with $\xi^0 = 0$, $\xi^s = \psi_{g_0}(g_1)$, and the other Lie algebra elements implicitly defined by

$$0 = h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial g}(c_i h) T_{\exp(\xi(c_i h))} L_{g_0} \cdot T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}} \tilde{l}_{\nu,s}(c_i) \right.$$
$$\left. + \frac{1}{h} \frac{\partial L}{\partial \dot{g}}(c_i h) T^2_{\exp(\xi(c_i h))} L_{\exp(\xi(c_i h))} \cdot T^2_e L_{\exp(\xi(c_i h))} \cdot \mathrm{ddexp}_{\mathrm{ad}_{\xi(c_i h)}} \dot{\tilde{l}}_{\nu,s}(c_i) \right] ,$$

for $\nu = 1, \ldots, s - 1$, and where

$$\mathrm{ddexp}_w = \sum_{n=0}^{\infty} \frac{w^n}{(n+2)!} .$$

This expression for the higher-order discrete Lagrangian, together with the discrete Euler–Lagrange equation,

$$D_2 L_d(g_0, g_1) + D_1 L_d(g_1, g_2) = 0 ,$$

yields a **higher-order Lie group variational integrator**.

2.3. **Higher-Order Discrete Euler–Poincaré Equations.** In this section, we will apply discrete Euler–Poincaré reduction (see, for example, Marsden et al. (1999)) to the Lie group variational integrator we derived previously, to construct a higher-order generalization of discrete Euler–Poincaré reduction.

2.3.1. *Reduced Discrete Lagrangian.* We first proceed by computing an expression for the reduced discrete Lagrangian in the case when the Lagrangian is G-invariant. Recall that our discrete Lagrangian uses G-equivariant interpolation, which, when combined with the G-invariance of the Lagrangian, implies that the discrete Lagrangian is G-invariant as well. We compute the reduced discrete Lagrangian,

$$l_d(g_0^{-1} g_1) \equiv L_d(g_0, g_1)$$
$$= L_d(e, g_0^{-1} g_1)$$
$$= \underset{\xi^\nu \in \mathfrak{g}; \xi^0 = 0; \xi^s = \log(g_0^{-1} g_1)}{\mathrm{ext}} h \sum_{i=1}^{s} b_i L\Big(L_e \exp(\xi(c_i h)),$$
$$T_{\exp(\xi(c_i h))} L_e \cdot T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}} (\dot{\xi}(c_i h)) \Big)$$
$$= \underset{\xi^\nu \in \mathfrak{g}; \xi^0 = 0; \xi^s = \log(g_0^{-1} g_1)}{\mathrm{ext}} h \sum_{i=1}^{s} b_i L\Big(\exp(\xi(c_i h)), T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}} (\dot{\xi}(c_i h)) \Big) .$$

Setting $\xi^0 = 0$, and $\xi^s = \log(g_0^{-1} g_1)$, we can solve the stationarity conditions for the other Lie algebra elements $\{\xi^\nu\}_{\nu=1}^{s-1}$ using the following implicit system of equations,

$$0 = h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial g}(c_i h) T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}} \tilde{l}_{\nu,s}(c_i) \right.$$
$$\left. + \frac{1}{h} \frac{\partial L}{\partial \dot{g}}(c_i h) T^2_e L_{\exp(\xi(c_i h))} \cdot \mathrm{ddexp}_{\mathrm{ad}_{\xi(c_i h)}} \dot{\tilde{l}}_{\nu,s}(c_i) \right]$$

where $\nu = 1, \ldots, s - 1$.

This expression for the reduced discrete Lagrangian is not fully satisfactory however, since it involves the Lagrangian, as opposed to the reduced Lagrangian. If we revisit the expression for the reduced discrete Lagrangian,

$$l_d(g_0^{-1} g_1) = \underset{\xi^\nu \in \mathfrak{g}; \xi^0 = 0; \xi^s = \log(g_0^{-1} g_1)}{\mathrm{ext}} h \sum_{i=1}^{s} b_i L\Big(\exp(\xi(c_i h)), T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}} (\dot{\xi}(c_i h)) \Big) ,$$

we find that by G-invariance of the Lagrangian, each of the terms in the summation,

$$L\Big(\exp(\xi(c_i h)), T_e L_{\exp(\xi(c_i h))} \cdot \mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}}(\dot\xi(c_i h)) \Big),$$

can be replaced by

$$l\Big(\mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}}(\dot\xi(c_i h)) \Big),$$

where $l : \mathfrak{g} \to \mathbb{R}$ is the **reduced Lagrangian** given by

$$l(\eta) = L(L_{g^{-1}}g, TL_{g^{-1}}\dot g) = L(e, \eta),$$

where $\eta = TL_{g^{-1}}\dot g \in \mathfrak{g}$.

From this observation, we have an expression for the reduced discrete Lagrangian in terms of the reduced Lagrangian,

$$l_d(g_0^{-1}g_1) = \underset{\xi^\nu \in \mathfrak{g};\, \xi^0 = 0;\, \xi^s = \log(g_0^{-1}g_1)}{\mathrm{ext}} h \sum_{i=1}^{s} b_i l\Big(\mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}}(\dot\xi(c_i h)) \Big).$$

As before, we set $\xi^0 = 0$, and $\xi^s = \log(g_0^{-1}g_1)$, and solve the stationarity conditions for the other Lie algebra elements $\{\xi^\nu\}_{\nu=1}^{s-1}$ using the following implicit system of equations,

$$0 = h \sum_{i=1}^{s} b_i \left[\frac{\partial l}{\partial \eta}(c_i h)\, \mathrm{ddexp}_{\mathrm{ad}_{\xi(c_i h)}} \dot{\tilde{l}}_{\nu,s}(c_i) \right],$$

where $\nu = 1, \dots, s-1$.

2.3.2. *Discrete Euler–Poincaré Equations.* As shown above, we have constructed a higher-order reduced discrete Lagrangian that depends on

$$f_{kk+1} \equiv g_k g_{k+1}^{-1}.$$

We will now recall the derivation of the discrete Euler–Poincaré equations, introduced in Marsden et al. (1999). The variations in f_{kk+1} induced by variations in g_k, g_{k+1} are computed as follows,

$$\delta f_{kk+1} = -g_k^{-1}\delta g_k g_k - 1 g_{k+1} + g_k^{-1}\delta g_{k+1}$$
$$= TR_{f_{kk+1}}(-g_k^{-1}\delta g_k + \mathrm{Ad}_{f_{kk+1}} g_{k+1}\delta g_{k+1}).$$

Then, the variation in the discrete action sum is given by

$$\delta \mathbb{S} = \sum_{k=0}^{N-1} l_d'(f_{kk+1})\delta f_{kk+1}$$
$$= \sum_{k=0}^{N-1} l_d'(f_{kk+1}) TR_{f_{kk+1}}(-g_k^{-1}\delta g_k + \mathrm{Ad}_{f_{kk+1}} g_{k+1}\delta g_{k+1})$$
$$= \sum_{k=1}^{N-1} \left[l_d'(f_{k-1k}) TR_{f_{k-1k}} \mathrm{Ad}_{f_{k-1k}} - l_d'(f_{kk+1}) TR_{f_{kk+1}} \right] \vartheta_k,$$

with variations of the form $\vartheta_k = g_k^{-1}\delta g_k$. In computing the variation of the discrete action sum, we have collected terms involving the same variations, and used the fact that $\vartheta_0 = \vartheta_N = 0$. This yields the **discrete Euler–Poincaré equation**,

$$l_d'(f_{k-1k}) TR_{f_{k-1k}} \mathrm{Ad}_{f_{k-1k}} - l_d'(f_{kk+1}) TR_{f_{kk+1}} = 0, \qquad\qquad k = 1, \dots, N-1.$$

For ease of reference, we will recall the expressions from the previous discussion that define the **higher-order reduced discrete Lagrangian**,

$$l_d(f_{kk+1}) = h \sum_{i=1}^{s} b_i l\Big(\mathrm{dexp}_{\mathrm{ad}_{\xi(c_i h)}}(\dot\xi(c_i h)) \Big),$$

where

$$\xi(\xi^\nu;\tau h) = \sum_{\kappa=0}^{s}\xi^\kappa \tilde{l}_{\kappa,s}(\tau)\,,$$

and

$$\xi^0 = 0\,,$$
$$\xi^s = \log(f_{kk+1})\,,$$

and the remaining Lie algebra elements $\{\xi^\nu\}_{\nu=1}^{s-1}$, are defined implicitly by

$$0 = h\sum_{i=1}^{s} b_i\left[\frac{\partial l}{\partial \eta}(c_i h)\,\mathrm{ddexp}_{\mathrm{ad}_{\xi(c_i h)}}\dot{\tilde{l}}_{\nu,s}(c_i)\right]\,,$$

for $\nu = 1,\dots,s-1$, and where

$$\mathrm{ddexp}_w = \sum_{n=0}^{\infty}\frac{w^n}{(n+2)!}\,.$$

When the discrete Euler–Poincaré equation is used in conjunction with the higher-order reduced discrete Lagrangian, we obtain the **higher-order Euler–Poincaré equations**.

2.4. **Example: Lie Group Velocity Verlet.** We will now construct a Lie group analogue of the velocity Verlet method for the free rigid body.The velocity Verlet method can be derived from the context of discrete mechanics by considering the following discrete Lagrangian,

$$L_d(q_k, q_{k+1}) = \frac{h}{2}\left[L\left(q_k, \frac{q_{k+1}-q_k}{h}\right) + L\left(q_{k+1}, \frac{q_{k+1}-q_k}{h}\right)\right]\,,$$

which corresponds to using a piecewise linear interpolant, and the trapezoidal rule to approximate the integral.

In the case of the free rigid body, the Lagrangian is given by,

$$L(R,\dot{R}) = \frac{1}{2}\Omega J\Omega^T = \frac{1}{2}\mathrm{tr}\big[S(\Omega)J_d S(\Omega)^T\big]\,.$$

Here, J_d is a modified moment of inertia that is related to the usual moment of inertia by the relations, $J_d = \frac{1}{2}(\mathrm{tr}[J]\,I_{3\times 3} - 2J)$, and $J = \mathrm{tr}[J_d]\,I_{3\times 3} - J_d$. From the kinematic relation $S(\Omega) = R^T\dot{R}$, we have that,

$$S(\Omega_k) = R_k^T\dot{R}_k \approx R_k\frac{R_{k+1}-R_k}{h} = \frac{1}{h}(F_k - I_{3\times 3}),$$

where $F_k = R_k^T R_{k+1}$. Then, the discrete Lagrangian for the velocity Verlet method applied to the free rigid body is given by,

$$\begin{aligned}
L_d(R_k, R_{k+1}) &= 2\cdot\frac{h}{2}\frac{1}{2}\frac{1}{h^2}\mathrm{tr}\big[(F_k - I_{3\times 3})^T J_d(F_k - I_{3\times 3})\big]\\
&= \frac{1}{2h}\mathrm{tr}\big[(F_k - I_{3\times 3})(F_k - I_{3\times 3})^T J_d\big]\\
&= \frac{1}{h}\mathrm{tr}\big[(I_{3\times 3} - F_k)J_d\big]\,,
\end{aligned}$$

where in the second to last equality, we used the fact that $\mathrm{tr}[AB] = \mathrm{tr}[BA]$, and in the last equality, we used the fact that F_k is an orthogonal matrix, J_d is symmetric, and $\mathrm{tr}[AB] = \mathrm{tr}\big[B^T A^T\big]$.

Recall that $\frac{\partial R^T}{\partial R}\cdot\delta R = -R^T(\delta R)R^T$. Furthermore, the variation of R_k is given by,

$$\delta R_k = R_k\eta_k,$$

where $\eta_k \in \mathfrak{so}(3)$ is a variation represented by a skew-symmetric matrix and vanishes at $k = 0$ and $k = N$. We may now compute the constrained variation of $F_k = R_k^T R_k$, which yields,

$$\delta F_k = \delta R_k^T R_{k+1} + R_k^T \delta R_{k+1} = \eta_k R_k^T R_{k+1} + R_k^T R_{k+1} \eta_{k+1} = -\eta_k F_k + F_k \eta_{k+1}.$$

Define the discrete action sum to be

$$\mathfrak{S}_d = \sum_{k=0}^{N-1} L_d(R_k, F_k).$$

Taking constrained variations of F_k yields,

$$\delta \mathfrak{S}_d = \sum_{k=0}^{N-1} \frac{1}{h} \left\{ \mathrm{tr}[-\eta_{k+1} J_d F_k] + \mathrm{tr}[\eta_k F_k J_d] \right\}.$$

Using the fact that the variations η_k vanish at the endpoints, we may reindex the sum to obtain,

$$\delta \mathfrak{S}_d = \sum_{k=1}^{N-1} \frac{1}{h} \mathrm{tr}[\eta_k (F_k J_d - J_d F_{k-1})].$$

The discrete Hamilton's principle states that the variation of the discrete action sum should be zero for all variations that vanish at the endpoints. Since η_k is an arbitrary skew-symmetric matrix, for the discrete action sum to be zero, it is necessary for $(F_k J_d - J_d F_{k-1})$ to be symmetric, which is to say that,

$$F_{k+1} J_d - J_d F_{k+1}^T - J_d F_k + F_k^T J_d = 0.$$

This implicit equation for F_{k+1} in terms of F_k, together with the reconstruction equation $R_{k+1} = R_k F_k$, yields the Lie group analogue of the velocity Verlet method.

In practice, in time marching the numerical solution, we need to solve the above equation for $F_{k+1} \in SO(3)$ given F_k. This equation is linear in F_{k+1}, but it is implicit due to the nonlinear constraint $F_{k+1}^T F_{k+1} = I_{3\times3}$. Since $J_d F_k - F_k^T J_d$ is a skew-symmetric matrix, it may be represented as $S(g)$, where $g \in \mathbb{R}^3$, which reduces the equation to the form,

$$(2.1) \qquad\qquad\qquad\qquad F J_d - J_d F^T = S(g).$$

We now introduce two iterative approaches to solve (2.1) numerically.

Exponential map. An element of a Lie group can be expressed as the exponential of an element of its Lie algebra, so $F \in SO(3)$ can be expressed as an exponential of $S(f) \in \mathfrak{so}(3)$ for some vector $f \in \mathbb{R}^3$. The exponential can be written in closed form, using Rodrigues' formula,

$$(2.2) \qquad\qquad F = \exp S(f) = I_{3\times3} + \frac{\sin \|f\|}{\|f\|} S(f) + \frac{1 - \cos \|f\|}{\|f\|^2} S(f)^2.$$

Substituting (2.2) into (2.1), we obtain

$$S(g) = \frac{\sin \|f\|}{\|f\|} S(Jf) + \frac{1 - \cos \|f\|}{\|f\|^2} S(f \times Jf).$$

Thus, (2.1) is converted into the equivalent vector equation $g = G(f)$, where $G : \mathbb{R}^3 \mapsto \mathbb{R}^3$ is given by

$$(2.3) \qquad\qquad G(f) = \frac{\sin \|f\|}{\|f\|} Jf + \frac{1 - \cos \|f\|}{\|f\|^2} f \times Jf.$$

We use the Newton method to solve $g = G(f)$, which gives the iteration

$$(2.4) \qquad\qquad f_{i+1} = f_i + \nabla G(f_i)^{-1} (g - G(f_i)).$$

We iterate until $\|g - G(f_i)\| < \epsilon$ for a small tolerance $\epsilon > 0$. The Jacobian $\nabla G(f)$ in (2.4) can be expressed as

$$\nabla G(f) = \frac{\cos\|f\|\,\|f\| - \sin\|f\|}{\|f\|^3} Jff^T + \frac{\sin\|f\|}{\|f\|} J$$
$$+ \frac{\sin\|f\|\,\|f\| - 2(1 - \cos\|f\|)}{\|f\|^4} (f \times Jf) f^T$$
$$+ \frac{1 - \cos\|f\|}{\|f\|^2} \{-S(Jf) + S(f)J\}.$$

Cayley transformation. Similarly, given $f_c \in \mathbb{R}^3$, the Cayley transformation is a local diffeomorphism that maps $S(f_c) \in \mathfrak{so}(3)$ to $F \in \mathrm{SO}(3)$, where

(2.5) $$F = \mathrm{cay}\, S(f_c) = (I_{3\times 3} + S(f_c))(I_{3\times 3} - S(f_c))^{-1}.$$

Substituting (2.5) into (2.1), we obtain a vector equation $G_c(f_c) = 0$ equivalent to (2.1)

(2.6) $$G_c(f_c) = g + g \times f_c + (g^T f_c)f_c - 2Jf_c = 0,$$

and its Jacobian $\nabla G_c(f_c)$ is written as

$$\nabla G_c(f_c) = S(g) + (g^T f_c)I_{3\times 3} + f_c g^T - 2J.$$

Then, (2.6) is solved by using Newton's iteration (2.4), and the rotation matrix is obtained by the Cayley transformation.

For both methods, numerical experiments show that 2 or 3 iterations are sufficient to achieve a tolerance of $\epsilon = 10^{-15}$. Numerical iteration with the Cayley transformation is a faster by a factor of 4-5 due to the simpler expressions in the iteration. It should be noted that since $F = \exp S(f)$ or $F = \mathrm{cay}\, S(f_c)$, it is automatically a rotation matrix, even when the equation $g = G(f)$ is not satisfied to machine precision.

These computational approaches are distinguished from solving the implicit equation (2.1) with 9 variables and 6 constraints. In the next section, we will consider a more involved numerical example for a system of rigid bodies interacting under their mutual gravitational potential.

3. Lie Group Variational Integrators for the Full Body Problem

The full body problem in orbital mechanics treats the dynamics of non-spherical rigid bodies in space interacting under their mutual potential. Since the mutual gravitational potential of distributed rigid bodies depends on both the position and the attitude of the bodies, the translational and the rotational dynamics are coupled in the full body problem. For example, the orbital motion and the attitude dynamics of a very large spacecraft in the Earth's gravity field are coupled, and the dynamics of a binary asteroid pair, with non-spherical mass distributions of the bodies, involves coupled orbital and attitude dynamics. Recently, interest in the full body problem has increased, as it is estimated that up to 16% of near-earth asteroids are binaries (Margot et al., 2002).

After introducing the continuous formulation of the full body model, we will construct a Lie group velocity Verlet method for this system, as introduced in §2.4. We then discuss in detail some of its numerical conservation properties, and compare its performance with other second-order methods.

3.1. **Full body models.** Maciejewski (1995) presented the continuous equations of motion for the full body problem in Hamiltonian form without providing a formal derivation. Here, we formulate the problem in terms of Hamilton's variational principle using the Lagrangian formalism. We then discretize the Lagrangian, and compute the proper form for the variations of Lie group elements in the configuration space, which leads to a systematic derivation of the discrete equations of motion.

3.2. **Inertial coordinates.** The configuration space of a rigid body is $\mathrm{SE}(3) = \mathbb{R}^3 \, \text{\textcircled{s}} \, \mathrm{SO}(3)$, where $\mathrm{SO}(3)$ denotes the group of 3×3 orthogonal matrices with unit determinant, and $\text{\textcircled{s}}$ represents a semi-direct product. We derive continuous equations of motion for n rigid bodies. We define an inertial frame and a body-fixed frame for each body, and assume that the origin of the i-th body-fixed frame is located at the center of mass of the i-th body.

For the i-th body, the position of the center of mass in the inertial frame, and the attitude, which is a rotation matrix from the body-fixed frame to the inertial frame, are represented by $(x_i, R_i) \in \mathrm{SE}(3)$. The translational velocity in the inertial frame and the angular velocity in the body-fixed frame are represented by $v_i, \Omega_i \in \mathbb{R}^3$. The subscript i denotes the i-th rigid body. The kinematic equations are given by

$$
\dot{x}_i = v_i \tag{3.1}
$$

$$
\dot{R}_i = R_i S(\Omega_i), \tag{3.2}
$$

where $S(\cdot) : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ is the isomorphism between the Lie algebra $\mathfrak{so}(3)$, which represents 3×3 skew-symmetric matrices, and $\mathbb{R}^3$ defined by the condition that $S(x)y = x \times y$ for any $x, y \in \mathbb{R}^3$. The mass and the moment of inertia matrix of the i-th body is denoted by $m_i \in \mathbb{R}$ and $J_i \in \mathbb{R}^{3 \times 3}$, respectively. We construct a nonstandard moment of inertia matrix $J_{d_i} \in \mathbb{R}^{3 \times 3}$ by

$$
J_{d_i} = \int_{\mathcal{B}_i} \rho_i \rho_i^T \, dm_i, \tag{3.3}
$$

where $\rho_i \in \mathbb{R}^3$ is the position of a mass element of the i-th body in its body-fixed frame. It can be shown that the standard moment of inertia matrix $J_i = \int_{\mathcal{B}_i} S(\rho_i)^T S(\rho_i) dm_i \in \mathbb{R}^{3 \times 3}$ is related to the nonstandard moment of inertia matrix by the following properties.

$$
J_i = \mathrm{tr}[J_{d_i}] \, I_{3 \times 3} - J_{d_i}, \tag{3.4}
$$

$$
S(J_i \Omega_i) = S(\Omega_i) J_{d_i} + J_{d_i} S(\Omega_i), \tag{3.5}
$$

for any $\Omega_i \in \mathbb{R}^3$. Conversely, one can obtain the nonstandard moment of inertia from the standard momentum of inertia from the following relation,

$$
J_{d_i} = \frac{1}{2} \mathrm{tr}[J_i] \, I_{3 \times 3} - J_i \,. \tag{3.6}
$$

The linear momentum in the inertial frame and the angular momentum in the body-fixed frame are denoted by $\gamma_i = m_i v_i$ and $\Pi_i = J_i \Omega_i \in \mathbb{R}^3$, respectively, for the i-th body.

Lagrangian. Given $(x_i, R_i) \in \mathrm{SE}(3)$, the inertial position of a mass element of the i-th body is given by $x_i + R_i \rho_i$, where $\rho_i \in \mathbb{R}^3$ denotes the position of the mass element in the body-fixed frame. Then, the kinetic energy of the i-th body $\mathcal{B}_i$ can be written as

$$
T_i = \frac{1}{2} \int_{\mathcal{B}_i} \| \dot{x}_i + \dot{R}_i \rho_i \|^2 \, dm_i.
$$

Using the fact that $\int_{\mathcal{B}_i} \rho_i dm_i = 0$ and (3.2), the kinetic energy T_i can be rewritten in terms of the nonstandard moment of inertia matrix as

$$
T_i(\dot{x}_i, \Omega_i) = \frac{1}{2} \int_{\mathcal{B}_i} \| \dot{x}_i \|^2 + \| S(\Omega_i) \rho_i \|^2 \, dm_i,
$$

$$
= \frac{1}{2} m_i \| \dot{x}_i \|^2 + \frac{1}{2} \mathrm{tr}\big[S(\Omega_i) J_{d_i} S(\Omega_i)^T \big], \tag{3.7}
$$

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

$$(3.8) \qquad U(x_1, \ldots, x_n, R_1, \ldots, R_n) = -\frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \int_{\mathcal{B}_i} \int_{\mathcal{B}_j} \frac{G dm_i dm_j}{\|x_i + R_i \rho_i - x_j - R_j \rho_j\|},$$

where G is the universal gravitational constant.

Then, the Lagrangian for n rigid bodies, $L : \mathrm{TSE}(3)^n \mapsto \mathbb{R}$, is given by

$$(3.9)$$
$$L(x_1, \dot{x}_1, R_1, \Omega_1, \ldots, x_n, \dot{x}_n, R_n, \Omega_n) = \sum_{i=1}^{n} \frac{1}{2} m_i \|\dot{x}_i\|^2 + \frac{1}{2} \mathrm{tr}\big[S(\Omega_i) J_{d_i} S(\Omega_i)^T\big] - U(x_1, \ldots, x_n, R_1 \ldots R_n).$$

The action integral is defined to be

$$(3.10) \qquad \mathfrak{G} = \int_{t_0}^{t_f} L(x_1, \dot{x}_1, R_1, \Omega_1, \ldots, x_n, \dot{x}_n, R_n, \Omega_n) \, dt.$$

Discrete Lagrangian. In continuous time, the structure of the kinematic equation $\dot{R}_i = R_i S(\Omega_i)$ ensures that R_i evolves on $\mathrm{SO}(3)$ automatically. Here, we introduce a new variable $F_{i_k} \in \mathrm{SO}(3)$ defined such that $R_{i_{k+1}} = R_{i_k} F_{i_k}$, i.e.

$$(3.11) \qquad F_{i_k} = R_{i_k}^T R_{i_{k+1}}.$$

Thus, F_{i_k} represents the relative attitude between two integration steps, and by requiring that $F_{i_k} \in \mathrm{SO}(3)$, we guarantee that R_{i_k} evolves on $\mathrm{SO}(3)$ automatically. This is a consequence of the fact that the Lie group is closed under the group operation of matrix multiplication.

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

$$(3.12) \qquad S(\Omega_{i_k}) = R_{i_k}^T \dot{R}_{i_k} \approx R_{i_k}^T \frac{R_{i_{k+1}} - R_{i_k}}{h} = \frac{1}{h}(F_{i_k} - I_{3\times 3}).$$

The velocity $\dot{x}_{i_k}$ can be approximated simply by $(x_{i_{k+1}} - x_{i_k})/h$. Using these approximations of the angular and linear velocity, the kinetic energy of the ith body given in (3.7) can be approximated as

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

A discrete Lagrangian $L_d : \mathrm{SE}(3)^n \times \mathrm{SE}(3)^n \mapsto \mathbb{R}$ is constructed such that it approximates a segment of the action integral (3.10),

$$(3.13) \quad L_d = \sum_{i=1}^{n} \frac{1}{2h} m_i \|x_{i_{k+1}} - x_{i_k}\|^2 + \frac{1}{h} \mathrm{tr}[(I_{3\times 3} - F_{i_k}) J_{d_i}] - \frac{h}{2} U(x_{1_k}, \ldots, R_{n_k}) - \frac{h}{2} U(x_{1_{k+1}}, \ldots, R_{n_{k+1}}).$$

This discrete Lagrangian is self-adjoint (Hairer et al., 2006), and self-adjoint numerical integration methods have even order, so we are guaranteed that the resulting integration method is at least second-order accurate.

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

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

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

$$\delta R_{i_k} = R_{i_k} \eta_{i_k}, \tag{3.14}$$

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

$$\begin{aligned}
\delta F_{i_k} &= \delta R_{i_k}^T R_{i_{k+1}} + R_{i_k}^T \delta R_{i_{k+1}}, \\
&= -\eta_{i_k} R_{i_k}^T R_{i_{k+1}} + R_{i_k}^T R_{i_{k+1}} \eta_{i_{k+1}}, \\
&= -\eta_{i_k} F_{i_k} + F_{i_k} \eta_{i_{k+1}}.
\end{aligned} \tag{3.15}$$

Discrete Hamilton's principle. To obtain the discrete equations of motion in Lagrangian form, we compute the variation of the discrete Lagrangian from (3.14) and (3.15) to give

$$\begin{aligned}
\delta L_d = \sum_{i=1}^{n} \frac{1}{h} m_i (x_{i_{k+1}} - x_{i_k})^T (\delta x_{i_{k+1}} - \delta x_{i_k}) + \frac{1}{h} \operatorname{tr}\big[(\eta_{i_k} F_{i_k} - F_{i_k} \eta_{i_{k+1}}) J_{d_i} \big] \\
- \frac{h}{2} \left(\frac{\partial U_k}{\partial x_{i_k}}^T \delta x_{i_k} + \frac{\partial U_{k+1}}{\partial x_{i_{k+1}}}^T \delta x_{i_{k+1}} \right) + \frac{h}{2} \operatorname{tr}\left[\eta_{i_k} R_{i_k}^T \frac{\partial U_k}{\partial R_{i_k}} + \eta_{i_{k+1}} R_{i_{k+1}}^T \frac{\partial U_{k+1}}{\partial R_{i_{k+1}}} \right],
\end{aligned} \tag{3.16}$$

where $U_k = U(x_{1_k}, \ldots, R_{n_k})$ denotes the value of the potential at $t = kh + t_0$.

Define the action sum as

$$\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(x_{1_k}, x_{1_{k+1}}, R_{1_k}, F_{1_k}, \ldots, x_{n_k}, x_{n_{k+1}}, R_{n_k}, F_{n_k}). \tag{3.17}$$

The discrete action sum $\mathfrak{G}_d$ approximates the action integral (3.10), because the discrete Lagrangian approximates a segment of the action integral. Substituting (3.16) into (3.17), the variation of the action sum is given by

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

Using the fact that δx_{i_k} and η_{i_k} vanish at $k = 0$ and $k = N$, we can reindex the summation, which is the discrete analogue of integration by parts, to yield

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

Hamilton's principle states that $\delta \mathfrak{G}_d$ should be zero for all possible variations $\delta x_{i_k} \in \mathbb{R}^3$ and $\eta_{i_k} \in \mathfrak{so}(3)$ that vanish at the endpoints. Therefore, the expression in the first brace should be zero, and since η_{i_k} is skew-symmetric, the expression in the second brace should be symmetric.

Discrete equations of motion. We obtain *the discrete equations of motion for the full body problem, in Lagrangian form,* for bodies $i \in (1, 2, \cdots, n)$ as

$$\frac{1}{h} (x_{i_{k+1}} - 2x_{i_k} + x_{i_{k-1}}) = -h \frac{\partial U_k}{\partial x_{i_k}}, \tag{3.18}$$

$$(3.19) \qquad \frac{1}{h}\left(F_{i_{k+1}}J_{d_i} - J_{d_i}F_{i_{k+1}}^T - J_{d_i}F_{i_k} + F_{i_k}^T J_{d_i}\right) = hS(M_{i_{k+1}}),$$

$$(3.20) \qquad R_{i_{k+1}} = R_{i_k}F_{i_k},$$

where $M_{i_k} \in \mathbb{R}^3$ is given by

$$(3.21) \qquad M_{i_k} = r_{i_1} \times u_{i_1} + r_{i_2} \times u_{i_2} + r_{i_3} \times u_{i_3},$$

where $r_{i_p}, u_{i_p} \in \mathbb{R}^{1\times3}$ are pth row vectors of R_{i_k} and $\frac{\partial U_k}{\partial R_{i_k}}$, respectively. Given initial conditions $(x_{i_0}, R_{i_0}, x_{i_1}, R_{i_1})$, we can obtain x_{i_2} from (3.18). Then, F_{i_0} is computed from (3.20), and F_{i_1} can be obtained by solving the implicit equation (3.19). Finally, R_{i_2} is found from (3.20). This yields an update map $(x_{i_0}, R_{i_0}, x_{i_1}, R_{i_1}) \mapsto (x_{i_1}, R_{i_1}, x_{i_2}, R_{i_2})$, and this process can be repeated.

As discussed above, equations (3.18) through (3.20) defines a discrete Lagrangian map that updates x_{i_k} and R_{i_k}. The discrete Legendre transformation relates the configuration variables x_{i_k}, R_{i_k} and the corresponding momenta γ_{i_k}, Π_{i_k}. This induces a discrete Hamiltonian map that is equivalent to the discrete Lagrangian map.

The discrete equations of motion for the full body problem, in Hamiltonian form, can be written for bodies $i \in (1, 2, \cdots, n)$ as

$$(3.22) \qquad x_{i_{k+1}} = x_{i_k} + \frac{h}{m_i}\gamma_{i_k} - \frac{h^2}{2m_i}\frac{\partial U_k}{\partial x_{i_k}},$$

$$(3.23) \qquad \gamma_{i_{k+1}} = \gamma_{i_k} - \frac{h}{2}\frac{\partial U_k}{\partial x_{i_k}} - \frac{h}{2}\frac{\partial U_{k+1}}{\partial x_{i_{k+1}}},$$

$$(3.24) \qquad hS(\Pi_{i_k} + \frac{h}{2}M_{i_k}) = F_{i_k}J_{d_i} - J_{d_i}F_{i_k}^T,$$

$$(3.25) \qquad \Pi_{i_{k+1}} = F_{i_k}^T\Pi_{i_k} + \frac{h}{2}F_{i_k}^T M_{i_k} + \frac{h}{2}M_{i_{k+1}}\cdot,$$

$$(3.26) \qquad R_{i_{k+1}} = R_{i_k}F_{i_k}.$$

Given $(x_{i_0}, \gamma_{i_0}, R_{i_0}, \Pi_{i_0})$, we can find x_{i_1} from (3.22). Solving the implicit equation (3.24) yields F_{i_0}, and R_{i_1} is computed from (3.26). Then, (3.23) and (3.25) gives γ_{i_1}, and Π_{i_1}. This defines the discrete Hamiltonian map, $(x_{i_0}, \gamma_{i_0}, R_{i_0}, \Pi_{i_0}) \mapsto (x_{i_1}, \gamma_{i_1}, R_{i_1}, \Pi_{i_1})$, and this process can be repeated.

3.3. **Properties of the Lie group variational integrator.** Since the LGVI is obtained by discretizing Hamilton's principle, it is symplectic and preserves the structure of the configuration space, SE(3), as well as the relevant geometric features of the full two rigid body problem, and the conserved first integrals of total linear and angular momenta. The total energy exhibits small bounded oscillations about its initial value, but there is no tendency for the mean of the oscillation in the total energy to drift (increase or decrease) from the initial value for exponentially long times.

The LGVI preserves the group structure. By using the computational approach described in §2.4, the matrices F_{i_k} representing the change in relative attitude are guaranteed to be rotation matrices. The group operation of the Lie group SO(3) is matrix multiplication. Since the rotation matrices R_{i_k} are updated using the group operation, they automatically evolve on SO(3) without constraints or reprojection. Therefore, the orthogonal structure of the rotation matrices is preserved, and the attitude of each rigid body is determined accurately and globally without the need to use local charts (parameterizations) such as Euler angles or quaternions. These exact geometric properties of the discrete flow not only generate improved qualitative behavior, but also allow for accurate long-time simulation.

This geometrically exact numerical integration method yields a highly efficient and accurate computational algorithm for the full rigid body problem. For arbitrary shaped rigid bodies such as binary asteroids, there is a large burden in computing the mutual gravitational forces and moments, so the number of force and moment evaluations should be minimized. We have seen that the LGVI requires only one such evaluation

per integration step, the minimum number of evaluations consistent with the presented LGVI having second-order accuracy (because it is a self-adjoint method). Within the LGVI, implicit equations must be solved at each time step to determine the matrix-multiplication updates for rotation matrices. However the LGVI is only weakly implicit in the sense that the iteration for each implicit equation is independent of the much more costly gravitational force and moment computation. The computational load to solve each implicit equation is negligible; only two or three iterations are typically required. This is addressed in §2.4 by expressing F_{i_k} as the exponential function of an element of the Lie algebra $\mathfrak{so}(3)$. Altogether, the entire method could be considered *almost explicit.*

The LGVI is a fixed step size integrator, but all of the properties above are independent of the step size. Consequently, we can achieve the same level of accuracy while choosing a larger step size as compared to other numerical integrators of the same order.

All of these features are revealed by numerical simulations in §3.4 and in the work by Fahnestock et al. (2006). In §3.4, the LGVI is compared with other second-order geometric integrators: a symplectic Runge–Kutta method and a Lie group method. In Fahnestock et al. (2006), the LGVI is directly compared with the 7(8)th order Runge–Kutta–Fehlberg method (RK78) for two octahedral rigid bodies. It is shown that the LGVI requires 8 times less computational load than RK78 for similar error measures, and the accuracy of the LGVI is maintained for exponentially long time. The trajectories computed using RK78 are unreliable for the long time simulation of the full two rigid body dynamics.

3.4. **Numerical simulations.** We simulate the dynamics of two simple dumbbell bodies acting under their mutual gravity. Each dumbbell model consists of two equal rigid spheres and a rigid massless connecting rod. This dumbbell rigid body model has a simple closed form for the mutual gravitational potential given by

$$U(X, R) = - \sum_{p,q=1}^{2} \frac{Gm_1 m_2/4}{\left\| X + \rho_{2_p} + R\rho_{1_q} \right\|},$$

where G is the universal gravitational constant, $m_i \in \mathbb{R}$ is the total mass of the ith dumbbell, and $\rho_{i_p} \in \mathbb{R}^3$ is a vector from the origin of the body-fixed frame to the pth sphere of the ith dumbbell in the ith body-fixed frame. The vectors $\rho_{i_1} = [l_i/2, 0, 0]^T$, $\rho_{i_2} = -\rho_{i_1}$, where l_i is the length between the two spheres. Mass, length and time dimensions are normalized.

The mass and length of the second dumbbell are twice that of the first dumbbell. The other simulation parameters are chosen such that the total linear momentum in the inertial frame is zero and the relative motion between two bodies are near-elliptic orbits. The trajectories of the dumbbell bodies are shown in Figure 2.

We compare the computational properties of the Lie group variational integrator (LGVI) with other second-order numerical integration methods; an explicit Runge–Kutta method (RK), a symplectic Runge–Kutta method (SRK), and a Lie group method (LGM). One of the distinct features of the LGVI is that it preserves both the symplectic property and the Lie group structure of the full rigid body dynamics. A comparison can be made between the LGVI and other integration methods that preserve either none or one of these properties: an integrator that does not preserve any of these properties (RK), a symplectic integrator that does not preserve the Lie group structure (SRK), and a Lie group integrator that does not preserve symplecticity (LGM). These methods are implemented by an explicit mid-point rule, an implicit mid-point rule, and the Crouch-Grossman method presented in Hairer et al. (2006) for the continuous equations of motion, respectively. For the LGVI, the discrete equations of motion given by (3.18) through (3.21) 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(a) shows the computed total energy response over 30 seconds with an integration step size $h = 0.002$ sec. For the LGVI, the total energy is nearly constant, and there is no tendency to drift, while
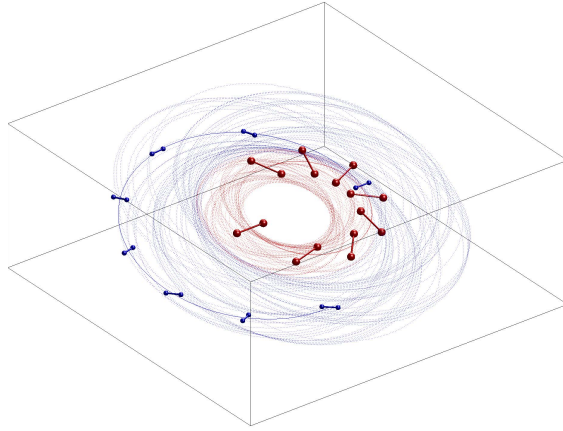
Figure 2: Trajectories of two dumbbell bodies in the inertial frame (The initial orbit is shown with solid lines and snapshots of dumbbell body maneuver.)

the other integrators fail to preserve the total energy. This can be observed in Figure 3(b), where the mean total energy deviations are shown for varying integration step sizes. It is seen that the total energy errors of the SRK method is close to the RK method, but the total energy error of the LGVI is smaller by several orders of magnitude. Figure 3(c) shows the mean orthogonality errors. The LGVI and the LGM conserve the orthogonal structure at an error level of 10^{-10}, while the RK and the SRK do not.

These computational comparisons suggest that for numerical integration of Hamiltonian systems evolving on a Lie group, such as full body problems, it is critical to preserve both the symplectic property and the Lie group structure. For the RK and the SRK, the orthogonality error in the rotation matrix corrupts the attitude of the rigid bodies. The accumulation of this attitude degradation causes significant errors in the computation of the gravitational forces and moments, which are dependent upon the position and the attitude, and affects 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 features of the full rigid body dynamics concurrently. Consequently, the LGVI yields numerical trajectories which are more qualitatively accurate, and the qualitative advantages of the LGVI become more pronounced as the length of the simulation is increased.

Computational efficiency is compared in Figure 3(d), where CPU times of all methods are shown for varying step sizes. The SRK requires the largest CPU time, since it involves the 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 the expensive force and moment computations at each step. The LGVI requires only one evaluation at each step in addition to the solution of an implicit equation (3.24). The implicit equation can be solved efficiently using the computational approach described in §2.4 and hence it takes less time than the evaluation of the forces and moments. 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 superior computational efficiency of the LGVI.
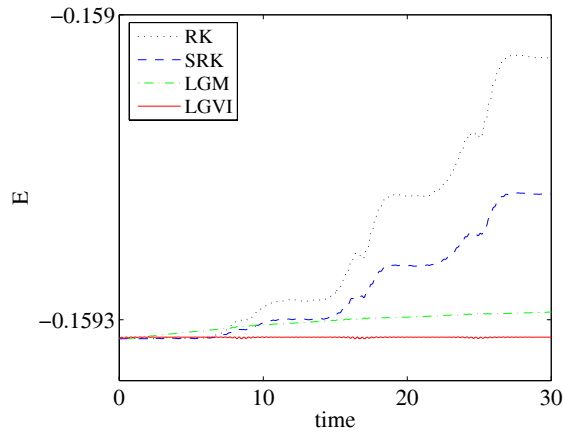
(a) Computed total energy for 30 seconds
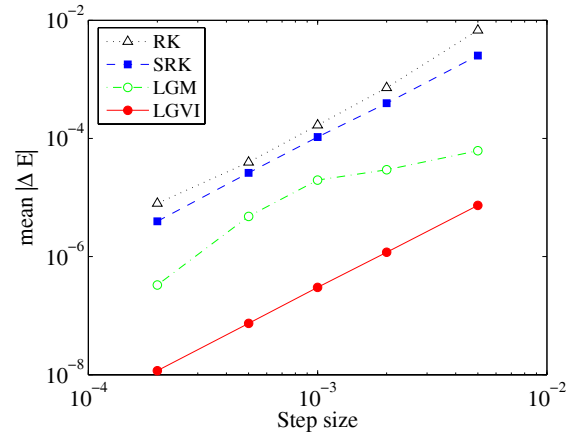
(b) Mean total energy error $|E - E_0|$ vs. step size
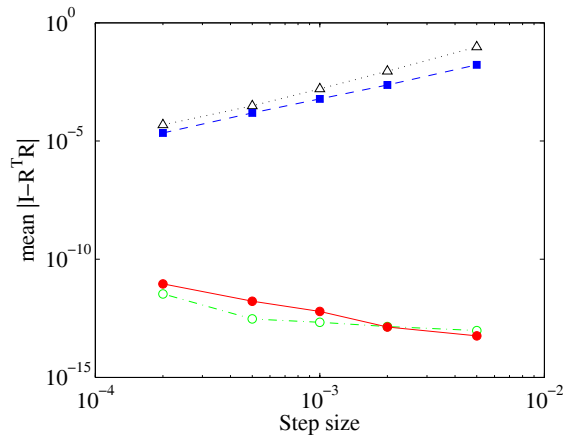
(c) Mean orthogonality error $\|I - R^T R\|$ vs. step size
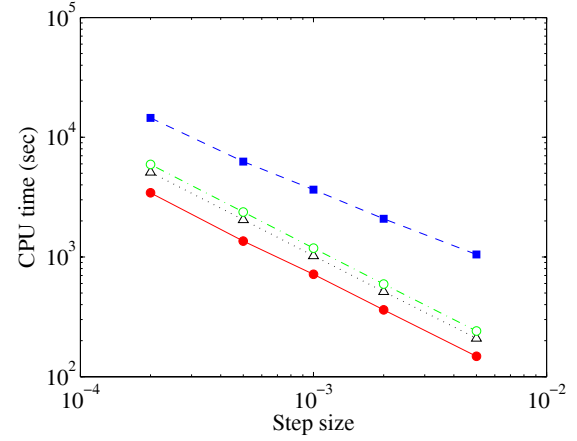
(d) CPU time vs. step size

Figure 3: Computational properties of explicit Runge–Kutta (RK), symplectic Runge–Kutta (SRK), Lie group method (LGM), and Lie group variational integrator (LGVI).

In summary, from Figures 3(b) and 3(d), we see that the LGVI requires 16 times less CPU time than the LGM, 35 times less CPU time than the RK, and 98 times less CPU time than the SRK, to achieve a similar total energy error in this computational example for the full body problem.

4. Discrete Optimal Control on Lie Groups

In this section, we apply the Lie group variational integrators which have been introduced in the previous section to the problem of optimal control on Lie groups. We will first derive first-order optimality conditions for discrete optimal solution, and then describe an efficient shooting based method for solving the discrete optimal control problem.

Our approach to discretizing the optimal control problem is in contrast to traditional techniques such as collocation, wherein the continuous equations of motion are imposed as constraints at a set of collocation points. In our approach, modeled after Junge et al. (2005), the discrete equations of motion are derived from a discrete variational principle, and this induces constraints on the configuration at each discrete time step.

This approach yields discrete dynamics that are more faithful to the continuous equations of motion, and consequently yields more accurate numerical solutions to the optimal control problem. This feature is extremely important in computing accurate (sub)optimal trajectories for long-term spacecraft attitude maneuvers.

For the purpose of numerical simulation, the corresponding discrete optimal control problem is posed on the discrete state space as a two-stage discrete variational problem. In the first step, we derive the discrete dynamics for the rigid body in the sense of Lie group variational integrators. These discrete equations are then imposed as constraints to be satisfied by the extremal solutions to the discrete optimal control problem, and we obtain the discrete extremal solutions in terms of the given terminal states.

We formulate an optimal control problem for a rigid body on SE(3) assuming that control forces and moments are applied during the maneuver. Necessary conditions for optimality are developed and computational approaches are presented to solve the corresponding two-point boundary value problem.

4.1. Lie Euler variational integrator.
In order to simplify the form of the necessary conditions for the optimal control problem, we consider a first-order variant of the Lie group variational integrator. Define a discrete Lagrangian L_d as

$$(4.1) \qquad L_d(R_k, F_k) = \frac{1}{h}\mathrm{tr}[(I_{3\times 3} - F_k)\, J_d] - hU(R_{k+1}).$$

This discrete Lagrangian is a first-order approximation of the integral of the continuous Lagrangian over one integration step. Therefore, the action sum, $\mathfrak{G}_d = \sum_{k=0}^{N-1} L_d(R_k, F_k)$, which is defined to be the summation of the discrete Lagrangian, approximates the action integral. Taking a variation of the action sum, we obtain the discrete equations of motion using the discrete Lagrange–d'Alembert principle. The variation of a rotation matrix can be expressed using the exponential of a Lie algebra element:

$$R_k^\epsilon = R_k e^{\epsilon \eta_k},$$

where $\epsilon \in \mathbb{R}$ and $\eta_k \in \mathfrak{so}(3)$ is the variation expressed as a skew-symmetric matrix. Thus, the infinitesimal variation is given by $\delta R_k = R_k \eta_k$. The Lagrange–d'Alembert principle states that the following equation is satisfied for all possible variations $\eta_k \in \mathfrak{so}(3)$.

$$(4.2) \qquad \delta \sum_{k=0}^{N-1} \frac{1}{h}\mathrm{tr}[(I_{3\times 3} - F_k)\, J_d] - hU(R_{k+1}) - \sum_{k=0}^{N-1} \frac{h}{2}\mathrm{tr}[\eta_{k+1} S(Bu_{k+1})] = 0.$$

Here, the second summation approximates the virtual work done by the external forces. Using the expression of the infinitesimal variation of a rotation matrix and using the fact that the variations vanish at the end

points, the above equation can be written as

$$\sum_{k=1}^{N-1} \text{tr}\left[\eta_k \left\{\frac{1}{h}\left(F_k J_d - J_d F_{k-1}\right) + h R_k^T \frac{\partial U}{\partial R_k} - \frac{h}{2} S(Bu_k)\right\}\right] = 0.$$

Since the above expression should be zero for all possible variations $\eta_k \in \mathfrak{so}(3)$, the expression in the braces should be symmetric. Then, *the discrete equations of motion in Lagrangian form* are given by

$$(4.3) \qquad \frac{1}{h}\left(F_{k+1} J_d - J_d F_k - J_d F_{k+1}^T + F_k^T J_d\right) = hS(M_{k+1}) + hS(Bu_{k+1}),$$

$$(4.4) \qquad\qquad\qquad\qquad R_{k+1} = R_k F_k.$$

Using the discrete version of the Legendre transformation, *the discrete equations of motion in Hamiltonian form* are given by

$$(4.5) \qquad\qquad\qquad hS(\Pi_k) = F_k J_d - J_d F_k^T,$$

$$(4.6) \qquad\qquad\qquad R_{k+1} = R_k F_k,$$

$$(4.7) \qquad\qquad\qquad \Pi_{k+1} = F_k^T \Pi_k + h\left(M_{k+1} + Bu_{k+1}\right).$$

Given (R_k, Π_k), we can obtain F_k by solving (4.5), and R_{k+1} is obtained by (4.6). Finally, Π_{k+1} is updated by (4.7). This yields a map $(R_k, \Pi_k) \mapsto (R_{k+1}, \Pi_{k+1})$, and this process can be repeated. The only implicit part is solving (4.5). We can express (4.5) in terms of a Lie algebra element $S(f_k) = \text{logm}(F_k) \in \mathfrak{so}(3)$, and find $f_k \in \mathbb{R}^3$ numerically by a Newton iteration. The relative attitude F_k is obtained by the exponential map: $F_k = e^{S(f_k)}$. Therefore we are guaranteed that F_k is a rotation matrix.

The order of the variational integrator is equal to the order of the corresponding discrete Lagrangian. Consequently, the above Lie group variational integrator is of first-order since (4.1) is a first-order approximation. While higher-order variational integrators can be obtained by modifying (4.1), we use the first-order integrator because it yields a compact form for the necessary conditions that preserves the geometry; these necessary conditions are developed in §4.4. Also, in §4.5, we will see that while our method is formally first-order, it shadows the numerical trajectory of a second-order method.

4.2. **Problem formulation.** An optimal impulsive control problem is formulated as a maneuver of a rigid body from a given initial configuration $(R_0, x_0, \Pi_0, \gamma_0)$ to a desired configuration $(R_N^d, x_N^d, \Pi_N^d, \gamma_N^d)$ during the given maneuver time N. Control inputs are parameterized by their value at each time step. The performance index is the square of the weighted l_2 norm of the control inputs.

$$\text{given: } (x_0, \gamma_0, R_0, \Pi_0),\ (x_N^d, \gamma_N^d, R_N^d, \Pi_N^d),\ N,$$

$$\min_{u_{k+1}} \mathcal{J} = \sum_{k=0}^{N-1} \frac{h}{2}(u_{k+1}^f)^T W_f u_{k+1}^f + \frac{h}{2}(u_{k+1}^m)^T W_m u_{k+1}^m,$$

$$\text{such that } (x_N, \gamma_N, R_N, \Pi_N) = (x_N^d, \gamma_N^d, R_N^d, \Pi_N^d),$$

$$\text{subject to discrete equations of motion,}$$

where $W_f, W_m \in \mathbb{R}^{3\times3}$ are symmetric positive definite matrices. Here, we use our Lie Euler variational integrator (4.5)-(4.7), as it yields a compact form for the necessary conditions, which are developed in §4.4.

4.3. **Sensitivity derivatives.**

Variational model. The variation of $g_k = (R_k, x_k) \in \text{SE}(3)$ can be expressed in terms of a Lie algebra $\eta_k \in \mathfrak{se}(3)$ and the exponential map as

$$g_k^\epsilon = g_k \exp \epsilon \eta_k.$$

The corresponding infinitesimal variation is given by

$$\delta g_k = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} g_k \exp \epsilon \eta_k = \mathrm{T}_e L_{g_k} \cdot \eta_k.$$

Using homogeneous coordinates (Murray et al., 1993), the above equation is written as

$$\begin{bmatrix} \delta R_k & \delta x_k \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} R_k & x_k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S(\zeta_k) & \chi_k \\ 0 & 0 \end{bmatrix},$$

(4.8)
$$= \begin{bmatrix} R_k S(\zeta_k) & R_k \chi_k \\ 0 & 0 \end{bmatrix},$$

where $\zeta_k, \chi_k \in \mathbb{R}^3$ so that $(S(\zeta_k), \chi_k) \in \mathfrak{se}(3)$. This gives an expression for the infinitesimal variation in terms of the Lie algebra. Then, small perturbations from the given discrete trajectory on $\mathrm{T}^*\mathrm{SE}(3)$ can be written as

(4.9)
$$x_k^\epsilon = x_k + \epsilon \delta x_k,$$

(4.10)
$$\gamma_k^\epsilon = \gamma_k + \epsilon \delta \gamma_k,$$

(4.11)
$$\Pi_k^\epsilon = \Pi_k + \epsilon \delta \Pi_k,$$

(4.12)
$$R_k^\epsilon = R_k + \epsilon R_k S(\zeta_k) + \mathcal{O}(\epsilon^2),$$

where $\delta x_k, \delta \gamma_k, \delta \Pi_k, \zeta_k \in \mathbb{R}^3$.

Since the force due to the potential depends on the position and the attitude, its variation can be written as

$$\delta f_k = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} f_k(x_k + \epsilon \delta x_k, R_k + \epsilon R_k S(\zeta_k)).$$

Since this operation is linear in δx_k and ζ_k, we can express δf_k as

(4.13)
$$\delta f_k = \mathcal{F}_{x_k} \delta x_k + \mathcal{F}_{R_k} \zeta_k,$$

where $\mathcal{F}_{x_k}, \mathcal{F}_{R_k} \in \mathbb{R}^{3\times3}$. Similarly, the variation of the moment due to the potential can be written as

(4.14)
$$\delta M_k = \mathcal{M}_{x_k} \delta x_k + \mathcal{M}_{R_k} \zeta_k,$$

where $\mathcal{M}_{x_k}, \mathcal{M}_{R_k} \in \mathbb{R}^{3\times3}$. Since $F_k = R_k^T R_{k+1}$, the infinitesimal variation δF_k is given by

$$\delta F_k = \delta R_k^T R_{k+1} + R_k^T \delta R_{k+1},$$
$$= -S(\zeta_k) F_k + F_k S(\zeta_{k+1}).$$

We can also express $\delta F_k = F_k S(\xi_k)$ for $\xi_k \in \mathbb{R}^3$, using (4.8). Using the property $S(R^T x) = R^T S(x) R$ for all $R \in \mathrm{SO}(3)$ and $x \in \mathbb{R}^3$, we obtain

(4.15)
$$\xi_k = -F_k^T \zeta_k + \zeta_{k+1}.$$

Linearized equations of motion. Substituting the variation model (4.9)–(4.14) and the constrained variation (4.15) into the equations of motion (3.22)–(3.26), and ignoring higher-order terms, the linearized equation of motion can be written as

(4.16)
$$z_{k+1} = A_k z_k,$$

where $z_k = [\delta x_k; \delta \gamma_k; \zeta_k; \delta \Pi_k] \in \mathbb{R}^{12}$, and the matrix $A_k \in \mathbb{R}^{12\times12}$ can be suitably defined. The solution of (4.16) is given by

(4.17)
$$z_N = \left(\prod_{k=0}^{N-1} A_k \right) z_0 = \Phi z_0,$$

where $\Phi \in \mathbb{R}^{12 \times 12}$ represents the sensitivity derivatives of the terminal state with respect to the initial state on SE(3).

4.4. **Necessary conditions for optimality.** Define an augmented performance index as

$$\mathcal{J}_a = \sum_{k=0}^{N-1} \frac{h}{2}(u_{k+1}^f)^T W^f u_{k+1}^f + \frac{h}{2}(u_{k+1}^m)^T W^m u_{k+1}^m + \lambda_k^{1,T}\left\{-x_{k+1} + x_k + \frac{h}{m}\gamma_k\right\}$$

$$+ \lambda_k^{2,T}\left\{-\gamma_{k+1} + \gamma_k + hf_{k+1} + hu_{k+1}^f\right\} + \lambda_k^{3,T} S^{-1}\left(\operatorname{logm}(F_k - R_k^T R_{k+1})\right)$$

$$+ \lambda_k^{4,T}\left\{-\Pi_{k+1} + F_k^T \Pi_k + h\left(M_{k+1} + u_{k+1}^m\right)\right\},$$

where $\lambda_k^i \in \mathbb{R}^3$ is a Lagrange multiplier corresponding to first-order expressions of the discrete equations of motion. The constraint (3.24) is considered implicitly using a constrained variation. The infinitesimal variation of $\mathcal{J}_a$ is

$$\delta\mathcal{J}_a = \sum_{k=0}^{N-1} h\delta u_{k+1}^{f,T} W_f u_{k+1}^f + h\delta u_{k+1}^{m,T} W_m u_{k+1}^m + \lambda_k^{1,T}\left\{-\delta x_{k+1} + \delta x_k + \frac{h}{m}\delta\gamma_k\right\}$$

$$+ \lambda_k^{2,T}\left\{-\delta\gamma_{k+1} + \delta\gamma_k + h\delta f_{k+1} + h\delta u_{k+1}^f\right\} + \lambda_k^{3,T}\left\{\xi_k + F_k^T \zeta_k - \zeta_{k+1}\right\}$$

(4.18)
$$+ \lambda_k^{4,T}\left\{-\delta\Pi_{k+1} + \delta F_k^T \Pi_k + F_k^T \delta\Pi_k + h\delta M_{k+1} + h\delta u_{k+1}^m\right\}$$

Instead of taking a variation of the matrix logarithm, the constrained variation (4.15) is applied.

We can find an expression for ξ_k using (4.5). Substituting the expressions and (4.13), (4.14) into the above equation and using the fact that the variations vanish at $k = 0, N$, we obtain

$$\delta\mathcal{J}_a = \sum_{k=1}^{N-1} h\delta u_k^{f,T}\left\{W_f u_k^f + \lambda_{k-1}^2\right\} + h\delta u_k^{m,T}\left\{W_m u_k^m + \lambda_{k-1}^4\right\}$$

$$+ \delta x_k^T\left\{-\lambda_{k-1}^1 + A^{11,T}\lambda_k^1 + A_k^{21,T}\lambda_k^2 + A_k^{41,T}\lambda_k^4\right\} + \delta\gamma_k^T\left\{-\lambda_{k-1}^2 + A^{12,T}\lambda_k^1 + A_k^{22,T}\lambda_k^2 + A^{42,T}\lambda_k^4\right\}$$

$$+ \zeta_k^T\left\{-\lambda_{k-1}^3 + A^{23,T}\lambda_k^2 + A_k^{33,T}\lambda_k^3 + A^{43,T}\lambda_k^4\right\} + \delta\Pi_k^T\left\{-\lambda_{k-1}^4 + A_k^{24,T}\lambda_k^2 + A_k^{34,T}\lambda_k^3 + A_k^{44,T}\lambda_k^4\right\},$$

where $A_k^{ij} \in \mathbb{R}^{3 \times 3}$ are 3×3 blocks of the matrix A_k presented in (4.16). For example,

$$A_k^{21} = h\mathcal{F}_{x_{k+1}}, \qquad\qquad\qquad A_k^{41} = h\mathcal{M}_{x_{k+1}},$$

$$A_k^{22} = I_{3\times3} + \frac{h^2}{m}\mathcal{F}_{x_{k+1}}, \qquad\qquad A_k^{42} = \frac{h^2}{m}\mathcal{M}_{x_{k+1}},$$

$$A_k^{33} = F_k^T, \qquad\qquad\qquad A_k^{43} = h\mathcal{M}_{R_{k+1}} A_k^{33},$$

$$A_k^{34} = hF_k^T \left\{\operatorname{tr}[F_k J_d]\, I_{3\times3} - F_k J_d\right\}^{-1}, \qquad A_k^{44} = F_k^T + h\mathcal{M}_{R_{k+1}} A_k^{34} + S(F_k^T \Pi_k) A_k^{34}.$$

Since $\delta\mathcal{J}_a = 0$ for all variations, we obtain necessary conditions for optimality as follows.

(4.19)
$$x_{k+1} = x_k + \frac{h}{m}\gamma_k,$$

(4.20)
$$\gamma_{k+1} = \gamma_k + hf_{k+1} + hu_{k+1}^f,$$

(4.21)
$$hS(\Pi_k) = F_k J_d - J_d F_k^T,$$

(4.22)
$$R_{k+1} = R_k F_k,$$

(4.23)
$$\Pi_{k+1} = F_k^T \Pi_k + hM_{k+1} + hu_{k+1}^m,$$

$$(4.24) \qquad\qquad u_{k+1}^f = -W_f^{-1}\lambda_k^2,$$

$$(4.25) \qquad\qquad u_{k+1}^m = -W_m^{-1}\lambda_k^4,$$

$$(4.26) \qquad\qquad \lambda_k = A_{k+1}^T\lambda_{k+1},$$

where $\lambda_k = [\lambda_k^1; \lambda_k^2; \lambda_k^3; \lambda_k^4] \in \mathbb{R}^{12}$. In the above equations, the only implicit part is (4.21). For a given initial condition $(R_0, x_0, \Pi_0, \gamma_0)$ and λ_0, we can find F_0 by solving (4.21). Then, R_1, x_1 is obtained by (4.22),(4.19), and the control input u_1^f, u_1^m is obtained by (4.24),(4.25). γ_1, Π_1 can be obtained by (4.20),(4.23), since f_1, M_1 depend only on x_1, R_1. Now we computed $(R_1, x_1, \Pi_1, \gamma_1)$. We solve (4.21) to find F_1. Finally, λ_1 can be obtained by (4.26), since A_1 are functions of $x_1, \gamma_1, R_1, \Pi_1, F_1$. This yields a map $\{(R_0, x_0, \Pi_0, \gamma_0), \lambda_0\} \mapsto \{(R_1, x_1, \Pi_1, \gamma_1), \lambda_1\}$, and this process can be repeated.

4.5. **Lie Euler and Symplectic Equivalence.** While the Lie group method we have constructed for the optimal control problem is formally first-order, we will show that it is symplectically equivalent to a second-order method, and consequently shadows the numerical trajectory generated by a second-order Lie group variational integrator. In this way, our optimal control algorithm is expressed in a form that can be efficiently solved using a shooting based approach to the two-point boundary value problem, while retaining the qualitative accuracy associated with a second-order method.

Notice that the discrete Lagrangian adopted in this section is obtained by approximating the velocity as a constant over the timestep h, and by approximating the integral in time by $\int_{t_1}^{t_2} f(t)dt \approx (t_2 - t_1)f(t_1)$. This is to say that the action integral is approximated by its left Riemann sum. In the Lie group setting, the constant angular velocity approximation corresponds to the condition,

$$R_{k+1} = R_k \exp(h\Omega_k)$$

or equivalently,

$$\Omega_k = \frac{1}{h}\log(R_k^{-1}R_{k+1}).$$

When we let $G = \mathbb{R}^n$, and we adopt the notation $(q, v) \in T\mathbb{R}^n$, we obtain,

$$v_k = \frac{q_{k+1} - q_k}{h},$$

which is a usual finite-difference approximation for the velocity. Consider then a Lagrangian of the form,

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

Approximating the action integral from 0 to h using a constant velocity approximation, yields,

$$\int_0^h L(q(t), v(t))dt \approx \int_0^h L\Big(q(t), \frac{q_{k+1} - q_k}{h}\Big)dt \approx hL\Big(q_k, \frac{q_{k+1} - q_k}{h}\Big).$$

We then choose as our discrete Lagrangian,

$$L_d(q_k, q_{k+1}) = hL\Big(q_k, \frac{q_{k+1} - q_k}{h}\Big) = h\left[\frac{1}{2}\Big(\frac{q_{k+1} - q_k}{h}\Big)^T M\Big(\frac{q_{k+1} - q_k}{h}\Big) - V(q_k)\right].$$

The discrete Euler–Lagrange equations,

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

yields,

$$M\Big(\frac{q_k - q_{k-1}}{h}\Big) - M\Big(\frac{q_{k+1} - q_k}{h}\Big) - h\frac{\partial V}{\partial q}(q_k) = 0,$$

which induces an implicit update map $(q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$. To obtain the corresponding Hamiltonian update map, we push-forward this algorithm to T^*Q by using the discrete fiber derivative $\mathbb{F}L_d : Q \times Q \to T^*Q$, which takes $(q_k, q_{k+1}) \mapsto (q_{k+1}, D_2 L_d(q_k, q_{k+1}))$. In particular, we have that,

$$p_{k+1} = D_2 L_d(q_k, q_{k+1}) = M\left(\frac{q_{k+1} - q_k}{h}\right),$$

which implies

$$(4.27) \qquad\qquad q_{k+1} = q_k + h M^{-1} p_{k+1}.$$

This allows us to rewrite the discrete Euler–Lagrange equations as,

$$p_k - p_{k+1} - h \frac{\partial V}{\partial q}(q_k) = 0,$$

or equivalently,

$$(4.28) \qquad\qquad p_{k+1} = p_k - h \frac{\partial V}{\partial q}(q_k).$$

Now, (4.27) and (4.28) are precisely the symplectic Euler method applied to the corresponding Hamiltonian vector field, as we shall see.

The corresponding Hamiltonian is given by,

$$H(q, p) = \frac{1}{2} p^T M^{-1} p + V(q).$$

Hamilton's equations yield,

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix} = \begin{pmatrix} M^{-1} p \\ -\frac{\partial V}{\partial q} \end{pmatrix}.$$

The symplectic Euler method has the form,

$$q_{k+1} = q_k + h\dot{q}(q_k, p_{k+1}),$$
$$p_{k+1} = p_k + h\dot{p}(q_k, p_{k+1}),$$

which yields,

$$q_{k+1} = q_k + h M^{-1} p_{k+1},$$
$$p_{k+1} = p_k + h\left(-\frac{\partial V}{\partial q}(q_k)\right),$$

which is precisely what we obtained in (4.27) and (4.28). This demonstrates that our method is the generalization of the symplectic Euler method to Lie groups, which has important numerical consequences. While symplectic Euler is formally first-order accurate, it is symplectically equivalent (Littell et al., 1997; Suzuki, 1993) to the second-order accurate Störmer–Verlet method (Hairer et al., 2003). This means that one can obtain the Störmer–Verlet method F_{SV} by conjugating the symplectic Euler method F_{E} with a symplectic transformation T,

$$F_{\mathrm{SV}} = T F_{\mathrm{E}} T^{-1}.$$

In particular, numerical trajectories of symplectic Euler will shadow numerical trajectories obtained using Störmer–Verlet. Consider the implications of this symplectic equivalence for our discrete optimal control problem. Let the boundary conditions be specified by q_0, q_N, and assume that we use Störmer–Verlet to propagate the solution, then the boundary condition is expressed as, $q_N = F_{\mathrm{SV}}^N q_0 = (T F_{\mathrm{E}} T^{-1})^N q_0 = T F_{\mathrm{E}}^N T^{-1} q_0$, which is equivalent to $\tilde{q}_N = T^{-1} q_N = F_{\mathrm{E}}^N T^{-1} q_0 = F_{\mathrm{E}}^N \tilde{q}_0$. This implies that if we preprocess the boundary conditions q_0, q_N, to obtain $\tilde{q}_0 = T^{-1} q_0, \tilde{q}_N = T^{-1} q_N$, we could use symplectic Euler at the internal stages to propagate the states and costates, and then postprocess them to obtain the trajectory one would have obtained by using Störmer–Verlet.

In practice, the shadowing result imparts the symplectic Euler method with the same desirable qualitative properties as Störmer–Verlet, and it is not necessary to postprocess the numerical solutions in order to achieve accurate results. Since on an appropriate choice of charts, our Lie symplectic Euler method reduces to symplectic Euler in coordinates, it follows that there is a corresponding second-order Lie Störmer–Verlet method that our method is symplectically equivalent to.

4.6. **Computational Approach.** The necessary conditions for optimality are given by a two-point boundary problem on $T^*SE(3)$ and its dual. This problem is to find the optimal discrete flow, multiplier, and control inputs to satisfy the equations of motion (4.19)–(4.23), optimality conditions (4.24),(4.25), multiplier equations (4.26), and boundary conditions simultaneously.

We use a neighboring extremal method (Bryson and Ho, 1975). A nominal solution satisfying all of the necessary conditions except the boundary conditions is chosen. The unspecified initial multiplier is updated by successive linearization so as to satisfy the specified terminal boundary conditions in the limit. This is also referred to as a shooting method. The main advantage of the neighboring extremal method is that the number of iteration variables is minimized. It is equal to the dimension of the equations of motion. In other approaches, the initial guess of control input history or multiplier variables are iterated, so the number of optimization parameters are proportional to the number of discrete time steps.

The difficulty is that the extremal solutions are sensitive to small changes in the unspecified initial multiplier values. The nonlinearity also make it hard to construct an accurate estimate of sensitivity, and it may yields numerical ill-conditioning. Therefore, it is important to compute the sensitivities accurately to apply the neighboring extremal method.

Here the optimality conditions (4.24) and (4.25) are substituted into the equations of motion and the multiplier equations. The sensitivities of the specified terminal boundary conditions with respect to the unspecified initial multiplier conditions is obtained by a linear analysis.

Similar to (4.16), the linearized equations of motion can be written as

$$(4.29) \qquad z_{k+1} = A_k z_k + \mathcal{A}^{12} \delta \lambda_k,$$

where $\mathcal{A}_k^{12} = -h \mathrm{diag}[0, W_f^{-1}, 0, W_m^{-1}] \in \mathbb{R}^{12 \times 12}$. We can linearize the multiplier equations (4.26) to obtain

$$(4.30) \qquad \delta \lambda_k = \mathcal{A}_{k+1}^{21} z_{k+1} + A_{k+1}^T \delta \lambda_{k+1},$$

where $\mathcal{A}_{k+1}^{21} \in \mathbb{R}^{12 \times 12}$ can be defined properly. The solution of the linear equations (4.29) and (4.30) can be obtained as

$$\begin{bmatrix} z_N \\ \delta \lambda_N \end{bmatrix} = \begin{bmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{bmatrix} \begin{bmatrix} z_0 \\ \delta \lambda_0 \end{bmatrix},$$

where $\Psi^{ij} \in \mathbb{R}^{12 \times 12}$.

For the given two-point boundary value problem $z_0 = 0$ since the initial condition is fixed, and λ_N is free. Then, we obtain

$$(4.31) \qquad z_N = \Psi_{12} \delta \lambda_0$$

Thus, the matrix Ψ_{12} represents the sensitivity of the specified terminal boundary conditions with respect to the unspecified initial multipliers. Using this sensitivity, an initial guess of the unspecified initial conditions is iterated to satisfy the specified terminal conditions in the limit.

Any type of Newton iteration can be applied to this problem using the sensitivity derivatives as gradients. We use a line search with backtracking algorithm, referred to as Newton-Armijo iteration in Kelley (1995). The procedure is summarized as follows.

1: Guess an initial multiplier λ_0.
2: Find $\Pi_k, R_k, \lambda_k^1, \lambda_k^2$ using (4.19)–(4.26).
3: Compute the error in satisfaction of the terminal boundary condition;
   Error $= \| z_N \|$.

4: Set $\text{Error}^t = \text{Error}$, $\;i = 1$.
5: **while** $\text{Error} > \epsilon_S$.
6:      Find a line search direction; $D = \Psi_{12}^{-1}$.
7:      Set $c = 1$.
8:          **while** $\text{Error}^t > (1 - 2\alpha c)\text{Error}$
9:          Choose a trial initial multiplier $\lambda_0^t = \lambda_0 + cDz_N$.
10:         Find $\Pi_k, R_k, \lambda_k^1, \lambda_k^2$ using (4.19)–(4.26).
11:         Compute the error $\text{Error}^t = \|z_N^t\|$.
12:         Set $c = c/10$, $\;i = i + 1$.
13:         **end while**
14:     Set $\lambda_0 = \lambda_0^t$, $\text{Error} = \text{Error}^t$. (accept the trial)
15: **end while**

Here i is the number of iterations, and $\epsilon_S, \alpha \in \mathbb{R}$ are a stopping criterion and a scaling factor, respectively. The outer loop finds a search direction by computing the sensitivity derivatives, and the inner loop performs a line search to find the largest step size $c \in \mathbb{R}$ along the search direction. The error in satisfaction of the terminal boundary condition is determined at each inner iteration.

4.7. **Numerical Examples.** We study an optimal orbital transfer problem to increase an orbital inclination by $60\,\deg$, and an orbital capture problem to the reference circular orbit. The corresponding boundary conditions are given by

(i) Orbital inclination change

$$x_0 = [1, 0, 0], \qquad\qquad x_N^d = [-0.353, 0.353, 0.866],$$

$$R_0 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad R_N^d = - \begin{bmatrix} 0.70 & -0.35 & -0.61 \\ 0.70 & 0.35 & 0.61 \\ 0 & 0.86 & -0.5 \end{bmatrix},$$

$$\gamma_0 = m[0, 0.983, 0], \quad \gamma_N^d = m[-0.695, -0.695, 0],$$

$$\Pi_0 = J[0, 0, 0.983], \quad \Pi_N^d = J[0, 0, 0.983].$$

(ii) Orbital capture

$$x_0 = [-5.196, -3, -1], \qquad\qquad x_N^d = [1, 0, 0],$$

$$R_0 = \begin{bmatrix} 0.88 & -0.22 & -0.42 \\ 0.07 & 0.94 & -0.33 \\ 0.46 & 0.26 & 0.84 \end{bmatrix}, \quad R_N^d = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\gamma_0 = m[0.983, 0, 0], \qquad\qquad \gamma_N^d = m[0, 0.983, 0],$$

$$\Pi_0 = J[0, 0, 0], \qquad\qquad \Pi_N^d = J[0, 0, 0.983].$$

Figures 4 and 5 show the optimized spacecraft maneuver and the control input history. For each case, the performance indices are 13.03 and 20.90, and the maximum violations of the constraint are 3.35×10^{-13} and 3.26×10^{-13}, respectively.

Figures 4.(b) and 5.(b) show the violation of the terminal boundary condition according to the number of iterations in a logarithm scale. 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 solution is sufficiently close to the local minimum at around the 20th iteration. These convergence results are consistent with the quadratic convergence rates expected of Newton methods with accurately computed gradients.

The neighboring extremal method, also referred to as the shooting method, is numerically efficient in the sense that the number of optimization parameters is minimized. But, this approach is prone to numerical
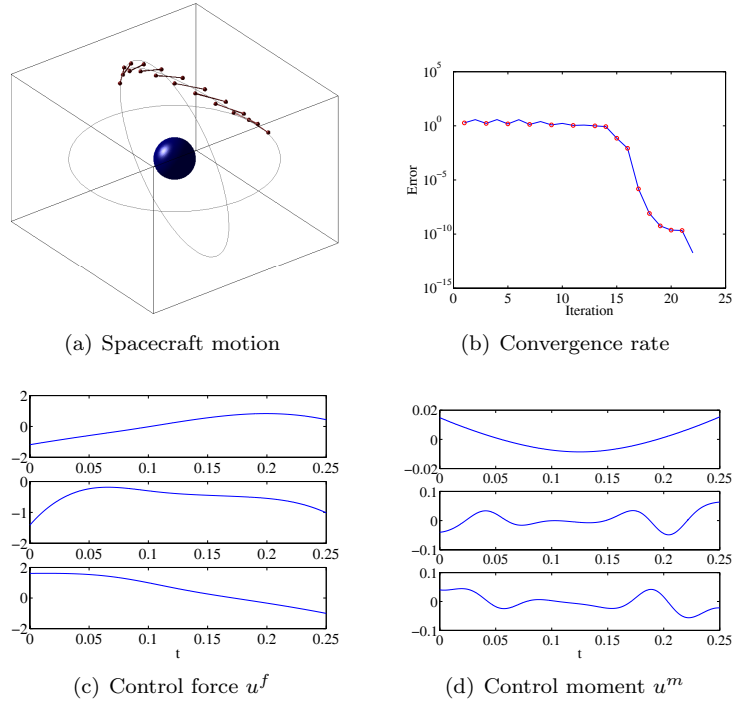
(a) Spacecraft motion

(b) Convergence rate

(c) Control force u^f

(d) Control moment u^m

Figure 4: Optimal control: Orbital inclination change

(a) Spacecraft motion

(b) Convergence rate

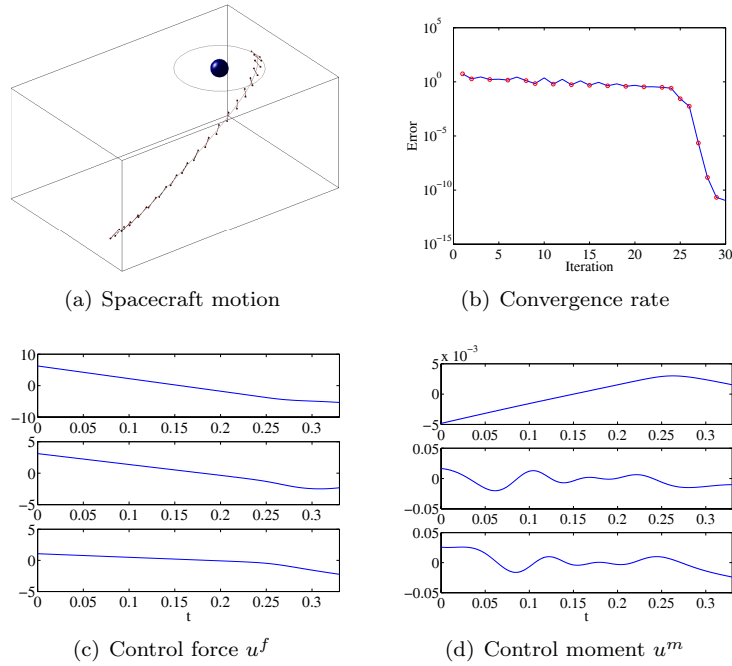(c) Control force u^f

(d) Control moment u^m

Figure 5: Optimal control: Orbital capture

ill-conditioning (Betts, 2001). A small change in the initial multiplier can cause highly nonlinear behavior of the terminal attitude and angular momentum. It is difficult to compute the gradient for Newton iterations accurately, and consequently, the numerical error may not converge. However, the numerical examples presented in this paper show excellent numerical convergence properties. This is because the proposed computational algorithms on SE(3) are geometrically exact and numerically accurate.

The dynamics of a rigid body arises from Hamiltonian mechanics, which have neutral stability, and its adjoint system is also neutrally stable. The proposed Lie group variational integrator and the discrete multiplier equations, obtained from variations expressed in the Lie algebra, preserve the neutral stability property numerically. As a consequence, the sensitivity derivatives are computed accurately.

5. Conclusions

In this paper, we have introduced a synthesis of Lie group methods and variational integrators, to yield Lie group variational integrators of arbitrarily high-order. They are an example of a more general class of methods referred to as generalized Galerkin variational integrators that were introduced in Leok (2004), which also include spatio-temporally adaptive, multiscale, and pseudospectral variational integrators.

Lie group variational integrators are geometrically exact; they preserve the momenta and symplectic form of the continuous dynamics, exhibit good energy properties, and they also conserve the geometry of the configuration space. They provide a numerically efficient computational approach especially for the full body problems in the sense that they require only one evaluation of mutual gravity forces and moments per step. The exact geometric properties of the discrete flow not only yields improved qualitative behavior, but also allow for accurate long-time simulation. The numerical example demonstrates that the simultaneous preservation of the Lie group structure and symplecticity is necessary for good energy stability, due to the dependence of the potential term on the relative attitude.

A discrete optimal control problem is also formulated in a manner that includes the use of Lie group variational integrators as a means of discretely imposing the equations of motion. Since the numerical trajectory automatically evolves on the Lie group, as opposed to being constrained with Lagrange multipliers to do so, the resulting optimal control algorithm is exceptionally efficient. By computing the sensitivity derivatives accurately at the level of the Lie algebra, one avoids the numerical ill-conditioning typically encountered in solving two-point boundary value problems. As demonstrated in the numerical example, once the basin of attraction is reached, the numerical optimization converges very rapidly.

In summary, Lie group variational integrators are a class of geometric integrators that, by virtue of their excellent numerical efficiency and geometric structure-preserving properties, are particularly appropriate for long-time simulations of rigid body models arising in celestial and astrodynamics. They also serve as the basis of optimal control algorithms capable of efficiently solving the challenging problem of continuously actuated optimal control on Lie groups.

References

G. Benettin and A. Giorgilli. On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms. *J. Stat. Phys.*, 74:1117–1143, 1994.

J. T. Betts. *Practical Methods for Optimal Control Using Nonlinear Programming*. SIAM, 2001.

A. M. Bloch, I. I. Hussein, M. Leok, and A. K. Sanyal. Discrete optimal control on Lie groups. 2006. Preprint, available at `http://www.math.purdue.edu/~mleok/pdf/BlHuLeSa-NM-06.pdf`.

A. E. Bryson and Y.-C. Ho. *Applied Optimal Control*. Hemisphere Publishing Corporation, 1975.

E. G. Fahnestock, T. Lee, M. Leok, N. H. McClamroch, and D. J. Scheeres. Polyhedral potential and variational integrator computation of the full two body problem. In *AIAA/AAS Astrodynamics Specialist Meeting*, August 2006. AIAA-2006-6289.

E. Hairer, C. Lubich, and G. Wanner. Geometric numerical integration illustrated by the Störmer-Verlet method. *Acta Numer.*, 12:399–450, 2003.

E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, second edition, 2006.

I. I. Hussein, M. Leok, A. K. Sanyal, and A.M. Bloch. A discrete variational integrator for optimal control problems in $SO(3)$. In *Proceedings of the IEEE Conference on Decision and Control*, pages 6636–6641, 2006.

A. Iserles, H. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. In *Acta Numerica*, volume 9, pages 215–365. Cambridge University Press, 2000.

O. Junge, J. E. Marsden, and S. Ober-Blöbaum. Discrete mechanics and optimal control. In *IFAC Congress*, Praha, 2005.

C. T. Kelley. *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.

T. Lee, M. Leok, and N. H. McClamroch. A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum. In *Proceedings of the IEEE Conference on Control Applications*, pages 962–967, 2005.

T. Lee, N. H. McClamroch, and M. Leok. Attitude maneuvers of a rigid spacecraft in a circular orbit. In *Proceedings of the American Control Conference*, pages 1742–1747, 2006a.

T. Lee, N. H. McClamroch, and M. Leok. Optimal control of a rigid body using geometrically exact computations on $SE(3)$. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2710–2715, 2006b.

T. Lee, A. K. Sanyal, M. Leok, and N. H. McClamroch. Deterministic global attitude estimation. In *Proceedings of the IEEE Conference on Decision and Control*, pages 3174–3179, 2006c.

T. Lee, M. Leok, and N. H. McClamroch. Lie group variational integrators for the full body problem. *Computer Methods in Applied Mechanics and Engineering*, 2007a. To appear, available at `http://arxiv.org/math.NA/0508365`.

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*, 2007b. To appear, available at `http://www.math.purdue.edu/~mleok/pdf/LeLeMc2006_CMDA.pdf`.

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*, 2007c. Submitted, available at `http://www.math.purdue.edu/~mleok/pdf/LeLeMc2007_oac.pdf`.

T. Lee, M. Leok, A. K. Sanyal, and N. H. McClamroch. Global attitude estimation using single direction measurement. In *Proceedings of the American Control Conference*, 2007d. To appear, available at `http://arxiv.org/math.OC/0609481`.

T. Lee, N. H. McClamroch, and M. Leok. Optimal attitude control for a rigid body with symmetry. In *Proceedings of the American Control Conference*, 2007e. To appear, available at `http://arxiv.org/math.OC/0609482`.

T. Lee, N. H. McClamroch, and M. Leok. A combinatorial optimal control problem for spacecraft formation reconfiguration. In *Proceedings of the IEEE Conference on Decision and Control*, 2007f. Submitted, available at `http://arxiv.org/math/0702738`.

B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 2004.

M. Leok. Generalized Galerkin variational integrators. 2004. Preprint, available at `http://arxiv.org/math.NA/0508360`.

T. R. Littell, R. D. Skeel, and M. Zhang. Error analysis of symplectic multiple time stepping. *SIAM J. Numer. Anal.*, 34(5):1792–1807, 1997.

A. J. Maciejewski. Reduction, relative equilibria and potential in the two rigid bodies problem. *Celestial Mechanics and Dynamical Astronomy*, 63:1–28, 1995.

J. L. Margot, M. C. Nolan, L. A. M. Benner, S. J. Ostro, R. F. Jurgens, J. D. Giorgini, M. A. Slade, and D. B. Campbell. Binary Asteroids in the Near-Earth object population. *Science*, 296:1445–1448, 2002.

J. E. Marsden and M. West. Discrete mechanics and variational integrators. In *Acta Numerica*, volume 10, pages 317–514. Cambridge University Press, 2001.

J. E. Marsden, S. Pekarsky, and S. Shkoller. Discrete Euler–Poincaré and Lie–Poisson equations. *Nonlinearity*, 12(6):1647–1662, 1999.

R. M. Murray, Z. Li, and S. S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1993.

A. K. Sanyal, T. Lee, M. Leok, and N. H. McClamroch. Global optimal attitude estimation using uncertainty ellipsoids. *Systems and Control Letters*, 2006. Submitted, available at `http://arxiv.org/math.OC/0606083`.

D. J. Scheeres, E. G. Fahnestock, S. J. Ostro, J.-L. Margot, L. A. M. Benner, S. B. Broschart, J. Bellerose, J. D. Giorgini, M. C. Nolan, C. Magri, P. Pravec, P. Scheirich, R. Rose, R. F. Jurgens, E. M. De Jong, and S. Suzuki. Dynamical configuration of binary near-Earth asteroid (66391) 1999 KW4. *Science*, 314 (5803):1280–1283, 2006.

R. D. Skeel, G. Zhang, and T. Schlick. A family of symplectic integrators: stability, accuracy, and molecular dynamics applications. *SIAM J. Sci. Comput.*, 18(1):203–222, 1997.

G. Y. Sussman and J. Wisdom. Chaotic evolution of the solar system. *Science*, 257(5066):56–62, 1992.

M. Suzuki. Improved Trotter-like formula. *Phys. Lett. A*, 180(3):232–234, 1993.

Y. -F. Tang. Formal energy of a symplectic scheme for Hamiltonian systems and its applications (I). *Computers Math. Applic.*, pages 31–39, 1994.

Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907, USA.

E-mail address: `mleok@math.purdue.edu`