Reduction of Sufficient Conditions for Optimal Control Problems With Subgroup Symmetry

Andy Borum, *Student Member, IEEE*, and Timothy Bretl, *Member, IEEE*

***Abstract*—In this paper, we apply symmetry reduction techniques from geometric mechanics to sufficient conditions for local optimality in optimal control problems. After reinterpreting some previous results for left-invariant problems on Lie groups, we focus on optimal control problems with subgroup symmetry. For these problems, the necessary conditions for optimality can be simplified by exploiting symmetries so as to reduce the number of variables needed to describe trajectories of the system. We show that sufficient conditions for optimality, based on the non-existence of conjugate points, can be simplified in an analogous way to the necessary conditions. We demonstrate these simplifications by analyzing an optimal control problem that models a spinning top in a gravitational field, and we give particular attention to the example of an axisymmetric sleeping top. The results we derive in this paper allow us to determine which trajectories of a sleeping top are locally optimal solutions of the optimal control problem, which is a new result that has not appeared in previous literature.**

***Index Terms*—Conjugate points, Lie groups, optimal control, sufficient conditions, symmetry reduction.**

I. Introduction

CONSIDER an optimal control problem whose state takes values on a Lie group. The Pontryagin maximum principle associates to this optimal control problem a Hamiltonian system that evolves on the cotangent bundle of the Lie group [1]. Geometric mechanics provides tools for studying such Hamiltonian systems, and a main theme in mechanics is simplifying Hamiltonian systems by exploiting symmetries [2]. Trajectories of a Hamiltonian system with symmetries evolve on spaces of lower dimension than the original Hamiltonian system. The equations of motion of the original Hamiltonian system can often be simplified by working in coordinates for these lower dimensional spaces. These same simplifications can be obtained for optimal control problems by exploiting symmetries in the necessary conditions for optimality.

Manuscript received April 25, 2016; revised October 18, 2016; accepted October 28, 2016. Date of publication November 14, 2016; date of current version June 26, 2017. This work was supported by the NSF under Grant No. IIS-1320519 and by the NSF-GRFP under Grant No. DGE-1144245. Recommended by Associate Editor M. Margaliot.

The authors are with the Department of Aerospace Engineering, University of Illinois, Urbana, Illinois 61801, USA (e-mail: borum2@illinois.edu; tbretl@illinois.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2016.2628638

The focus of this paper is on applying symmetry reduction to sufficient conditions for optimality, which has received far less attention than reduction of necessary conditions. The sufficient conditions we use rely on the non-existence of conjugate points, and symmetries allow us to simplify the computations for finding conjugate points. We focus on a class of optimal control problems on Lie groups whose associated Hamiltonian functions are invariant under the left action of a subgroup of the Lie group. To account for the symmetries, we apply Lie-Poisson reduction by stages to the Hamiltonian system associated with the optimal control problem [3]. We show that the simplifications that appear in the necessary conditions for optimality also appear in the sufficient conditions.

To illustrate the application of the reduced sufficient conditions, we analyze an optimal control problem on the Lie group $SO(3)$ whose cost function is the Lagrangian of a spinning top in a gravitational field. We analytically determine which trajectories of an axisymmetric sleeping top are locally optimal solutions of this optimal control problem. Similar results for a top without gravity have appeared in previous literature. However, the inclusion of gravity in this analysis, which is made possible by the reduced sufficient conditions we derive, is a new result that has not appeared in previous literature.

We begin in Section II by covering related work from geometric mechanics and optimal control. In Section III, we state necessary and sufficient conditions for optimal control problems on smooth manifolds. In Section IV, we recall reduced necessary and sufficient conditions for left-invariant optimal control problems, and we study an optimal control problem that models a spinning top. Then, in Section V, we provide a reinterpretation of the results in Section IV that makes clear the connection between the reduced necessary and sufficient conditions for left-invariant problems. In Sections VI and VII, we derive reduced necessary and sufficient conditions, respectively, for optimal control problems with subgroup symmetry. In Section VIII, we apply these results to a spinning top in a gravitational field. Closing remarks are given in Section IX. Readers may benefit from reviewing the work in [4], which provides a more detailed account of the topics we recall in Section IV, and from reading [5], [6], which provide additional details on some of the computations in Sections VI and VII.

II. Related Work

Hamiltonian systems with symmetries have been studied extensively in the field of geometric mechanics [2]. Symmetry reduction allows these systems to be studied in a quotient space of reduced dimension. When a Hamiltonian system is invariant with respect to a group action, and the symmetry group contains a normal subgroup, reduction can first be carried out by the normal subgroup and then by the complement of the normal subgroup. This procedure, called reduction by stages [3], can often be applied when the Hamiltonian system evolves on the cotangent bundle of the semidirect product of a Lie group and a vector space [7]–[9]. This is the case that we will examine in this paper.

We focus on Hamiltonian systems that are left-invariant. Finite dimensional mechanical systems with symmetries, such as the heavy spinning top and underwater vehicle dynamics, often fall into this category [8], [10]. Infinite dimensional systems with symmetries can be right-invariant, such as compressible fluids, magnetohydrodynamics, and three-dimensional elasticity [8]. Some systems are both left and right-invariant and evolve on spaces called centered semidirect products [11].

Symmetry reduction techniques can be applied to the necessary conditions provided by Pontryagin's maximum principle for optimal control problems. Grizzle and Marcus showed that symmetry allows optimal feedback laws to be decomposed into two components, with one component depending upon the symmetry, and the other component depending upon a lower dimensional optimization problem [12]. Symmetries in the maximization condition of the maximum principle were studied by van der Schaft [13], whereas Echeverría-Enríquez *et al.* studied symmetries in optimal control from a presymplectic viewpoint [14]. Principal connections in optimal control problems with symmetries were explored by Ohsawa [15], de León *et al.* applied results for vakonomic systems with symmetries to optimal control [16], and Martínez derived a reduced maximum principle in terms of Lie algebroids [17].

In the case when the state of the optimal control problem takes values on a Lie group, Lie-Poisson reduction can be applied if the associated Hamiltonian function is invariant (left or right-invariant) [18]. This reduction decouples the costate trajectory in Pontryagin's maximum principle from the state of the system. Examples of invariant control problems on Lie groups include motion planning problems for aircraft [19], [20], autonomous underwater vehicles [21], Euler's elastica [22], the Kirchhoff elastic rod [4], conflict resolution in differential games [23], biological models of collective motion [24], and time-optimal control of quantum systems [25].

As an alternative to the maximum principle, a Lagrangian approach can be taken to exploit symmetries in the necessary conditions for optimality [26], [27]. These approaches have been applied to higher order variational problems, with applications to optimal control of underactuated systems [28], [29]. Lagrangian systems on semidirect product spaces have previously been studied [7], [30], and Gupta applied these results to optimal control problems on semidirect products [31]. Optimal control on semidirect products was also studied by Gay-Balmaz and Ratiu using a Clebsch formulation [32].

While symmetry reduction has been applied to necessary conditions for optimality, less attention has been given to the role of symmetries in sufficient conditions. Sufficient conditions in terms of conjugate points can sometimes be computed if the Hamiltonian system associated with the optimal control problem is integrable, e.g., rigid body motion [33], Euler's elastica [22], and some sub-Riemannian geometry problems [34], where the symmetries can simplify these computations. When determining the optimality of geodesics on a Riemannian or sub-Riemannian manifold with a left-invariant metric, comparison theorems can be used to bound conjugate points [35]. For left-invariant optimal control problems on Lie groups, it has been shown that conjugate points can be computed using the reduced system provided by Lie-Poisson reduction [4]. However, the connection between this result and the procedure for deriving reduced necessary conditions for optimality was not explored in [4].

The results in this paper generalize those in a conference paper by the authors [5]. In the conference paper, we considered optimal control problems on matrix Lie groups whose cost functions depended upon a symmetry breaking term (such as gravity) that was decoupled from the control input. The results in this paper can be applied to general Lie groups. Furthermore, we do not impose the decoupled structure on the cost function in this paper. Also, we give particular attention to establishing a clear relationship between simplifications in the necessary and sufficient conditions, which was not explored in the conference paper.

III. Optimal Control on Smooth Manifolds

In this section, we recall a few results from geometric optimal control. First, in Section III-A, we review some notation from differential geometry. Then, in Section III-B, we state a geometric version of Pontryagin's maximum principle [36]. In Section III-C, we give a sufficient optimality condition based on the theory of conjugate points. In later sections, we will specialize these optimality conditions for optimal control problem with certain symmetry properties.

A. Smooth Manifolds

For a smooth manifold M, denote the set of all smooth real-valued functions on M by $C^\infty(M)$ and the set of all smooth vector fields on M by $\mathfrak{X}(M)$. Let $v \cdot df$ and $\langle w, v \rangle$ denote the actions of a tangent vector $v \in T_m M$ on a function $f \in C^\infty(M)$ and a tangent covector $w \in T_m^* M$ on v, respectively. The function $X[f] \in C^\infty(M)$ denotes the action of a vector field $X \in \mathfrak{X}(M)$ on a function $f \in C^\infty(M)$, and satisfies

$$X[f](m) = X(m) \cdot df$$

for all $m \in M$. For $X, Y \in \mathfrak{X}(M)$, the Jacobi-Lie bracket produces the vector field $[X, Y]$ that satisfies

$$[X, Y][f] = X[Y[f]] - Y[X[f]]$$

for all $f \in C^\infty(M)$. The pushforward of a smooth map $F : M \to N$, where N is a smooth manifold, is the linear map $T_m F : T_m M \to T_{F(m)} N$ that satisfies

$$T_m F(v) \cdot df = v \cdot d(f \circ F)$$

for all $v \in T_m M$ and $f \in C^\infty(N)$. The pullback of F at $m \in M$ is the dual map $T_m^* F : T_{F(m)}^* N \to T_m^* M$ that satisfies

$$\langle T_m^* F(w), v \rangle = \langle w, T_m F(v) \rangle$$

for all $v \in T_m M$ and $w \in T_{F(m)}^* N$. If there exists a non-zero $v \in T_m M$ such that $T_m F(v) = 0$, then we say F is degenerate at $m \in M$. The canonical symplectic form on $T^* M$ is

$$\Omega = \sum_{i=1}^{n} dq^i \wedge dp_i,$$

where (q, p) are local coordinates on $T^* M$ and $n = \dim M$. The Poisson bracket generated by the canonical symplectic form on $T^* M$ is denoted by $\{\cdot, \cdot\} : C^\infty(T^* M) \times C^\infty(T^* M) \to C^\infty(T^* M)$ and satisfies

$$\{f, g\} = \Omega(X_f, X_g)$$

for all $f, g \in C^\infty(T^* M)$, where X_f satisfies

$$\Omega(X_f(m), v) = v \cdot df(m)$$

for all $m \in M$ and $v \in T_m M$. We call X_f the Hamiltonian vector field of $f \in C^\infty(T^* M)$. Finally, let $\pi : T^* M \to M$ denote the projection map $\pi(m, w) = m$ for all $w \in T_m^* M$.

B. Necessary Conditions

We now consider an optimal control problem whose state takes values on a smooth manifold M. Let $g : M \times U \to \mathbb{R}$ and $f : M \times U \to TM$ be smooth maps where $U \subset \mathbb{R}^m$ for some $m > 0$. Consider the optimal control problem

$$\begin{aligned} \underset{q,u}{\text{minimize}} \quad & \int_0^{t_f} g(q(t), u(t))\, dt \\ \text{subject to} \quad & \dot{q}(t) = f(q(t), u(t)) \quad \text{for all} \quad t \in [0, t_f] \\ & q(0) = q_0, \qquad q(t_f) = q_f \end{aligned} \tag{1}$$

for some fixed $t_f > 0$, where $q_0, q_f \in M$ are fixed and $(q, u) : [0, t_f] \to M \times U$. Necessary conditions for (q, u) to be a local optimum of (1) are provided by Pontryagin's maximum principle [36]. To apply the maximum principle, we define the parameterized Hamiltonian $\widehat{H} : T^* M \times \mathbb{R} \times U \to \mathbb{R}$ by

$$\widehat{H}(q, p, k, u) = \langle p, f(q, u) \rangle - k g(q, u),$$

where $p \in T_q^* M$. Theorem 1 provides necessary conditions that local optima of (1) must satisfy.

Theorem 1 (Necessary Conditions) Suppose *(q, u):* $[0, t_f] \to M \times U$ is a local optimum of (1). Then, there exists $k \geq 0$ and $p : [0, t_f] \to T_{q(t)}^* M$ such that (q, p) is an integral curve of the time-varying Hamiltonian vector field X_H, where $H : T^* M \times \mathbb{R} \to \mathbb{R}$ is given by $H(q, p, t) = \widehat{H}(q, p, k, u(t))$, and (q, p) satisfies

$$H(q(t), p(t), t) = \max_{u \in U} \widehat{H}(q(t), p(t), k, u) \tag{2}$$

for all $t \in [0, t_f]$. If $k = 0$, then $p(t) \neq 0$ for all $t \in [0, t_f]$.

Proof: See Theorem 12.10 in [1]. ■

The integral curve (q, p) is called an abnormal extremal when $k = 0$ and a normal extremal otherwise. If $k \neq 0$, we may assume $k = 1$. We call (q, u) abnormal if it is the projection of an abnormal extremal. We call (q, u) normal if it is the projection of a normal extremal and it is not abnormal.

C. Sufficient Conditions

The conditions in Theorem 1 are necessary for a trajectory (q, u) to be a local optimum of (1). Second order conditions are needed to ensure (q, u) is indeed a local minimum. Theorem 2 provides sufficient optimality conditions based on the non-existence of conjugate points.

Theorem 2 (Sufficient Conditions) Suppose *(q,p):* $[0, t_f] \to T^* M$ is a normal extremal of (1) and $\partial^2 \widehat{H} / \partial u^2 < 0$ in a neighborhood of the curve (q, p). Assume that the maximized Hamiltonian function

$$H(q, p) = \max_{u \in U} \widehat{H}(q, p, 1, u) \tag{3}$$

is defined and smooth on $T^* M$. Also assume that X_H is complete and that there exists no other integral curve (q', p') of X_H satisfying $q'(t) = q(t)$ for all $t \in [0, t_f]$. Let $\varphi_t : T^* M \to T^* M$ be the flow of X_H and define the endpoint map $\phi_t : T_{q_0}^* M \to M$ by $\phi_t(w) = \pi \circ \varphi_t(q_0, w)$. Define $u : [0, t_f] \to U$ so $u(t)$ is the unique maximizer of (3) at $(q(t), p(t))$. Then (q, u) is a local optimum if there exists no $t \in (0, t_f]$ for which ϕ_t is degenerate at $p(0)$.

Proof: See Theorem 21.8 in [1]. ■

A time at which ϕ_t is degenerate is called a conjugate time, and the endpoint map ϕ_t is degenerate when its Jacobian matrix is singular. To compute the integral curves (q, p) in Theorem 1 or establish non-degeneracy of the endpoint map ϕ_t in Theorem 2, we could introduce local coordinates on $T^* M$. Integral curves could then be found by solving Hamilton's canonical equations

$$\dot{q}^i = H_{p_i} \qquad \dot{p}_i = -H_{q^i}, \tag{4}$$

where (q^i, p_i) are local coordinates on $T^* M$ with $i = 1, \ldots, n = \dim M$, and subscripts denote partial derivatives.

In order to establish local optimality of an integral curve, let $\mathbf{J}(t)$ denote the Jacobian matrix of the endpoint map ϕ_t in this coordinate system, i.e., $\mathbf{J}(t)$ is the Jacobian matrix of the state $q(t)$ with respect to the initial costate $p(0)$, and let $\mathbf{M}(t)$ denote the Jacobian matrix of the costate $p(t)$ with respect to the initial value of the costate $p(0)$. These matrices can be found by solving the time-varying matrix differential equations

$$\dot{\mathbf{J}} = (H_{qp})\mathbf{J} + (H_{pp})\mathbf{M} \qquad \dot{\mathbf{M}} = -(H_{qq})\mathbf{J} - (H_{pq})\mathbf{M} \tag{5}$$

with the initial conditions $\mathbf{M}(0) = I$ and $\mathbf{J}(0) = 0$. The endpoint map ϕ_t is degenerate if $\det(\mathbf{J}(t)) = 0$.

Note that in (4), the evolution of the costate p could depend upon both the state q and the costate p, depending upon the structure of the Hamiltonian. Similarly, the evolution of q could depend upon both q and p. Analogously, note that in (5), the evolution of the matrix $\mathbf{M}$ could depend upon both $\mathbf{M}$ and $\mathbf{J}$. The evolution of the matrix $\mathbf{J}$ could also depend upon both $\mathbf{M}$ and $\mathbf{J}$. Also observe that the coefficient matrices in (5) could be a function of both q and p. In the following sections, we will show that symmetries allow us to decouple some of these differential equations and that this decoupling occurs in analogous ways in the necessary and sufficient conditions.

IV. Left-Invariant Optimal Control Problems

Theorems 1 and 2 provide coordinate-free conditions that local solutions of (1) must satisfy. As described in the previous section, these conditions can be evaluated by introducing local coordinates on T^*M. However, if the Hamiltonian function (3) possesses symmetries, we can use these symmetries to simplify these computations by reducing the number of variables needed to describe trajectories of the system.

In this section, we review some results for the case when the Hamiltonian function (3) is left-invariant. We begin by recalling some facts about Lie Groups in Section IV-A. In Sections IV-B and IV-C, we give reduced statements of the necessary and sufficient conditions for optimality in Theorems 1 and 2, respectively. Then, in Section IV-D, we consider a left-invariant optimal control problem on the Lie group $SO(3)$

A. Lie Groups

Let G be an n-dimensional Lie group with identity element $e \in G$. Let $\mathfrak{g} = T_eG$ be the Lie algebra associated with G and $\mathfrak{g}^* = T_e^*G$ its dual. For any $q \in G$, define the left translation map $L_q : G \to G$ by

$$L_q(r) = qr$$

for all $r \in G$. A function $H \in C^\infty(T^*G)$ is left-invariant if

$$H(r, T_r^* L_q(w)) = H(s, w) \tag{6}$$

for all $w \in T_s^*G$ and $q, r, s \in G$ satisfying $s = L_q(r)$. For any $\zeta \in \mathfrak{g}$, let X_ζ be the vector field that satisfies

$$X_\zeta(q) = T_e L_q(\zeta)$$

for all $q \in G$. Define the Lie bracket $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ by

$$[\zeta, \eta] = [X_\zeta, X_\eta](e)$$

for all $\zeta, \eta \in \mathfrak{g}$. For any $\zeta \in \mathfrak{g}$, the adjoint operator $\mathrm{ad}_\zeta : \mathfrak{g} \to \mathfrak{g}$ is defined by the Lie bracket

$$\mathrm{ad}_\zeta(\eta) = [\zeta, \eta],$$

and the coadjoint operator $\mathrm{ad}_\zeta^* : \mathfrak{g}^* \to \mathfrak{g}^*$ is given by its dual map and determined by

$$\langle \mathrm{ad}_\zeta^*(\mu), \eta \rangle = \langle \mu, \mathrm{ad}_\zeta(\eta) \rangle$$

for all $\eta \in \mathfrak{g}$ and $\mu \in \mathfrak{g}^*$. The functional derivative of $h \in C^\infty(\mathfrak{g}^*)$ at $\mu \in \mathfrak{g}^*$ is the element $\delta h/\delta\mu \in \mathfrak{g}$ that satisfies

$$\lim_{s \to 0} \frac{h(\mu + s\delta\mu) - h(\mu)}{s} = \left\langle \delta\mu, \frac{\delta h}{\delta \mu} \right\rangle$$

for all $\delta\mu \in \mathfrak{g}^*$. Let $\{X_1, \ldots, X_n\}$ be a basis for $\mathfrak{g}$ and let $\{X^1, \ldots, X^n\}$ be the dual basis for $\mathfrak{g}^*$ that satisfies $\langle X^i, X_j \rangle = \delta^i{}_j$ for $i, j \in \{1, \ldots, n\}$, where $\delta^i{}_j$ is the Kronecker delta. We write ζ^i to denote the i^{th} component of $\zeta \in \mathfrak{g}$ with respect to this basis. For $i, j \in \{1, \ldots, n\}$, define the structure constants $C_{ij}^k \in \mathbb{R}$ for our choice of basis by

$$[X_i, X_j] = \sum_{k=1}^n C_{ij}^k X_k. \tag{7}$$

B. Left-Invariant Necessary Conditions

We now revisit the statement of necessary conditions for the optimal control problem (1) in the case where the smooth manifold M is a Lie group G and the Hamiltonian function H is left-invariant under the cotangent lift of left translations. Theorem 1 implies the existence of an integral curve (q,p) in the cotangent bundle T^*G. The following theorem implies the existence of a corresponding integral curve μ in $\mathfrak{g}^*$.

Theorem 3 (Reduction of Necessary Conditions) Suppose *(q, u):* $[0, t_f] \to M \times U$ is a local optimum of (1). Assume the time-varying Hamiltonian function $H : T^*M \times \mathbb{R} \to \mathbb{R}$ defined in Theorem 1 is both smooth and left-invariant for all $t \in [0, t_f]$, and denote the restriction of H to $\mathfrak{g}^*$ by $h = H|_{\mathfrak{g}^* \times [0, t_f]}$. Then, the integral curve $(q, p) : [0, t_f] \to T^*M$ described in Theorem 1 satisfies

$$p(t) = T_{q(t)}^* L_{q(t)^{-1}}(\mu(t)) \qquad \dot{q} = X_{\delta h/\delta\mu}(q) \tag{8}$$

for all $t \in [0, t_f]$, where $\mu : [0, t_f] \to \mathfrak{g}^*$ is the solution of the Lie-Poisson equations

$$\dot{\mu} = \mathrm{ad}_{\delta h/\delta \mu}^*(\mu) \tag{9}$$

with initial condition $\mu(0) = T_e^* L_{q_0}(p(0))$.

Proof: See Theorem 13.4.4 in [2]. ■

Since $\mathfrak{g}^*$ is a vector space, the trajectory μ described by (9) can be evaluated by solving a system of ordinary differential equations. Taking $\mu_1(t), \ldots, \mu_n(t)$ as coordinates of $\mu(t)$, (9) is equivalent to (see [18])

$$\dot{\mu}_i = -\sum_{j=1}^n \sum_{k=1}^n C_{ij}^k \frac{\delta h}{\delta \mu_j} \mu_k. \tag{10}$$

C. Left-Invariant Sufficient Conditions

We now revisit the sufficient conditions in Theorem 2 for left-invariant optimal control problems. As shown in [4], nondegeneracy of the endpoint map ϕ_t can be established by working with the variables μ_i for $i = 1, \ldots, n$ from the reduced necessary conditions in Theorem 3.

Theorem 4 (Reduction of Sufficient Conditions) Suppose *(q,p):* $[0, t_f] \to T^*M$ is a normal extremal of (1), and assume the conditions in Theorem 2 hold. Also assume the Hamiltonian function $H : T^*M \to \mathbb{R}$ defined in Theorem 2 is left-invariant, and let $h = H|_{\mathfrak{g}^*}$ be the restriction of H to $\mathfrak{g}^*$. Let μ be the solution of (9) with initial condition $\mu(0) = T_e^* L_{q_0}(p(0))$, and define the matrices $\mathbf{F}, \mathbf{G}, \mathbf{H} \in \mathbb{R}^{n \times n}$ by

$$\mathbf{F}^i{}_j = -\frac{\partial}{\partial \mu_j} \sum_{r=1}^n \sum_{s=1}^n C_{ir}^s \frac{\delta h}{\delta \mu_r} \mu_s$$

$$\mathbf{G}^i{}_j = \frac{\partial}{\partial \mu_j} \frac{\delta h}{\delta \mu_i} \qquad \mathbf{H}^i{}_j = -\sum_{r=1}^n \frac{\delta h}{\delta \mu_r} C_{rj}^i.$$

Solve the (linear, time-varying) matrix differential equations

$$\dot{\mathbf{M}} = \mathbf{F}\mathbf{M} \qquad \dot{\mathbf{J}} = \mathbf{G}\mathbf{M} + \mathbf{H}\mathbf{J} \tag{11}$$

with initial conditions $\mathbf{M}(0) = I$ and $\mathbf{J}(0) = 0$. Define $u : [0, t_f] \to U$ as in Theorem 2. Then (q, u) is a local optimum if there exists no $t \in (0, t_f]$ for which $\det(\mathbf{J}(t)) = 0$.

Proof: See Theorem 4 in [4]. ■

Compared with the necessary and sufficient conditions in Theorems 1 and 2, we see that the conditions in Theorems 3 and 4 have some advantages. Whereas the differential equations for the state q and the costate p were possibly coupled in Theorem 1, as shown in (4), the Lie-Poisson equations (9) in Theorem 3 governing the reduced costate μ are decoupled from the state q. Similarly, the differential equation (11) in Theorem 4 for the matrix $\mathbf{M}$ is decoupled from $\mathbf{J}$, whereas they were coupled in Theorem 2, as shown in (5). In Section V, we will further explore the decouplings that occur through reduction of the necessary and sufficient conditions.

D. The Torque-Free Spinning Top

To demonstrate the application of the conditions in Theorems 3 and 4, consider a spinning top that does not experience external torques. The motion of the top corresponds to a trajectory on the matrix Lie group $SO(3)$ that extremizes (but does not necessarily minimize) the top's action functional. From a mechanics viewpoint, we are often concerned with finding equations of motion and solving them as an initial value problem. We are not typically concerned with finding trajectories that minimize a system's action functional subject to given boundary conditions. However, to show how Theorems 3 and 4 can be applied, we will search for trajectories of the spinning top that satisfy given boundary conditions and minimize the top's action functional. We also note that problems similar to the one considered in this section have previously been studied in the context of optimal attitude control of spacecraft and satellites [37].

The optimal control problem that corresponds to the spinning top is given by

$$\begin{aligned} &\underset{q,u}{\text{minimize}} \int_0^{t_f} \left(\frac{1}{2} \sum_{i=1}^{3} c_i {u^i}^2 \right) dt \\ &\text{subject to } \dot{q} = q \left(\sum_{i=1}^{3} u^i X_i \right) \\ &q(0) = q_0, \qquad q(t_f) = q_f \end{aligned} \tag{12}$$

for some fixed $q_0, q_f \in SO(3)$ and $t_f > 0$, where $(q, u)\colon [0, t_f] \to SO(3) \times \mathbb{R}^3$. The matrices X_i are defined by $X_i = \widehat{e}_i$, where e_i are the standard basis elements of $\mathbb{R}^3$ and $\widehat{\ }\colon \mathbb{R}^3 \to \mathfrak{so}(3)$ is the map satisfying $\widehat{a}b = a \times b$ for all $a, b \in \mathbb{R}^3$. Since $SO(3)$ is a matrix Lie group, we have used $q\zeta$ to denote the left action of $q \in G = SO(3)$ on $\zeta \in \mathfrak{g} = \mathfrak{so}(3)$ in the dynamic constraints in (12), where the Lie algebra $\mathfrak{so}(3)$ is the set of all 3×3 skew-symmetric matrices. The positive constants c_1, c_2, and c_3 play the role of weights in the cost function and correspond to the moments of inertia of the top. The control input u is the angular velocity of the top, and the integrand in the cost function in (12) is the kinetic energy of the top. For each $q \in SO(3)$, the right hand side of the dynamic constraint in (12) spans the tangent space $T_q SO(3)$, and the system is therefore controllable [1].

Applying Theorem 1 gives that local extrema of (12) correspond to integral curves of the Hamiltonian vector field X_H, where $H : T^*SO(3) \to \mathbb{R}$ is defined by

$$\widehat{H}(q, p, k, u) = \left\langle p, q \left(\sum_{i=1}^{3} u^i X_i \right) \right\rangle - \frac{k}{2} \left(\sum_{i=1}^{3} c_i {u^i}^2 \right)$$

and

$$H(q, p) = \max_u \widehat{H}(q, p, k, u).$$

In the abnormal case (that is, $k = 0$), $\widehat{H}$ is extremized in u when $p = 0$. Therefore, by Theorem 1, there are no abnormal extremals. In the normal case, if we take $k = 1$, then the maximum is achieved when

$$u^i = c_i^{-1} \langle p, qX_i \rangle \tag{13}$$

for $i \in \{1, 2, 3\}$. This is indeed a maximum since

$$\partial^2 \widehat{H} / \partial u^2 = -\mathrm{diag}(c_1, c_2, c_3) < 0.$$

The maximized Hamiltonian function is then

$$H(q, p) = \frac{1}{2} \sum_{i=1}^{3} c_i^{-1} \langle p, qX_i \rangle^2.$$

Note that for any $p \in T_q^* SO(3)$ and $q, g, r \in SO(3)$ satisfying $q = gr$, we have

$$\begin{aligned} H(r, T_r^* L_g(p)) &= \frac{1}{2} \sum_{i=1}^{3} c_i^{-1} \langle T_r^* L_g(p), g^{-1} qX_i \rangle^2 \\ &= \frac{1}{2} \sum_{i=1}^{3} c_i^{-1} \langle p, g \left(g^{-1} qX_i \right) \rangle^2 \\ &= \frac{1}{2} \sum_{i=1}^{3} c_i^{-1} \langle p, qX_i \rangle^2 \\ &= H(q, p). \end{aligned} \tag{14}$$

Therefore, H is left-invariant and we can apply Theorem 3. The reduced Hamiltonian on $\mathfrak{so}^*(3)$ is given by

$$h(\mu) = H(e, \mu) = \frac{1}{2} \sum_{i=1}^{3} c_i^{-1} \mu_i^2.$$

The Lie-Poisson equations (9) for the reduced Hamiltonian h are given by

$$\dot{\mu} = \mu \times u, \tag{15}$$

where $u^i = c_i^{-1} \mu_i$. In this case, the coadjoint operator in (9) is the cross product after an identification of $\mathfrak{so}(3)$ with $\mathbb{R}^3$ using the map $\widehat{\ }\colon \mathbb{R}^3 \to \mathfrak{so}(3)$.

Candidate solutions of (12) are obtained by finding an initial value of $\mu(0) \in \mathfrak{so}^*(3) \cong \mathbb{R}^3$ that places $q(t_f)$ at q_f. Such solutions are only guaranteed to be extrema of (12). It is clear that that $\mu \in \mathfrak{g}^*$ (and hence $p \in T^*SO(3)$) is uniquely determined by (q, u), and in this case, X_H is complete. Therefore, we may apply Theorem 4 to determine which extrema are actually local minima.

Computing the matrices $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{H}$ in Theorem 4 (and defining $c_{ij} = (c_i^{-1} - c_j^{-1})$) gives

$$\mathbf{F} = \begin{bmatrix} 0 & c_{32}\mu_3 & c_{32}\mu_2 \\ c_{13}\mu_3 & 0 & c_{13}\mu_1 \\ c_{21}\mu_2 & c_{21}\mu_1 & 0 \end{bmatrix} \tag{16}$$

$$\mathbf{G} = \operatorname{diag}(c_1^{-1}, c_2^{-1}, c_3^{-1}) \qquad \mathbf{H} = -\widehat{u}.$$

After finding $\mu(0) \in \mathfrak{so}^*(3)$ that places $q(t_f)$ at q_f, (11) can be solved with the initial conditions $\mathbf{M}(0) = I$ and $\mathbf{J}(0) = 0$. If $\det(\mathbf{J}(t)) \neq 0$ for all $t \in (0, t_f]$, then the solution corresponding to this choice of $\mu(0) \in \mathfrak{so}^*(3)$ is a local minimum of (12).

We now consider a top that is axisymmetric with $c_2{=}c_3{=}1$. With these parameters, setting $\mu_2 = \mu_3 = 0$ and letting μ_1 be arbitrary gives a fixed point of the system (15). This fixed point corresponds to the top rotating about its axis of symmetry. Solving the linear system (11) at this fixed point, which now becomes time-invariant, and computing the determinant of $\mathbf{J}(t)$ gives

$$\det(\mathbf{J}(t)) = \frac{4t}{c_1 \mu_1^2} \sin^2\left(\frac{\mu_1}{2} t\right).$$

We see that if $|\mu_1 t_f| < 2\pi$, then this trajectory of the top is locally optimal, since $\det(\mathbf{J}(t)) > 0$ for all $t \in (0, t_f]$. If $|\mu_1 t_f| > 2\pi$, this trajectory of the top is not locally optimal.

These results for the axisymmetric top are consistent with previous studies of conjugate points in rigid body motion, such as [33], in which the conjugate locus for an axisymmetric body was computed, and [38], in which conjugate points for a sleeping but non-axisymmetric body were computed. In Section VIII, we derive similar results for a top in a gravitational field, which have not appeared in previous literature.

V. Reinterpretation of the Optimality Conditions for Left-Invariant Problems

The key insight in Theorem 3 is that by taking an integral curve (q, p) of X_H and left-translating the costate to the identity to obtain $\mu(t) = T_e^* L_{q(t)}(p(t))$, we find that μ satisfies the ordinary differential equation (9), which is decoupled from the state q. This same procedure was used to prove Theorem 4 in [4], although this connection between the reduced necessary and sufficient conditions was not made explicit. In this section, we explicitly connect the results in Theorems 3 and 4.

Recall from Theorem 2 that we need to determine if the map $\phi_t : T_{q_0}^* G \to G$ is degenerate at $p(0)$ for some $t \in (0, t_f]$. In other words, for each $t \in (0, t_f]$, we need to determine if the image of the pushforward $T_{p(0)}\phi_t$ spans the tangent space $T_{q(t)}G$. Following the approach in Theorem 3 to decouple the state and the costate, we pre- and post-compose $T_{p(0)}\phi_t$ with left-translation from and to the identity, respectively. Then we evaluate this map at each $X^j \in \mathfrak{g}^*$, which produces the Lie algebra element

$$\eta_j(t) = T_{q(t)} L_{q(t)^{-1}} \left(T_{p(0)}\phi_t \left(T_{q_0}^* L_{q_0^{-1}}(X^j)\right)\right). \tag{17}$$

After defining the matrix $\mathbf{J}^i{}_j(t) = \eta_j^i(t)$, we can check for degeneracy of the endpoint map ϕ_t by checking the determinant of $\mathbf{J}(t)$. As shown in Theorem 4, the matrix $\mathbf{J}(t)$ can be computed by solving a matrix differential equation that only depends upon the reduced costate μ. Just as the costate p can be reconstructed from the reduced costate μ using (8) in Theorem 3, the Jacobian of the endpoint map ϕ_t can be reconstructed by left-translating each $\eta_j(t)$ to $q(t)$ for $j = 1, \ldots, n$, i.e., by $T_e L_{q(t)}(\eta_j(t))$. The reconstruction of $\eta_j(t)$ provides a variation along the curve $q(t)$ in G, and such variations have been used to establish first order necessary conditions in variational problems with symmetries [39].

At the end of Section III-C, we saw that when working in local coordinates, the differential equations (4) were coupled in a similar way to the matrix differential equations (5). Now note the similarities between the differential equations (8) and (9) in Theorem 3 and the matrix differential equations (11) in Theorem 4. The evolution of the covector μ is decoupled from the state q in (9), whereas the evolution of q, given by (8), depends on both q and μ. Analogously, the matrix $\mathbf{M}$ is decoupled from $\mathbf{J}$ in (11), whereas the evolution of $\mathbf{J}$ depends upon both $\mathbf{J}$ and $\mathbf{M}$. Furthermore, the coefficient matrices in (11) depend only upon the reduced costate μ. These simplifications were derived by applying the reduction procedure in Theorem 3 to the sufficient conditions in Theorem 2. We therefore call the conditions in Theorem 4 reduced sufficient conditions.

The procedure for exploiting symmetries in necessary and sufficient conditions for optimality is outlined in Fig. 1, along with the corresponding theorems in this paper. In this figure, solid arrows denote standard results for optimal control problems (such as those in Section III) that can be applied to problems without symmetry. The dotted arrows denote previous work on symmetry in necessary conditions for optimality, some of which is described in Section II. Our focus is on deriving reduced sufficient conditions and equating the reduced and unreduced sufficient conditions, i.e., the dashed arrows, which has received less consideration in previous literature than the necessary conditions. In the remainder of this paper, we extend the results in this section by exploring the analogous simplifications in reduced necessary and sufficient conditions for problems that are not left-invariant, but are invariant under a subgroup of the Lie group G.

VI. Necessary Conditions for Problems With Subgroup Symmetry

In Section IV, we assumed that the Hamiltonian function provided by the maximum principle was left-invariant under an action of the Lie Group G. In this section, we consider the case when the Hamiltonian is left-invariant with respect to a subgroup of G. As was done in Theorem 3, we will derive a

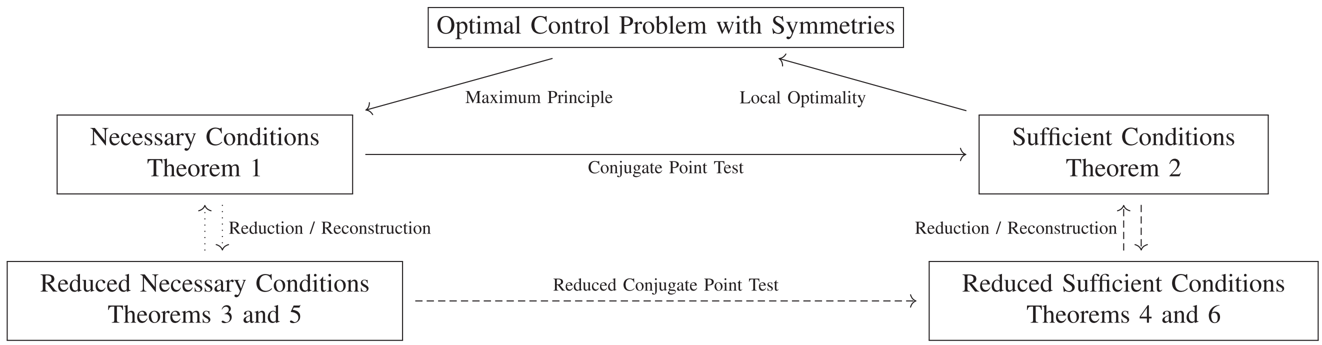

Fig. 1. The procedure for applying symmetry reduction to optimal control problems. Solid arrows represent standard results from optimal control theory, such as those in Section III. Dotted arrows represent previous work, covered in Section II, on reduction of necessary conditions. Dashed arrows represent the focus of this paper, which is reduction of sufficient conditions.

reduced Hamiltonian system whose integral curves correspond to integral curves of the Hamiltonian vector field X_H on T^*G.

In Section VI-A, we motivate the need to consider optimal control problems with subgroup symmetry by examining a generalization of the optimal control problem in Section IV-D. Then, in Section VI-B, we review semidirect products and Lie group representations. In Section VI-C, we give reduced necessary conditions for optimality when the Hamiltonian function is left-invariant with respect to a subgroup of G.

A. The Heavy Spinning Top

Consider again the optimal control problem (12) (i.e., the same dynamic constraints and boundary conditions as (12)), but now with the cost function

$$g(q,u) = \frac{1}{2}\sum_{i=1}^{3} c_i u^{i2} + \chi_0(q\nu), \tag{18}$$

where $\nu \in \mathbb{R}^3$ is a constant vector and $\chi_0 : \mathbb{R}^3 \to \mathbb{R}$ is a linear map. Since the map χ_0 is linear, it can be represented by a constant three-dimensional row vector. This cost function is the Lagrangian of a spinning top in a gravitational field, where ν is a vector pointing from the fixed point of the top to the top's center of mass, and χ_0 points in the direction of gravity and has magnitude equal to the weight of the top. The control input u and the constants c_i still have the same interpretation as in Section IV-D, i.e., u is the angular velocity of the top, the constants c_i are the moments of inertia of the top, and the first term in the cost function (18) is the kinetic energy of the top. The second term in the cost function (18), which did not appear in the problem (12), is the negative of the top's gravitational potential energy.

The Hamiltonian function depends upon the two parameters ν (which is a vector) and χ_0 (which is a linear map). As we will see in Section VI-C, the parameter χ_0 will become important when we apply symmetry reduction to this system. Therefore, to denote the dependence of the Hamiltonian function on the parameter χ_0, we will denote the Hamiltonian by H_{χ_0}. Applying Theorem 1 gives that local extrema of (12) with the cost function (18) correspond to integral curves of the Hamiltonian vector field $X_{H_{\chi_0}}$, where

$$\widehat{H}_{\chi_0}(q,p,k,u) = \left\langle p, q\left(\sum_{i=1}^{3} u^i X_i\right)\right\rangle - kg(q,u)$$

and

$$H_{\chi_0}(q,p) = \max_u \widehat{H}_{\chi_0}(q,p,k,u).$$

The abnormal case for this problem is identical to the abnormal case in Section IV-D, so there are no abnormal extremals. In the normal case, when $k=1$, the maximum is again given by (13). The maximized Hamiltonian function is then

$$H_{\chi_0}(q,p) = \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\langle p, qX_i\rangle^2 - \chi_0(q\nu). \tag{19}$$

Using the computations in (14), note that for any $p \in T_q^* SO(3)$ and $q,g,r \in SO(3)$ satisfying $q = gr$, we have

$$H_{\chi_0}(r, T_r^* L_g(p)) = \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\langle p, qX_i\rangle^2 - \chi_0(g^{-1}q\nu).$$

We see that H is left-invariant under the elements of $SO(3)$ that satisfy $\chi_0 g^{-1} = \chi_0$. These elements form a subgroup of G, called the isotropy group of χ_0. In the remainder of this section, we will give necessary conditions for problems with this subgroup symmetry property.

B. Semidirect Products

Let V be an l-dimensional vector space and let $\rho : G \to GL(V)$ be a left representation of G on V, i.e., ρ is a smooth group homomorphism that assigns to each $g \in G$ a linear map $\rho(g) : V \to V$ satisfying

$$\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$$

for all $g_1, g_2 \in G$. The associated left and right representations of G on V^*, denoted ρ_* and ρ^*, respectively, are

$$\rho_*(g) = \left[\rho(g^{-1})\right]^* \qquad \rho^*(g) = \left[\rho(g)\right]^*,$$

where $[\,]^*$ denotes the dual transformation. The induced Lie algebra representation $\rho' : \mathfrak{g} \to \mathrm{End}[V]$ of $\zeta \in \mathfrak{g}$ satisfies

$$\rho'(\zeta)(v) = \frac{d}{dt}\left[\rho(\exp(t\zeta))(v)\right]|_{t=0}$$

for all $v \in V$, where $\exp : \mathfrak{g} \to G$ is the exponential map. Denote by G_χ the isotropy group of $\chi \in V^*$, i.e.,

$$G_\chi = \{g \in G | \rho^*(g)\chi = \chi\}. \tag{20}$$

Let $S = G \times V$ be the semidirect product of G and V with multiplication and inversion given by

$$(g_1, v_1)(g_2, v_2) = (g_1 g_2, v_1 + \rho(g_1)v_2)$$

$$(g_1, v_1)^{-1} = (g_1^{-1}, -\rho(g_1^{-1})v_1)$$

for all $g_1, g_2 \in G$ and $v_1, v_2 \in V$. The Lie algebra of S is $\mathfrak{s} = \mathfrak{g} \times V$ with the Lie bracket

$$[(\zeta_1, v_1), (\zeta_2, v_2)] = ([\zeta_1, \zeta_2], \rho'(\zeta_1)v_2 - \rho'(\zeta_2)v_1)$$

for all $\zeta_1, \zeta_2 \in \mathfrak{g}$ and $v_1, v_2 \in V$. The left action of S on T^*S is given by

$$T^*_{(r,z)} L_{(q,u)}(s, w, v, \chi) = (r, T^*_r L_q(w), z, \rho^*(q)\chi) \tag{21}$$

for all $u, v, z \in V$, $\chi \in V^*$, $w \in T^*_s G$, and $q, s, r \in G$ satisfying $s = L_q(r)$ and $z = u + \rho(q^{-1})v$ [8].

C. Reduction of Necessary Conditions

We now consider the statement of necessary conditions in Theorem 1 in the case when the Hamiltonian function is left-invariant under the action of a subgroup of G. In many situations, the Hamiltonian function depends upon a parameter in the dual of some vector space, and the subgroup under which the Hamiltonian is left-invariant is the isotropy group of this parameter. For such systems, Theorem 5 provides necessary conditions similar to those in Theorem 3.

Before stating Theorem 5, we provide a motivation for the results contained in the theorem. The key idea of the theorem is to embed a Hamiltonian system with subgroup symmetry within an extended Hamiltonian system that is left-invariant. This embedding procedure has been applied to many problems in geometric mechanics with subgroup symmetry [7]–[9], and Theorem 5 applies this idea to the necessary conditions given in Theorem 1. To see how this is done, suppose a Hamiltonian function depends smoothly on a parameter $\chi_0 \in V^*$ and is left-invariant under the action of G_{χ_0} on T^*G, so that (6) holds when $q \in G_{\chi_0}$ (recall from (20) that G_{χ_0} is the isotropy group of χ_0). We denote the Hamiltonian by $H_{\chi_0} : T^*G \to \mathbb{R}$ to note the dependence on $\chi_0 \in V^*$.

The procedure for applying reduction to such Hamiltonian systems is to consider the augmented Hamiltonian function $H : T^*S \to \mathbb{R}$ defined by $H(q, p, v, \chi) = H_\chi(q, p)$, where $T^*S = T^*G \times V \times V^*$. Since $H_\chi(q, p)$ is independent of the variable $v \in V$, we ignore the V component of the left action of S on T^*S and define H to be constant in the variable $v \in V$ [7]. We then show that $H : T^*S \to \mathbb{R}$ is left-invariant under the action of S, i.e., using (21), we show that

$$H(r, T^*_r L_q(w), v, \rho^*(q)\chi) = H(s, w, v, \chi) \tag{22}$$

for all $v \in V$, $\chi \in V^*$, $w \in T^*_s G$, and $q, r, s \in G$ satisfying $s = L_q(r)$. An example of this computation for a specific Hamiltonian function is shown in (35) of Section VIII.

The original Hamiltonian system on T^*G, with Hamiltonian H_{χ_0}, is now embedded within an extended Hamiltonian system on T^*S, with Hamiltonian H. Since the Hamiltonian function H is left-invariant, we can apply reduction to the Hamiltonian system on T^*S. Note that if (22) holds and $q \in G_{\chi_0}$, then $\chi_0 = \rho^*(q)\chi_0$ by (20) and

$$\begin{aligned} H_{\chi_0}(r, T^*_r L_q(w)) &= H(r, T^*_r L_q(w), v, \chi_0) \\ &= H(r, T^*_r L_q(w), v, \rho^*(q)\chi_0) \\ &= H(s, w, v, \chi_0) \\ &= H_{\chi_0}(s, w) \end{aligned}$$

for all $w \in T^*_s G$, $r, s \in G$, and $q \in G_{\chi_0}$ satisfying $s = L_q(r)$. Therefore (22) implies that H_{χ_0} is left-invariant under the action of G_{χ_0} on T^*G.

If (22) holds, then the family of Hamiltonians $\{H_\chi | \chi \in V^*\}$ induces a reduced Hamiltonian h on $\mathfrak{s}^*$. As shown in the following theorem, the existence of an integral curve (μ, χ) in $\mathfrak{s}^*$ implies the existence of a corresponding integral curve (q, p) of $X_{H_{\chi_0}}$ in the cotangent bundle T^*G.

Theorem 5 (Semidirect Product Reduction of Necessary Conditions) Suppose *(q, u):* $[0, t_f] \to M \times U$ is a local optimum of (1). Assume the time-varying Hamiltonian function defined in Theorem 1, which we now denote by $H_{\chi_0} : T^*G \times [0, t_f] \to \mathbb{R}$, is smooth and depends smoothly on the parameter $\chi_0 \in V^*$. In addition, let $S = G \times V$ be the semidirect product between G and V, and suppose that the Hamiltonian function $H : T^*S \times [0, t_f] \to \mathbb{R}$, defined by $H(q, p, v, \chi, t) = H_\chi(q, p, t)$, is left-invariant under the action of S for all $t \in [0, t_f]$. Denote the restriction of H to $\mathfrak{s}^*$ by $h = H|_{\mathfrak{s}^* \times [0, t_f]}$. Then, the integral curve $(q, p) : [0, t_f] \to T^*M$ described in Theorem 1 satisfies

$$p(t) = T^*_{q(t)} L_{q(t)^{-1}}(\mu(t)) \qquad \dot{q} = X_{\delta h/\delta\mu}(q) \tag{23}$$

for all $t \in [0, t_f]$, where $(\mu, \chi) : [0, t_f] \to \mathfrak{s}^*$ is the solution of

$$\dot{\mu} = \mathrm{ad}^*_{\delta h/\delta\mu}(\mu) - \left(\rho'_{\delta h/\delta\chi}\right)^* \chi \tag{24}$$

$$\dot{\chi} = \rho'\left(\delta h/\delta\mu\right)^* \chi \tag{25}$$

with initial conditions $\mu(0) = T^*_e L_{q_0}(p(0))$ and $\chi(0) = \rho^*(q_0)\chi_0$, where $\rho'_{\delta h/\delta\chi} : \mathfrak{g} \to V$ satisfies

$$\rho'_{\delta h/\delta\chi}(\zeta) = \rho'(\zeta)\frac{\delta h}{\delta\chi}$$

for all $\zeta \in \mathfrak{g}$.

Proof: See Theorem 3.4 in [8]. ■

As was the case in Theorem 3, writing (24) and (25) in coordinates allows us to find μ and χ by solving a system of ordinary differential equations. From (10), we know the structure of the coadjoint term in (24). Next, since $\chi \in V^*$ and $\dim(V) = l$, we may represent χ as an l-dimensional row vector. We then have

$$\left(\rho'_{\delta h/\delta\chi}\right)^* (\chi)(\cdot) = \chi\left(\rho'(\cdot)\frac{\delta h}{\delta\chi}\right) \in \mathfrak{g}^*.$$

Therefore, we have

$$\dot{\mu}_i = -\sum_{j=1}^{n}\sum_{k=1}^{n} C_{ij}^k \frac{\delta h}{\delta\mu_j}\mu_k - \chi\left(\rho'\left(X_i\right)\frac{\delta h}{\delta\chi}\right).$$

Expanding the second term in the above expression gives

$$\dot{\mu}_i = -\sum_{j=1}^{n}\sum_{k=1}^{n} C_{ij}^k \frac{\delta h}{\delta\mu_j}\mu_k - \sum_{j=1}^{l}\sum_{k=1}^{l}\chi_j\left[\rho'(X_i)\right]^j{}_k\frac{\delta h}{\delta\chi_k}. \quad (26)$$

We can also write (25) in coordinates as

$$\dot{\chi}_i = \sum_{j=1}^{l}\chi_j\left[\rho'\left(\delta h/\delta\mu\right)\right]^j{}_i. \quad (27)$$

Also note that from (23) we have

$$\frac{d}{dt}\rho(q) = \rho(q)\rho'(\delta h/\delta\mu),$$

and therefore

$$\begin{aligned}\frac{d}{dt}\left(\rho(q)^*\chi_0\right) &= \frac{d}{dt}\left(\chi_0(\rho(q))\right)\\ &= \chi_0\left(\frac{d}{dt}\rho(q)\right)\\ &= \chi_0\left(\rho(q)\rho'(\delta h/\delta\mu)\right)\\ &= \rho'(\delta h/\delta\mu)^*\chi_0\left(\rho(q)\right)\\ &= \rho'(\delta h/\delta\mu)^*\left(\rho(q)^*\chi_0\right).\end{aligned}$$

This shows that

$$\chi(t) = \rho(q(t))^*\chi_0 \quad (28)$$

solves (25) with the correct initial condition.

We now have two ways of finding integral curves of the Hamiltonian vector field $X_{H_{\chi_0}}$. We could solve for the reduced variables (μ, χ) using the differential equations (24) and (25), and then reconstruct the trajectory (q, p) using (23). This is analogous to the result in Theorem 3, where we first solved for the reduced variable μ and then reconstructed the trajectory (q, p). Now, due to the subgroup symmetry, we have to keep track of the extra reduced variable χ. Alternatively, we could substitute the expression in (28) for χ into the differential equation (24). This would explicitly show how the subgroup symmetry of the problem couples the reduced costate μ with the state q. In the next section, we will show that the sufficient conditions can be computed in two alternative ways that are analogous to the necessary conditions.

Before moving on, we make one note about the notation used in this section. In the differential equation (24) in Theorem 5, we have a term of the form $\left(\rho'_v\right)^*\chi$ with $v \in V$ and $\chi \in V^*$. In previous work, the diamond operator was used to denote this function [7], i.e., $\left(\rho'_v\right)^*\chi = v \diamond \chi$. Readers should keep this notation in mind when comparing Theorem 5 to previous results in geometric mechanics. However, when we state sufficient conditions for problems with subgroup symmetry in Section VII and when we prove these conditions in the appendices, it will be more convenient to work with the notation we have used in Theorem 5.

VII. Sufficient Conditions for Problems With Subgroup Symmetry

In the previous section, we found reduced necessary conditions for optimal control problems with subgroup symmetry. In this section, we give reduced sufficient conditions for such problems. We do this by deriving a system of matrix differential equations, similar to those in (11), that can be evaluated to establish non-degeneracy of the endpoint map ϕ_t from Theorem 2. The reduced sufficient conditions rely on the gradients of the state q and the reduced variables μ and χ with respect to the initial value of μ at $t = 0$. Formulas for computing these gradients are derived in Section VII-A. We then state the reduced sufficient conditions in Section VII-B. In Section VII-C, we compare the structure of the sufficient conditions with the necessary conditions found in Theorem 5.

A. Computation of the State and Costate Gradients

We now derive a set of differential equations for computing the gradients of the state q and the reduced variables μ and χ with respect to the initial value of μ at time $t = 0$. These gradients will be used to establish the reduced sufficient conditions in Section VII-B. In this section, we will use $\Phi_t : \mathfrak{s}^* \to \mathfrak{s}^*$ to denote the flow of the system (24)–(25), i.e., Φ_t maps an initial condition $(\mu(0), \chi(0)) \in \mathfrak{s}^*$ to $(\mu(t), \chi(t)) \in \mathfrak{s}^*$. Also recall that $\{X_1, \ldots, X_n\}$ is a basis for the Lie algebra $\mathfrak{g}$ and $\{X^1, \ldots, X^n\}$ is the corresponding dual basis for $\mathfrak{g}^*$. These bases are used in Lemmas 1 and 2. We first compute the gradients of the reduced variables μ and χ.

Lemma 1 Suppose *(q,p)*:$[0, t_f] \to T^*M$ is a normal extremal of (1), and assume the conditions in Theorem 2 hold. Also assume the Hamiltonian function defined in Theorem 2, which we now denote by $H_{\chi_0} : T^*G \to \mathbb{R}$, depends smoothly on the parameter $\chi_0 \in V^*$. In addition, let $S = G \times V$ be the semidirect product between G and V, and suppose that the Hamiltonian function $H : T^*S \to \mathbb{R}$, defined by $H(q, p, v, \chi) = H_\chi(q, p)$, is left-invariant under the action of S. Denote the restriction of H to $\mathfrak{s}^*$ by $h = H|_{\mathfrak{s}^*}$.

Let Φ_t be the flow of (24)–(25), and let $(\mu(t), \chi(t)) = \Phi_t(T_e^* L_{q_0}(p(0)), \rho^*(q_0)\chi_0)$. Define the matrices $\mathbf{F}, \mathbf{L} \in \mathbb{R}^{n\times n}$, $\mathbf{N}, \mathbf{P} \in \mathbb{R}^{n\times l}$, $\mathbf{R} \in \mathbb{R}^{l\times n}$, and $\mathbf{S} \in \mathbb{R}^{l\times l}$ by

$$\mathbf{F}^i{}_j = -\frac{\partial}{\partial\mu_j}\sum_{r=1}^{n}\sum_{s=1}^{n} C_{ir}^s\frac{\delta h}{\delta\mu_r}\mu_s$$

$$\mathbf{L}^i{}_j = -\frac{\partial}{\partial\mu_j}\sum_{r=1}^{l}\sum_{s=1}^{l}\chi_r\left[\rho'(X_i)\right]^r{}_s\frac{\delta h}{\delta\chi_s}$$

$$\mathbf{N}^i{}_j = -\frac{\partial}{\partial\chi_j}\sum_{r=1}^{n}\sum_{s=1}^{n} C_{ir}^s\frac{\delta h}{\delta\mu_r}\mu_s$$

$$\mathbf{P}^i{}_j = -\frac{\partial}{\partial\chi_j}\sum_{r=1}^{l}\sum_{s=1}^{l}\chi_r\left[\rho'(X_i)\right]^r{}_s\frac{\delta h}{\delta\chi_s}$$

$$\mathbf{R}^i{}_j = \frac{\partial}{\partial \mu_j} \sum_{r=1}^{l} \chi_r [\rho'(\delta h/\delta\mu)]^r{}_i$$

$$\mathbf{S}^i{}_j = \frac{\partial}{\partial \chi_j} \sum_{r=1}^{l} \chi_r [\rho'(\delta h/\delta\mu)]^r{}_i.$$

Solve the (linear, time-varying) matrix differential equations

$$\dot{\mathbf{M}} = (\mathbf{F} + \mathbf{L})\mathbf{M} + (\mathbf{N} + \mathbf{P})\mathbf{K} \tag{29}$$

$$\dot{\mathbf{K}} = \mathbf{R}\mathbf{M} + \mathbf{S}\mathbf{K} \tag{30}$$

with initial conditions $\mathbf{M}(0) = I$ and $\mathbf{K}(0) = 0$. Then

$$\begin{bmatrix} \mathbf{M}(t) \\ \mathbf{K}(t) \end{bmatrix} = \left(\nabla_{\mu_0} \Phi_t(\mu_0, \rho^*(q_0)\chi_0) \right)\Big|_{\mu_0 = T_e^* L_{q_0}(p(0))}. \tag{31}$$

Proof: See Appendix A. ■

The matrices $\mathbf{M}(t)$ and $\mathbf{K}(t)$ are the gradients of the reduced variables $\mu(t)$ and $\chi(t)$ with respect to the initial condition $\mu(0) = T_e^* L_{q_0}(p(0))$. In other words, if $a \in \mathbb{R}^n$ is the coordinate representation of $T_e^* L_{q_0}(p(0))$ so that $\sum_{i=1}^n a_i X^i = T_e^* L_{q_0}(p(0))$, then $\mathbf{M}^i{}_j(t)$ is the gradient of $\mu_i(t)$ with respect to a_j. Similarly, $\mathbf{K}^i{}_j(t)$ is the gradient of $\chi_i(t)$ with respect to a_j. Using Lemma 1, we can now compute the gradients of the state trajectory.

Lemma 2 Suppose *(q, p):* $[0, t_f] \to T^*M$ is a normal extremal of (1), and assume that the conditions in Lemma 1 hold. Define the maps $\mu(t)$ and $\chi(t)$ as in Lemma 1, and define the endpoint map $\phi_t : T_{q_0}^* G \to G$ as in Theorem 2 for the Hamiltonian vector field $X_{H_{\chi_0}}$. Next, define the matrices $\mathbf{G}, \mathbf{H} \in \mathbb{R}^{n\times n}$ and $\mathbf{T} \in \mathbb{R}^{n\times l}$ by

$$\mathbf{G}^i{}_j = \frac{\partial}{\partial \mu_j}\frac{\delta h}{\delta \mu_i} \quad \mathbf{H}^i{}_j = -\sum_{r=1}^{n} \frac{\delta h}{\delta \mu_r} C^i_{rj} \quad \mathbf{T}^i{}_j = \frac{\partial}{\partial \chi_j}\frac{\delta h}{\delta \mu_i}.$$

Solve the (linear, time-varying) matrix differential equation

$$\dot{\mathbf{J}} = \mathbf{G}\mathbf{M} + \mathbf{T}\mathbf{K} + \mathbf{H}\mathbf{J} \tag{32}$$

with initial condition $\mathbf{J}(0) = 0$, where $\mathbf{M}$ and $\mathbf{K}$ solve the matrix differential equations (29) and (30) in Lemma 1. Then the j^{th} column of $\mathbf{J}(t)$ gives the coordinate representation of $\eta_j(t)$ in (17) with respect to the basis $\{X_1, \ldots, X_n\}$ of the Lie algebra $\mathfrak{g}$.

Proof: See Appendix B. ■

Let $a \in \mathbb{R}^n$ again be the coordinate representation of $T_e^* L_{q_0}(p(0))$. The Lie algebra element $\eta_j(t)$ is the gradient of the state $q(t)$ with respect to a_j after left-translation to the identity $e \in G$. The columns of the matrix $\mathbf{J}(t)$ are therefore coordinate representations of the gradients of $q(t)$ with respect to a in terms of the Lie algebra basis $\{X_1, \ldots, X_n\}$.

B. Reduction of Sufficient Conditions

We can now state the reduced sufficient conditions, which establish a correspondence between times when the matrix $\mathbf{J}(t)$ is singular and times when the endpoint map ϕ_t is degenerate.

Theorem 6 (Semidirect Product Reduction of Sufficient Conditions) Suppose *(q,p):* $[0, t_f] \to T^*M$ is a normal extremal of (1), and assume the conditions in Lemma 1 hold. Solve the matrix differential equations in Lemmas 1 and 2 to find the matrix function $\mathbf{J} : [0, t_f] \to \mathbb{R}^{n\times n}$. Define $u : [0, t_f] \to U$ as in Theorem 2. Then (q, u) is a local optimum if there exists no $t \in (0, t_f]$ for which $\det(\mathbf{J}(t)) = 0$.

Proof: See Appendix C. ■

We conclude that non-degeneracy of the endpoint map ϕ_t, and therefore local optimality, can be established by solving the matrix differential equations (29), (30), and (32). We began in Section V with the goal of exploring the connections in Fig. 1 denoted by the dashed lines. For systems with subgroup symmetry, Theorem 6 establishes these connections.

C. Comparison of the Necessary and Sufficient Conditions

Recall from Section VI-C that we could solve for the reduced variables μ and χ and the state q in two ways. First, we could solve the differential equations (24) and (25), which are decoupled from q, to find μ and χ. Then, using (23), we could solve for q. Analogously, in the reduced sufficient conditions, we can first solve the matrix differential equations (29) and (30), which are decoupled from $\mathbf{J}$, to find $\mathbf{M}$ and $\mathbf{K}$. Then, using (32), we can solve for $\mathbf{J}$.

Alternatively, we could use the solution of (25), given by (28), to eliminate the variable χ in (24). However, this couples the reduced costate μ and the state q. The same is possible for the sufficient conditions. To see this, let a again be the coordinate representation of $\mu_0 = T_e^* L_{q_0}(p(0))$, and recall from Theorem 2 that ϕ_t is the endpoint map sending $p(0) \in T_{q_0}^* G$ to $\phi_t(p(0)) = q(t)$. Using the definition of $\mathbf{J}$, the gradient of $\rho(\phi_t(p(0)))$ with respect to a_j is

$$\frac{\partial}{\partial a_j}\rho(\phi_t(p(0))) = \rho(q(t)) \left(\sum_{s=1}^{n} \rho'(X_s)\mathbf{J}^s{}_j(t) \right).$$

Therefore, using the solution $\chi(t) = \chi_0 \rho(q(t))$ given in (28), we can write the gradient of $\chi(t)$ with respect to a_j as

$$\begin{aligned} \frac{\partial}{\partial a_j}\chi(t) &= \chi_0 \left(\frac{\partial}{\partial a_j}\rho(\phi_t(p(0))) \right) \\ &= \chi_0 \rho(q(t)) \left(\sum_{s=1}^{n} \rho'(X_s)\mathbf{J}^s{}_j(t) \right) \\ &= \sum_{s=1}^{n} \left(\rho'(X_s)\right)^* \chi(t)\mathbf{J}^s{}_j(t), \end{aligned}$$

where we have used the notation $\partial\chi/\partial a_j$ to denote the gradient of the χ component of the flow of (24)–(25) with respect to a_j. Writing this expression in coordinates of χ gives

$$\frac{\partial}{\partial a_j}\chi_i(t) = \sum_{r=1}^{l}\sum_{s=1}^{n} \chi_r(t)[\rho'(X_s)]^r{}_i \mathbf{J}^s{}_j(t).$$

In Lemma 3, we show that the above expression can be used to construct the solution of the differential equation (30).

Lemma 3: Define the matrix $\mathbf{K} \in \mathbb{R}^{l\times n}$ by

$$\mathbf{K}^i{}_j = \sum_{r=1}^{l}\sum_{s=1}^{n} \chi_r[\rho'(X_s)]^r{}_i \mathbf{J}^s{}_j. \tag{33}$$

TABLE I
COMPUTATIONS AND COUPLING IN NECESSARY AND SUFFICIENT CONDITIONS FOR OPTIMAL CONTROL PROBLEMS WITH SYMMETRY

Symmetry Type	Necessary Conditions	Sufficient Conditions	Coupling of Equations
No Symmetry (working in local coordinates)	$\dot{p} = -H_q$ $\dot{q} = H_p$	$\dot{\mathbf{M}} = -H_{pq}\,\mathbf{M} - H_{qq}\,\mathbf{J}$ $\dot{\mathbf{J}} = H_{pp}\,\mathbf{M} + H_{qp}\,\mathbf{J}$	p and q can be coupled. $\mathbf{M}$ and $\mathbf{J}$ can be coupled. Coefficient matrices in sufficient conditions may depend upon p and q.
Left-Invariant	$\dot{\mu} = \mathrm{ad}^*_{\delta h/\delta\mu}(\mu)$ $\dot{q} = X_{\delta h/\delta\mu}(q)$	$\dot{\mathbf{M}} = \mathbf{FM}$ $\dot{\mathbf{J}} = \mathbf{GM} + \mathbf{HJ}$	μ is decoupled from q. $\mathbf{M}$ is decoupled from $\mathbf{J}$. Coefficient matrices in sufficient conditions only depend upon μ.
Subgroup Symmetry	$\dot{\mu} = \mathrm{ad}^*_{\delta h/\delta\mu}(\mu) - \left(\rho'_{\delta h/\delta\chi}\right)^*\chi$ $\dot{\chi} = \rho'(\delta h/\delta\mu)^*\chi$ $\dot{q} = X_{\delta h/\delta\mu}(q)$ or $\dot{\mu} = \mathrm{ad}^*_{\delta h/\delta\mu}(\mu) - \left(\rho'_{\delta h/\delta\chi}\right)^*\chi$ $\chi(t) = \rho(q(t))^*\chi_0$ $\dot{q} = X_{\delta h/\delta\mu}(q)$	$\dot{\mathbf{M}} = (\mathbf{F}+\mathbf{L})\mathbf{M} + (\mathbf{N}+\mathbf{P})\mathbf{K}$ $\dot{\mathbf{K}} = \mathbf{RM} + \mathbf{SK}$ $\dot{\mathbf{J}} = \mathbf{GM} + \mathbf{TK} + \mathbf{HJ}$ or $\dot{\mathbf{M}} = (\mathbf{F}+\mathbf{L})\mathbf{M} + (\mathbf{N}+\mathbf{P})\mathbf{K}$ $\mathbf{K} = \mathbf{UJ}$ $\dot{\mathbf{J}} = \mathbf{GM} + \mathbf{TK} + \mathbf{HJ}$	μ and χ are decoupled from q. $\mathbf{M}$ and $\mathbf{K}$ are decoupled from $\mathbf{J}$. Coefficient matrices in sufficient conditions only depend upon μ and χ. χ is a function of q, which couples μ and q. $\mathbf{K}$ is a function of $\mathbf{J}$, which couples $\mathbf{M}$ and $\mathbf{J}$. Coefficient matrices in sufficient conditions only depend upon μ and χ.

Then $\mathbf{K}$ solves the matrix differential equation (30) with the initial condition $\mathbf{K}(0) = 0$.

Proof: See Appendix D. ■

We can now use the solution for $\mathbf{K}$ in the matrix differential equations (29) and (32). This gives an alternative system of matrix differential equations that can be solved to find the matrix $\mathbf{J}$, as shown in Theorem 7.

Theorem 7 (Alternative Test for Conjugate Points) Suppose $(q, p) : [0, t_f] \to T^*M$ is a normal extremal of (1), and assume the conditions in Theorem 6 hold. Define the matrix $\mathbf{U} \in \mathbb{R}^{l\times n}$ by

$$\mathbf{U}^i{}_j = \sum_{r=1}^{l} \chi_r \left[\rho'(X_j)\right]^r{}_i.$$

Solve the (linear, time-varying) matrix differential equations

$$\begin{aligned}\dot{\mathbf{M}} &= (\mathbf{F}+\mathbf{L})\mathbf{M} + (\mathbf{N}+\mathbf{P})\mathbf{UJ}\\ \dot{\mathbf{J}} &= \mathbf{GM} + (\mathbf{TU}+\mathbf{H})\mathbf{J}\end{aligned} \tag{34}$$

with the initial conditions $\mathbf{M}(0) = I$ and $\mathbf{J}(0) = 0$. Define $u : [0, t_f] \to U$ as in Theorem 2. Then (q, u) is a local optimum if there exists no $t \in (0, t_f]$ for which $\det(\mathbf{J}(t)) = 0$.

Proof: From Lemma 3, we have $\mathbf{K} = \mathbf{UJ}$. The differential equations (34) are obtained by directly substituting this expression for $\mathbf{K}$ into the differential equations (29) and (32). ■

Just as we were able to eliminate the reduced variable χ in the differential equation (24) in the necessary conditions, we can eliminate the matrix $\mathbf{K}$ in the matrix differential equations (29) and (32) in the sufficient conditions. In the necessary conditions, the elimination of χ results in a coupling between the reduced costate μ and the state q. In the sufficient conditions, the elimination of $\mathbf{K}$ results in a coupling between the matrices $\mathbf{M}$ and $\mathbf{J}$.

Table I summarizes the computations and the coupling between equations in the necessary and sufficient conditions for optimal control problems without symmetry, left-invariant problems, and problems with subgroup symmetry. For systems with subgroup symmetry, we see that the variables μ, χ, and q play analogous roles in the necessary conditions as the matrices $\mathbf{M}$, $\mathbf{K}$, and $\mathbf{J}$ in the sufficient conditions, respectively. Applying symmetry reduction provides similar simplifications in both the necessary and sufficient conditions.

VIII. CONJUGATE POINTS IN THE HEAVY TOP

In this section, we return to the optimal control problem (12) with the augmented cost function (18), which models a spinning top in a gravitational field. We first apply the results in Theorem 5 to obtain necessary conditions for optimality in Section VIII-A. In Section VIII-B, we apply the reduced sufficient conditions in Theorems 6 and 7, which give two equivalent ways of establishing local optimality. In Section VIII-C, we compute conjugate points in the axisymmetric sleeping top and determine which of these trajectories are locally optimal solutions of the optimal control problem.

A. Necessary Conditions for the Heavy Top

Recall that the Hamiltonian function H_{χ_0} for the heavy spinning top is given by (19). Extending H_{χ_0} to be a function on $T^*G \times V \times V^* = T^*SO(3) \times \mathbb{R}^3 \times \mathbb{R}^{3*}$ gives

$$H(q, p, v, \chi) = \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\langle p, qX_i\rangle^2 - \chi(q\nu).$$

For any $v \in V$, $\chi \in V^*$, $p \in T_q^*SO(3)$, and $g, q, r \in SO(3)$ satisfying $q = gr$, we have

$$\begin{aligned} &H(r, T_r^*L_g(p), v, \rho^*(g)\chi)\\ &= \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\langle T_r^*L_g(p), g^{-1}qX_i\rangle^2 - \rho^*(g)\chi(g^{-1}q\nu)\\ &= \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\langle p, gg^{-1}qX_i\rangle^2 - \chi(gg^{-1}q\nu)\\ &= \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\langle p, qX_i\rangle^2 - \chi(q\nu)\\ &= H(q, p, v, \chi).\end{aligned} \tag{35}$$

Therefore H is left-invariant under the action of S. This implies that H_{χ_0} is left-invariant under the action of G_{χ_0}, which simply means that H_{χ_0} is left-invariant under rotations around the gravity vector. As a consequence, we can apply Theorem 5. The reduced Hamiltonian on $\mathfrak{s}^*$ is given by

$$h(\mu, \chi) = H(e, \mu, v, \chi) = \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\mu_i^2 - \chi_i \nu^i.$$

The necessary conditions in Theorem 5 give that μ and χ satisfy (24) and (25), which are equivalent to

$$\dot{\mu} = \mu \times u + \chi \times \nu \qquad \dot{\chi} = \chi \times u, \tag{36}$$

where $u^i = c_i^{-1}\mu_i$. The solution for χ, given by (28), is $\chi(t) = \chi_0 q(t)$. We see that $\chi(t)$ gives the direction of the gravity vector rotated into the local coordinate frame at $q(t)$.

B. Sufficient Conditions for Optimality

We now apply the sufficient conditions in Theorem 6. Computing the matrices $\mathbf{F}$, $\mathbf{L}$, $\mathbf{N}$, $\mathbf{P}$, $\mathbf{R}$, and $\mathbf{S}$ from Lemma 1, we see that $\mathbf{F}$ is identical to (16), $\mathbf{L}$ and $\mathbf{N}$ are both zero matrices, $\mathbf{S}$ is identical to $\mathbf{H}$ in (16), and

$$\mathbf{P} = -\widehat{\nu} \qquad \mathbf{R} = \begin{bmatrix} 0 & -\chi_3/c_2 & \chi_2/c_3 \\ \chi_3/c_1 & 0 & -\chi_1/c_3 \\ -\chi_2/c_1 & \chi_1/c_2 & 0 \end{bmatrix}.$$

Computing the matrices in Lemma 2, we see that $\mathbf{G}$ and $\mathbf{H}$ are identical to (16), and $\mathbf{T} = 0$. Therefore, to find the matrix $\mathbf{J}$ and establish non-degeneracy of the endpoint map ϕ_t, we must solve the matrix differential equations (29), (30), and (32) with the initial conditions $\mathbf{M}(0) = I$, $\mathbf{K}(0) = 0$, and $\mathbf{J}(0) = 0$. Alternatively, we can apply Theorem 7 and solve the matrix differential equations (34). Computing the matrix $\mathbf{U}$ in Lemma 7 gives $\mathbf{U} = \widehat{\chi}$. For a given solution of (36), these matrix differential equations can be solved to determine if the solution is a local optimum of the optimal control problem.

C. The Axisymmetric Sleeping Top

We now consider a sleeping axisymmetric top in a gravitational field with the parameters in the cost function (18) satisfying $c_2 = c_3 = 1$ and $v = [1\ 0\ 0]^T$. The top is said to be sleeping when its axis of symmetry is aligned with the direction of gravity, and the top is rotating about this axis. These trajectories correspond to the fixed points of the system (36) given by $\mu_2 = \mu_3 = \chi_2 = \chi_3 = 0$, where μ_1 and χ_1 are arbitrary. At these fixed points, the matrix differential equations (34) are linear time-invariant equations.

After solving the linear time-invariant system (34) at these fixed points, we can compute $\det(\mathbf{J}(t))$. If $\chi_1 - \mu_1^2/4 > 0$, then $\det(\mathbf{J}(t))$ simplifies to

$$\det(\mathbf{J}(t)) = \frac{t}{c_1\left(\chi_1 - \frac{\mu_1^2}{4}\right)} \sinh^2\left(t\sqrt{\chi_1 - \frac{\mu_1^2}{4}}\right).$$

If the term $\chi_1 - \mu_1^2/4 < 0$, then $\det(\mathbf{J}(t))$ simplifies to

$$\det(\mathbf{J}(t)) = \frac{t}{c_1\left(\frac{\mu_1^2}{4} - \chi_1\right)} \sin^2\left(t\sqrt{\frac{\mu_1^2}{4} - \chi_1}\right).$$

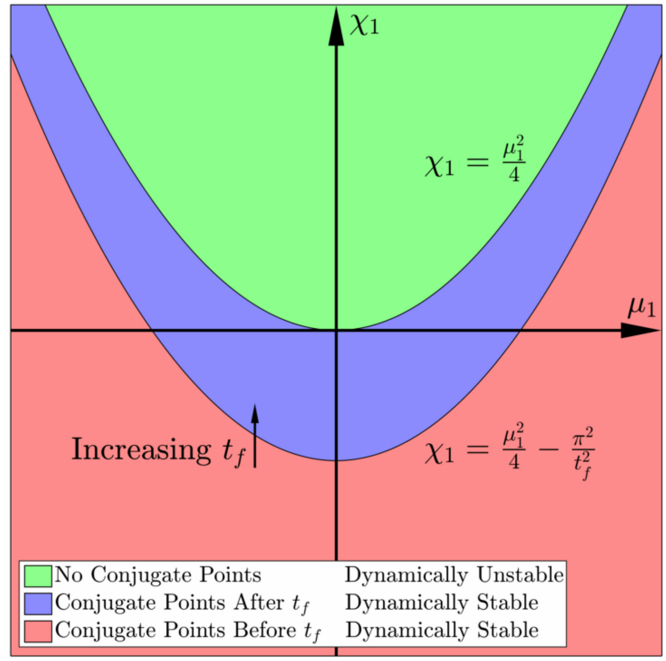

Fig. 2. Regions of optimal and non-optimal fixed points in the μ_1 - χ_1 plane corresponding to trajectories of the sleeping top.

If $\chi_1 - \mu_1^2/4 = 0$, then $\det(\mathbf{J}(t))$ simplifies to

$$\det(\mathbf{J}(t)) = \frac{t^3}{c_1}.$$

We see that if $\chi_1 \geq \mu_1^2/4$, then $\det(\mathbf{J}(t)) > 0$ for all $t > 0$. Therefore, the sleeping top with $\chi_1 \geq \mu_1^2/4$ is a local optimum of the optimal control problem for arbitrarily large final times t_f. If $\chi_1 < \mu_1^2/4$, then the first conjugate point occurs at

$$t = \pi\left(\frac{\mu_1^2}{4} - \chi_1\right)^{-1/2}. \tag{37}$$

We conclude that the sleeping top with $\chi_1 < \mu_1^2/4$ is a local optimum if the final time t_f satisfies

$$t_f < \pi\left(\frac{\mu_1^2}{4} - \chi_1\right)^{-1/2}.$$

These expressions for conjugate points in the heavy sleeping top have not appeared in previous work, and our ability to derive them relies on the results we proved in Section VII.

We end by discussing an interesting link between conjugate points and dynamic stability of the sleeping top. The boundary between locally optimal and non-optimal fixed points in the μ_1–χ_1 plane is given by (37) with $t = t_f$. Rearranging this expression gives

$$\chi_1 = \frac{\mu_1^2}{4} - \frac{\pi^2}{t_f^2}.$$

This boundary between locally optimal and non-optimal fixed points is shown in Fig. 2 in the μ_1 - χ_1 plane. As t_f increases, the boundary between locally optimal and non-optimal fixed points approaches $\chi_1 = \mu_1^2/4$, which is also shown in Fig. 2.

In particular, $\chi_1 = \mu_1^2/4$ is the boundary in the μ_1 - χ_1 plane between fixed points that are locally optimal for arbitrarily large $t_f > 0$ and fixed points that lose optimality for some finite t_f. For a given $t_f > 0$, these boundaries allow us to partition the μ_1 - χ_1 plane into three regions. The red region in Fig. 2 corresponds to trajectories of the sleeping top that have conjugate points before t_f, and are therefore non-optimal. The blue region corresponds to trajectories that have conjugate points after t_f, and are therefore locally optimal. The green region corresponds to locally optimal trajectories that have no conjugate points for all time.

The curve $\chi_1 = \mu_1^2/4$ is also the boundary between dynamically stable and unstable trajectories of the sleeping top, as denoted in Fig. 2 [40]. This result agrees with the kinetic instability theorem of Kelvin and Tait [41], which states that a trajectory of a conservative system without conjugate points for arbitrarily large t_f is unstable [42].

IX. Conclusion

We have applied Lie-Poisson reduction by stages to geometric optimal control problems on Lie groups with subgroup symmetry. After providing reduced necessary conditions for optimality, we derived reduced sufficient conditions for optimality based on the non-existence of conjugate points. Whereas the general necessary and sufficient conditions in Section III were coordinate-free conditions, the reduced necessary and sufficient conditions were stated in terms of coordinate formulas and rely on solutions of ordinary differential equations, and evaluating these coordinate formulas did not depend upon local coordinates on the Lie group. These results were then applied to a geometric optimal control problem that models the motion of a spinning top in a gravitational field. Using Theorems 5, 6, and 7, we derived new results on conjugate points in the axisymmetric sleeping top.

The results in this paper suggest that by exploiting symmetries, sufficient conditions for optimality can be simplified in an analogous way to necessary conditions. This was shown by comparing the computations needed to evaluate the reduced necessary and sufficient conditions. In particular, the two alternative ways of evaluating the necessary conditions for problems with subgroup symmetry lead to two analogous ways of evaluating the sufficient conditions, as shown in Table I. A deeper understanding of this connection could be obtained by comparing the quotient spaces on which the reduced necessary and sufficient conditions evolve, rather than comparing coordinate formulas. For the left-invariant problems considered in Section IV, the unreduced necessary conditions evolve on T^*G, while the reduced necessary conditions evolve on a quotient space that is diffeomorphic to $\mathfrak{g}^*$. Future work could analyze the corresponding quotient spaces on which the reduced sufficient conditions evolve.

Other extensions include the consideration of more general approaches to symmetry group reduction. In this paper, we focused on the case when the state of the optimal control problem evolves on a Lie group and the symmetry group of the system is the isotropy group of a parameter. However, reduction by stages, and more generally symmetry group reduction, can be applied to systems with less structure, i.e., where the state takes values on a smooth manifold [3]. Exploring the connection between reduction of the necessary and sufficient conditions for optimality in more general optimal control problems is an interesting direction for future work.

Appendix A
Proof of Lemma 1

We use the notation $\partial\mu_i/\partial a_j$ and $\partial\chi_i/\partial a_j$ to denote the μ_i and χ_i components, respectively, of the gradient of the flow of (24)–(25), i.e., the right side of the expression (31). Defining $\mathbf{M}^i{}_j = \partial\mu_i/\partial a_j$ and $\mathbf{K}^i{}_j = \partial\chi_i/\partial a_j$ and using (26), we find

$$\begin{aligned}
\frac{d}{dt}\mathbf{M}^i{}_j &= \frac{\partial\dot{\mu}_i}{\partial a_j}\\
&= -\frac{\partial}{\partial a_j}\sum_{r=1}^{n}\sum_{s=1}^{n} C_{ir}^s \frac{\delta h}{\delta\mu_r}\mu_s\\
&\quad -\frac{\partial}{\partial a_j}\sum_{r=1}^{l}\sum_{s=1}^{l}\chi_r\left[\rho'(X_i)\right]^r{}_s\frac{\delta h}{\delta\chi_s}\\
&= \sum_{k=1}^{n}\left(-\frac{\partial}{\partial\mu_k}\sum_{r=1}^{n}\sum_{s=1}^{n} C_{ir}^s\frac{\delta h}{\delta\mu_r}\mu_s\right)\frac{\partial\mu_k}{\partial a_j}\\
&\quad + \sum_{k=1}^{l}\left(-\frac{\partial}{\partial\chi_k}\sum_{r=1}^{n}\sum_{s=1}^{n} C_{ir}^s\frac{\delta h}{\delta\mu_r}\mu_s\right)\frac{\partial\chi_k}{\partial a_j}\\
&\quad + \sum_{k=1}^{n}\left(-\frac{\partial}{\partial\mu_k}\sum_{r=1}^{l}\sum_{s=1}^{l}\chi_r\left[\rho'(X_i)\right]^r{}_s\frac{\delta h}{\delta\chi_s}\right)\frac{\partial\mu_k}{\partial a_j}\\
&\quad + \sum_{k=1}^{l}\left(-\frac{\partial}{\partial\chi_k}\sum_{r=1}^{l}\sum_{s=1}^{l}\chi_r\left[\rho'(X_i)\right]^r{}_s\frac{\delta h}{\delta\chi_s}\right)\frac{\partial\chi_k}{\partial a_j}\\
&= \sum_{k=1}^{n}\left(\mathbf{F}^i{}_k + \mathbf{L}^i{}_k\right)\mathbf{M}^k{}_j + \sum_{k=1}^{l}\left(\mathbf{N}^i{}_k + \mathbf{P}^i{}_k\right)\mathbf{K}^k{}_j.
\end{aligned}$$

It is clear that $\mathbf{M}^i{}_j(0) = \delta^i{}_j$, so we have verified (29). A similar calculation, using (27), shows that (30) holds.

Appendix B
Proof of Lemma 2

We need the following lemmas before proving Lemma 2.

Lemma 4: Let $q : W \to G$ be a smooth map, where $W \subset \mathbb{R}^2$ is simply connected. Denote its partial derivatives $\zeta : W \to \mathfrak{g}$ and $\eta : W \to \mathfrak{g}$ by

$$\begin{aligned}
\zeta(t,\epsilon) &= T_{q(t,\epsilon)}L_{q(t,\epsilon)^{-1}}\left(\frac{\partial q(t,\epsilon)}{\partial t}\right)\\
\eta(t,\epsilon) &= T_{q(t,\epsilon)}L_{q(t,\epsilon)^{-1}}\left(\frac{\partial q(t,\epsilon)}{\partial \epsilon}\right).
\end{aligned} \tag{38}$$

Then

$$\frac{\partial \zeta}{\partial \epsilon} - \frac{\partial \eta}{\partial t} = [\zeta, \eta]. \tag{39}$$

Conversely, if there exist smooth maps ζ and η satisfying (39), then there exists a smooth map q satisfying (38).

Proof: See Proposition 5.1 in [43]. ■

Lemma 5: Let $\alpha, \beta, \gamma \in \mathfrak{g}$ and suppose $\gamma = [\alpha, \beta]$. Then

$$\gamma^k = \sum_{r=1}^{n}\sum_{s=1}^{n} \alpha^r \beta^s C^k_{rs}.$$

Proof: See Lemma 2 in [4]. ■

We can now prove Lemma 2. For $j \in \{1, \dots, n\}$, define $\eta_j(t) \in \mathfrak{g}$ as in (17). Also let $\zeta(t) \in \mathfrak{g}$ be

$$\zeta(t) = T_{q(t)} L_{q(t)^{-1}}\left(\dot q(t)\right) = \frac{\delta h}{\delta \mu},$$

where the second equality follows from (23) in Theorem 5. From Lemma 4, we have

$$\dot\eta_j = \frac{\partial \zeta}{\partial a_j} - [\zeta, \eta_j] = \frac{\partial}{\partial a_j}\frac{\delta h}{\delta \mu} - \left[\frac{\delta h}{\delta \mu}, \eta^j\right].$$

After defining $\mathbf{J}^i{}_j(t) = \eta^i_j(t)$, so that the j^{th} column of $\mathbf{J}$ is the coordinate representation of $\eta_j(t)$, the previous equation can be written in coordinates using Lemma 5:

$$\begin{aligned}
\dot{\mathbf{J}}^i_j &= \dot\eta^i_j \\
&= \sum_{k=1}^{n}\left(\frac{\partial}{\partial \mu_k}\frac{\delta h}{\delta \mu_i}\right)\frac{\partial \mu_k}{\partial a_j} + \sum_{k=1}^{l}\left(\frac{\partial}{\partial \chi_k}\frac{\delta h}{\delta \mu_i}\right)\frac{\partial \chi_k}{\partial a_j} \\
&\quad + \sum_{k=1}^{n}\left(-\sum_{r=1}^{n}\frac{\delta h}{\delta \mu_r} C^i_{rk}\right)\eta^k_j \\
&= \sum_{k=1}^{n}\mathbf{G}^i{}_k \mathbf{M}^k{}_j + \sum_{k=1}^{l}\mathbf{T}^i{}_k \mathbf{K}^k{}_j + \sum_{k=1}^{n}\mathbf{H}^i{}_k \mathbf{J}^k{}_j,
\end{aligned} \tag{40}$$

where we used $\mathbf{M}^i{}_j = \partial \mu_i / \partial a_j$ and $\mathbf{K}^i{}_j = \partial \chi_i / \partial a_j$ from Lemma 1. It is clear that $\mathbf{J}^i{}_j = 0$, so we have verified (32).

Appendix C
Proof of Theorem 6

Define the smooth map $\sigma : \mathbb{R}^n \to T^*_{q_0} G$ by

$$\sigma(a) = T^*_{q_0} L_{q_0^{-1}}\left(\sum_{i=1}^{n} a_i P_i\right).$$

This expression also defines $\sigma : \mathbb{R}^n \to T_{p_0}(T^*_{q_0} G)$ if we identify $T^*_{q_0} G$ with $T_{p_0}(T^*_{q_0} G)$ in the usual way. Given $p_0 = \sigma(a)$ for some $a \in \mathbb{R}^n$, there exists non-zero $\lambda \in T_{p_0}(T^*_{q_0} G)$ satisfying $T_{p_0}\phi_t(\lambda) = 0$ if and only if there exists non-zero $s \in \mathbb{R}^n$ satisfying $T_{\sigma(a)}\phi_t(\sigma(s)) = 0$. Now observe that

$$T_{\sigma(a)}\phi_t(\sigma(s)) = \sum_{j=1}^{n} s_j \left(T_{\sigma(a)}\phi_t\left(T^*_{q_0} L_{q_0^{-1}}\left(X^j\right)\right)\right),$$

where, recall, $\{X^1, \dots, X^n\}$ is a basis for $\mathfrak{g}^*$. By left translation, $T_{\sigma(a)}\phi_t(\sigma(s)) = 0$ if and only if

$$0 = \sum_{j=1}^{n} s_j T_{q(t)} L_{q(t)^{-1}}\left(T_{\sigma(a)}\phi_t\left(T^*_{q_0} L_{q_0^{-1}}\left(X^j\right)\right)\right). \tag{41}$$

With $\eta_j(t) \in \mathfrak{g}$ as defined in (17), the above expression is equivalent to $0 = \sum_{j=1}^{n} s_j \eta_j(t)$. From Lemma 2, $\mathbf{J}(t)$ satisfies $\mathbf{J}^i{}_j(t) = \eta^i_j(t)$, i.e., the j^{th} column of $\mathbf{J}(t)$ is the coordinate representation of $\eta^j(t)$ with respect to the basis $\{X_1, \dots, X_n\}$. Then, (41) holds for some $s \neq 0$ if and only if $\det(\mathbf{J}(t)) = 0$. Therefore ϕ_t is degenerate at p_0 if and only if $\det(\mathbf{J}(t)) = 0$. The result follows by application of Theorem 2.

Appendix D
Proof of Lemma 3

Since $\mathbf{J}(0) = 0$, it is clear that $\mathbf{K}(0) = 0$. Taking the time derivative of (33) gives

$$\dot{\mathbf{K}}^i_j = \sum_{r=1}^{l}\sum_{s=1}^{n}\left(\dot\chi_r [\rho'(X_s)]^r{}_i \mathbf{J}^s{}_j + \chi_r [\rho'(X_s)]^r{}_i \dot{\mathbf{J}}^s_j\right).$$

Using (27) and (40), we have

$$\begin{aligned}
\dot{\mathbf{K}}^i_j &= \sum_{r=1}^{l}\sum_{s=1}^{n}\sum_{k=1}^{l} \chi_k [\rho'(\delta h/\delta\mu)]^k{}_r [\rho'(X_s)]^r{}_i \mathbf{J}^s{}_j \\
&\quad + \sum_{r=1}^{l}\sum_{s=1}^{n}\sum_{k=1}^{n} \chi_r [\rho'(X_s)]^r{}_i \left(\frac{\partial}{\partial \mu_k}\frac{\delta h}{\delta \mu_s}\right)\mathbf{M}^k{}_j \\
&\quad + \sum_{r=1}^{l}\sum_{s=1}^{n}\sum_{k=1}^{l} \chi_r [\rho'(X_s)]^r{}_i \left(\frac{\partial}{\partial \chi_k}\frac{\delta h}{\delta \mu_s}\right)\mathbf{K}^k{}_j \\
&\quad - \sum_{r=1}^{l}\sum_{s=1}^{n}\sum_{k=1}^{n} \chi_r [\rho'(X_s)]^r{}_i \left(\sum_{p=1}^{n}\frac{\delta h}{\delta \mu_p} C^s_{pk}\right)\mathbf{J}^k{}_j.
\end{aligned} \tag{42}$$

We will analyze each row in the above expression individually. The second row of (42) is equivalent to

$$\begin{aligned}
&\sum_{r=1}^{l}\sum_{s=1}^{n}\sum_{k=1}^{n} \chi_r [\rho'(X_s)]^r{}_i \left(\frac{\partial}{\partial \mu_k}\frac{\delta h}{\delta \mu_s}\right)\mathbf{M}^k{}_j \\
&= \sum_{k=1}^{n}\left(\frac{\partial}{\partial \mu_k}\sum_{r=1}^{l}\chi_r\left(\sum_{s=1}^{n}[\rho'(X_s)]^r{}_i \frac{\delta h}{\delta \mu_s}\right)\right)\mathbf{M}^k{}_j \\
&= \sum_{k=1}^{n}\left(\frac{\partial}{\partial \mu_k}\sum_{r=1}^{l}\chi_r [\rho'(\delta h/\delta \mu)]^r{}_i\right)\mathbf{M}^k{}_j \\
&= \sum_{k=1}^{n}\mathbf{R}^i{}_k \mathbf{M}^k{}_j.
\end{aligned}$$

Next, consider the third row of (42):

$$\sum_{r=1}^{l}\sum_{s=1}^{n}\sum_{k=1}^{l}\chi_r\left[\rho'(X_s)\right]^r{}_i\left(\frac{\partial}{\partial\chi_k}\frac{\delta h}{\delta\mu_s}\right)\mathbf{K}^k{}_j$$

$$=\sum_{k=1}^{l}\left(\sum_{r=1}^{l}\chi_r\frac{\partial}{\partial\chi_k}\left(\sum_{s=1}^{n}\left[\rho'(X_s)\right]^r{}_i\frac{\delta h}{\delta\mu_s}\right)\right)\mathbf{K}^k{}_j$$

$$=\sum_{k=1}^{l}\left(\sum_{r=1}^{l}\chi_r\frac{\partial}{\partial\chi_k}\left[\rho'(\delta h/\delta\mu)\right]^r{}_i\right)\mathbf{K}^k{}_j.$$

Now consider the fourth row of (42). Using the definition of the structure constants in (7), we have

$$\sum_{r=1}^{l}\sum_{s=1}^{n}\sum_{k=1}^{n}\chi_r\left[\rho'(X_s)\right]^r{}_i\left(\sum_{p=1}^{n}\frac{\delta h}{\delta\mu_p}C^s_{pk}\right)\mathbf{J}^k{}_j$$

$$=\sum_{k=1}^{n}\sum_{r=1}^{l}\sum_{p=1}^{n}\chi_r\frac{\delta h}{\delta\mu_p}\left(\sum_{s=1}^{n}C^s_{pk}\left[\rho'(X_s)\right]^r{}_i\right)\mathbf{J}^k{}_j$$

$$=\sum_{k=1}^{n}\sum_{r=1}^{l}\sum_{p=1}^{n}\chi_r\frac{\delta h}{\delta\mu_p}\left[\rho'([X_p,X_k])\right]^r{}_i[\mathbf{J}]^k{}_j.$$

By the definition of the Lie bracket, we have

$$\rho'([X_p,X_k])=\rho'(X_p)\rho'(X_k)-\rho'(X_k)\rho'(X_p)$$

and

$$\left[\rho'([X_p,X_k])\right]^r{}_i=\sum_{s=1}^{l}\left[\rho'(X_p)\right]^r{}_s\left[\rho'(X_k)\right]^s{}_i-\left[\rho'(X_k)\right]^r{}_s\left[\rho'(X_p)\right]^s{}_i.$$

We now have that the fourth row of (42) is equivalent to

$$\sum_{k=1}^{n}\sum_{r=1}^{l}\sum_{p=1}^{n}\sum_{s=1}^{l}\chi_r\frac{\delta h}{\delta\mu_p}\left[\rho'(X_p)\right]^r{}_s\left[\rho'(X_k)\right]^s{}_i\mathbf{J}^k{}_j$$

$$-\sum_{k=1}^{n}\sum_{r=1}^{l}\sum_{p=1}^{n}\sum_{s=1}^{l}\chi_r\frac{\delta h}{\delta\mu_p}\left[\rho'(X_k)\right]^r{}_s\left[\rho'(X_p)\right]^s{}_i\mathbf{J}^k{}_j$$

$$=\sum_{k=1}^{n}\sum_{r=1}^{l}\sum_{s=1}^{l}\chi_r\left[\rho'(\delta h/\delta\mu)\right]^r{}_s\left[\rho'(X_k)\right]^s{}_i\mathbf{J}^k{}_j$$

$$-\sum_{k=1}^{n}\sum_{r=1}^{l}\sum_{s=1}^{l}\chi_r\left[\rho'(\delta h/\delta\mu)\right]^s{}_i\left[\rho'(X_k)\right]^r{}_s\mathbf{J}^k{}_j.$$

Since the fourth line in (42) begins with a minus sign, we see that the first term in the above expression cancels with the first line in (42). We are left with

$$\sum_{k=1}^{n}\sum_{r=1}^{l}\sum_{s=1}^{l}\chi_r\left[\rho'(\delta h/\delta\mu)\right]^s{}_i\left[\rho'(X_k)\right]^r{}_s\mathbf{J}^k{}_j$$

$$=\sum_{s=1}^{l}\left[\rho'(\delta h/\delta\mu)\right]^s{}_i\left(\sum_{r=1}^{l}\sum_{k=1}^{n}\chi_r\left[\rho'(X_k)\right]^r{}_s\mathbf{J}^k{}_j\right).$$

Using (33), this expression is equivalent to

$$\sum_{s=1}^{l}\left[\rho'(\delta h/\delta\mu)\right]^s{}_i\mathbf{K}^s{}_j.$$

Combining these calculations, we have that

$$\dot{\mathbf{K}}^i_j=\sum_{k=1}^{n}\mathbf{R}^i{}_k\mathbf{M}^k{}_j+\sum_{k=1}^{l}\left[\rho'(\delta h/\delta\mu)\right]^k{}_i\mathbf{K}^k{}_j+\sum_{k=1}^{l}\left(\sum_{r=1}^{l}\chi_r\frac{\partial}{\partial\chi_k}\left[\rho'(\delta h/\delta\mu)\right]^r{}_i\right)\mathbf{K}^k{}_j$$

$$=\sum_{k=1}^{n}\mathbf{R}^i{}_k\mathbf{M}^k{}_j+\sum_{k=1}^{l}\frac{\partial}{\partial\chi_k}\left(\sum_{r=1}^{l}\chi_r\left[\rho'(\delta h/\delta\mu)\right]^r{}_i\right)\mathbf{K}^k{}_j$$

$$=\sum_{k=1}^{n}\mathbf{R}^i{}_k\mathbf{M}^k{}_j+\sum_{k=1}^{n}\mathbf{S}^i{}_k\mathbf{K}^k{}_j.$$

We conclude that $\mathbf{K}$ solves the differential equation (30).

Acknowledgment

The authors would like to thank R. Loja Fernandes and I. Marcut for many helpful discussions.

References

[1] A. A. Agrachev and Y. L. Sachkov, *Control Theory from the Geometric Viewpoint*. Berlin, Germany: Springer, 2004.

[2] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*. New York, NY: Springer, 1999.

[3] J. E. Marsden, G. Misiolek, J. P. Ortega, M. Perlmutter, and T. S. Ratiu, *Hamiltonian Reduction by Stages*. Berlin, Germany: Springer, 2003.

[4] T. Bretl and Z. McCarthy, "Quasi-static manipulation of a Kirchhoff elastic rod based on a geometric analysis of equilibrium configurations," *Int. J. Robot. Res.*, vol. 33, no. 1, pp. 48–68, 2014.

[5] A. Borum and T. Bretl, "Geometric optimal control for symmetry breaking cost functions," in *Proc. 53rd IEEE Conf. Decision Control*, 2014, pp. 5855–5861.

[6] A. Borum, "Optimal control problems on Lie groups with symmetry breaking cost functions," M.S. thesis, University of Illinois at Urbana-Champaign, 2014.

[7] D. D. Holm, J. E. Marsden, and T. S. Ratiu, "The Euler-Poincaré equations and semidirect products with applications to continuum theories," *Adv. Math.*, vol. 137, no. 1, pp. 1–81, 1998.

[8] J. E. Marsden, T. S. Ratiu, and A. Weinstein, "Semidirect products and reduction in mechanics," *Trans. Amer. Math. Soc.*, vol. 281, no. 1, pp. 147–177, 1984.

[9] J. E. Marsden, T. S. Ratiu, and A. Weinstein, "Reduction and Hamiltonian structures on duals of semidirect product Lie algebras," *Contemp. Math*, vol. 28, pp. 55–100, 1984.

[10] N. E. Leonard and J. E. Marsden, "Stability and drift of underwater vehicle dynamics: Mechanical systems with rigid motion symmetry," *Physica D*, vol. 105, pp. 130–162, 1997.

[11] L. Colombo and H. O. Jacobs, "Lagrangian mechanics on centered semi-direct products," *Geometry, Mech., Dynamics, Fields Inst. Commun.*, vol. 74, pp. 167–184, 2015.

[12] J. Grizzle and S. Marcus, "Optimal control of systems possessing symmetries," *IEEE Trans. Autom. Control*, vol. 29, no. 11, pp. 1037–1040, 1984.

[13] A. J. van der Schaft, "Symmetries in optimal control," *SIAM J. Control Optim.*, vol. 25, no. 2, pp. 245–259, 1987.

[14] A. Echeverría-Enríquez, J. Marín-Solano, M. C. M. Lecanda, and N. Román-Roy, "Geometric reduction in optimal control theory with symmetries," *Rep. Math. Phys.*, vol. 52, no. 1, pp. 89–113, 2003.

[15] T. Ohsawa, "Symmetry reduction of optimal control systems and principal connections," *SIAM J. Control Optim.*, vol. 51, no. 1, pp. 96–120, 2013.

[16] M. de León, J. Cortés, D. M. de Diego, and S. Martínez, "General symmetries in optimal control," *Rep. Math. Phys.*, vol. 53, no. 1, pp. 55–78, 2004.
[17] E. Martínez, "Reduction in optimal control theory," *Rep. Math. Phys.*, vol. 53, no. 1, pp. 79–90, 2004.
[18] P. S. Krishnaprasad, "Optimal control and Poisson reduction," *Institute for Systems Research Technical Report*, T.R. 93-97, 1993.
[19] J. Biggs, W. Holderbaum, and V. Jurdjevic, "Singularities of optimal control problems on some 6-D Lie groups," *IEEE Trans. Autom. Control*, vol. 52, no. 6, pp. 1027–1038, 2007.
[20] G. Walsh, R. Montgomery, and S. Sastry, "Optimal path planning on matrix Lie groups," in *Proc. 33rd IEEE Conf. Decision Control*, 1994, pp. 1258–1263.
[21] N. E. Leonard and P. S. Krishnaprasad, "Motion control of drift-free, left-invariant systems on lie groups," *IEEE Trans. Autom. Control*, vol. 40, no. 9, pp. 1539–1554, 1995.
[22] Y. Sachkov, "Conjugate points in the Euler elastic problem," *J. Dyn. Control Syst.*, vol. 14, no. 3, pp. 409–439, 2008.
[23] J. Zhang and S. Sastry, "Aircraft conflict resolution: Lie-Poisson reduction for game on SE(2)," in *Proc. 40th IEEE Conf. Decision Control*, 2001, pp. 1663–1668.
[24] E. W. Justh and P. S. Krishnaprasad, "Optimality, reduction and collective motion," *Proc. R. Soc. A*, vol. 471, 20140606, 2015.
[25] U. Boscain and P. Mason, "Time minimal trajectories for a spin 1/2 particle in a magnetic field," *J. Math. Phys.*, vol. 47, 062101, 2006.
[26] W. S. Koon and J. E. Marsden, "Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction," *SIAM J. Control Optim.*, vol. 35, no. 3, pp. 901–929, 1997.
[27] L. Colombo and D. M. de Diego, "Optimal control of underactuated mechanical systems with symmetries," *Discret. Contin. Dyn. S.*, Supplement, pp. 149–158, 2013.
[28] L. Colombo and D. M. de Diego, "Higher-order variational problems on Lie groups and optimal control applications," *J. Geom. Mech.*, vol. 6, no. 4, pp. 451–478, 2014.
[29] F. Gay-Balmaz, D. D. Holm, and T. S. Ratiu, "Higher order Lagrange-Poincaré and Hamilton-Poincaré reductions," *Bull. Braz. Math. Soc. (N.S.)*, vol. 42, no. 4, pp. 579–606, 2011.
[30] H. Cendra, J. E. Marsden, and T. S. Ratiu, *Lagrangian Reduction by Stages*. Memoirs of the American Mathematical Society, 2001.
[31] R. Gupta, "Analytical and numerical methods for optimal control problems on manifolds and Lie groups," Ph.D. dissertation, University of Michigan, 2016.
[32] F. Gay-Balmaz and T. S. Ratiu, "Clebsch optimal control formulation in mechanics," *J. Geom. Mech.*, vol. 3, no. 1, pp. 41–79, 2011.
[33] L. Bates and F. Fassò, "The conjugate locus for the Euler top I. The axisymmetric case," *Int. Math. Forum.*, vol. 2, no. 43, pp. 2109–2139, 2007.
[34] U. Boscain and F. Rossi, "Invariant Carnot-Caratheodory metrics on S^3, $SO(3)$, $SL(2)$, and lens spaces," *SIAM J. Control Optim.*, vol. 47, no. 4, pp. 1851–1878, 2008.
[35] D. Barilari and L. Rizzi, "Comparison theorems for conjugate points in sub-Riemannian geometry," *ESAIM Contr. Optim. Ca.*, vol . 22, no. 2, pp. 439–472, 2016.
[36] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*. New York, NY: Wiley, 1962.
[37] K. Spindler, "Optimal control on Lie groups with applications to attitude control," *Math. Control Signals Syst.*, vol. 11, no. 3, pp. 197–219, 1998.
[38] K. Suzuki, Y. Watanabe, and T. Kambe, "Geometrical analysis of free rotation of a rigid body," *J. Phys. A: Math. Gen.*, vol. 31, pp. 6073–6080, 1998.
[39] H. Cendra, J. E. Marsden, S. Pekarsky, and T. S. Ratiu, "Variational principles for Lie-Poisson and Hamilton-Poincaré equations," *Mosc. Math. J.*, vol. 3, no. 3, pp. 833–867, 2003.
[40] D. Lewis, T. Ratiu, J. C. Simo, and J. E. Marsden, "The heavy top: A geometric treatment," *Nonlinearity*, vol. 5, no. 1, pp. 1–48, 1992.
[41] W. Thomson and P. G. Tait, *Treatise on Natural Philosophy: Part 1*. Cambridge, U.K.: Cambridge University Press, 1912.
[42] J. G. Papastavridis, "On a Lagrangean action based kinetic instability theorem of Kelvin and Tait," *Int. J. Eng. Sci.*, vol. 24, no. 1, pp. 1–17, 1986.
[43] A. Bloch, P. S. Krishnaprasad, J. E. Marsden, and R. S. Ratiu, "The Euler-Poincaré equations and double bracket dissipation," *Commun. Math. Phys.*, vol. 175, no. 1, pp. 1–42, 1996.

Andy Borum (S'14) received the B.S. degree in engineering science and mechanics and the B.S. degree in mathematics from Virginia Tech, Blacksburg, VA, in 2012 and the M.S. degree in aerospace engineering from the University of Illinois at Urbana-Champaign, in 2014.

He is currently a National Science Foundation Graduate Research Fellow in the Department of Aerospace Engineering, University of Illinois at Urbana-Champaign.

Timothy Bretl (M'16) received the B.S. degree in engineering and the B.A. degree in mathematics from Swarthmore College, Swarthmore, PA, in 1999 and the M.S. and Ph.D. degrees in aeronautics and astronautics from Stanford University, Stanford, CA, in 2000 and 2005, respectively.

Subsequently, he was a Postdoctoral Fellow in the Department of Computer Science, Stanford University. Since 2006, he has been with the University of Illinois at Urbana-Champaign, where he is an Associate Professor of Aerospace Engineering and a Research Associate Professor in the Coordinated Science Laboratory. Dr. Bretl received the National Science Foundation Faculty Early Career Development Award in 2010.