

Geometric optimal control for symmetry breaking cost functions

Andy D. Borum and Timothy Bretl

Abstract—We consider an optimal control problem defined on a Lie group whose associated Hamiltonian function is left-invariant under the action of a subgroup of the Lie group. Necessary conditions for optimality are derived using Lie-Poisson reduction for semidirect products, which allows us to study the Hamiltonian system in a space of lower dimension. Our main contribution is a reduced sufficient condition for optimality that relies on the nonexistence of conjugate points. We derive coordinate formulae for computing conjugate points in the reduced Hamiltonian system, and we relate these conjugate points to local optimality in the original optimal control problem. These conditions are applied to an optimal control problem that can be used to model either a kinematic airplane or a Kirchhoff elastic rod in a gravitational field.

I. INTRODUCTION

Consider an optimal control problem whose state evolves on a smooth manifold. The necessary conditions for optimality provided by Pontryagin's maximum principle [1] relate solutions of this optimal control problem to integral curves of a Hamiltonian vector field. When the associated Hamiltonian function has certain symmetry properties, tools from Poisson geometry can be used to study the Hamiltonian system in a space of reduced dimension [2]. These symmetry group reduction techniques are widely applied to Hamiltonian systems in classical mechanics [3].

When the Hamiltonian system evolves on the cotangent bundle of a Lie group and the Hamiltonian function is invariant under the left action of the Lie group, Lie-Poisson reduction can be applied. The dynamics of the system can then be studied by considering a reduced Hamiltonian system which evolves on the dual Lie algebra of the Lie group. Lie-Poisson reduction has been used to derive reduced necessary conditions for left-invariant optimal control problems [4], [5], [6], [7]. Less focus has been given to applying symmetry reduction techniques to sufficient conditions for optimality. A reduced test for conjugate points in left-invariant optimal control problems is given in [4].

In some systems, the Hamiltonian function is not left-invariant under the action of the entire Lie group, but is invariant under the action of a subgroup of the Lie group. This issue can sometimes be resolved by embedding the problem in a larger semidirect product space in which the system becomes left-invariant [8], [9], [10]. A classic example of such a system is the heavy spinning top. Other systems to which this method has been applied include compressible

fluids, magnetohydrodynamics, elasticity, and plasma physics [9]. In this paper, we apply semidirect product reduction to necessary and sufficient conditions for optimal control problems on Lie groups. After applying Pontryagin's maximum principle to the optimal control problem, we assume that the Hamiltonian function is left-invariant under the action of a subgroup of the Lie group. Applying Lie-Poisson reduction for semidirect products to these optimal control problems reduces the associated Hamiltonian system, which originally evolved on the cotangent bundle of the Lie group, to the dual Lie algebra of a semidirect product.

Our main contribution is a sufficient condition for optimality which relies on the nonexistence of conjugate points. We provide coordinate formulae for computing conjugate points by establishing non-degeneracy of the exponential map of the reduced Hamiltonian system. We show that the absence of conjugate points in the reduced system implies local optimality in the original system. While geometric statements of necessary and sufficient conditions for optimality (such as those stated in Section II) are, in principle, all we need to find optimal solutions, they do not provide coordinate formulae for computing solutions. One advantage of working in the reduced space is that optimal trajectories can be found by solving a system of ordinary differential equations, and conjugate points can be computed by solving a system of matrix differential equations. After stating the necessary and sufficient conditions, we apply them to a geometric optimal control problem on $SE(3)$ with broken symmetry. This optimal control problem can be used to model a kinematic airplane [11], [12] or a Kirchhoff elastic rod [4], [13] in a gravitational field.

We review the general theory of optimal control on manifolds in Section II. In Section III, Lie-Poisson reduction for semidirect products is applied to the necessary conditions for optimality provided by Pontryagin's maximum principle. In Section IV, we derive a test for conjugate points in the reduced system and relate this test for optimality to the original system. The applications described above are treated in Section V, and closing remarks are given in Section VI.

II. GEOMETRIC OPTIMAL CONTROL

In this section, we review the framework for characterizing solutions of geometric optimal control problems. We begin by recalling a few facts about smooth manifolds. Let M be a smooth manifold, $C^\infty(M)$ be the set of all smooth real-valued functions on M , and $\mathfrak{X}(M)$ be the set of all smooth vector fields on M . Denote the action of a tangent vector $v \in T_m M$ on a function $f \in C^\infty(M)$ by $v \cdot f$ and the action of a tangent covector $w \in T_m^* M$ on a tangent

This work was supported by the NSF under Grant No. IIS-1320519. The work of A. Borum was supported by the NSF-GRFP under Grant No. DGE-1144245.

The authors are with the Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA, {borum2, tbretl}@illinois.edu

vector $v \in T_m M$ by $\langle w, v \rangle$. The action of a vector field $X \in \mathfrak{X}(M)$ on a function $f \in C^\infty(M)$ produces the function $X[f] \in C^\infty(M)$ that satisfies

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

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

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

for all $f \in C^\infty(M)$. If N is a smooth manifold and $F: M \rightarrow N$ is a smooth map, then the pushforward and pullback of F at $m \in M$ are the maps $T_m F: T_m M \rightarrow T_{F(m)} N$ and $T_m^* F: T_{F(m)}^* N \rightarrow T_m^* M$, respectively, that satisfy

$$T_m F(v) \cdot f = v \cdot (f \circ F), \quad \langle T_m^* F(w), v \rangle = \langle w, T_m F(v) \rangle$$

for all $v \in T_m M$, $w \in T_{F(m)}^* N$ and $f \in C^\infty(N)$. We say F is degenerate at $m \in M$ if there exists non-zero $v \in T_m M$ such that $T_m F(v) = 0$. It is equivalent that the Jacobian matrix of any coordinate representation of F at m has zero determinant. The Poisson bracket generated by the canonical symplectic form on T^*M is

$$\{\cdot, \cdot\}: C^\infty(T^*M) \times C^\infty(T^*M) \rightarrow C^\infty(T^*M)$$

The Hamiltonian vector field of $H \in C^\infty(T^*M)$ is the unique vector field $X_H \in \mathfrak{X}(T^*M)$ that satisfies

$$X_H[K] = \{K, H\}$$

for all $K \in C^\infty(T^*M)$. We use $\pi: T^*M \rightarrow M$ to denote the projection map $\pi(w, m) = m$ for all $w \in T_m^* M$.

We now state necessary and sufficient conditions for optimal control problems on smooth manifolds. Assume $g: M \times U \rightarrow \mathbb{R}$ and $f: M \times U \rightarrow TM$ are smooth maps where $U \subset \mathbb{R}^m$ for some $m > 0$. Consider the optimal control problem

$$\begin{aligned} & \text{minimize}_{q, u} \int_0^T g(q(t), u(t)) dt \\ & \text{subject to } \dot{q}(t) = f(q(t), u(t)) \text{ for all } t \in [0, T] \\ & \quad q(0) = q_0, \quad q(T) = q_1 \end{aligned} \quad (1)$$

for some fixed $T > 0$, where q_0 and $q_1 \in M$ and $(q, u): [0, T] \rightarrow M \times U$. Define the parameterized Hamiltonian function $\hat{H}: T^*M \times \mathbb{R} \times U \rightarrow \mathbb{R}$ by

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

where $p \in T_q^* M$. The following theorem is a geometric statement of Pontryagin's maximum principle [1].

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

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

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

Proof: See Theorem 12.10 in [14]. \blacksquare

We call the integral curve (p, q) in Theorem 1 an abnormal extremal when $k = 0$ and a normal extremal otherwise. When $k \neq 0$, we may simply 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.

Theorem 2: (Sufficient Conditions) Suppose $(p, q): [0, T] \rightarrow T^*M$ is a normal extremal of (1). Define $H \in C^\infty(M)$ by

$$H(p, q) = \max_{u \in U} \hat{H}(p(t), q(t), 1, u) \quad (3)$$

assuming the maximum exists and $\partial^2 \hat{H} / \partial u^2 < 0$. Define $u: [0, T] \rightarrow U$ so $u(t)$ is the unique maximizer of (3) at $(p(t), q(t))$. Assume that X_H is complete and that there exists no other integral curve (p', q') of X_H satisfying $q(t) = q'(t)$ for all $t \in [0, T]$. Let $\varphi: \mathbb{R} \times T^*M \rightarrow T^*M$ be the flow of X_H and define the endpoint map $\phi_t: T_{q(0)}^* M \rightarrow M$ by $\phi_t(w) = \pi \circ \varphi(t, w, q(0))$. Then (q, u) is a local optimum of (1) if and only if there exists no $t \in (0, T]$ for which ϕ_t is degenerate at $p(0)$.

Proof: See Theorem 21.8 in [14]. \blacksquare

III. LIE GROUPS, SEMIDIRECT PRODUCTS, AND REDUCTION OF NECESSARY CONDITIONS

While the geometric statements of necessary and sufficient conditions in Theorems 1 and 2 are, in principle, all we need to find solutions of the optimal control problem (1), it is not clear yet how to compute the integral curves (p, q) or how to establish non-degeneracy of the endpoint map ϕ_t . Coordinate formulae for performing these computations are provided in [4] in the case when M is a Lie group G and the Hamiltonian function (2) is left-invariant under the action of G . In this section, we consider the case when the Hamiltonian is left-invariant with respect to a subgroup of G . We then apply Lie-Poisson reduction for semidirect products (Theorem 3.4 in [9]) to the necessary conditions given by Theorem 1.

We begin by recalling some notation for Lie groups and semidirect products. Let G be a Lie group with identity element $e \in G$. Let $\mathfrak{g} = T_e G$ and $\mathfrak{g}^* = T_e^* G$. For any $q \in G$, define the left translation map $L_q: G \rightarrow G$ by

$$L_q(r) = qr$$

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

$$H(T_r^* L_q(w), r) = H(w, s) \quad (4)$$

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 $[\cdot, \cdot]: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ by

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

for all $\zeta, \eta \in \mathfrak{g}$. For any $\zeta \in \mathfrak{g}$, the adjoint operator $\text{ad}_\zeta : \mathfrak{g} \rightarrow \mathfrak{g}$ and the coadjoint operator $\text{ad}_\zeta^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ satisfy

$$\text{ad}_\zeta(\eta) = [\zeta, \eta] \quad \langle \text{ad}_\zeta^*(\mu), \eta \rangle = \langle \mu, \text{ad}_\zeta(\eta) \rangle$$

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

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

for all $\delta\mu \in \mathfrak{g}^*$. Let $\{X_1, \dots, X_n\}$ be a basis for \mathfrak{g} and let $\{P_1, \dots, P_n\}$ be the dual basis for \mathfrak{g}^* that satisfies

$$\langle P_i, X_j \rangle = \delta_{ij}$$

for $i, j \in \{1, \dots, n\}$, where δ_{ij} 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 $i, j \in \{1, \dots, n\}$, define the structure constants $C_{ij}^k \in \mathbb{R}$ for our choice of basis by

$$[X_i, X_j] = \sum_{k=1}^n C_{ij}^k X_k \quad (5)$$

Now let V be a vector space and let $\rho : G \rightarrow \text{Aut}(V)$ 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 $\rho(g) : V \rightarrow V$ 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

$$\rho_*(g) = [\rho(g^{-1})]^* \quad \rho^*(g) = [\rho(g)]^*$$

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

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

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

$$G_\chi = \{g \in G | \rho_*(g)\chi = \chi\}$$

Let $S = G \times_\rho V$ 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 $\mathfrak{s} = \mathfrak{g} \times_{\rho'} V$ 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

$$T_{(q,u)}^*(w, s, v, \chi) = (T_{qs}^* L_{q^{-1}}(w), qs, u + \rho(q)v, \rho_*(q)\chi)$$

for all $u, v \in V$, $\chi \in V^*$, $w \in T_s^*G$, and $q, s \in G$ [9].

We now revisit the statement of necessary conditions in Theorem 1 in the case when the Hamiltonian function

(2) depends smoothly on a parameter $\chi_0 \in V^*$ and is left-invariant under the action of G_{χ_0} on T^*G , i.e., (4) holds when $q \in G_{\chi_0}$. We denote the Hamiltonian by $H_{\chi_0} : T^*G \rightarrow \mathbb{R}$ to note the dependence on $\chi_0 \in V^*$. The procedure for applying Lie-Poisson reduction to such Hamiltonian systems is to consider the Hamiltonian function $H : T^*S \rightarrow \mathbb{R}$ defined by $H(p, q, v, \chi) = H_\chi(p, q)$, where $T^*S = T^*G \times V \times V^*$. Since $H_\chi(p, q)$ 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 [8]. We then show that $H : T^*S \rightarrow \mathbb{R}$ is left-invariant under the action of S , i.e. we show that

$$H(T_r^* L_q(w), r, v, \rho^*(q)\chi) = H(w, s, v, \chi) \quad (6)$$

for all $v \in V$, $\chi \in V^*$, $w \in T_s^*G$, and $q, r, s \in G$ satisfying $s = L_q(r)$. Note that if (6) holds and $q \in G_{\chi_0}$, then $\chi_0 = \rho^*(q)\chi_0$ (since G_{χ_0} is a subgroup) and

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

for all $w \in T_s^*G$, $r, s \in G$, and $q \in G_{\chi_0}$ satisfying $s = L_q(r)$. Therefore (6) implies that H_{χ_0} is left-invariant under the action of G_{χ_0} on T^*G .

If (6) holds, then the family of Hamiltonians $\{H_\chi | \chi \in V^*\}$ induces a reduced Hamiltonian h on \mathfrak{s}^* , and the existence of an integral curve (μ, χ) in \mathfrak{s}^* implies the existence of a corresponding integral curve (p, q) of $X_{H_{\chi_0}}$ in the cotangent bundle T^*G . This result is important because \mathfrak{s}^* is a vector space and so (μ, χ) can be readily computed by solving a system of ordinary differential equations.

Theorem 3: (Semidirect Product Reduction of Necessary Conditions) Suppose the smooth time-varying Hamiltonian function $H_{\chi_0} : T^*G \times [0, T] \rightarrow \mathbb{R}$ depends smoothly on the parameter $\chi_0 \in V^*$ and is left-invariant under the action of G_{χ_0} on T^*G for all $t \in [0, T]$. In addition, assume that the Hamiltonian is left-invariant under the action of S when defined as a function on $T^*S \times [0, T]$ for all $t \in [0, T]$. The family of Hamiltonians $\{H_\chi | \chi \in V^*\}$ induces a Hamiltonian function h on $\mathfrak{s}^* \times [0, T]$, defined by

$$h(T_e^* L_q(p), \rho^*(q)\chi, t) = H_\chi(p, q, t) \quad (7)$$

Given $p_0 \in T_{q_0}^*G$, let $(\mu, \chi) : [0, T] \rightarrow \mathfrak{s}^*$ be the solution of

$$\dot{\mu} = \text{ad}_{\delta h / \delta \mu}^*(\mu) - (\rho'_{\delta h / \delta \chi})^* \chi \quad (8)$$

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

with initial conditions $\mu(0) = T_e^* L_{q_0}(p_0)$ and $\chi(0) = \rho^*(q_0)\chi_0$, and where $\rho'_{\delta h / \delta \chi} : \mathfrak{g} \rightarrow V$ is given by $\rho'_{\delta h / \delta \chi}(\zeta) = \rho'(\zeta) \frac{\delta h}{\delta \chi}$. The integral curve $(p, q) : [0, T] \rightarrow T^*G$ of $X_{H_{\chi_0}}$ with initial condition $p(0) = p_0$ satisfies

$$p(t) = T_{q(t)}^* L_{q(t)^{-1}}(\mu(t))$$

for all $t \in [0, T]$, where q is the solution of

$$\dot{q} = X_{\delta h / \delta \mu}(q)$$

with initial condition $q(0) = q_0$. The evolution of $\chi \in V^*$ is given by

$$\chi(t) = \rho^*(q(t))\chi_0 \quad (9)$$

Proof: See Theorem 3.4 of [9]. ■

IV. SEMIDIRECT PRODUCT REDUCTION OF SUFFICIENT CONDITIONS

In this section, we revisit our statement of sufficient conditions for (1) in the case when the Hamiltonian in (2) satisfies the conditions in Theorem 3. We will lose some of the generality of the previous section and now restrict ourselves to dealing only with matrix Lie groups. Assume that the vector space V is Euclidean space with dimension such that multiplication qv is well defined for all matrices $q \in G$ and vectors $v \in V$, and let the representation $\rho : G \rightarrow \text{Aut}(V)$ be defined by $\rho(q)(v) = qv$. We will denote the actions of G and \mathfrak{g} on V and V^* by concatenation. We will also assume that the Hamiltonian in (2) has the form

$$H_{\chi_0}(p, q) = K(p, q) + U(\chi_0, q) \quad (10)$$

This happens when f in (1) is independent of χ_0 and g has the form $g(q, u) = g_1(q, u) + g_2(q, \chi_0)$, where g_1 is independent of χ_0 . Thus the symmetry breaking term appears in the cost function. Note that, in this case, $\delta h/\delta \mu$ is independent of χ and $\delta h/\delta \chi$ is independent of μ . (This fact will be used in the proof of Theorem 4). We begin with two lemmas that will be required to show our main result.

Lemma 1: Let $q : U \rightarrow G$ be a smooth map, where $U \subset \mathbb{R}^2$ is simply connected. Denote its partial derivatives $\zeta : U \rightarrow \mathfrak{g}$ and $\eta : U \rightarrow \mathfrak{g}$ by

$$\begin{aligned} \zeta(t, \epsilon) &= T_{q(t, \epsilon)} L_{q(t, \epsilon)^{-1}} \left(\frac{\partial q(t, \epsilon)}{\partial t} \right) \\ \eta(t, \epsilon) &= T_{q(t, \epsilon)} L_{q(t, \epsilon)^{-1}} \left(\frac{\partial q(t, \epsilon)}{\partial \epsilon} \right) \end{aligned} \quad (11)$$

Then

$$\frac{\partial \zeta}{\partial \epsilon} - \frac{\partial \eta}{\partial t} = [\zeta, \eta] \quad (12)$$

Conversely, if there exist smooth maps ζ and η satisfying (12), then there exists a smooth map q satisfying (11).

Proof: See Proposition 5.1 in [15]. ■

Lemma 2: Let $\alpha, \beta, \gamma \in \mathfrak{g}$ and suppose $\gamma = [\alpha, \beta]$. Then

$$\gamma_k = \sum_{r=1}^n \sum_{s=1}^n \alpha_r \beta_s C_{rs}^k$$

Proof: This is obtained from the definition in (5). ■

Theorem 4 provides coordinate formulae to test the non-degeneracy of the endpoint map ϕ_t that was defined geometrically in Theorem 2. These formulae can be evaluated by solving a system of linear, time-varying matrix differential equations, something that is easy to do using modern numerical methods.

Theorem 4: (Semidirect Product Reduction of Sufficient Conditions) Suppose that $H_{\chi_0} \in C^\infty(T^*G)$ satisfies the conditions in Theorem 3, has the form given in (10), and that $X_{H_{\chi_0}}$ is complete. Let h be the Hamiltonian function on \mathfrak{s}^* induced by the family of Hamiltonians $\{H_\chi | \chi \in V^*\}$

and let $\varphi : \mathbb{R} \times T^*G \rightarrow T^*G$ be the flow of $X_{H_{\chi_0}}$. Given $q_0 \in G$, define the endpoint map $\phi_t : T_{q_0}^*G \rightarrow G$ by $\phi_t(p) = \pi \circ \varphi(t, p, q_0)$. Given $p_0 \in T_{q_0}^*G$, let $a \in \mathbb{R}^n$ be the coordinate representation of $T_e^*L_{q_0}(p_0)$, i.e.

$$T_e^*L_{q_0}(p_0) = \sum_{i=1}^n a_i P_i$$

Solve the ordinary differential equations

$$\dot{\mu}_i = - \sum_{j=1}^n \sum_{k=1}^n C_{ij}^k \frac{\delta h}{\delta \mu_j} \mu_k - \chi \left(X_i \frac{\delta h}{\delta \chi} \right) \quad (13)$$

with initial conditions $\mu_i(0) = a_i$ for $i \in \{1, \dots, n\}$, and where χ satisfies

$$\chi(t)(v) = \chi_0(\phi_t(p_0)v) \quad (14)$$

for all $v \in V$. Define the time-varying matrices **F**, **G**, **H**, **K**, and **L** $\in \mathbb{R}^{n \times n}$ by

$$\begin{aligned} [\mathbf{F}]_{ij} &= - \frac{\partial}{\partial \mu_j} \sum_{r=1}^n \sum_{s=1}^n C_{ir}^s \frac{\delta h}{\delta \mu_r} \mu_s, & [\mathbf{G}]_{ij} &= \frac{\partial}{\partial \mu_j} \frac{\delta h}{\delta \mu_i}, \\ [\mathbf{H}]_{ij} &= - \sum_{r=1}^n \frac{\delta h}{\delta \mu_r} C_{rj}^i, & [\mathbf{K}]_{ij} &= \chi \left(X_j X_i \frac{\delta h}{\delta \chi} \right), \\ [\mathbf{L}]_{ij} &= \sum_{k=1}^n \chi \left(X_i \frac{\partial}{\partial \chi_k} \frac{\delta h}{\delta \chi} \right) \chi \left(X_j^{(k)} \right) \end{aligned}$$

where $X_j^{(k)}$ denotes the k^{th} column of the matrix X_j . Solve the (linear, time-varying) matrix differential equations

$$\dot{\mathbf{M}} = \mathbf{F}\mathbf{M} - (\mathbf{K} + \mathbf{L})\mathbf{J} \quad \dot{\mathbf{J}} = \mathbf{G}\mathbf{M} + \mathbf{H}\mathbf{J} \quad (15)$$

with initial conditions $\mathbf{M}(0) = \mathbf{I}$ and $\mathbf{J}(0) = 0$. The endpoint map ϕ_t is degenerate at p_0 if and only if $\det(\mathbf{J}(t)) = 0$ for some $t \in (0, T]$.

Proof: Define the smooth map $\sigma : \mathbb{R}^n \rightarrow T_{q_0}^*G$ by

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

This same expression defines $\sigma : \mathbb{R}^n \rightarrow T_{p_0}(T_{q_0}^*G)$ 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 $T_{\sigma(a)}\phi_t(\sigma(s)) = 0$. Define the smooth map $q : [0, T] \times \mathbb{R}^n \rightarrow G$ by $q(t, a) = \phi_t \circ \sigma(a)$. Noting that

$$\frac{\partial q(t, a)}{\partial a_j} = T_{\sigma(a)}\phi_t \left(T_{q_0}^*L_{q_0^{-1}}(P_j) \right)$$

for $j \in \{1, \dots, n\}$, we have

$$T_{\sigma(a)}\phi_t(\sigma(s)) = \sum_{j=1}^n s_j \frac{\partial q(t, a)}{\partial a_j}$$

By left translation, $T_{\sigma(a)}\phi_t(\sigma(s)) = 0$ if and only if

$$0 = \sum_{j=1}^n s_j T_{q(t, a)} L_{q(t, a)^{-1}} \left(\frac{\partial q(t, a)}{\partial a_j} \right) \quad (16)$$

For each $j \in \{1, \dots, n\}$, let

$$\eta^j(t, a) = T_{q(t, a)} L_{q(t, a)^{-1}} \left(\frac{\partial q(t, a)}{\partial a_j} \right) \quad (17)$$

Define $\mathbf{J} : [0, T] \rightarrow \mathbb{R}^{n \times n}$ so that $\mathbf{J}(t)$ has entries

$$[\mathbf{J}]_{ij} = \eta_i^j(t, a) \quad (18)$$

i.e. the j th column of $\mathbf{J}(t)$ is the coordinate representation of $\eta^j(t, a)$ with respect to $\{X_1, \dots, X_n\}$. Then, (16) 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 $\det(\mathbf{J}(t)) = 0$.

It remains to show that $\mathbf{J}(t)$ can be computed as described in the theorem. Taking $\mu_1(t), \dots, \mu_n(t)$ as coordinates of $\mu(t)$, it is easy to verify that (see [2])

$$\text{ad}_{\delta h / \delta \mu}^* \mu = - \sum_{j=1}^n \sum_{k=1}^n C_{ij}^k \frac{\delta h}{\delta \mu_j} \mu_k \quad (19)$$

We also have that $(\rho'_{\delta h / \delta \chi})^* \chi \in \mathfrak{g}^*$ and for each $\zeta \in \mathfrak{g}$

$$(\rho'_{\delta h / \delta \chi})^* \chi(\zeta) = \chi \left(\rho'(\zeta) \frac{\delta h}{\delta \chi} \right) = \chi \left(\zeta \frac{\delta h}{\delta \chi} \right) \quad (20)$$

Therefore, since $\{X_1, \dots, X_n\}$ is a basis for \mathfrak{g} , the i th component of $(\rho'_{\delta h / \delta \chi})^* \chi$ in terms of the dual basis $\{P_1, \dots, P_n\}$ is given by

$$\chi \left(X_i \frac{\delta h}{\delta \chi} \right) \quad (21)$$

Using (19)-(21), we see that (8) and (13) are equivalent. Also, noting that $\phi_t(p_0) = q(t)$, we see that (9) and (14) are equivalent. We extend each coordinate function to $\mu_i : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$, so $\mu_i(t, a)$ solves (13) with initial condition $\mu_i(0, a) = a_i$. Define $\mathbf{M} : [0, T] \rightarrow \mathbb{R}^{n \times n}$ by

$$[\mathbf{M}(t)]_{ij} = \frac{\partial \mu_i}{\partial a_j}$$

The computations performed in Appendix 1 show that $[\mathbf{M}(t)]_{ij}$ can be computed by solving

$$[\dot{\mathbf{M}}]_{ij} = \sum_{k=1}^n [\mathbf{F}]_{ik} [\mathbf{M}]_{kj} - \sum_{k=1}^n ([\mathbf{K}]_{ik} + [\mathbf{L}]_{ik}) [\mathbf{J}]_{kj}$$

It is clear that $[\mathbf{M}(0)]_{ij} = \delta_{ij}$, so we have verified the first equation in (15). Next, define

$$\zeta(t, a) = T_{q(t, a)} L_{q(t, a)^{-1}} \left(\frac{\partial q(t, a)}{\partial t} \right)$$

We have

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

from Lemma 1 and Theorem 3. We write this equation in coordinates by application of Lemma 2.

$$\begin{aligned} [\dot{\mathbf{J}}]_{ij} &= \dot{\eta}_i^j \\ &= \sum_{k=1}^n \left(\frac{\partial}{\partial \mu_k} \frac{\delta h}{\delta \mu_i} \right) \frac{\partial \mu_k}{\partial a_j} + \sum_{k=1}^n \left(- \sum_{r=1}^n \frac{\delta h}{\delta \mu_r} C_{rk}^i \right) \eta_k^j \\ &= \sum_{k=1}^n [\mathbf{G}]_{ik} [\mathbf{M}]_{kj} + \sum_{k=1}^n [\mathbf{H}]_{ik} [\mathbf{J}]_{kj} \end{aligned}$$

It is clear that $[\mathbf{J}(0)]_{ij} = 0$, so we have verified (15). ■

V. THE KINEMATIC AIRPLANE AND THE HEAVY KIRCHHOFF ELASTIC ROD

In this section, we consider a geometric optimal control problem which can be used to model two different systems. First consider a kinematic airplane that flies at a constant speed [11], [12]. Three control inputs are used to yaw, pitch and roll the aircraft. The position and orientation of the airplane at time t is described by an element of the Lie group $SE(3)$, which has the Lie algebra $\mathfrak{se}(3)$ and dual Lie algebra $\mathfrak{se}^*(3)$. Consider the basis $\{X_1, \dots, X_6\}$ of $\mathfrak{se}(3)$ given by

$$\begin{aligned} X_1 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & X_2 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & X_3 &= \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ X_4 &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & X_5 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & X_6 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

If the aircraft flies forward at a constant unit speed for time T , the trajectory of the aircraft is given by a continuous map $q : [0, T] \rightarrow SE(3)$ which satisfies

$$\dot{q} = q(u_1 X_1 + u_2 X_2 + u_3 X_3 + X_4)$$

for all $t \in [0, T]$, where $u : [0, T] \rightarrow U = \mathbb{R}^3$ is the control input. In [12], the problem of finding a trajectory connecting two given points in $SE(3)$ and locally minimizing the sum of the squared control inputs was considered. Using this cost function, the resulting Hamiltonian function is left-invariant. We consider a similar cost function, however we add a term which accounts for gravity. Thus, we now want to find a trajectory that minimizes a combination of the sum of the squared control inputs and the vertical height of the aircraft. Therefore, we consider the optimal control problem

$$\begin{aligned} &\text{minimize}_{q, u} \int_0^T \left(\frac{1}{2} \sum_{i=1}^3 c_i u_i^2 + W \chi_0(d(q)) \right) dt \\ &\text{subject to} \quad \dot{q} = q(u_1 X_1 + u_2 X_2 + u_3 X_3 + X_4) \\ &\quad q(0) = q_0 \quad q(T) = q_1 \end{aligned} \quad (22)$$

for some q_0 and $q_1 \in SE(3)$ and $T > 0$, where c_1, c_2 , and c_3 are constants, W is the weight of the aircraft, $\chi_0^T = [\bar{g} \ 0]^T \in \mathbb{R}^4$ ($\bar{g}^T \in \mathbb{R}^3$ is a unit vector pointing in the opposite direction of gravity), and $d : SE(3) \rightarrow \mathbb{R}^4$ maps the 4×4 matrix $q \in SE(3)$ to the last column of q , i.e., $\begin{bmatrix} R & b \\ 0 & 1 \end{bmatrix} \mapsto \begin{bmatrix} b \\ 1 \end{bmatrix}$, where $R \in SO(3)$ and $b \in \mathbb{R}^3$. If gravity points in the downward direction, we choose $\chi_0 = [0 \ 0 \ 1 \ 0]$. In the notation from Section 3, we have chosen $G = SE(3)$ and $V = \mathbb{R}^4$. The gravity term breaks the $SE(3)$ symmetry, and the methods used in [12] cannot be applied. However, this problem fits into the framework in Sections 3 and 4, and Theorems 3 and 4 can be used to find optimal trajectories.

This same optimal control problem models equilibrium configurations of a Kirchhoff elastic rod under the force of gravity. Here T is the length of the rod, c_1 is the torsional stiffness, c_2 and c_3 are the bending stiffnesses, and W is the weight of the rod per unit length. An analysis of a Kirchhoff rod without the effect of gravity is performed in [4], in which Theorems 1 and 2 from this paper are applied

to (22) (with the gravity term neglected) to derive necessary and sufficient conditions for a configuration of the rod to be a local minimum of the elastic potential energy.

We will now analyze (22) using the tools developed in this paper. Applying Theorem 1 gives that normal (q, u) correspond to integral curves of the Hamiltonian vector field $X_{H_{\chi_0}}$, where H_{χ_0} is defined by

$$\widehat{H}_{\chi_0}(p, q, k, u) = \langle p, q(u_1 X_1 + u_2 X_2 + u_3 X_3 + X_4) \rangle - k \left(\frac{c_1}{2} u_1^2 + \frac{c_2}{2} u_2^2 + \frac{c_3}{2} u_3^2 + W_{\chi_0}(d(q)) \right)$$

and

$$H_{\chi_0}(p(t), q(t), t) = \max_u \widehat{H}_{\chi_0}(p(t), q(t), 1, u)$$

This maximum is achieved when

$$u_i = c_i^{-1} \langle p, q X_i \rangle$$

for $i \in \{1, 2, 3\}$. This is indeed a maximum since $\partial^2 \widehat{H} / \partial u^2 = -\text{diag}(c_1, c_2, c_3) < 0$. The maximized Hamiltonian function is then

$$H_{\chi_0}(p, q) = \frac{1}{2} \sum_{i=1}^3 c_i^{-1} \langle p, q X_i \rangle^2 + \langle p, q X_4 \rangle - W_{\chi_0}(d(q))$$

Extending H_{χ_0} to be a function on $T^*SE(3) \times V \times V^*$ gives

$$H(p, q, v, \chi) = \frac{1}{2} \sum_{i=1}^3 c_i^{-1} \langle p, q X_i \rangle^2 + \langle p, q X_4 \rangle - W_{\chi}(d(q))$$

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

$$\begin{aligned} & H(T_r^* L_g(p), r, v, \rho^*(g)\chi) \\ &= \frac{1}{2} \sum_{i=1}^3 c_i^{-1} \langle T_r^* L_g(p), g^{-1} q X_i \rangle^2 \\ & \quad + \langle T_r^* L_g(p), g^{-1} q X_4 \rangle - W_{\rho^*(g)} \chi(d(g^{-1}q)) \\ &= \frac{1}{2} \sum_{i=1}^3 c_i^{-1} \langle p, g g^{-1} q X_i \rangle^2 \\ & \quad + \langle p, g g^{-1} q X_4 \rangle - W_{\chi}(g g^{-1} d(q)) \\ &= \frac{1}{2} \sum_{i=1}^3 c_i^{-1} \langle p, q X_i \rangle^2 + \langle p, q X_4 \rangle - W_{\chi}(d(q)) \\ &= H(p, q, v, \chi) \end{aligned}$$

So 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 translations orthogonal to the gravity vector and rotations around the gravity vector. As a consequence, we can apply Theorem 3. The reduced Hamiltonian on \mathfrak{s}^* is given by

$$h(\mu, \chi) = \frac{1}{2} \sum_{i=1}^3 c_i^{-1} \mu_i^2 + \mu_4 - W_{\chi_4}$$

Applying (13) gives

$$\begin{aligned} \dot{\mu}_1 &= u_3 \mu_2 - u_2 \mu_3 & \dot{\mu}_4 &= u_3 \mu_5 - u_2 \mu_6 + W_{\chi_1} \\ \dot{\mu}_2 &= \mu_6 + u_1 \mu_3 - u_3 \mu_1 & \dot{\mu}_5 &= u_1 \mu_6 - u_3 \mu_4 + W_{\chi_2} \\ \dot{\mu}_3 &= -\mu_5 + u_2 \mu_1 - u_1 \mu_2 & \dot{\mu}_6 &= u_2 \mu_4 - u_1 \mu_5 + W_{\chi_3} \end{aligned}$$

where $u_i = c_i^{-1} \mu_i$. Treating $\chi(t)$ as a row vector, (14) gives

$$\chi(t)^T = q(t)^T \chi_0^T$$

Solutions of (22) are obtained by finding an initial value of μ (which, from Theorem 4, is equivalent to finding $a \in \mathbb{R}^6$) which places $q(T)$ at q_1 . This can be done using a numerical shooting method. Such solutions are only guaranteed to be extrema of (22). It can be shown that (q, u) is abnormal if and only if $u_2 = u_3 = 0$, and that $\mu \in \mathfrak{g}^*$ (and hence $p \in T^*SE(3)$) is uniquely determined by (q, u) . It is also clear that $X_{H_{\chi_0}}$ is complete. Therefore, if $u_2 \neq 0$ and $u_3 \neq 0$, we may apply Theorem 4 to determine which extrema are actually local minima. Computing the matrices **F**, **G**, **H**, **K**, and **L** (and defining $c_{ij} = (c_i^{-1} - c_j^{-1})$) gives

$$\mathbf{F} = \begin{bmatrix} 0 & c_{32} \mu_3 & c_{32} \mu_2 & 0 & 0 & 0 \\ c_{13} \mu_3 & 0 & c_{13} \mu_1 & 0 & 0 & 1 \\ c_{21} \mu_2 & c_{21} \mu_1 & 0 & 0 & -1 & 0 \\ 0 & -c_2^{-1} \mu_6 & c_3^{-1} \mu_5 & 0 & c_3^{-1} \mu_3 & -c_2^{-1} \mu_2 \\ c_1^{-1} \mu_6 & 0 & -c_3^{-1} \mu_4 & -c_3^{-1} \mu_3 & 0 & c_1^{-1} \mu_1 \\ -c_1^{-1} \mu_5 & c_2^{-1} \mu_4 & 0 & c_2^{-1} \mu_2 & -c_1^{-1} \mu_1 & 0 \end{bmatrix}$$

$$\mathbf{G} = \text{diag}(c_1^{-1}, c_2^{-1}, c_3^{-1}, 0, 0, 0)$$

$$\mathbf{H} = \begin{bmatrix} 0 & c_3^{-1} \mu_3 & -c_2^{-1} \mu_2 & 0 & 0 & 0 \\ -c_3^{-1} \mu_3 & 0 & c_1^{-1} \mu_1 & 0 & 0 & 0 \\ c_2^{-1} \mu_2 & -c_1^{-1} \mu_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c_3^{-1} \mu_3 & -c_2^{-1} \mu_2 \\ 0 & 0 & 1 & -c_3^{-1} \mu_3 & 0 & c_1^{-1} \mu_1 \\ 0 & -1 & 0 & c_2^{-1} \mu_2 & -c_1^{-1} \mu_1 & 0 \end{bmatrix}$$

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & W_{\chi_3} & -W_{\chi_2} & 0 & 0 & 0 \\ -W_{\chi_3} & 0 & W_{\chi_1} & 0 & 0 & 0 \\ W_{\chi_2} & -W_{\chi_1} & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{L} = 0$$

After using a shooting method to find $a \in \mathbb{R}^6$ which places $q(T)$ at q_1 , (15) can be solved numerically with the initial conditions $\mathbf{M}(0) = \mathbf{I}$ and $\mathbf{J}(0) = 0$. If $\det(\mathbf{J}(t)) = 0$ for some $t \in (0, T]$, then the solution corresponding to this choice of $a \in \mathbb{R}^6$ is not a local minimum of (22). Note from (18) that $\mathbf{J}(T)$ provides the gradients of $q(T)$ with respect to $a \in \mathbb{R}^6$. These gradients can be used to improve the convergence of the shooting method described above.

VI. CONCLUSIONS

We have applied tools from Lie-Poisson reduction for semidirect products to geometric optimal control problems with broken symmetry. After deriving reduced necessary conditions for optimality, we provided a sufficient test for optimality based on conjugate points in the reduced system. While the general necessary and sufficient conditions in Section II are stated in terms of coordinate-free geometric results, the reduced necessary and sufficient conditions are stated in terms of coordinate formulae and rely on solutions of ordinary differential equations, which can be solved numerically. These results were then applied to a geometric optimal control problem which can be used to model either a kinematic airplane or a Kirchhoff elastic rod in a gravitational field. Semidirect product reduction is a special case of a more general reduction procedure known as reduction by stages [16]. The results in this paper could be extended by considering these more general approaches to symmetry group reduction.

APPENDIX 1

In this appendix, we show how to compute $[\mathbf{M}(t)]_{ij}$. Differentiating (13), we find

$$\begin{aligned} [\dot{\mathbf{M}}]_{ij} &= \frac{\partial}{\partial t} \frac{\partial \mu_i}{\partial a_j} = \frac{\partial}{\partial a_j} \frac{\partial \mu_i}{\partial t} \\ &= \frac{\partial}{\partial a_j} \left[\left(- \sum_{r=1}^n \sum_{s=1}^n C_{ir}^s \frac{\delta h}{\delta \mu_r} \mu_s \right) - \chi \left(X_i \frac{\delta h}{\delta \chi} \right) \right] \end{aligned}$$

Using (14), the equivalence of $\phi_t(p_0)$ and $q(t)$, and the fact that $\frac{\delta h}{\delta \mu}$ is independent of χ , we have

$$\begin{aligned} [\dot{\mathbf{M}}]_{ij} &= \sum_{k=1}^n - \frac{\partial}{\partial \mu_k} \left(\sum_{r=1}^n \sum_{s=1}^n C_{ir}^s \frac{\delta h}{\delta \mu_r} \mu_s \right) \frac{\partial \mu_k}{\partial a_j} \\ &\quad - \frac{\partial}{\partial a_j} \chi \left(X_i \frac{\delta h}{\delta \chi} \right) \\ &= \sum_{k=1}^n [\mathbf{F}]_{ik} [\mathbf{M}]_{kj} - \frac{\partial}{\partial a_j} \left(\chi_0 \left(q(t) X_i \frac{\delta h}{\delta \chi} \right) \right) \end{aligned} \quad (23)$$

Since χ_0 is a constant linear function on V , we can rewrite the second term in the last equation as

$$\chi_0 \left(\left(\frac{\partial}{\partial a_j} q(t) \right) X_i \frac{\delta h}{\delta \chi} \right) + \chi_0 \left(q(t) X_i \left(\frac{\partial}{\partial a_j} \frac{\delta h}{\delta \chi} \right) \right) \quad (24)$$

Now using (17), we have

$$\frac{\partial}{\partial a_j} q(t) = q(t) \left(\sum_{k=1}^n X_k \eta_k^j(t, a) \right) \quad (25)$$

Thus

$$\begin{aligned} &\chi_0 \left(\left(\frac{\partial}{\partial a_j} q(t) \right) X_i \frac{\delta h}{\delta \chi} \right) \\ &= \chi_0 \left(q(t) \left(\sum_{k=1}^n X_k \eta_k^j(t, a) \right) X_i \frac{\delta h}{\delta \chi} \right) \\ &= \chi \left(\sum_{k=1}^n X_k X_i \frac{\delta h}{\delta \chi} \eta_k^j(t, a) \right) \\ &= \sum_{k=1}^n \chi \left(X_k X_i \frac{\delta h}{\delta \chi} \right) \eta_k^j(t, a) \\ &= \sum_{k=1}^n [\mathbf{K}]_{ik} [\mathbf{J}]_{kj} \end{aligned} \quad (26)$$

where, in the second to last equality, we have used the fact that χ is linear. Next, since $\frac{\delta h}{\delta \chi}$ is independent of μ , we have

$$\frac{\partial}{\partial a_j} \frac{\delta h}{\delta \chi} = \sum_{k=1}^n \left(\frac{\partial}{\partial \chi_k} \frac{\delta h}{\delta \chi} \right) \frac{\partial \chi_k}{\partial a_j}$$

From (23)-(26), we have already shown that

$$\frac{\partial}{\partial a_j} \chi(v) = \sum_{k=1}^n \chi(X_k v) \eta_k^j(t, a)$$

for any $v \in V$. With $X_k^{(i)}$ denoting the i^{th} column of X_k , the i^{th} component of $\partial \chi / \partial a_j$ is

$$\frac{\partial \chi_i}{\partial a_j} = \sum_{k=1}^n \chi \left(X_k^{(i)} \right) \eta_k^j(t, a)$$

We now have

$$\begin{aligned} &\chi_0 \left(q(t) X_i \left(\frac{\partial}{\partial a_j} \frac{\delta h}{\delta \chi} \right) \right) \\ &= \chi_0 \left(q(t) X_i \left(\sum_{k=1}^n \left(\frac{\partial}{\partial \chi_k} \frac{\delta h}{\delta \chi} \right) \frac{\partial \chi_k}{\partial a_j} \right) \right) \\ &= \sum_{k=1}^n \chi \left(X_i \frac{\partial}{\partial \chi_k} \frac{\delta h}{\delta \chi} \right) \frac{\partial \chi_k}{\partial a_j} \\ &= \sum_{k=1}^n \chi \left(X_i \frac{\partial}{\partial \chi_k} \frac{\delta h}{\delta \chi} \right) \sum_{r=1}^n \chi \left(X_r^{(k)} \right) \eta_r^j(t, a) \\ &= \sum_{r=1}^n \left(\sum_{k=1}^n \chi \left(X_i \frac{\partial}{\partial \chi_k} \frac{\delta h}{\delta \chi} \right) \chi \left(X_r^{(k)} \right) \right) \eta_r^j(t, a) \\ &= \sum_{r=1}^n [\mathbf{L}]_{ir} [\mathbf{J}]_{rj} \end{aligned}$$

Combining these computations, we see that

$$[\dot{\mathbf{M}}]_{ij} = \sum_{k=1}^n [\mathbf{F}]_{ik} [\mathbf{M}]_{kj} - \sum_{k=1}^n ([\mathbf{K}]_{ik} + [\mathbf{L}]_{ik}) [\mathbf{J}]_{kj}$$

VII. ACKNOWLEDGMENTS

The authors would like to thank Rui Loja Fernandes and Ioan Marcu for many helpful discussions.

REFERENCES

- [1] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*. John Wiley, 1962.
- [2] J.E. Marsden and T.S. Ratiu, *Introduction to Mechanics and Symmetry*, Springer, New York, 1999.
- [3] R. Abraham and J.E. Marsden, *Foundations of Mechanics*, Addison-Wesley, New York, 2nd ed., 1978.
- [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] E.W. Justh and P.S. Krishnaprasad, Optimal natural frames, *Commun. Inf. Syst.*, vol. 11, no. 1, pp. 17-34, 2011.
- [6] P.S. Krishnaprasad, Optimal control and Poisson reduction, *Inst. Systems Research*, T.R. 93-87, Univ. of Maryland, 1993.
- [7] T. Ohsawa, Symmetry reduction of optimal control systems and principal connections, *SIAM J. Control Optim.*, vol. 51, no. 1, pp. 96-120, 2013.
- [8] 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, pp. 1-81, 1998.
- [9] 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.
- [10] J.E. Marsden, T.S. Ratiu, and A. Weinstein, Reduction and Hamiltonian structures on duals of semidirect product Lie algebras, *Contemp. Math.*, vol. 28, no. 55-99, 1984.
- [11] J. Biggs, W. Holderbaum, and V. Jurdjevic, Singularities of optimal control problems on some 6-D Lie groups, *IEEE Trans. Automat. Control*, vol. 52, no. 6, pp. 1027-1038, 2007.
- [12] G. Walsh, R. Montgomery, and S. Sastry, Optimal path planning on matrix Lie groups, In *IEEE Conf. on Decision and Control*, 1994.
- [13] V. Jurdjevic, *Geometric Optimal Control*, Cambridge University Press, 1997.
- [14] A.A. Agrachev and Y.L. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer, Berlin, 2004.
- [15] A. Bloch, P.S. Krishnaprasad, J.E. Marsden, and R.S. Ratiu, The Euler-Poincaré equations and double bracket dissipation, *Comm. Math. Phys.*, vol. 175, no. 1, pp. 1-42, 1996.
- [16] J.E. Marsden, G. Misiulek, J.P. Ortega, M. Perlmutter, and T.S. Ratiu, *Hamiltonian Reduction by Stages*, Springer Verlag, 2003.