OPTIMAL CONTROL PROBLEMS ON LIE GROUPS WITH SYMMETRY BREAKING COST FUNCTIONS

BY

ANDY D. BORUM

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Aerospace Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Adviser:

Professor Timothy Bretl

ABSTRACT

In this thesis, we consider smooth optimal control systems that evolve on Lie groups. Pontryagin's maximum principle allows us to search for local solutions of the optimal control problem by studying an associated Hamiltonian dynamical system. When the associated Hamiltonian function possess symmetries, we can often study the Hamiltonian system in a vector space whose dimension is lower than the original system. We apply these symmetry reduction techniques to optimal control problems on Lie groups for which the associated Hamiltonian function is left-invariant under the action of a subgroup of the Lie group. Necessary conditions for optimality are derived by applying Lie-Poisson reduction for semidirect products, a previously developed method of symmetry group reduction in the field of geometric mechanics. Our main contribution is a reduced sufficient condition for optimality that relies on the nonexistence of conjugate points. Coordinate formulae are derived 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 optimality conditions are then applied to an example optimal control problem on the Lie group $SE(3)$ that exhibits symmetries with respect to $SE(2)$, a subgroup of $SE(3)$. This optimal control problem can be used to model either a kinematic airplane, i.e. a rigid body moving at a constant speed whose angular velocities can be controlled, or a Kirchhoff elastic rod in a gravitational field.

To my sister, parents, and grandparents.

ACKNOWLEDGMENTS

This thesis would not have been possible without the help of many people. I would first like to thank my adviser, Professor Timothy Bretl, for his guidance and support. My weekly meetings with Professor Bretl always result in challenging problems and new ideas. Second, I would like to thank past and present members of the Bretl Research Group (Olivier Roussel, Dennis Matthews, Aadeel Akhtar, Jamie Norton, Miles Johnson, Navid Aghasadegi, Joe Degol, Mary Nguyen, Xinke Deng, David Hanley, Jessica Mullins, Sean Yen, Jacob Wagner, Kevin Chen) for making our lab such a fun place to work. I am thankful to my research advisers from Virginia Tech, Professors Raymond Plaut, David Dillard, and Douglas Holmes, as we have continued to collaborate on many interesting research problems. Thanks to Professor Rui Loja Fernandes, Professor Yuliy Baryshnikov, and Dr. Ioan Marcut for many helpful discussions. I would also like to acknowledge support I received from the National Science Foundation Graduate Research Fellowship Program under grant number DGE-1144245 and support from the University of Illinois Aerospace Engineering Department through the Stillwell Fellowship. Finally, thanks to my family. My sister, mother, and grandmother have provided endless support during the last two years.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The use of differential geometry to study control systems has led to many insightful results, particularly in the field of optimal control [28]. In this thesis, we use these geometric techniques to study optimal control problems whose state evolves on a Lie group. The necessary conditions for optimality provided by Pontryagin's maximum principle [26] relate solutions of this optimal control problem to integral curves of a Hamiltonian vector field. Thus, finding solutions of a geometric optimal control problem involves studying a Hamiltonian dynamical system on a smooth manifold. The field of geometric mechanics provides many techniques for studying such systems.

One of the main tools used in geometric mechanics is reduction, whereby symmetries of a dynamical system are used to reduce the dimension of the system [23]. These symmetry group reduction techniques are widely applied to Hamiltonian systems in classical mechanics [1, 3]. One particular type of symmetry group reduction is Lie-Poisson reduction, in which a Hamiltonian system evolves on the cotangent bundle of a Lie group, and the Hamiltonian function is invariant under either the left or right action of the Lie group. In this case, the dynamics of the system can be studied by considering a reduced Hamiltonian system which evolves on the dual Lie algebra of the Lie group. After this reduction is performed, the stability of the original system can be studied using tools such as the Energy-Casimir method or the Energy-Momentum method [22, 27].

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, i.e the symmetry has been broken. This issue can sometimes be resolved by embedding the problem in a larger semidirect product space in which the system becomes left-invariant [13, 20, 21]. 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 [20].

In optimal control theory, various types of symmetry group reduction techniques have been applied to the conditions for optimality [10, 12, 16, 9, 25, 29], including Lie-Poisson reduction for left-invariant systems [8, 15, 17, 31]. 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 [8].

1.1 Main Results

We apply semidirect product reduction to the 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 Hamiltonain 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 given in Chapter 2) are, in principle, all we need to find optimal solutions, they do not provide coordinate formulate 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 [4, 31] or a Kirchhoff elastic rod [8, 14] in a gravitational field. Without the effects of gravity, this system is left-invariant under the action of $SE(3)$. However, when gravity is included in the analysis, the symmetry is broken and the system is left-invariant under the action of $SE(2)$, which is a subgroup of $SE(3)$. While this example focuses on a mechanical system, further motivation for the work in this thesis comes from the use of symmetries in quantum optimal control problems, such as the contrast imaging problem in nuclear magnetic resonance [6].

The work in this thesis extends the results that appeared in a previous conference paper [7]. In that paper, the necessary and sufficient conditions described above were stated for matrix Lie groups. In this thesis, these conditions are stated and proved for general Lie groups. More generally, this work builds upon the work in [8], in which sufficient conditions are given for left-invariant geometric optimal control problems.

1.2 Outline of the Thesis

We review the general theory of optimal control on manifolds in Chapter 2. In Chapter 3, we review the application of Lie-Poisson reduction to left-invariant optimal control problems on Lie groups. This leads to reduced necessary and sufficient conditions for optimality. Then, in Chapter 4, reduction for semidirect products is applied to the necessary conditions for optimality provided by Pontryagin's maximum principle. We also 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 Chapter 5, and closing remarks are given in Chapter 6.

CHAPTER 2

OPTIMAL CONTROL ON MANIFOLDS

In this chapter, we review the framework for characterizing solutions of geometric optimal control problems. In Section 2.1, we recall some definitions from differential geometry. Introductory material on smooth manifolds that is not covered here can be found in any differential geometry text, e.g. Lee [18]. Using the language developed in Section 2.1, we state a geometric version of Pontryagin's maximum principle [26] in Section 2.2. The maximum principle allows us to search for extrema of an optimal control problem by analyzing integral curves of a Hamiltonian vector field.

The maximum principle provides necessary conditions for optimality. In Section 2.3, we state a sufficient condition for optimality based on the theory of conjugate points. This sufficient condition is a generalization of Jacobi's condition in the calculus of variations [11]. Proofs of the necessary and sufficient conditions we state in this chapter can be found in Agrachev and Sachkov [2]. In later chapters, we will analyze these necessary and sufficient conditions for optimality under certain symmetry assumptions on the optimal control problem.

2.1 Review of Smooth Manifolds

We begin by recalling some notation regarding smooth manifolds. Let M be a smooth manifold. The set of all smooth real-valued functions on M is denoted by $C^\infty(M)$, and the set of all smooth vector fields on M is denoted by $\mathfrak{X}(M)$. The action of a tangent vector $v \in T_mM$ on a function $f \in C^\infty(M)$ is $v \cdot f$, and the action of a tangent covector $w \in T_m^*M$ on a tangent vector $v \in T_mM$ is $\langle w, v\rangle$. 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 $F \colon M \to N$ is a smooth map, then the pushforward of F at $m \in M$ is the linear map
$T_m F \colon T_m M \to T_{F(m)} N$ that satisfies

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

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 $T_m^* F \colon T_{F(m)}^* N \to T_m^* M$ that satisfies

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

for all $v \in T_m M$ and $w \in T_{F(m)}^* N$. 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

$$\{\cdot, \cdot\} \colon C^\infty(T^* M) \times C^\infty(T^* M) \to C^\infty(T^* M)$$

The co-tangent bundle $T^* M$ together with the bracket $\{\cdot, \cdot\}$ is a Poisson manifold. 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 $K \in C^\infty(T^* M)$. Finally, let $\pi \colon T^* M \to M$ denote the projection map $\pi(w, m) = m$ for all $w \in T_m^* M$.

2.2 Necessary Conditions for Optimality

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

$$\begin{aligned} &\underset{q,u}{\text{minimize}} && \int_0^T g(q(t), u(t))\, dt \\ &\text{subject to} && \dot{q}(t) = f(q(t), u(t)) \quad \text{for all } t \in [0, T] \\ &&& q(0) = q_0, \qquad q(T) = q_1 \end{aligned} \tag{2.1}$$

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

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

where $p \in T_q^*M$.

Theorem 1 is a geometric statement of Pontryagin's maximum principle [26] and provides a set of necessary conditions that all local optima of (2.1) must satisfy.

Theorem 1. (Necessary Conditions) *Suppose $(q_{opt}, u_{opt})\colon [0, T] \to M \times U$ is a local optimum of (2.1). Then, there exists $k \geq 0$ and an integral curve $(p, q)\colon [0, T] \to T^*M$ of the time-varying Hamiltonian vector field X_H, where $H\colon T^*M \times \mathbb{R} \to \mathbb{R}$ is given by $H(p, q, t) = \widehat{H}(p, q, k, u_{opt}(t))$, that satisfies $q(t) = q_{opt}(t)$ and*

$$H(p(t), q(t), t) = \max_{u \in U} \widehat{H}(p(t), q(t), k, u) \tag{2.2}$$

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 [2]. □

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.

2.3 Sufficient Conditions for Optimality

When the conditions in Theorem 1 are satisfied by a normal (q, u), this trajectory is an stationary point of the cost function in (2.1). 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* $(p, q)\colon [0, T] \to T^*M$ *is a normal extremal of* (2.1)*. Define* $H \in C^\infty(M)$ *by*

$$H(p, q) = \max_{u \in U} \widehat{H}(p(t), q(t), 1, u) \tag{2.3}$$

assuming the maximum exists and $\partial^2 \widehat{H}/\partial u^2 < 0$*. Define* $u\colon [0, T] \to U$ *so* $u(t)$ *is the unique maximizer of* (2.3) *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* $\varphi\colon \mathbb{R} \times T^*M \to T^*M$ *be the flow of* X_H *and define the endpoint map* $\phi_t\colon T^*_{q(0)}M \to M$ *by* $\phi_t(w) = \pi \circ \varphi(t, w, q(0))$*. Then* (q, u) *is a local optimum of (2.1) 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 [2]. $\square$

CHAPTER 3

LIE-POISSON REDUCTION OF LEFT-INVARIANT OPTIMAL CONTROL PROBLEMS

While the necessary and sufficient conditions in Theorems 1 and 2 characterize solutions of the optimal control problem (2.1), 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 [8] in the case when M is a Lie group G and the Hamiltonian function (2.2) is left-invariant under the action of G. In this chapter, we review the results in [8] for finding solutions of (2.1) when the optimal control problem is left-invariant. We begin by recalling some facts about Lie Groups in Section 3.1. We then give reduced statements of the necessary and sufficient conditions for optimality in Sections 3.2 and 3.3, respectively. Unlike the conditions in Theorems 1 and 2, these reduced conditions can be evaluated by solving a system of ordinary differential equations.

3.1 Review of Lie Groups

Let G be a Lie group with identity element $e \in G$. Let $\mathfrak{g} = T_eG$ and $\mathfrak{g}^* = T_e^*G$. For any $q \in G$, define the left translation map $L_q\colon G \to G$ by

$$L_q(r) = qr$$

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

$$H(T_r^* L_q(w), r) = H(w, s) \tag{3.1}$$

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 $[\cdot, \cdot]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ by

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

for all $\zeta, \eta \in \mathfrak{g}$. For any $\zeta \in \mathfrak{g}$, the adjoint operator $\mathrm{ad}_\zeta\colon \mathfrak{g} \to \mathfrak{g}$ satisfies

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

and the coadjoint operator $\mathrm{ad}_\zeta^*\colon \mathfrak{g}^* \to \mathfrak{g}^*$ satisfies

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

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 $\{P_1, \ldots, P_n\}$ be the dual basis for $\mathfrak{g}^*$ that satisfies

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

for $i, j \in \{1, \ldots, n\}$, where δ_{ij} is the Kronecker delta. We write ζ_i to denote the ith 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

$$[X_i, X_j] = \sum_{k=1}^{n} C_{ij}^k X_k \tag{3.2}$$

The following two lemmas will be used to prove our main results in Chapter 4.

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

$$\begin{aligned}\zeta(t,\epsilon) &= T_{q(t,\epsilon)}L_{q(t,\epsilon)^{-1}}\left(\frac{\partial q(t,\epsilon)}{\partial t}\right)\\ \eta(t,\epsilon) &= T_{q(t,\epsilon)}L_{q(t,\epsilon)^{-1}}\left(\frac{\partial q(t,\epsilon)}{\partial \epsilon}\right)\end{aligned} \tag{3.3}$$

Then

$$\frac{\partial \zeta}{\partial \epsilon} - \frac{\partial \eta}{\partial t} = [\zeta, \eta] \tag{3.4}$$

Conversely, if there exist smooth maps ζ *and* η *satisfying* (3.4)*, then there exists a smooth map* q *satisfying* (3.3)*.*

Proof. See Proposition 5.1 in [5]. □

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 of the structure constants in (3.2). □

3.2 Reduction of the Necessary Conditions

Now we revisit the statement of necessary conditions for the optimal control problem (2.1) in the case where the smooth manifold M is a Lie group G and where the Hamiltonian function H is left-invariant under the action of G. Theorem 1 implies the existence of a particular integral curve (p, q) in the cotangent bundle T^*G. The following theorem implies the existence of a corresponding integral curve μ in the dual Lie algebra $\mathfrak{g}^*$.

Theorem 3. (Reduction of Necessary Conditions) *Suppose the time-varying Hamiltonian function* $H\colon T^*G \times [0, T] \to \mathbb{R}$ *is both smooth and left-invariant for all* $t \in [0, T]$*. Denote the restriction of* H *to* $\mathfrak{g}^*$ *by* $h = H|_{\mathfrak{g}^* \times [0,T]}$*. Given* $p_0 \in T^*_{q_0}G$*, let* $\mu\colon [0, T] \to \mathfrak{g}^*$ *be the solution of*

$$\dot{\mu} = \mathrm{ad}^*_{\delta h/\delta \mu}(\mu) \tag{3.5}$$

with initial condition $\mu(0) = T_e^* L_{q_0}(p_0)$*. The integral curve* $(p, q)\colon [0, T] \to T^*G$ *of* X_H *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 $q(0) = q_0$*.*

Proof. See the proof of Theorem 13.4.4 in [23]. □

Since $\mathfrak{g}^*$ is a vector space, the trajectory μ described by (3.5) can be evaluated by solving a system of ordinary differential equations, as shown in the following corollary.

Corollary 1. *Suppose that* $H \in C^\infty(T^*G)$ *satisfies the conditions in Theorem 3 and that* X_H *is complete. Given* $q_0 \in G$ *and* $p_0 \in T_{q_0}^*G$*, 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

$$\dot{\mu}_i = -\sum_{j=1}^{n}\sum_{k=1}^{n} C_{ij}^k \frac{\delta h}{\delta \mu_j}\mu_k \tag{3.6}$$

with initial conditions $\mu_i(0) = a_i$ *for* $i \in \{1, \ldots, n\}$*. Now let* $q\colon [0, 1] \to G$ *be the solution of*

$$\dot{q} = X_{\delta h/\delta\mu}(q) \tag{3.7}$$

with initial condition $q(0) = q_0$*. Next, define*

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

for all $t \in [0, T]$*. Then the integral curve of* X_H *with initial conditions* $p(0) = p_0$ *and* $q(0) = q_0$ *is* $(p, q)\colon [0, T] \to T^*G$*.*

Proof. Taking $\mu_1(t), \ldots, \mu_n(t)$ as coordinates of $\mu(t)$, it is easy to verify that (see [23])

$$\mathrm{ad}^*_{\delta h/\delta\mu}(\mu) = -\sum_{j=1}^{n}\sum_{k=1}^{n} C^k_{ij}\frac{\delta h}{\delta \mu_j}\mu_k \tag{3.8}$$

We conclude that (3.5) and (3.6) are equivalent. □

3.3 Reduction of the Sufficient Conditions

We now revisit the statement of sufficient conditions for the optimal control problem (2.1). Reduction of these conditions provides coordinate formulae to test the non-degeneracy of the endpoint map ϕ_t defined in Theorem 2.

Theorem 4. (Reduction of Sufficient Conditions) *Suppose that $H \in C^\infty(T^*G)$ is left-invariant and that X_H is complete. Let $h = H|_{g^*}$ be the restriction of H to $\mathfrak{g}^*$ and let $\varphi\colon \mathbb{R} \times T^*G \to T^*G$ be the flow of X_H. Give $q_0 \in G$, define the endpoint map $\phi_t\colon T^*_{q_0}G \to G$ 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 $T^*_e L_{q_0}(p_0)$, and let μ be the solution of (3.6) with initial conditions $\mu_i(0) = a_i$ for $i \in \{1, \ldots, n\}$. Define the matrices* $\mathbf{F}$, $\mathbf{G}$, $\mathbf{H} \in \mathbb{R}^{n\times n}$ *as follows:*

$$\begin{aligned}
\left[\mathbf{F}\right]_{ij} &= -\frac{\partial}{\partial \mu_j}\sum_{r=1}^{n}\sum_{s=1}^{n} C^s_{ir}\frac{\delta h}{\delta \mu_r}\mu_s \\
\left[\mathbf{G}\right]_{ij} &= \frac{\partial}{\partial \mu_j}\frac{\partial h}{\partial \mu_i} \\
\left[\mathbf{H}\right]_{ij} &= -\sum_{r=1}^{n}\frac{\delta h}{\delta \mu_r}C^i_{rj}
\end{aligned}$$

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

$$\dot{\mathbf{M}} = \mathbf{F}\mathbf{M} \tag{3.9}$$

$$\dot{\mathbf{J}} = \mathbf{G}\mathbf{M} + \mathbf{H}\mathbf{J} \tag{3.10}$$

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

Proof. See the proof of Theorem 4 in [8]. □

Corollary 1 and Theorem 4 show that when the Hamiltonian function is left-invariant under the action of G, the geometric necessary and sufficient conditions in Theorems 1 and 2 can be evaluated by solving a system of ordinary differential equations.

CHAPTER 4

REDUCTION OF OPTIMAL CONTROL PROBLEMS WITH BROKEN SYMMETRY

In chapter 3, we assumed that the Hamiltonian function provided by the maximum principle was left-invariant under the action of the Lie Group G. In this chapter, 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 show a correspondence between integral curve (p, q) in the cotangent bundle T^*G and curves μ in the dual Lie algebra $\mathfrak{g}^*$. As before, these curves can be computed by solving a system of ordinary differential equations. Furthermore, we will derive a system of matrix differential equations, similar to those in (3.9)-(3.10), that can be evaluated to establish non-degeneracy of the endpoint map ϕ_t from Theorem 2.

We begin with a review of semidirect products and Lie group representations in Section 4.1. Further information on Lie groups and their representations can be found in Varadarajan [30]. Then, in Sections 4.2 and 4.3, we derive reduced necessary and sufficient conditions for optimality when the Hamiltonian function is left-invariant with respect to a subgroup of G.

4.1 Review of Semidirect Products

Let V be a vector space and let $\rho\colon G \to \mathrm{Aut}(V)$ 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 $\rho(g)\colon V \to V$ 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

$$\rho_*(g) = \left[\rho(g^{-1})\right]^* \qquad \rho^*(g) = [\rho(g)]^* \tag{4.1}$$

where $[\,]^*$ denotes the dual transformation. The induced Lie algebra representation $\rho'\colon \mathfrak{g} \to \mathrm{End}[V]$ 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 $\exp\colon \mathfrak{g} \to G$ is the exponential map. Denote by G_χ the isotropy group of $\chi \in V^*$, i.e.,

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

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

$$T^*_{(q,u)}(w, s, v, \chi) = \left(T^*_{qs}L_{q^{-1}}(w), L_q(s), u + \rho(q)v, \rho_*(q)\chi\right) \tag{4.2}$$

for all $u, v \in V$, $\chi \in V^*$, $w \in T^*_s G$, and $q, s \in G$ [20].

The following lemma will be used in Section 4.3 to compute conjugate points in systems with broken symmetry.

Lemma 3. *Let $q\colon I \to G$ be a smooth map, where $I \subset \mathbb{R}$ is connected. Denote its derivative $\eta\colon I \to \mathfrak{g}$ by*

$$\eta(\epsilon) = T_{q(\epsilon)}L_{q(\epsilon)^{-1}}\left(\frac{\partial q(\epsilon)}{\partial \epsilon}\right)$$

Then

$$\frac{\partial}{\partial \epsilon}\rho\left(q(\epsilon)\right) = \rho\left(q(\epsilon)\right)\rho'\left(\eta(\epsilon)\right) \tag{4.3}$$

Proof. From the definition of $\eta(\epsilon)$, we have

$$\frac{\partial q(\epsilon)}{\partial \epsilon} = T_e L_{q(\epsilon)}\left(\eta(\epsilon)\right)$$

Now consider the function $g\colon I \times \mathbb{R} \to G$ given by

$$g(\epsilon, s) = L_{q(\epsilon)}\exp\left(\eta(\epsilon)s\right)$$

It is clear that $g(\epsilon, 0) = q(\epsilon)$. Now observe that

$$\begin{aligned}\frac{\partial}{\partial s}g(\epsilon, s) &= \frac{\partial}{\partial s}\left(L_{q(\epsilon)}\exp\left(\eta(\epsilon)s\right)\right)\\ &= T_{\exp(\eta(\epsilon)s)}L_{q(\epsilon)}\left(\frac{\partial}{\partial s}\exp\left(\eta(\epsilon)s\right)\right)\end{aligned}$$

Therefore, at $s = 0$ we have

$$\begin{aligned}\frac{\partial}{\partial s}g(\epsilon, s)|_{s=0} &= T_e L_{q(\epsilon)}\left(\frac{\partial}{\partial s}\exp\left(\eta(\epsilon)s\right)|_{s=0}\right)\\ &= T_e L_{q(\epsilon)}\left(\eta(\epsilon)\right)\\ &= \frac{\partial q(\epsilon)}{\partial \epsilon}\end{aligned}$$

Since ρ is smooth, we have

$$\rho\left(q(\epsilon)\right) = \rho\left(g(\epsilon, s)\right)|_{s=0}$$

and

$$\frac{\partial}{\partial \epsilon}\rho\left(q(\epsilon)\right) = \frac{\partial}{\partial s}\rho\left(g(\epsilon, s)\right)|_{s=0}$$

Therefore

$$\begin{aligned}\frac{\partial}{\partial \epsilon}\rho\left(q(\epsilon)\right) &= \frac{\partial}{\partial s}\rho\left(L_{q(\epsilon)}\exp\left(\eta(\epsilon)s\right)\right)|_{s=0}\\ &= \frac{\partial}{\partial s}\left(\rho\left(q(\epsilon)\right)\rho\left(\exp\left(\eta(\epsilon)s\right)\right)\right)|_{s=0}\\ &= \rho\left(q(\epsilon)\right)\left(\frac{\partial}{\partial s}\rho\left(\exp\left(\eta(\epsilon)s\right)\right)|_{s=0}\right)\\ &= \rho\left(q(\epsilon)\right)\rho'\left(\eta(\epsilon)\right)\end{aligned}$$

We have verified (4.3). □

4.2 Reduction of the 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 on 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.

An example of this case is a heavy rigid body (i.e. a rigid body that experiences a constant gravitational force), whose configuration space is $G = SE(3)$. In this situation, the Hamiltonian of the rigid body is the sum of its kinetic and potential energy. The kinetic energy term is left-invariant under the action of $SE(3)$. However, the potential energy term is only left-invariant under the action of elements in $SE(3)$ that correspond to rotations about the direction of gravity and translations orthogonal to the direction of gravity. In this case, the vector space corresponds to the position of the body in $\mathbb{R}^3$, and the parameter on which the Hamiltonian depends is the linear function that maps the position of the rigid body to its potential energy.

Theorem 5 provides necessary conditions similar to those in Theorem 3 in the case described above, i.e. when the 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 (3.1) holds when $q \in G_{\chi_0}$). We denote the Hamiltonian by $H_{\chi_0} \colon T^*G \to \mathbb{R}$ 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 $H \colon T^*S \to \mathbb{R}$ 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 $v \in V$ [13]. We then show that $H \colon T^*S \to \mathbb{R}$ is left-invariant under the action of S, i.e. we show that

$$H(T_r^* L_q(w), r, v, \rho^*(q)\chi) = H(w, s, v, \chi) \tag{4.4}$$

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 (4.4) holds and $q \in G_{\chi_0}$, then $\chi_0 = \rho^*(q)\chi_0$ 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 (4.4) implies that H_{χ_0} is left-invariant under the action of G_{χ_0} on T^*G.

If (4.4) 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 (p, q) of $X_{H_{\chi_0}}$ in the cotangent bundle T^*G.

Theorem 5. (Semidirect Product Reduction of Necessary Conditions) *Suppose the time-varying Hamiltonian function $H_{\chi_0} \colon T^*G \times [0, T] \to \mathbb{R}$ is smooth, 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*

$$h(T_e^* L_q(p), \rho^*(q)\chi, t) = H_\chi(p, q, t) \tag{4.5}$$

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

$$\dot{\mu} = \mathrm{ad}^*_{\delta h/\delta \mu}(\mu) - \left(\rho'_{\delta h/\delta \chi}\right)^* \chi \tag{4.6}$$

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

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

$$\chi(t) = \rho^*(q(t))\chi_0 \tag{4.7}$$

Proof. See Theorem 3.4 of [20]. □

As was the case in Theorem 3 and Corollary 1, writing (4.6) in coordinates allows us to find μ by solving a system of ordinary differential equations, as shown in the following corollary.

Corollary 2. *Suppose that* $H_{\chi_0} \in C^\infty(T^*G)$ *satisfies the conditions in Theorem 5 and that* $X_{H_{\chi_0}}$ *is complete. Let* h *be the Hamiltonian function on* $\mathfrak{s}^*$ *induced by the family of Hamiltonians* $\{H_\chi | \chi \in V^*\}$*. Given* $q_0 \in G$ *and* $p_0 \in T^*_{q_0}G$*, let* $a \in \mathbb{R}^n$ *be the coordinate representation of* $T^*_e L_{q_0}(p_0)$*. Solve the ordinary differential equations*

$$\dot{\mu}_i = -\sum_{j=1}^{n}\sum_{k=1}^{n} C^k_{ij} \frac{\delta h}{\delta \mu_j}\mu_k - \chi\left(\rho'\left(X_i\right)\frac{\delta h}{\delta \chi}\right) \tag{4.8}$$

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

$$\chi(t)(v) = \chi_0\left(\rho(q(t))v\right) \tag{4.9}$$

for all $v \in V$ *and* q *solves*

$$\dot{q} = X_{\delta h/\delta \mu}(q) \tag{4.10}$$

with initial condition $q(0) = q_0$*. Next, define*

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

for all $t \in [0, T]$*. Then the integral curve of* $X_{H_{\chi_0}}$ *with initial conditions* $p(0) = p_0$ *and* $q(0) = q_0$ *is* $(p, q)\colon [0, T] \to T^*G$*.*

Proof. First, using (4.1), note that (4.7) and (4.9) are equivalent. Taking

$\mu_1(t), ..., \mu_n(t)$ as coordinates of $\mu(t)$, we saw in the proof of Corollary 1 that

$$\mathrm{ad}^*_{\delta h/\delta\mu}(\mu) = -\sum_{j=1}^{n}\sum_{k=1}^{n} C_{ij}^{k}\frac{\delta h}{\delta \mu_j}\mu_k \tag{4.11}$$

From the definition of $(\rho'_{\delta h/\delta\chi})^*\chi \in \mathfrak{g}^*$ in Theorem 5, we have that for each $\zeta \in \mathfrak{g}$,

$$(\rho'_{\delta h/\delta\chi})^*\chi\,(\zeta) = \chi\left(\rho'(\zeta)\frac{\delta h}{\delta\chi}\right) \tag{4.12}$$

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

$$\chi\left(\rho'\left(X_i\right)\frac{\delta h}{\delta\chi}\right) \tag{4.13}$$

Using (4.11)-(4.13), we see that (4.6) and (4.8) are equivalent. □

4.3 Reduction of the Sufficient Conditions

In this section, we revisit our statement of sufficient conditions for (2.1) in the case when the Hamiltonian function satisfies the conditions in Theorem 5. We will lose some of the generality of the previous section by assuming that the Hamiltonian function has the form

$$H_{\chi_0}(p,q) = K(p,q) + U(\chi_0, q) \tag{4.14}$$

This happens when f in (2.1) is independent of χ_0 and g in (2.1) 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 Lemma 4).

Before stating our main result, we begin with two lemmas that describe the computations needed to establish non-degeneracy of the endpoint map ϕ_t defined in Theorem 2.

Lemma 4. *Suppose that $H_{\chi_0} \in C^\infty(T^*G)$ satisfies the conditions in Theorem 5, has the form given in (4.14), and that $X_{H_{\chi_0}}$ is complete. Let h be the Hamiltonian function on $\mathfrak{s}^*$ induced by the family of Hamiltonians*

$\{H_\chi | \chi \in V^*\}$. Given $q_0 \in G$ and $p_0 \in T^*_{q_0}G$, let $a \in \mathbb{R}^n$ be the coordinate representation of $T^*_e L_{q_0}(p_0)$. Define the smooth maps $\mu_i \colon [0,T] \times \mathbb{R}^n \to \mathbb{R}$ and $q \colon [0,T] \times \mathbb{R}^n \to G$ so that $\mu(t,a)$ solves (4.8) and $q(t,a)$ solves (4.10) with initial conditions $\mu_i(0,a) = a_i$ and $q(0,a) = q_0$ for $i \in \{1,\dots,n\}$. Also, define $\chi \colon [0,T] \to V^*$ by (4.9).

Define the time-varying matrices $\mathbf{M}$ and $\mathbf{J} \colon [0,T] \to \mathbb{R}^{n\times n}$ by

$$\Big[\mathbf{M}(t)\Big]_{ij} = \frac{\partial \mu_i(t,a)}{\partial a_j} \qquad \Big[\mathbf{J}(t)\Big]_{ij} = \eta^j_i(t,a)$$

where

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

Then $\mathbf{M}$ satisfies the (linear, time-varying) matrix differential equation

$$\dot{\mathbf{M}} = \mathbf{F}\mathbf{M} - (\mathbf{K}{+}\mathbf{L})\mathbf{J} \tag{4.15}$$

with initial condition $\mathbf{M}(0) = I$, where the time-varying matrices $\mathbf{F}$, $\mathbf{K}$, and $\mathbf{L} \in \mathbb{R}^{n\times n}$ are defined by

$$\begin{aligned}
\Big[\mathbf{F}\Big]_{ij} &= -\frac{\partial}{\partial \mu_j} \sum_{r=1}^{n} \sum_{s=1}^{n} C^s_{ir} \frac{\delta h}{\delta \mu_r} \mu_s \\
\Big[\mathbf{K}\Big]_{ij} &= \chi \left(\rho'(X_j) \rho'(X_i) \frac{\delta h}{\delta \chi} \right) \\
\Big[\mathbf{L}\Big]_{ij} &= \sum_{k=1}^{n} \chi \left(\rho'(X_i) \frac{\partial}{\partial \chi_k} \frac{\delta h}{\delta \chi} \right) \chi \left(\rho'^{(k)}(X_j) \right)
\end{aligned}$$

where $\rho'^{(k)}(X_j)$ denotes the k^{th} column of the matrix representation of $\rho'(X_j)$.

Proof. Differentiating (4.8), we find

$$\begin{aligned}
\Big[\dot{\mathbf{M}}\Big]_{ij} &= \frac{\partial}{\partial t} \frac{\partial \mu_i}{\partial a_j} \\
&= \frac{\partial}{\partial a_j} \frac{\partial \mu_i}{\partial t} \\
&= \frac{\partial}{\partial a_j} \left[\left(-\sum_{r=1}^{n} \sum_{s=1}^{n} C^s_{ir} \frac{\delta h}{\delta \mu_r} \mu_s \right) - \chi \left(\rho'(X_i) \frac{\delta h}{\delta \chi} \right) \right]
\end{aligned}$$

Using (4.9) and the fact that $\frac{\delta h}{\delta \mu}$ is independent of χ, we have

$$\begin{aligned}
\left[\dot{\mathbf{M}}\right]_{ij} &= \sum_{k=1}^{n} -\frac{\partial}{\partial \mu_k}\left(\sum_{r=1}^{n}\sum_{s=1}^{n} C_{ir}^{s}\frac{\delta h}{\delta \mu_r}\mu_s\right)\frac{\partial \mu_k}{\partial a_j} - \frac{\partial}{\partial a_j}\chi\left(\rho'(X_i)\frac{\delta h}{\delta \chi}\right) \\
&= \sum_{k=1}^{n}\left[\mathbf{F}\right]_{ik}\left[\mathbf{M}\right]_{kj} - \frac{\partial}{\partial a_j}\left(\chi_0\left(\rho(q(t))\rho'(X_i)\frac{\delta h}{\delta \chi}\right)\right)
\end{aligned}$$

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

$$\chi_0\left(\left(\frac{\partial}{\partial a_j}\rho(q(t))\right)\rho'(X_i)\frac{\delta h}{\delta \chi}\right) + \chi_0\left(\rho(q(t))\rho'(X_i)\left(\frac{\partial}{\partial a_j}\frac{\delta h}{\delta \chi}\right)\right)$$

Now using Lemma 3, we have

$$\begin{aligned}
\frac{\partial}{\partial a_j}\rho(q(t)) &= \rho(q(t))\rho'(\eta(t)) \\
&= \rho(q(t))\left(\sum_{k=1}^{n}\rho'(X_k)\eta_k^j(t,a)\right)
\end{aligned}$$

Thus

$$\begin{aligned}
&\chi_0\left(\left(\frac{\partial}{\partial a_j}\rho(q(t))\right)\rho'(X_i)\frac{\delta h}{\delta \chi}\right) \\
&= \chi_0\left(\rho(q(t))\left(\sum_{k=1}^{n}\rho'(X_k)\eta_k^j(t,a)\right)\rho'(X_i)\frac{\delta h}{\delta \chi}\right) \\
&= \chi\left(\sum_{k=1}^{n}\rho'(X_k)\rho'(X_i)\frac{\delta h}{\delta \chi}\eta_k^j(t,a)\right) \\
&= \sum_{k=1}^{n}\chi\left(\rho'(X_k)\rho'(X_i)\frac{\delta h}{\delta \chi}\right)\eta_k^j(t,a) \\
&= \sum_{k=1}^{n}\left[\mathbf{K}\right]_{ik}\left[\mathbf{J}\right]_{kj}
\end{aligned}$$

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

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

We have already shown that

$$\frac{\partial}{\partial a_j}\chi(v) = \sum_{k=1}^{n} \chi(\rho'(X_k)v)\eta_k^j(t,a)$$

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

$$\frac{\partial \chi_i}{\partial a_j} = \sum_{k=1}^{n} \chi\left(\rho'^{(i)}(X_k)\right)\eta_k^j(t,a)$$

We now have

$$\begin{aligned}
&\chi_0\left(\rho(q(t))\rho'(X_i)\left(\frac{\partial}{\partial a_j}\frac{\delta h}{\delta \chi}\right)\right)\\
&= \chi_0\left(\rho(q(t))\rho'(X_i)\left(\sum_{k=1}^{n}\left(\frac{\partial}{\partial \chi_k}\frac{\delta h}{\delta \chi}\right)\frac{\partial \chi_k}{\partial a_j}\right)\right)\\
&= \sum_{k=1}^{n}\chi\left(\rho'(X_i)\frac{\partial}{\partial \chi_k}\frac{\delta h}{\delta \chi}\right)\frac{\partial \chi_k}{\partial a_j}\\
&= \sum_{k=1}^{n}\chi\left(\rho'(X_i)\frac{\partial}{\partial \chi_k}\frac{\delta h}{\delta \chi}\right)\sum_{r=1}^{n}\chi\left(\rho'^{(k)}(X_r)\right)\eta_r^j(t,a)\\
&= \sum_{r=1}^{n}\left(\sum_{k=1}^{n}\chi\left(\rho'(X_i)\frac{\partial}{\partial \chi_k}\frac{\delta h}{\delta \chi}\right)\chi\left(\rho'^{(k)}(X_r)\right)\right)\eta_r^j(t,a)\\
&= \sum_{r=1}^{n}\Big[\mathbf{L}\Big]_{ir}\Big[\mathbf{J}\Big]_{rj}
\end{aligned}$$

Combining these computations, we see that

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

It is clear that $\Big[\mathbf{M}(0)\Big]_{ij} = \delta_{ij}$, so we have verified (4.15). □

Lemma 5. *Suppose that the assumptions in Lemma 4 hold, and define the matrix functions* $\mathbf{M}$ *and* $\mathbf{J}$ *as in Lemma 4. Then* $\mathbf{J}$ *satisfies the (linear, time-varying) matrix differential equation*

$$\dot{\mathbf{J}} = \mathbf{G}\mathbf{M} + \mathbf{H}\mathbf{J} \tag{4.16}$$

with initial condition $\mathbf{J}(0) = 0$*, where the time-varying matrices* $\mathbf{G}$ *and* $\mathbf{H} \in \mathbb{R}^{n\times n}$ *are defined by*

$$\Big[\mathbf{G}\Big]_{ij} = \frac{\partial}{\partial \mu_j}\frac{\delta h}{\delta \mu_i}$$
$$\Big[\mathbf{H}\Big]_{ij} = -\sum_{r=1}^{n} \frac{\delta h}{\delta \mu_r} C^i_{rj}$$

Proof. Define

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

From Lemma 1, Theorem 5, and Lemma 4, we have

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

This equation can be written in coordinates by using Lemma 2.

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

It is clear that $\Big[\mathbf{J}(0)\Big]_{ij} = 0$, so we have verified (4.16). □

We can now state our main result.

Theorem 6. (Semidirect Product Reduction of Sufficient Conditions) *Suppose that* $H_{\chi_0} \in C^\infty(T^*G)$ *satisfies the conditions in Theorem 5, has the form given in* (4.14)*, 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* $\varphi \colon \mathbb{R} \times T^*G \to T^*G$ *be the flow of* $X_{H_{\chi_0}}$*. Given* $q_0 \in G$*, define the endpoint map* $\phi_t \colon T^*_{q_0}G \to G$ *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* $T^*_e L_{q_0}(p_0)$*, and let* μ *be the solution of* (4.8) *with initial conditions* $\mu_i(0) = a_i$ *for* $i \in \{1, \ldots, n\}$*. Solve the matrix differential equations in Lemmas 4 and 5 to find the matrix function* $\mathbf{J} \colon [0, T] \to \mathbb{R}^{6\times 6}$ *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 $\sigma : \mathbb{R}^n \to T^*_{q_0}G$ 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 $T_{\sigma(a)}\phi_t(\sigma(s)) = 0$. Define the smooth map $q : [0, T] \times \mathbb{R}^n \to G$ 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 $j \in \{1, ..., n\}$, 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

$$0 = \sum_{j=1}^{n} s_j T_{q(t,a)} L_{q(t,a)^{-1}} \left(\frac{\partial q(t,a)}{\partial a_j} \right) \tag{4.17}$$

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

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

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

$$\Big[\mathbf{J}\Big]_{ij} = \eta^j_i(t,a) \tag{4.18}$$

i.e. the jth column of $\mathbf{J}$(t) is the coordinate representation of $\eta^j(t,a)$ with respect to $\{X_1, ..., X_n\}$. Then, (4.17) 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 $\det(\mathbf{J}(t)) = 0$. □

As was the case in Chapter 3, non-degeneracy of the endpoint map ϕ_t can be established by solving the matrix differential equations (4.15) and (4.16).

CHAPTER 5

APPLICATION TO AN OPTIMAL CONTROL PROBLEM ON $SE(3)$

In this chapter, we apply the tools developed in Chapter 4 to an optimal control problem on the Lie group $SE(3)$. We use Theorems 5 and 6 to derive ordinary differential equations that characterize the local solutions of this optimal control problem.

5.1 The Kinematic Airplane and the Heavy Kirchhoff Elastic Rod

In this section, we consider a geometric optimal control problem that can be used to model two different systems. First consider a kinematic airplane that flies at a constant speed [4, 31]. 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 $\{X_1, ...X_6\}$ of $\mathfrak{se}(3)$ given by

$$X_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad X_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad X_3 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$X_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad X_5 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad X_6 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

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

which satisfies

$$\dot{q} = q\left(u_1X_1 + u_2X_2 + u_3X_3 + X_4\right)$$

for all $t \in [0,T]$, where $u : [0,T] \to U = \mathbb{R}^3$ is the control input. In [31], 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

$$\begin{aligned}
&\underset{q,u}{\text{minimize}} && \int_0^T \left(\frac{1}{2}\sum_{i=1}^{3} c_i u_i^2 + W\chi_0(d(q))\right) dt \\
&\text{subject to} && \dot{q} = q(u_1X_1 + u_2X_2 + u_3X_3 + X_4) \\
& && q(0) = q_0 \quad q(T) = q_1
\end{aligned} \tag{5.1}$$

for some fixed 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 $d : SE(3) \to \mathbb{R}^4$ maps the 4×4 matrix $q \in SE(3)$ to the last column of q, i.e.

$$d\left(\begin{bmatrix} R & b \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} b \\ 1 \end{bmatrix}$$

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 Chapter 4, we have chosen $G = SE(3)$ and $V = \mathbb{R}^4$. The gravity term breaks the $SE(3)$ symmetry, and the methods used in [31] cannot be applied. However, this problem fits into the framework in Chapter 4, and Theorems 5 and 6 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 [8], in which Theorems 3 and 4 from this paper are applied to (5.1) (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.

5.2 Necessary Conditions for Optimality

We now analyze (5.1) using the tools developed in this thesis. 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

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

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

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

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

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

The maximized Hamiltonian function is then

$$H_{\chi_0}(p, q) = \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\langle p, qX_i\rangle^2 + \langle p, qX_4\rangle - W\chi_0(d(q))$$

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

$$H(p, q, v, \chi) = \frac{1}{2}\sum_{i=1}^{3} c_i^{-1}\langle p, qX_i\rangle^2 + \langle p, qX_4\rangle - W\chi(d(q))$$

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

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

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

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

where χ_4 is the fourth entry of χ. To see this, observe that

$$\begin{aligned}
h(T_e^* L_q(p), \rho(q)^*\chi) &= \frac{1}{2}\sum_{i=1}^{3} c_i^{-1} \langle T_e^* L_q(p), X_i \rangle^2 \\
&\quad + \langle T_e^* L_q(p), X_4 \rangle - W\rho(q)^*\chi(d(e)) \\
&= \frac{1}{2}\sum_{i=1}^{3} c_i^{-1} \langle p, q X_i \rangle^2 + \langle p, q X_4 \rangle - W\chi(d(q)) \\
&= H(p, q, v, \chi)
\end{aligned}$$

so (4.5) is satisfied. Applying (4.8) 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, (4.9) gives

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

Carrying out this computation, we wee that $\chi(t)$ gives the direction of the gravity vector rotated into the local coordinate frame at $q(t)$. Also, $\chi_4(t)$ gives the vertical position component of $q(t)$. This explains why the reduced Hamiltonian h only depends on the fourth component of χ.

5.3 Sufficient Conditions for Optimality

Solutions of (5.1) are obtained by finding an initial value of $\mu(0)$ (which, from Corollary 2, 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 (5.1). The analysis in [8] shows 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 6 to determine which extrema are actually local minima.

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

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

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

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

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

$$\mathbf{L} = 0$$

After using a shooting method to find $a \in \mathbb{R}^6$ which places $q(T)$ at q_1, (4.15) and (4.16) can be solved numerically with the initial conditions $\mathbf{M}(0) = I$ 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 (5.1). Note from (4.18) 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.

CHAPTER 6

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 Chapter 2 were stated in terms of coordinate-free geometric results, the reduced necessary and sufficient conditions were 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 [24]. The results in this thesis could be extended by considering these more general approaches to symmetry group reduction. Furthermore, reduction for systems defined on the semidirect product of a Lie group and multiple vector spaces has previously been studied [19]. In this thesis, we only considered Hamiltonian systems that depend on one parameter from such a vector space. A Hamiltonian function that depends on two or more parameters may not possess a non-trivial symmetry group. However, the reduction procedure used in this thesis may lead to coordinate formulae for finding optimal solutions for these asymmetric systems.

REFERENCES

[1] R. Abraham and J.E. Marsden, *Foundations of Mechanics*, Addison-Wesley, New York, 2nd ed., 1978.

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

[3] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, New York, 1978.

[4] J. Biggs, W. Holderbaum, and V. Jurdjevic, Singularities of optimal control problems on some 6-D Lie groups, *IEEE Transactions on Automatic Control*, vol. 52, no. 6, pp. 1027-1038, 2007.

[5] A. Bloch, P.S. Krishnaprasad, J.E. Marsden, and R.S. Ratiu, The Euler-Poincaré equations and double bracket dissipation, *Communications in Mathematical Physics*, 175(1):1-42, 1996.

[6] B. Bonnard and D. Sugny, *Optimal Control with Applications in Space and Quantum Dynamics*, AIMS, Springfield, 2012.

[7] A.D. Borum and T. Bretl, Geometric optimal control for symmetry breaking cost functions, in *IEEE Conference on Decision and Control (CDC)*, 2014.

[8] T. Bretl and Z. McCarthy, Quasi-static manipulation of a Kirchhoff elastic rod based on a geometric analysis of equilibrium configurations, *International Journal of Robotics Research*, 33(1):48-68, 2014.

[9] M. de León, J. Cortés, D. Martín de Diego, and S. Martínez, General symmetries in optimal control, *Reports on Mathematical Physics*, 53(1):5578, 2004.

[10] A. Echeverría-Enríquez, J. Marín-Solano, M.C. Muñoz Lecanda, and N. Román-Roy, Geometric reduction in optimal control theory with symmetries, *Reports on Mathematical Physics*, 52(1):89113, 2003.

[11] I.M. Gelfand and S.V. Fomin, *Calculus of Variations*, Dover, Mineola, NY, 1963.

[12] J. Grizzle and S. Marcus, Optimal control of systems possessing symmetries, *IEEE Transactions on Automatic Control*, 29(11):10371040, 1984.

[13] D.D. Holm, J.E. Marsden, and T.S. Ratiu, The Euler-Poincaré equations and semidirect products with applications to continuum theories, *Advances in Mathematics*, 137:1-81, 1998.

[14] V. Jurdjevic, *Geometric Optimal Control*, Cambridge University Press, 1997.

[15] E.W. Justh and P.S. Krishnaprasad, Optimal natural frames, *Communications in Information and Systems*, 11(1):17-34, 2011.

[16] W.S. Koon and J.E. Marsden, Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction, *SIAM Journal on Control and Optimization*, 35(3):901929, 1997.

[17] P.S. Krishnaprasad, Optimal control and Poisson reduction, *Institute for Systems Research Technical Report*, T.R. 93-97, 1993.

[18] J.M. Lee, *Introduction to Smooth Manifolds*, Springer, New York, 2003.

[19] N.E. Leonard and J.E. Marsden, Stability and drift of underwater vehicle dynamics: mechanical systems with rigid motion symmetry, *Physica D*, 105:130-162, 1997.

[20] J.E. Marsden, T.S. Ratiu, and A. Weinstein, Semidirect products and reduction in mechanics, *Transactions of the American Mathematical Society*, 281(1):147-177, 1984.

[21] J.E. Marsden, T.S. Ratiu, and A. Weinstein, Reduction and Hamiltonian structures on duals of semidirect product Lie algebras, *Contemporary Mathematics*, 28:55-99, 1984.

[22] J.E. Marsden, *Lectures on Mechanics*, London Mathematical Society Lecture Note Series, 174, Cambridge University Press, 1992.

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

[24] J.E. Marsden, G. Misiolek, J.P. Ortega, M. Perlmutter, and T.S. Ratiu, *Hamiltonian Reduction by Stages*, Springer-Verlag, 2003.

[25] T. Ohsawa, Symmetry reduction of optimal control systems and principal connections, *SIAM Journal on Control and Optimization*, 51(1):96120, 2013

[26] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*. John Wiley, 1962.

[27] J.C. Simo, D. Lewis, and J.E. Marsden, Stability of relative-equilibria I: The reduced energy-momentum method, *Archive for Rational Mechanics and Analysis*, 115(1):15-59, 1991.

[28] H.J. Sussmann, Geometry and optimal control, In *Mathematical Control Theory*, J. Baillieul and J.C. Willems Eds., Springer-Verlag, 140-198, 1999.

[29] A.J. van der Schaft, Symmetries in optimal control, *SIAM Journal on Control and Optimization*, 25(2):245259, 1987.

[30] V.S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations*, Springer, New York, 1984.

[31] G. Walsh, R. Montgomery, and S. Sastry, Optimal path planning on matrix Lie groups, In *IEEE Conference on Decision and Control (CDC)*, 2:1258-1263, 1994.