NEIGHBORING EXTREMAL OPTIMAL CONTROL FOR MECHANICAL SYSTEMS ON RIEMANNIAN MANIFOLDS

ANTHONY M. BLOCH
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109-1043, USA

ROHIT GUPTA AND ILYA V. KOLMANOVSKY
Department of Aerospace Engineering
University of Michigan
Ann Arbor, MI 48109-2140, USA

(Communicated by Andrew Lewis)

Abstract. In this paper, we extend neighboring extremal optimal control, which is well established for optimal control problems defined on a Euclidean space (see, e.g., [8]) to the setting of Riemannian manifolds. We further specialize the results to the case of Lie groups. An example along with simulation results is presented.

1. Introduction. Neighboring extremal optimal control (NEOC) is well established for optimal control problems (OCPs) defined on a Euclidean space (see, e.g., [8]). NEOC provides a sensitivity-based fast correction to an optimal control for changes in the initial conditions and/or parameters, thereby providing a form of a local feedback and enhancing the real-world applicability of optimal control. The configuration space for most mechanical systems that perform large maneuvers is not a Euclidean space. For instance, the configuration space of a spacecraft modeled as a rigid body is the Lie group $SE(3) = SO(3) \ltimes \mathbb{R}^3$ (see, e.g., [21]). With this motivation, in this paper, we extend NEOC to OCPs for mechanical systems evolving on Riemannian manifolds. This extension and rigorous treatment of the underlying details represent the main contribution of this paper. We will first briefly discuss NEOC for OCPs defined on a Euclidean space. In what follows, we will suppress the explicit dependence of the state, costate and control trajectories on time unless otherwise necessary.

1.1. Neighboring extremal optimal control. We will first review some background material also covered in [20], where NEOC is discussed in the setting of $\mathbb{R}^n$. Consider a parameter dependent OCP, where the objective is to minimize a cost functional given by

$$
\min_{u(t)} J = K(x(T), p) + \int_0^T L(x(t), u(t), p) dt
$$

2010 Mathematics Subject Classification. Primary: 49J15, 49K40; Secondary: 37N35.

Key words and phrases. Lie groups, mechanical systems, neighboring extremal optimal control, optimal control, Riemannian manifolds.
subject to
\[ \dot{x}(t) = f(x(t), u(t), p), \quad x(0) = x_0, \quad (2) \]

where \( x(\cdot) \in AC([0, T], \mathbb{R}^n) \), \( u(\cdot) \in L^\infty([0, T], \mathbb{R}^m) \), \( p \in \mathbb{R}^l \) is a parameter, \( K : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R} \), \( L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R} \) and \( f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R} \) are functions of class \( C^2 \). Let \((x^*_p, u^*_p)\) be a solution for the OCP (1)-(2), where \( u^*_p(t) \) denotes the optimal control, which satisfies the Lagrange multiplier rule in a normal form (see, e.g., [6]). Let \( \lambda^*_p \) be the solution corresponding to \((x, u) = (x^*_p, u^*_p)\) of the following costate equation
\[ \dot{\lambda} = -H_x(x, u, \lambda, p), \quad \lambda(T) = K_x(x(T), p), \]
where \( \lambda(\cdot) \in AC([0, T], \mathbb{R}^n) \), \( H \) is the Hamiltonian and \( H(x, u, \lambda, p) := L(x, u, p) + \lambda^T f(x, u, p) \). Altogether, \((x^*_p, u^*_p, \lambda^*_p)\) satisfy the following necessary conditions for optimality
\[ \dot{x}(t) = f(x(t), u(t), p), \quad x(0) = x_0, \]
\[ \dot{\lambda}(t) = -H_x(x(t), u(t), \lambda(t), p), \quad \lambda(T) = K_x(x(T), p), \]
\[ 0 = H_u(x(t), u(t), \lambda(t), p). \quad (5) \]

Suppose there is a small variation in the initial condition and/or the parameter, and we would like to update the optimal control. Instead of solving the original OCP again, we employ a first-order approximation of the necessary conditions for optimality around the nominal trajectory. This approximation is based on the linearized relations (see, e.g., [8], [17], [18], [19])
\[ \delta \dot{x}(t) = \frac{\partial f}{\partial x} \delta x(t) + \frac{\partial f}{\partial u} \delta u(t) + \frac{\partial f}{\partial p} \delta p, \quad \delta x(0) = \delta x_0, \]
\[ \delta \dot{\lambda}(t) = -H_{xx} \delta x(t) - H_{xu} \delta u(t) - H_{x\lambda} \delta \lambda(t) - H_{x\upsilon} \delta \upsilon, \quad \delta \lambda(T) = K_{x\lambda} \delta x(T) + K_{x\upsilon} \delta \upsilon, \]
\[ 0 = H_{ux} \delta x(t) + H_{uu} \delta u(t) + H_{u\lambda} \delta \lambda(t) + H_{u\upsilon} \delta \upsilon. \quad (8) \]

Under the the second-order sufficient optimality condition (see, e.g., [17], [19]), (6)-(8) represents the optimality condition for the following OCP (see, e.g., [8], [17], [18], [19])
\[ \min_{\delta u(\cdot)} \delta^2 J = \frac{1}{2} \begin{bmatrix} \delta x(T) \delta p \end{bmatrix}^T \begin{bmatrix} K_{xx}(T) & K_{xp}(T) \\ K_{px}(T) & 0 \end{bmatrix} \begin{bmatrix} \delta x(T) \\ \delta p \end{bmatrix} + \frac{1}{2} \int_0^T \begin{bmatrix} \delta x(t) \\ \delta u(t) \\ \delta p \end{bmatrix}^T \begin{bmatrix} H_{xx}(t) & H_{xu}(t) & H_{x\lambda}(t) & H_{x\upsilon}(t) \\ H_{ux}(t) & H_{uu}(t) & H_{u\lambda}(t) & H_{u\upsilon}(t) \\ H_{px}(t) & H_{pu}(t) & 0 \end{bmatrix} \begin{bmatrix} \delta x(t) \\ \delta u(t) \\ \delta p \end{bmatrix} \, dt \]
\[ (9) \]

subject to the perturbed dynamics
\[ \delta \dot{x}(t) = \frac{\partial f}{\partial x} \delta x(t) + \frac{\partial f}{\partial u} \delta u(t) + \frac{\partial f}{\partial p} \delta p, \quad \delta x(0) = \delta x_0, \]
\[ (10) \]

where the matrices in the cost functional (9) and the Jacobian matrices in the dynamic constraint (10) are evaluated at the nominal trajectories. The optimal control for the OCP (9)-(10) is given by
\[ \delta u^*(t) = -H^{-1}_{uu}(t) \begin{bmatrix} H_{ux}(t) \delta x(t) + f_u^T(t) \delta \lambda(t) + H_{u\upsilon}(t) \delta \upsilon \end{bmatrix}, \quad (11) \]
where all partial derivative matrices are evaluated at the nominal trajectories and\( \delta \lambda(t) \) is a perturbation from \( \lambda^*(t) \), ultimately expressible in terms of \( \delta x(t) \) and \( \delta p \).

The updated control is now calculated as the sum of \( u^*(t) \) and \( \delta u^*(t) \) and can be used directly or to warm start an optimizer for parameter \( p + \delta p \). This is the basic idea behind NEOC. For a detailed description of NEOC see [8]. For a mathematically rigorous introduction to NEOC see [30].

**Remark 1.** The OCP (9)-(10) is known as the accessory minimum problem in the calculus of variations (see, e.g., [32]). If there is no variation in the initial condition, i.e., the initial condition remains fixed, then \( \delta x(0) = 0 \) and similarly, if there is no variation in the parameter, i.e., the parameter remains fixed, then \( \delta p = 0 \). Note that it is also possible to obtain the solution in the conventional NEOC setting (see, e.g., [8]), by adding \( p \) as a state, with \( \dot{p} = 0 \).

For \( (x^*_t(t), u^*_t(t)) \) to be a strong local minimizer for the OCP (1)-(2), the second-order sufficient condition (strengthened Legendre-Clebsch condition) requires that \( H_{uu}(t) > 0 \), for a.e. \( t \in [0, T] \) and conjugate points for the OCP (9)-(10) must not exist (Jacobi condition) (see, e.g., [30]). An indicator for the existence of conjugate points is that the Riccati equation associated with the OCP (9)-(10) has a finite escape time (see, e.g., [30]). The existence of a solution of the Riccati equation associated with the OCP (9)-(10) over the interval \( [0, T] \) is enough to rule out the existence of conjugate points. For a modern exposition on conjugate points see [2], [30]. For more on conjugate points for OCPs see [7], [8], [12], [13], [23], [25], [34], [35], [36].

In the extension of NEOC to the Riemannian manifold setting, we will be using a few concepts from Riemannian geometry. We refer the reader unfamiliar with Riemannian geometry to [16], [26]. Before proceeding further, we define some notation which will be used in the paper.

2. **Notation.** We denote an \( n \)-dimensional complete connected Riemannian manifold by \( Q \) and the Riemannian metric by \( \langle \cdot, \cdot \rangle \). For \( q \in Q \), the tangent space of \( Q \) at \( q \) is denoted by \( T_q Q \). The tangent bundle is denoted by \( TQ \). The cotangent space corresponding to \( T_q Q \) is denoted by \( T^*_q Q \). The tangent bundle is denoted by \( T^* Q \). For \( \bar{v}_q := (q, \bar{v}) \in TQ \), the tangent space of \( TQ \) at \( \bar{v}_q \) is denoted by \( T_{\bar{v}_q} TQ \). The natural pairing between \( T^*_q Q \) and \( T_q Q \) is denoted by \( \cdot (\cdot) \). The musical isomorphism associated with the Riemannian metric \( \langle \cdot, \cdot \rangle \) is denoted by \( \sharp \), where \( \cdot : T^* Q \to TQ \). The natural projection map is denoted by \( \pi \), where \( \pi : TQ \to Q \). The real vector space of all smooth vector fields is denoted by \( \mathfrak{X}(Q) \). The real vector space of all smooth covector fields is denoted by \( \mathfrak{X}^*(Q) \). The unique Levi-Civita connection is denoted by \( \nabla \). The covariant derivative is denoted by \( \frac{D}{\pi} \). The Lie bracket is denoted by \( [\cdot, \cdot] \). The curvature tensor of the connection \( \nabla \) is denoted by \( R(\cdot, \cdot) \). The exponential map is denoted by \( \exp \). The interior of a set is denoted by \( \text{int} \). The linear span of a set of vectors is denoted by \( \text{span} \).

The rest of the paper is organized as follows. In Section 3, we consider a particular OCP, which will be used to illustrate NEOC. In Section 4, the OCP is solved (only for the sake of completeness) using Lagrange multipliers and the corresponding variational equations are also derived. In Section 5, the OCP is solved (only for the sake of completeness) as a variational problem and the corresponding variational equation is also given. In Section 6, the results are specialized to the case when \( Q \) is a compact semisimple Lie group and an example along with simulation results...
is presented. Finally, in Section 7 we make some concluding remarks with possible directions for future research.

3. Optimal control problem. Let \{\mathcal{X}_i\}_{i=1}^n be smooth vector fields on \mathcal{Q}. For a given time interval \([0, T]\), it is assumed that the flow of each vector field in \{\mathcal{X}_i\}_{i=1}^n exists, for all \(t \in [0, T]\). Additionally, if \mathcal{Q} is compact, then each vector field in \{\mathcal{X}_i\}_{i=1}^n is complete (see, e.g., [22]). Consider the following OCP

\[
\min_{u(\cdot)} J = \frac{1}{2} \int_0^T \langle u(t), u(t) \rangle dt
\]

subject to

\[
\frac{dq}{dt}(t) = v(t), \quad q(0) = q_0, \quad q(T) = q_T,
\]

\[
\frac{dv}{dt}(t) = u(t), \quad v(0) = v_0, \quad v(T) = v_T,
\]

where \(q(\cdot) \in C^2([0, T], \mathcal{Q}), v(\cdot) \in C^1([0, T], T_q(\mathcal{Q}))\) and the \(n\)-tuple of control inputs \([u^1 \ldots u^n]^T\) take values in \(\mathbb{R}^n\). Note that in general, the \(n\)-tuple of control inputs \([u^1 \ldots u^n]^T\) are constrained to take values in the set \(\mathcal{U} \subset \mathbb{R}^n\) (nonempty, connected, with \(0 \in \text{int}(\mathcal{U})\) and also generally assumed to be compact and convex). In a more general setting, e.g., when admissible controls are only assumed to be measurable locally bounded mappings taking values in the set \(\mathcal{U}\), more technical assumptions are needed (see, e.g., [2], [10]) but we do not consider such a setting in this paper.

Remark 2. It is possible to generalize the idea presented in this paper to a cost functional, which has a more general form with a more complicated dynamic constraint (see, e.g., [21]). We choose to work with the cost functional (12) and the class of fully-actuated controlled mechanical systems for which the Lagrangian \(L : T\mathcal{Q} \rightarrow \mathbb{R}\) is given by \(L(v_q) = \frac{1}{2} \langle v_q, v_q \rangle\), where \(v_q \in T_q\mathcal{Q}\), as the solution for (P) has a nice geometric interpretation thereby helping to present the main idea of the paper clearly and avoid unnecessary mathematical complications. One can also extend the idea presented in this paper to the class of under-actuated controlled mechanical systems (with a more general Lagrangian), using the method of Lagrange multipliers (see, e.g., [6]) but we leave such an extension to future work. A somewhat related paper, which is similar in spirit to our paper, is [33]. However, [33] does not treat NEOC. In fact, (P) is equivalent to the well known Riemannian geodesic problem (see, e.g., [6]). The local existence and uniqueness of the solution for (P) follow from the theorems on local existence and uniqueness of the solution for ordinary differential equations. The equations of motion for the class of fully-actuated controlled mechanical systems with the Lagrangian defined above are given by

\[
\nabla q \dot{q} = \sum_{i=1}^n u^i X_i(q),
\]

where \(q : [0, T] \rightarrow \mathcal{Q}\). The vertical lift of a vector field \(X\) on \(\mathcal{Q}\) is the vector field \(X^{\text{lift}}\) on \(T\mathcal{Q}\) given by

\[
X^{\text{lift}}(v_q) = \left. \frac{d}{dt} (v_q + tX(q)) \right|_{t=0} \in T_{v_q}T\mathcal{Q},
\]
where $v_q \in T_q Q$. In local coordinates, (16) has a simple interpretation. Let $(q^1, \ldots, q^n)$ be the local coordinates for $Q$ and $(q^1, \ldots, q^n, v^1, \ldots, v^n)$ be the corresponding local coordinates for $TQ$. If $X = \sum_{i=1}^n X^i \frac{\partial}{\partial q^i}$, then $X^{\text{lift}} = \sum_{i=1}^n X^i \frac{\partial}{\partial v^i}$, where $(X^1, \ldots, X^n)$ are the component functions of $X$ in some given chart. We can now re-write (15) as follows

$$\dot{\gamma} = Z(\gamma) + \sum_{i=1}^n u^i X^i_{\text{lift}}(\gamma),$$

(17)

where $\gamma : [0, T] \to TQ$ and $Z$ is the geodesic spray associated with the connection $\nabla$. In local coordinates, $Z = \sum_{i=1}^n v^i \frac{\partial}{\partial q^i} - \sum_{i=1}^n \sum_{j,k=1}^n \Gamma^i_{jk} v^j v^k \frac{\partial}{\partial v^i}$. Note that $\gamma$ is the canonical lifting of $q$, i.e., $(\pi \circ \gamma)(t) = q(t)$. It is not difficult to see that (15) is equivalent to (17). Indeed, in local coordinates, (15) has the following form

$$\dot{q}^i + \sum_{j,k=1}^n \Gamma^i_{jk} \dot{q}^j \dot{q}^k = \sum_{i=1}^n u^i X_i(q), \quad i = 1, \ldots, n.$$

(18)

Observe that (18) is a system of second-order ordinary differential equations on $Q$, which is equivalent to a system of first-order ordinary differential equations on $TQ$ of the form

$$\dot{q}^i = v^i, \quad i = 1, \ldots, n,$$

(19)

$$\dot{v}^i = -\sum_{j,k=1}^n \Gamma^i_{jk} v^j v^k + \sum_{i=1}^n u^i X_i(q), \quad i = 1, \ldots, n.$$

(20)

The connection $\nabla$ induces an Ehresmann connection on $\pi : TQ \to Q$ such that, for all $v_q \in T_q Q$, there is a splitting of $T_{v_q} TQ$ into a horizontal subspace and a vertical subspace, i.e., $T_{v_q} TQ \cong H_{v_q}(TQ) \oplus V_{v_q}(\pi)$, where $H_{v_q}(TQ) \cong T_q Q$ and $V_{v_q}(\pi) \cong T_q Q$. Note that $H_{v_q}(TQ) = \text{span}\{\frac{\partial}{\partial q^i} - \sum_{j,k=1}^n \Gamma^i_{jk} v^j v^k \frac{\partial}{\partial v^i}\}_{i=1}^n$ and $V_{v_q}(\pi) = \text{span}\{\frac{\partial}{\partial v^i}\}_{i=1}^n$. It is now easy to verify that with respect to the above splitting, for all $v_q \in T_q Q$, $Z(v_q) \in H_{v_q}(TQ)$ and $X^{\text{lift}}(v_q) \in V_{v_q}(\pi)$. For more details, see [1], [3], [4], [5], [6], [11], [31]. In view of the above discussion, we note that (13)-(14) are equivalent to (17). Using the splitting of $T_{v_q} TQ$ discussed above, for all $r \in T_{v_q} TQ$, $r$ can be uniquely written as follows

$$r = r^h + r^v,$$

where $r^h \in H_{v_q}(TQ)$ and $r^v \in V_{v_q}(\pi)$. For all pairs $r_1, r_2 \in T_{v_q} TQ$, the Riemannian metric (Sasaki metric) on $TQ$ is obtained in terms of the Riemannian metric on $Q$ as follows

$$\langle \langle r_1, r_2 \rangle \rangle = \langle r^h_1, r^h_2 \rangle + \langle r^v_1, r^v_2 \rangle.$$

It is now easy to verify that (12) is well defined, since $\frac{1}{2} \int_0^T \langle \langle u(t), u(t) \rangle \rangle dt = \frac{1}{2} \int_0^T \langle u(t), u(t) \rangle dt$. For more details, see [28], [29], [31].

Before we proceed further, we introduce the concept of a variation (see, e.g., [1], [6], [10], [16], [24], [26]). Let $\Omega$ denote the set of all $C^2$ curves on $Q$ satisfying the boundary conditions (13)-(14). The set $\Omega$ is also referred to as the path space of $Q$ (see, e.g., [26]). For a curve $q(t) \in \Omega$, $T_{q(t)} \Omega$ is a vector space consisting of all $C^2$ vector fields $w(t)$ along $q(t)$ such that $w(0) = 0$ and $w(T) = 0$. 


Definition 3.1 ([26]). A one-parameter variation of a curve $q \in \Omega$ is a function $\bar{q} : (-\epsilon, \epsilon) \to \Omega$, for some $\epsilon > 0$ such that

(a) $\bar{q}(0) = q$.

(b) The map $q_\epsilon : [0, T] \times (-\epsilon, \epsilon) \to \mathcal{Q}$ defined by $q_\epsilon(t, \bar{\epsilon}) = \bar{q}(\epsilon)(t)$ is $C^2$ on $[0, T] \times (-\epsilon, \epsilon)$.

Note that a one-parameter variation of a curve $q(t) \in \Omega$ defined above is proper (see, e.g., [16]). The vector field $v(t) := \frac{\partial q}{\partial t}(t, 0)$ is the velocity vector field along $q(t)$ and the vector field $w(t) := \frac{\partial q}{\partial \epsilon}(t, 0)$ is the variation vector field associated with the one-parameter variation $q_\epsilon$ (see, e.g., [16], [26]). By setting $q_\epsilon(t, \epsilon) := \exp(\epsilon w(t))$, we obtain a one-parameter variation of a curve $q(t) \in \Omega$, where $w(t) \in T_{q(t)}\Omega$ (see, e.g., [16], [26]).

To derive NEOC for (P), we first obtain the nominal trajectory, by solving (P) using two methods. The first method is solving (P) using Lagrange multipliers and the second method is solving (P) as a variational problem.

4. Solution using Lagrange multipliers. We proceed by following the same procedure as given in [15] and defining the augmented cost functional as follows

$$J^\alpha = \int_0^T \left[ \frac{1}{2} \langle u, u \rangle + \lambda_1 \left( \frac{dq}{dt} - v \right) + \lambda_2 \left( \frac{Dv}{dt} - u \right) \right] dt,$$

where $\lambda_1(\cdot), \lambda_2(\cdot) \in C^1([0, T], T_{q(\cdot)}\mathcal{Q})$. We will now fix some notation.

4.1. Notation. For any smooth vector field $y = \sum_{i=1}^n y^i(t)X_i(q)$ along the curve $q$, with velocity vector field $v, \frac{Dy}{dt} = \sum_{i=1}^n \dot{y}^i(t)X_i(q) + \sum_{i=1}^n y^i(t)(\nabla_v X_i)(q)$, or in shorthand is written as $\frac{Dy}{dt} = \dot{y} + \nabla_v y$. Using this shorthand, $\frac{Dy}{dt}|_{t=0} = \delta y + \nabla_w y$. Similarly, for any smooth covector field $\alpha = \sum_{i=1}^n \alpha^i(t)\omega_i(q)$ along the curve $q$, with velocity vector field $v, \frac{Da}{dt} = \sum_{i=1}^n \dot{\alpha}^i(t)\omega_i(q) + \sum_{i=1}^n \alpha^i(t)(\nabla_v \omega_i)(q)$, or in shorthand is written as $\frac{Da}{dt} = \dot{\alpha} + \nabla_v \alpha$. Using this shorthand, $\frac{Da}{dt}|_{t=0} = \delta \alpha + \nabla_w \alpha$. For more details, see [15]. Before we proceed further, we need a few lemmas.

Lemma 4.1 ([15]). Given $y(\cdot) \in C^1([0, T], T_{q(\cdot)}\mathcal{Q})$ and $\alpha(\cdot) \in C^1([0, T], T_{q(\cdot)}^*\mathcal{Q})$, then

$$\int_0^T \alpha \left( \frac{Dy}{dt} \right) \; dt = \int_0^T \left[ \frac{Da}{dt}(\delta y) + \alpha(\nabla_\delta y) \right] \; dt.$$

Lemma 4.2 ([16], [26]). If the connection $\nabla$ is symmetric, then

$$\frac{D}{dt} \frac{\partial q_\epsilon}{\partial \epsilon} = \frac{D}{dt} \frac{\partial q_\epsilon}{\partial \epsilon}.$$

The necessary conditions for a normal extremal (see, e.g., [6]) for (P) are obtained by setting

$$\left. \frac{dJ^\alpha}{d\epsilon} \right|_{\epsilon=0} = 0,$$

where

$$J^\alpha_{\epsilon} = \int_0^T \left[ \frac{1}{2} \langle u_\epsilon, u_\epsilon \rangle + \lambda_1 \left( \frac{\partial q_\epsilon}{\partial t} - v_\epsilon \right) + \lambda_2 \left( \frac{Dv_\epsilon}{dt} - u_\epsilon \right) \right] dt.$$
The above condition, with the use of Lemmas 4.1–4.2, gives the following

\[
\frac{dJ}{dt}{\bigg|}_{\epsilon=0} = \int_0^T \left[ \langle u, \delta u + \nabla_w u \rangle + \lambda_1 \left( \frac{Dw}{dt} - \delta v - \nabla_w v \right) + \lambda_2 \left( \frac{Dv}{dt} + \nabla_w \frac{Dv}{dt} - \delta u - \nabla_w u \right) \right] dt
\]

\[
= \int_0^T \left[ \langle u, \nabla_w u \rangle + \langle u, \delta u \rangle + \lambda_1 \left( \frac{Dw}{dt} - \nabla_w v \right) - \lambda_1(\delta v) + \lambda_2 \left( \nabla_w \frac{Dv}{dt} - \nabla_w u \right) \right] dt
\]

\[
= \int_0^T \left[ -\frac{D\lambda_1}{dt}(w) - \lambda_1(\nabla_w v) + \lambda_2 \left( \nabla_w \frac{Dv}{dt} - \nabla_w u \right) + \langle u, \nabla_w u \rangle - \frac{D\lambda_2}{dt}(\delta v) - \lambda_1(\delta v) + \lambda_2(\nabla_\delta v) + \langle u, \delta u \rangle - \lambda_2(\delta u) \right] dt,
\]

where we have used integration by parts along with the fact that the one-parameter variation \( q \) is proper. We are now ready to state a theorem.

**Theorem 4.3** ([15]). A normal extremal for \( (P) \) satisfies the following equations

\[
\frac{dq}{dt} = v, \quad (22)
\]

\[
\frac{Dv}{dt} = u, \quad (23)
\]

\[
\frac{D\lambda_1}{dt} = -\lambda_1(\nabla v) + \lambda_2(\nabla \lambda_2^T), \quad (24)
\]

\[
\frac{D\lambda_2}{dt} = -\lambda_1 + \lambda_2(\nabla v), \quad (25)
\]

where \( u = \lambda_2^T \).

We assume that the nominal solution has been obtained for a fixed initial condition. Suppose there is a small variation in the initial condition and we would like to update the optimal control for \( (P) \). Instead of solving \( (P) \) from scratch, we employ NEOC as described previously. We will now fix some more notation.

**4.2. Notation.** In what follows, we use superscript \( n \) to denote the nominal trajectory and the corresponding vector and covector fields. The one-parameter variation of \( q^n(t) \) is denoted by \( q^n_{\epsilon} \). Note that the one-parameter variation of \( q^n(t) \) is not proper as there is a small variation in the initial condition. The vector field \( v^n(t) := \frac{\partial q^n}{\partial t}(t, 0) \) is the velocity vector field along \( q^n(t) \) and the vector field \( w^n(t) := \frac{\partial q^n}{\partial \epsilon}(t, 0) \) is the variation vector field associated with the one-parameter variation \( q^n_{\epsilon} \).

Employing the NEOC approach described previously, the variational equations for (22)-(25) are given as follows

\[
\frac{D}{\partial \epsilon} \left. \frac{\partial q^n}{\partial t} \right|_{\epsilon=0} = \left. \frac{Dv^n}{\partial t} \right|_{\epsilon=0}, \quad (26)
\]

\[
\frac{D}{\partial \epsilon} \left. \frac{Dv^n}{\partial t} \right|_{\epsilon=0} = \left. \frac{D\lambda_2^n}{\partial \epsilon} \right|_{\epsilon=0}, \quad (27)
\]
\[ \frac{D}{\partial \epsilon} \frac{D \lambda_{1,\epsilon}}{\partial t} \bigg|_{\epsilon=0} = \frac{D}{\partial \epsilon} \left( -\lambda_{1,\epsilon}^n (\nabla v^n_\epsilon) + \lambda_{2,\epsilon}^n (\nabla \lambda_{2,\epsilon}^n) \right) \bigg|_{\epsilon=0}, \] (28)

\[ \frac{D}{\partial \epsilon} \frac{D \lambda_{2,\epsilon}}{\partial t} \bigg|_{\epsilon=0} = \frac{D}{\partial \epsilon} \left( -\lambda_{1,\epsilon}^n + \lambda_{2,\epsilon}^n (\nabla v^n_\epsilon) \right) \bigg|_{\epsilon=0}. \] (29)

Note that the change in the control trajectory corresponding to the change in the initial condition is given by \( \frac{D \lambda_{2,\epsilon}^n}{\partial \epsilon} \bigg|_{\epsilon=0} \). Before we proceed further, we need a few lemmas.

**Lemma 4.4** ([16], [26]). Given any smooth vector field \( y \) along \( q_\epsilon \), then

\[ \frac{D}{\partial \epsilon} \frac{D y}{\partial t} - \frac{D}{\partial \epsilon} \frac{D y}{\partial \epsilon} = R \left( \frac{\partial q_\epsilon}{\partial \epsilon}, \frac{\partial q_\epsilon}{\partial t} \right) y. \]

**Remark 3.** Note that the definition of the curvature tensor of the connection \( \nabla \) used in this paper, differs by a negative sign from the one defined in [16], [26].

**Lemma 4.5** ([15]). Given \( y, z \in \mathfrak{X}(Q) \) and \( \alpha \in \mathfrak{X}^*(Q) \), then

\[ \frac{D}{\partial \epsilon} \alpha (\nabla z y) = \frac{D \alpha}{\partial \epsilon} (\nabla z y) + \alpha \left( \frac{D}{\partial \epsilon} (\nabla z y) - \nabla \frac{\partial \alpha}{\partial \epsilon} y \right). \]

**Remark 4.** Note that the expression, \( \frac{D}{\partial \epsilon} (\nabla z y) - \nabla \frac{\partial \alpha}{\partial \epsilon} y \) in Lemma 4.5 represents the second covariant derivative.

We are now ready to state two theorems.

**Theorem 4.6.** The variational equations (26)-(29) give the following two-point boundary value problem (TPBVP)

\[ \dot{w}^n = \delta v^n + [w^n, v^n], \] (30)

\[ \delta \dot{v}^n + \nabla_{w^n} \delta v^n + \nabla_{w^n} v^n + \nabla_{v^n} \dot{v}^n + \nabla_{v^n} \nabla_{w^n} v^n = \delta \lambda_{2}^{n}\epsilon + \nabla_{w^n}\lambda_{2}^{n}\epsilon, \] (31)

\[ (\delta \dot{\lambda}_{1}^{n} + \nabla_{w^n} \dot{\lambda}_{1}^{n} + \nabla_{\dot{v}^{n}} \lambda_{1}^{n} + \nabla_{v^n} \delta \lambda_{1}^{n} + \nabla_{w^n} \nabla_{v^n} \lambda_{1}^{n})(z) = (\delta \lambda_{1}^{n} - \nabla_{w^n} \lambda_{1}^{n}) (\nabla_{z} v^n) - \lambda_{1}^{n} (\nabla_{z} \delta v^n + \nabla_{w^n} \nabla_{z} v^n - \nabla_{w^n} \nabla_{\dot{v}^{n}} v^n) + (\delta \lambda_{2}^{n} + \nabla_{w^n} \lambda_{2}^{n} (\nabla_{z} \delta v^n + \nabla_{w^n} \lambda_{2}^{n} (\nabla_{z} v^n) + \nabla_{w^n} \nabla_{z} \lambda_{2}^{n} (\nabla_{z} v^n) + \nabla_{w^n} \nabla_{z} \lambda_{2}^{n} (\nabla_{w^n} v^n)), \] (32)

\[ (\delta \dot{\lambda}_{2}^{n} + \nabla_{w^n} \delta \lambda_{2}^{n} + \nabla_{\dot{v}^{n}} \lambda_{2}^{n} + \nabla_{v^n} \delta \lambda_{2}^{n} + \nabla_{w^n} \nabla_{v^n} \lambda_{2}^{n})(z) = (\delta \lambda_{1}^{n} - \nabla_{w^n} \lambda_{1}^{n}) (\nabla_{z} v^n) + (\delta \lambda_{2}^{n} + \nabla_{w^n} \lambda_{2}^{n} (\nabla_{z} v^n) + \lambda_{2}^{n} (\nabla_{z} \delta v^n + \nabla_{w^n} \nabla_{z} v^n - \nabla_{w^n} \nabla_{\dot{v}^{n}} v^n), \] (33)

where \( z \in \mathfrak{X}(Q) \).

**Proof.** Using Lemma 4.2, (26) can be re-written as follows

\[ \frac{D}{\partial \epsilon} \frac{D q_\epsilon^n}{\partial t} \bigg|_{\epsilon=0} = \frac{D}{\partial \epsilon} \frac{D v_\epsilon^n}{\partial \epsilon} \bigg|_{\epsilon=0}. \]

The above equation gives the following

\[ \dot{w}^n = \delta v^n + \nabla_{w^n} v^n - \nabla_{v^n} w^n. \]

Using the symmetry of the connection \( \nabla \), the above equation can be re-written as follows

\[ \dot{w}^n = \delta v^n + [w^n, v^n]. \]
Using Lemma 4.4, (27) can be re-written as follows

\[
\left. \frac{D}{\partial t} \frac{Dv^n}{\partial \epsilon} \right|_{\epsilon=0} + R(w^n, v^n)v^n = \left. \frac{D\lambda_2^{n^2}}{\partial \epsilon} \right|_{\epsilon=0},
\]

where \(R(w^n, v^n)v^n := \nabla_{w^n} \nabla_{v^n} v^n - \nabla_{v^n} \nabla_{w^n} v^n - \nabla_{[w^n, v^n]} v^n\). The above equation gives the following

\[
\delta \dot{v}^n + \nabla_{\dot{w}^n} v^n + \nabla_{w^n} \dot{v}^n + \nabla_{v^n} \dot{v}^n + \nabla_{v^n} \nabla_{w^n} v^n + R(w^n, v^n)v^n = \delta \lambda_2^{n^2} + \nabla_{w^n} \lambda_2^{n^2}.
\]

Substituting \(\ddot{w}^n = \delta v^n + [w^n, v^n]\) into the above equation, gives the following

\[
\delta \dot{v}^n + \nabla_{v^n} \delta v^n + \nabla_{\dot{v}^n} v^n + \nabla_{w^n} \dot{v}^n + \nabla_{w^n} \nabla_{v^n} v^n = \delta \lambda_2^{n^2} + \nabla_{w^n} \lambda_2^{n^2}.
\]

Similarly, the other two variational equations can be derived using Lemma 4.5. \(\square\)

**Theorem 4.7.** The variational equations (30)-(31) give the following Jacobi equation

\[
\ddot{w}^n + 2\nabla_{v^n} \dot{w}^n + \nabla_{\dot{w}^n} w^n + \nabla_{w^n} \nabla_{v^n} w^n + R(w^n, v^n)v^n = \delta \lambda_2^{n^2} + \nabla_{w^n} \lambda_2^{n^2}. \quad (34)
\]

**Proof.** Substituting (30) into (31), gives the following

\[
\ddot{w}^n - [\dot{w}^n, v^n] + \nabla_{v^n} \dot{w}^n - \nabla_{\dot{w}^n} w^n + \nabla_{w^n} v^n - \nabla_{[w^n, v^n]} v^n + \nabla_{w^n} \nabla_{v^n} v^n = \delta \lambda_2^{n^2} + \nabla_{w^n} \lambda_2^{n^2}.
\]

Using the symmetry of the connection \(\nabla\), the above equation can be re-written as follows

\[
\ddot{w}^n + \nabla_{v^n} \dot{w}^n - \nabla_{\dot{w}^n} w^n + \nabla_{\dot{w}^n} w^n - \nabla_{w^n} \dot{w}^n + \nabla_{v^n} \dot{w}^n + \nabla_{v^n} \nabla_{w^n} w^n - \nabla_{v^n} \nabla_{w^n} w^n + \nabla_{w^n} v^n - \nabla_{[w^n, v^n]} v^n + \nabla_{w^n} \nabla_{v^n} v^n = \delta \lambda_2^{n^2} + \nabla_{w^n} \lambda_2^{n^2}.
\]

Using the definition of the curvature tensor of the connection \(\nabla\), the above equation can be re-written as follows

\[
\ddot{w}^n + 2\nabla_{v^n} \dot{w}^n + \nabla_{v^n} w^n + \nabla_{w^n} \nabla_{v^n} w^n + R(w^n, v^n)v^n = \delta \lambda_2^{n^2} + \nabla_{w^n} \lambda_2^{n^2},
\]

where \(R(w^n, v^n)v^n := \nabla_{w^n} \nabla_{v^n} v^n - \nabla_{v^n} \nabla_{w^n} v^n - \nabla_{[w^n, v^n]} v^n\). \(\square\)

**Remark 5.** It should be noted that (34) plays a crucial role in determining conjugate points for \((P)\). It is also worthwhile to note that (34) corresponds to (3.3) in Theorem 4 of [9], where the case of a Lie group has been considered but not in a control theoretic setting.

5. **Solution as a variational problem.** We will follow the same procedure as given in [6]. Before we proceed further, we need a lemma.

**Lemma 5.1 ([16], [26]).** Given \(w, x, y, z \in \mathcal{X}(\mathcal{Q})\), then

\[
\langle R(x, y)z, w \rangle = \langle R(w, z)y, x \rangle.
\]

The necessary conditions for a normal extremal for \((P)\) are obtained by setting

\[
\left. \frac{dJ_\epsilon}{d\epsilon} \right|_{\epsilon=0} = 0,
\]

where

\[
J_\epsilon = \frac{1}{2} \int_0^T \left\langle \frac{D^2 q_\epsilon}{\partial t^2}, \frac{D^2 q_\epsilon}{\partial t^2} \right\rangle dt.
\]
The above condition, with the use of Lemmas 4.2, 4.4, 5.1, gives the following
\[
\frac{dJ}{dt} \bigg|_{t=0} = \int_0^T \left< Dv, \frac{D^2w}{dt^2} + R(w, v) \right> dt \\
= \int_0^T \left< \frac{D^3v}{dt^3} + R \left( \frac{Dv}{dt}, v \right), v, w \right> dt,
\]
where we have used integration by parts twice along with the fact that the one-parameter variation \( q_\epsilon \) is proper. We are now ready to state a theorem.

**Remark 6.** It is sometimes appropriate to assume that \( Q \) is parallelizable (see, e.g., [6]). This means that there exist smooth vector fields \( \{X_i\}_{i=1}^n \) on \( Q \) such that the vectors \( \{X_i(q)\}_{i=1}^n \) form an orthonormal basis for \( T_qQ \), for all \( q \in Q \). Equivalently, the assumption that \( Q \) is parallelizable means that \( TQ \) is a trivial bundle. The assumption that \( Q \) is parallelizable is restrictive in some sense but it is satisfied for the case of Lie groups (see, e.g., [22]), which are of special interest.

**Theorem 5.2** ([27]). A necessary condition for a curve \( q(\cdot) \in C^2([0,T], Q) \) to be a normal extremal for \((P)\) is that the velocity vector field \( v = \frac{dq}{dt} \) satisfies the following equation
\[
\frac{D^3v}{dt^3} + R \left( \frac{Dv}{dt}, v \right) v = 0. \tag{35}
\]

**Remark 7.** In [15], it has been shown that (22)-(25) are equivalent to (35). In the case when \( Q = \mathbb{R}^n \), with the standard inner product, the covariant derivative is the usual derivative and \( R = 0 \). We now see that (35) simplifies to the equation \( \frac{Dv}{dt} = 0 \), which shows that each coordinate function of a normal extremal \( q \) for \((P)\) is a cubic spline. So, Theorem 5.2 may be viewed as a generalization of cubic splines to the setting of Riemannian manifolds (see, e.g., [27]). Also, in the case, when one considers \((P)\) with multiple way points (see, e.g., [14]), NEOC can be used to update the “cubic splines”, with respect to the changes in data.

We do not give all the details, as they are similar to the previous section. The variational equation for (35) is given as follows
\[
\frac{D}{d\epsilon} \left( \frac{D^3v^\epsilon}{dt^3} + R \left( \frac{Dv^\epsilon}{dt}, v^\epsilon \right) v^\epsilon \right) \bigg|_{\epsilon=0} = 0. \tag{36}
\]
Note that the change in the control trajectory corresponding to the change in the initial condition is given by \( \frac{D^2q^\epsilon}{dt^2} \bigg|_{\epsilon=0} \). We will now specialize the results to the case of Lie groups.

6. **Application to Lie groups.** We will now present NEOC for OCPs for mechanical systems evolving on Lie groups but before proceeding, we will fix some more notation.

6.1. **Notation.** We will denote an \( n \)-dimensional compact semisimple Lie group by \( G \) and its Lie algebra by \( \mathfrak{g} \). The left translation map on \( G \) is denoted by \( L_g \) and the tangent map of \( L_g \) at \( h \in G \) is denoted by \( T_hL_g \). The dual space of \( \mathfrak{g} \) (space of linear functionals \( \alpha : \mathfrak{g} \to \mathbb{R} \)) is denoted by \( \mathfrak{g}^* \). The map \( \text{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \) is the adjoint representation of \( \mathfrak{g} \). For \( x \in \mathfrak{g} \), the adjoint action of \( x \) on \( \mathfrak{g} \) is given by the endomorphism \( \text{ad}_x : \mathfrak{g} \to \mathfrak{g} \), with \( \text{ad}_x(y) = [x,y] \), for all \( y \in \mathfrak{g} \). The map
ad*: g \rightarrow \mathfrak{gl}(\mathfrak{g}^*) is the coadjoint representation of g. For x \in g, the coadjoint action of x on g^* is given by the endomorphism ad^*: g^* \rightarrow g^*, with ad^*_x(\alpha) = \alpha([x,y]), for all y \in g, \alpha \in g^*. The inverse of the exponential map is denoted by log. The trace of a matrix is denoted by tr. The 2-sphere is denoted by S^2.

Given x, y and z are left invariant vector fields on G and given \alpha is a left invariant one-form on G, then \nabla_x y = \frac{1}{2}[x, y], \nabla_x \alpha = -\frac{1}{2} \text{ad}^*_x \alpha (see, e.g., [15]) and \nabla([x, y], z) = -\frac{1}{4} ([x, [y, z]]) (see, e.g., [14], [31]).

Remark 8. Note that the adjoint representation is equivalent to the coadjoint representation for semisimple Lie algebras.

We will still retain the same notation (P), in the case when Q = G. We are now ready to state a lemma.

Lemma 6.1 ([15]). A normal extremal for (P) satisfies the following equations

\dot{g} = T_L L_g(v),
\dot{v} = u, \quad (37)
\dot{\lambda}_1 = \text{ad}^*_v \lambda_1, \quad (39)
\dot{\lambda}_2 = -\lambda_1, \quad (40)

where u = \lambda^2_2.

We assume that the nominal solution has been obtained for a fixed initial condition. Suppose there is a small variation in the initial condition and we would like to update the optimal control for (P). Instead of solving (P) from scratch, we employ NEOC as described previously. The variational equations for (37)-(40) are given as follows

\dot{w} = \delta v + [w, v], \quad (41)
\dot{\delta v} = \delta \lambda^2_2, \quad (42)
\dot{\delta \lambda}_1 = \text{ad}^*_v \lambda^n_1 + \text{ad}^*_v \delta \lambda^n_1, \quad (43)
\dot{\delta \lambda}_2 = -\delta \lambda^n_1. \quad (44)

To illustrate NEOC for OCPs for mechanical systems evolving on Lie groups, we now consider an example, which is a slightly modified form of the example presented in [15].

6.2. Numerical example. Consider the following OCP

\min_{u(\cdot)} J = \frac{1}{2} \int_0^T \langle u(t), u(t) \rangle dt \quad (45)

subject to

\dot{Q}(t) = Q(t) \Omega_1(t), \quad Q(0) = Q_0, \quad Q(T) = Q_T, \quad (46)
\dot{\Omega}_1(t) = u(t), \quad \Omega_1(0) = \Omega_{10}, \quad \Omega_1(T) = \Omega_{1T}, \quad (47)

where Q(\cdot) \in C^2([0, T], \text{SO}(n)), \Omega_1(\cdot) \in C^1([0, T], \mathfrak{so}(n)) and \langle x, y \rangle = \text{tr}(x^T y), for all x, y \in \mathfrak{so}(n). A normal extremal for the OCP (45)-(47) satisfies the following equations (see [15])

\dot{Q} = Q \Omega_1. \quad (48)
\[ \dot{\Omega}_1 = \lambda_2, \]  
\[ \dot{\lambda}_1 = -\lambda_1 \Omega_1^T, \]  
\[ \dot{\lambda}_2 = -\frac{1}{2}(Q^T\lambda_1 - \lambda_1^TQ), \]  
where \( \lambda_1(\cdot), \lambda_2(\cdot) \in C^1([0, T], \mathfrak{s}\mathfrak{o}(n)) \) and the optimal control \( u^* = \lambda_2 \). By hypothesizing a solution of the form \( \lambda_1 = Q\Omega_2 \), with \( \Omega_2(\cdot) \in C^1([0, T], \mathfrak{s}\mathfrak{o}(n)) \), (48)-(51) give the following equations

\[ \dot{Q} = Q\Omega_1, \]  
\[ \dot{\Omega}_1 = \lambda_2, \]  
\[ \dot{\Omega}_2 = [\Omega_2, \Omega_1], \]  
\[ \dot{\lambda}_2 = -\Omega_2, \]  
which are in the form of (37)-(40). For more details, see [15]. We assume that the nominal solution has been obtained for a fixed initial condition \( [Q(0) \quad \Omega_1(0)]^T = [Q_0 \quad \Omega_{10}]^T \). Suppose there is a small variation in the initial condition, i.e., \( [Q(0) \quad \Omega_1(0)]^T = [Q_0 \quad \Omega_{10} + \Omega_{101}]^T \), where \( Q_0 \in \mathfrak{s}\mathfrak{o}(n) \) and \( \Omega_{10} \in \mathfrak{s}\mathfrak{o}(n) \). We would now like to update the optimal control for the OCP (45)-(47). Instead of solving the OCP (45)-(47) from scratch, we employ NEOC as described previously. The variational equations for (52)-(55) are given as follows

\[ \dot{w}^n = \delta \Omega^n_1 + [w^n, \Omega^n_1], \]  
\[ \delta \Omega^n_1 = \delta \lambda^n_2, \]  
\[ \delta \Omega^n_2 = [\delta \Omega^n_2, \Omega^n_1] + [\Omega^n_2, \delta \Omega^n_1], \]  
\[ \delta \lambda^n_2 = -\delta \Omega^n_2, \]  
with \( w^n(0) = \log(Q_0), \ w^n(T) = 0_{n \times n}, \ \delta \Omega^n_1(0) = \Omega_{10} \) and \( \delta \Omega^n_2(T) = 0_{n \times n} \). Note that the change in the control trajectory corresponding to the change in the initial condition is given by \( \delta \lambda^n_2 \).

**Remark 9.** Note that we have only considered an initial condition variation and to obtain the results shown in the subsequent figures, we have solved all the TPBVPs using fsolve.m.

We will now present simulation results for the case when \( n = 3 \), with \( T = 10 \) (sec) and the following data

\[ Q_0 = \exp(v_1^x), \ v_1 = [0.25 \ 0.5 \ 0.5]^T, \]  
\[ Q_T = \exp(v_3^x), \ v_2 = [0 \ 0 \ 0]^T, \]  
\[ \Omega_{10} = v_4^x, \ v_3 = [0.1 \ 0.1 \ 0.1]^T, \]  
\[ \Omega_{10} = v_4^x, \ v_4 = [0 \ 0 \ 0]^T, \]  
\[ \Omega_{10} = 0 \]  
\[ Q_T = \exp(v_6^x), \ v_5 = [0 \ 0 \ 0]^T, \]  
\[ \Omega_{10} = v_6^x, \ v_6 = [0.01 \ 0.01 \ 0.01]^T. \]

In the subsequent figure (Figure 3), the attitude maneuver is plotted on \( S^2 \), where the vectors \([x \ y \ z]^T\) corresponding to the first, second and third column of \( Q_0 \) are plotted in dashed-red, dashed-green and dashed-blue, respectively. Similarly, the vectors \([x \ y \ z]^T\) corresponding to the first, second and third column of \( Q_T \) are...
plotted in red, green and blue, respectively. For all other $Q(t), t \in (0, T)$, only the coordinates are shown in the corresponding colors.

Figure 1 shows the trajectories of $\Omega_1$ obtained by re-solving the OCP (45)-(47) and using NEOC. Figure 2 shows the trajectories of $u$ obtained by re-solving the OCP (45)-(47) and using NEOC. Figure 3 shows the attitude maneuver obtained by re-solving the OCP (45)-(47) and using NEOC. Figure 4 shows the error in the trajectories of $\Omega_1$ and $u$ obtained by re-solving the OCP (45)-(47) and using NEOC.

The time taken to re-solve the OCP (45)-(47) is 5.58 (sec) approximately and using NEOC is 3.58 (sec) approximately on a 3.4 GHz Intel Core i7-2600K desktop computer with 16 GB of RAM.

7. Conclusions and future work. In this paper, we extended NEOC, which is well established for OCPs defined on a Euclidean space to the setting of Riemannian manifolds. We further specialized the results to the case of Lie groups. We presented
an example, along with simulation results, which validates the ideas presented in this paper.

NEOC described in this paper only gives a prediction step and not a correction step. To improve the solution, a prediction step can be augmented by a correction step. In the future, we intend to extend the method presented in this paper to include a correction step as well along with the generalization to a more general cost function, with a more complicated dynamic constraint (such as the one considered in [21]).

**Acknowledgments.** The research of Anthony M. Bloch was supported by NSF grants DMS-0907949, DMS-1207693, DMS-1613819, INSPiRE-1343720 and the Simons Foundation. The research of Ilya V. Kolmanovsky was supported by NSF grants CMMI-1130160, CNS-1544844 and ECCS-1404814. We would also like to thank the referees for their valuable comments and suggestions.
REFERENCES

[1] R. Abraham and J. E. Marsden, *Foundations of Mechanics*, AMS Chelsea Publishing, 1978.
[2] A. A. Agrachev and Y. L. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer Science & Business Media, 2004.
[3] C. Altafini, Reduction by group symmetry of second order variational problems on a semidirect product of Lie groups with positive definite Riemannian metric, *ESAIM: Control, Optimization and Calculus of Variations*, 10 (2004), 526–548.
[4] M. Barbero-Liñán, *A Geometric Study of Abnormality in Optimal Control Problems for Control and Mechanical Control Systems*, PhD thesis, Technical University of Catalonia, 2008.
[5] M. Barbero-Liñán, Characterization of accessibility for affine connection control systems at some points with nonzero velocity, in *Proceedings of IEEE Conference on Decision and Control and European Control Conference*, 2011, 6526–6533.
[6] A. M. Bloch, *Nonholonomic Mechanics and Control*, Springer Science & Business Media, 2003.
[7] J. V. Breakwell and H. Yu-Chi, On the conjugate point condition for the control problem, *International Journal of Engineering Science*, 2 (1965), 565–579.
[8] A. E. Bryson, *Applied Optimal Control: Optimization, Estimation and Control*, CRC Press, 1975.
[9] F. Bullo, *Invariant Affine Connections and Controllability on Lie Groups*, Technical Report Final Project Report for CIT-CDS 141a, Control and Dynamical Systems, California Institute of Technology, 1995.
[10] F. Bullo and A. D. Lewis, *Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Simple Mechanical Control Systems*, Springer Science & Business Media, 2005.
[11] F. Bullo and A. D. Lewis, Reduction, linearization, and stability of relative equilibria for mechanical systems on Riemannian manifolds, *Acta Applicandae Mathematicae*, 99 (2007), 53–95.
[12] J.-B. Caillau, O. Cots and J. Gergaud, Differential continuation for regular optimal control problems, *Optimization Methods and Software*, 27 (2012), 177–196.
[13] N. Caroff and H. Frankowska, Conjugate points and shocks in nonlinear optimal control, *Transactions of the American Mathematical Society*, 348 (1996), 3133–3153.
[14] P. Crouch and F. Silva Leite, The dynamic interpolation problem: On Riemannian manifolds, Lie groups, and symmetric spaces, *Journal of Dynamical and Control Systems*, 1 (1995), 177–202.
[15] P. Crouch, F. Silva Leite and M. Camarinha, A second order Riemannian variational problem from a Hamiltonian perspective, 1998.
[16] M. P. do Carmo, *Riemannian Geometry*, Birkhäuser, 1992.
[17] A. L. Dontchev and W. W. Hager, Lipschitzian stability in nonlinear control and optimization, *SIAM Journal on Control and Optimization*, 31 (1993), 569–603.
[18] A. L. Dontchev, W. W. Hager, A. B. Poore and B. Yang, Optimality, stability, and convergence in nonlinear control, *Applied Mathematics and Optimization*, 31 (1995), 297–326.
[19] A. L. Dontchev and W. W. Hager, Lipschitzian stability for state constrained nonlinear optimal control, *SIAM Journal on Control and Optimization*, 36 (1998), 698–718.
[20] R. Gupta, A. M. Bloch and I. V. Kolmanovsky, Combined homotopy and neighboring extremal optimal control, *Optimal Control Applications and Methods*, (2016), to appear.
[21] R. V. Iyer, R. Holsapple and D. Domani, Optimal control problems on parallelizable Riemannian manifolds: Theory and applications, *ESAIM: Control, Optimisation and Calculus of Variations*, 12 (2006), 1–11.
[22] J. M. Lee, *Introduction to Smooth Manifolds*, Springer-Verlag, New York, 2003.
[23] P. D. Loewen and H. Zheng, Generalized conjugate points for optimal control problems, *Nonlinear Analysis: Theory, Methods & Applications*, 22 (1994), 771–791.
[24] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*, Springer Science & Business Media, 1999.
[25] P. M. Mereau and W. F. Powers, Conjugate point properties for linear quadratic problems, *Journal of Mathematical Analysis and Applications*, 55 (1976), 418–433.
[26] J. W. Milnor, *Morse Theory*, Princeton University Press, 1963.
[27] L. Noakes, G. Heinzinger and B. Paden, Cubic splines on curved spaces, *IMA Journal of Mathematical Control and Information*, 6 (1989), 465–473.
[28] S. Sasaki, On the differential geometry of tangent bundles of Riemannian manifolds I, *Tohoku Mathematical Journal, Second Series*, **10** (1958), 338–354.

[29] S. Sasaki, On the differential geometry of tangent bundles of Riemannian manifolds II, *Tohoku Mathematical Journal, Second Series*, **14** (1962), 146–155.

[30] H. Schättler and U. Ledzewicz, *Geometric Optimal Control: Theory, Methods and Examples*, Springer Science & Business Media, 2012.

[31] F. Silva Leite, M. Camarinha and P. Crouch, Elastic curves as solutions of Riemannian and sub-Riemannian control problems, *Mathematics of Control, Signals, and Systems*, **13** (2000), 140–155.

[32] J. L. Speyer and D. H. Jacobson, *Primer on Optimal Control Theory*, SIAM, 2010.

[33] D. R. Tyner and A. D. Lewis, Geometric jacobian linearization and LQR theory, *Journal of Geometric Mechanics*, **2** (2010), 397–440.

[34] V. Zeidan and P. Zezza, The conjugate point condition for smooth control sets, *Journal of Mathematical Analysis and Applications*, **132** (1988), 572–589.

[35] V. Zeidan and P. Zezza, Conjugate points and optimal control: Counterexamples, *IEEE Transactions on Automatic Control*, **34** (1989), 254–256.

[36] V. Zeidan, The riccati equation for optimal control problems with mixed state-control constraints: Necessity and sufficiency, *SIAM Journal on Control and Optimization*, **32** (1994), 1297–1321.

Received October 2015; revised June 2016.

E-mail address: abloch@umich.edu
E-mail address: rohitgpt@umich.edu
E-mail address: ilya@umich.edu