2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

Optimal control of under-actuated systems with application to Lie groups

I. I. Hussein A. M. Bloch

***Abstract*— In this paper we study a class of optimal control problems known as the τ-elastic variational problem for second order, under-actuated systems. After introducing and stating the problem, we derive the necessary optimality conditions using two approaches. The first approach is purely variational where the resulting necessary conditions are represented by a single fourth order differential equation. In the second approach, we use the Lagrange multiplier technique. In this case, the necessary conditions are represented by a set of four first order differential equations. We show that the two results are equivalent. Finally, we further specialize the result for the compact semi-simple Lie group case and use $SO(3)$ as an example. We also make some remarks on the $SE(3)$ case, which is the subject of current research.**

I. Introduction

In this paper we use differential geometric techniques on Riemannian manifolds to obtain necessary conditions for a class of optimal control problems. This class of optimal control problems is known in the mathematical literature as τ-elastic variational problems (see [1] and [2] and references therein for more on the τ-elastic variational problem). Interest in the τ-elastic variational problem is two-fold.

The first is pure interest in the mathematical and theoretical implications of this problem as the resulting necessary conditions represent elastic curves that deviate from geodesic curves joining the boundary points. Secondly, the authors are generally interested in applying the results to multi-spacecraft, especially dual spacecraft, formation flying for imaging applications. As will be seen in the next section, the cost functional in the τ-elastic variational problem is a weighted sum of fuel expenditure and the relative speed between the spacecraft pair. In interferometric imaging, relative speed is inversely proportional to the attained signal-to-noise ratio (see [3] and [4] and references therein). Hence, the optimal control problem is suitable for the motion path planning problem of a two-spacecraft formation. Modeling the formation as a pair of fully-actuated point particles has been treated in [2] and [5]. One may also model the two-spacecraft as a pair of rigid bodies evolving in three-dimensional space and, hence, as a system in $SE(3) \times SE(3)$ (two copies of the three-dimensional special Euclidean group). This however is a much harder problem and is subject to current investigation.

This research is supported by NSF grants DMS-0103895, DMS-0305837 and CMS-0408542.

Islam Hussein, Aerospace Engineering, University of Michigan, Ann Arbor, ihussein@umich.edu.

Anthony Bloch, Professor of Mathematics, Mathematics Department, University of Michigan, Ann Arbor, abloch@umich.edu.

We focus our attention on the case where the system is under-actuated, that is, when the control vector spans a subspace of the tangent space at a point on the manifold. Optimal kinematic control problems on Riemannian manifolds with under-actuated systems are known in the literature as the sub-Riemannian optimal control problem ([6]). In [6], the authors study a restricted version of the cost functional we consider in this paper (setting $\tau = 0$). Moreover, the authors in that paper consider systems satisfying first order (that is, kinematic) differential equations. We, on the other hand, study second order (dynamic) systems.

Here is how the paper is organized. In Section (II), we state the problem and describe it in more detail. In Section (III), we provide some preliminary definitions, facts and lemmas. In Section (IV), we derive the necessary conditions for the general problem. This is done following two approaches. One results in a single fourth order differential equation and the second in four first order differential equations as necessary conditions. We show that these results are indeed equivalent. In Section (V), we specialize the result to the compact semi-simple Lie group case, where we relate current results to those obtain previously in the literature. We use the $SO(3)$ Lie group case as an example. We conclude the paper with a summary and final remarks on future research in Section (VI).

II. Problem Statement

Let M be a smooth ($\mathcal{C}^\infty$) n-dimensional *Riemannian manifold* with the *Riemannian metric* denoted by $\langle \cdot, \cdot \rangle$ for a point $\mathbf{q} \in M$. Thus the length of a tangent vector $\mathbf{v} \in T_{\mathbf{q}}M$ is denoted by $\|\mathbf{v}\| = \langle \mathbf{v}, \mathbf{v} \rangle^{1/2}$, where $T_{\mathbf{q}}M$ is the tangent space of M at $\mathbf{q}$. The *Riemannian connection* on M, denoted ∇, is a mapping that assigns to any two smooth vector fields $\mathbf{X}$ and $\mathbf{Y}$ in M a new vector field, $\nabla_{\mathbf{X}}\mathbf{Y}$. For the properties of ∇, we refer the reader to [7], [8], [1], and [9]. We take ∇ to be the Levi-Civita connection. Γ_{ij}^k denote the connections coefficients. The operator $\nabla_{\mathbf{X}}$, which assigns to every vector field $\mathbf{Y}$ the vector field $\nabla_{\mathbf{X}}\mathbf{Y}$, is called the *covariant derivative of* $\mathbf{Y}$ *with respect to* $\mathbf{X}$. We will denote by $[\mathbf{X}, \mathbf{Y}]$ the *Lie bracket* of the vector fields $\mathbf{X}$ and $\mathbf{Y}$ and is defined by the identity: $[\mathbf{X}, \mathbf{Y}]f = \mathbf{X}(\mathbf{Y}f) - \mathbf{Y}(\mathbf{X}f)$. Given vector fields $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ on M, define the vector field $R(\mathbf{X}, \mathbf{Y})\mathbf{Z}$ by the identity

$$R(\mathbf{X}, \mathbf{Y})\mathbf{Z} = \nabla_{\mathbf{X}}\nabla_{\mathbf{Y}}\mathbf{Z} - \nabla_{\mathbf{Y}}\nabla_{\mathbf{X}}\mathbf{Z} - \nabla_{[\mathbf{X},\mathbf{Y}]}\mathbf{Z}. \quad (2.1)$$

R is trilinear in $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ and is thus a tensor of type $(1, 3)$, which is called the *curvature tensor* of M.

We will use $\mathrm{D}/\mathrm{d}t$ and $\nabla_{\mathbf{v}}$ to denote the *covariant time derivative*. The manifold M is assumed to be parallelizable.

That is, there exists vector fields $\mathbf{X}_1(\mathbf{q}), \ldots, \mathbf{X}_n(\mathbf{q}) \in T_{\mathbf{q}}M$ at each point $\mathbf{q} \in M$ such that $\langle \mathbf{X}_i, \mathbf{X}_j \rangle = \delta_{ij}$ for all $\mathbf{q} \in M$, where δ is the Kronecker delta. In this paper we take $\mathbf{X}_i$ to be the standard basis $\mathbf{X}_i = \partial_i = \frac{\partial}{\partial q_i}$, where $\mathbf{q} = (q_1, q_2, \ldots, q_n)$.

In this paper we consider systems that satisfy dynamics of the form:

$$\begin{aligned} \frac{\mathrm{D}\mathbf{q}}{\mathrm{d}t}(t) &= \frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t}(t) = \mathbf{v}(t) \\ \frac{\mathrm{D}\mathbf{v}}{\mathrm{d}t}(t) &= \mathbf{u}(t), \end{aligned} \tag{2.2}$$

where we now view $\mathbf{q} : [0,T] \to M$ as a curve on M, $\mathbf{v}(t) \in T_{\mathbf{q}(t)}M$ and $\mathbf{u}(t) \in TT_{\mathbf{q}(t)}M$. In this paper, we are interested in the situation where $\mathbf{u}(t)$ is given by

$$\mathbf{u}(t) = \sum_{i=1}^{m} u_i(t)\mathbf{X}_i(\mathbf{q}(t)), \tag{2.3}$$

where $m < n$. Thus, we have:

$$\langle \mathbf{u}(t), \mathbf{u}(t) \rangle = \sum_{i=1}^{m} u_i^2(t). \tag{2.4}$$

$m = n$ corresponds to the fully actuated system, whereas $m < n$ corresponds to the under-actuated system. Different versions of the the optimal control problem with $m = n$ have been treated in the past. See, for example, [10], [11], [12], [12] and [5], [13], [2]. The case where $m < n$ has been treated in [14] and [6] for kinematic systems. In Section (I), we briefly described how our present work differs from that in [6]. We now state the optimal control problem.

Problem II.1. *Find critical values of*

$$\mathcal{J}^\tau(\mathbf{q}) = \frac{1}{2}\int_0^T \left[\langle \mathbf{u}, \mathbf{u}\rangle + \tau^2 \langle \mathbf{v}, \mathbf{v}\rangle\right] \mathrm{d}t, \tag{2.5}$$

over the set Ω of $\mathcal{C}^1$-paths $\mathbf{q}$ on M, satisfying

- *the* dynamic constraints *(2.2),*
- *$\mathbf{q}(t)$ is smooth for all $t \in [0,T]$,*
- boundary conditions

$$\begin{aligned} \mathbf{q}(0) = \mathbf{q}_0 \quad \mathbf{q}(T) = \mathbf{q}_T \\ \mathbf{v}(0) = \mathbf{v}_0 \quad \mathbf{v}(T) = \mathbf{v}_T, \end{aligned} \tag{2.6}$$

- *and the* motion constraints

$$\left\langle \frac{\mathrm{D}\mathbf{q}}{\mathrm{d}t}, \mathbf{X}_i(\mathbf{q}) \right\rangle = k_i, \; i = 1, \ldots, l \; (l < n) \tag{2.7}$$

for $\mathbf{X}_i$, $i = 1, \ldots, n$, linearly independent vector fields in some neighborhood of $\mathbf{q}(t)$ and given constants k_i, $i = 1, \ldots, l$.

Though not studied in the present paper, the above cost function is motivated by optimal path planning for dual spacecraft interferometric imaging formations (see [4], [2] and [5]). Our interest in the above cost function arises because in interferometric imaging applications, not only are we interested in minimizing fuel expenditure, but, also, in executing the maneuver with the smallest possible speed. While minimizing fuel expenditure is obvious, minimum speed trajectories are desired in interferometric imaging because the light collectors' speed and image quality (namely, achievable signal-to-noise ratio) are reciprocal; the larger the collectors' speeds are ("shorter exposure time"), the worse the image becomes, and vice versa (see [3] and [2]). This is analogous to exposure time in conventional photography, where longer exposure times (without spoiling the photographic film) result in more photon arrivals and a better image. The results presented in Section (IV) are general enough to be applied to the multi-spacecraft interferometric imaging problem, where the formation can be treated as either a set of two point particles or two bodies. In the latter case, the formation evolves on $SE(3)$, which is a non-compact Lie group and is subject of current research.

The compact Lie group case is also of interest in that many previous results (e.g., [6]) study the optimal control problem for Lie groups. It is interesting and instructive to compare our current generalized results with those appearing in previous literature. We do some of this in Section (V).

When $m < n$ and the control variables belong not to TTM but to TM and directly control the system speeds as opposed to the accelerations, the problem is known at the Sub-Riemannian kinematic optimal control problem. See for example [14], [6] and [9] and references therein for the treatment of these kinematic control problems. In this paper we restrict our attention to the dynamic (second order) version of the problem.

Moreover, we may ignore the motion constraints (2.7). As shown in [2], these constraints are automatically appended to the final expression for the necessary conditions corresponding to the unconstrained problem. Here again, once the necessary conditions are obtained for systems with a potential field, the motion constraints are simply appended to the necessary conditions.

III. Preliminaries

In this section, we first introduce the notion of a variational vector field. This allows us to introduce the necessary notation and tools to derive the necessary conditions for the optimal control problem (II.1).

Let Ω be the set of all $\mathcal{C}^1$ piecewise smooth curves $\mathbf{q} : [0,T] \to M$ in M satisfying the boundary conditions (2.6). The set Ω is called the *admissible set*. For the class of $\mathcal{C}^1$ curves in Ω we introduce the *$\mathcal{C}^1$ piecewise smooth one-parameter variation* of a curve $\mathbf{q} \in \Omega$ by

$$\begin{aligned} \mathbf{q}_\epsilon : [0,T] \times (-\epsilon, \epsilon) &\to M \\ (t, u) &\to \mathbf{q}(t, \epsilon) = \mathbf{q}_\epsilon(t). \end{aligned}$$

A vector field $\mathbf{Y}$ along a variation $\mathbf{q}_\epsilon$ is defined as the mapping that assigns to each $(t, \epsilon) \in [0,T] \times (-\rho, \rho)$ a tangent vector $\mathbf{Y}(t,\epsilon) \in T_{\mathbf{q}_\epsilon}M$. For example, the vector fields $\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}$ and $\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial t}$ are defined by

$$\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon} f = \frac{\mathrm{D}}{\partial \epsilon}(f \circ \mathbf{q}_\epsilon) \text{ and } \frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial t} f = \frac{\mathrm{D}}{\partial t}(f \circ \mathbf{q}_\epsilon),$$

respectively, where f is a $\mathcal{C}^\infty$ real-valued function on M. With $\epsilon = 0$, the vector fields $\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}$ and $\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial t}$ are now restricted to $\mathbf{q}$ and the $\mathcal{C}^1$ piecewise smooth vector field along $\mathbf{q}$, $\mathbf{v}(t) := \frac{\mathrm{D}}{\partial t}\mathbf{q}_\epsilon(t, 0)$, is the *velocity vector field along* $\mathbf{q}$. On the other hand, the $\mathcal{C}^1$ piecewise smooth vector field $\mathbf{W}_t = \mathbf{W}(t) := \frac{\mathrm{D}}{\partial \epsilon}\mathbf{q}_\epsilon(t, 0) \in T_{\mathbf{q}}\Omega$ is called the *variational vector field* associated with $\mathbf{q}_\epsilon$ along $\mathbf{q}$.

The one-parameter variation $\mathbf{q}_\epsilon$ is characterized infinitesimally by the vector space $T_{\mathbf{q}_\epsilon}\Omega$ by setting $\mathbf{q}_\epsilon(t) = \exp_{\mathbf{q}(t)}(\epsilon \mathbf{W}_t)$, where $\exp_{\mathbf{q}(t)}$ is the exponential map on M. $\mathbf{q}_\epsilon$ is said to be *admissible* if, for each $\epsilon \in (-\rho, \rho)$, the curve $\mathbf{q}_\epsilon$ satisfies the boundary conditions

$$\begin{aligned}
\mathbf{q}_\epsilon(t,0) &= \mathbf{q}(t), \\
\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}(t,0) &= \mathbf{W}_t \\
\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}(0,0) &= \frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}(T,0) = 0 \qquad (3.1)\\
\frac{\mathrm{D}}{\mathrm{d}t}\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}(t,0) &= \frac{\mathrm{D}}{\mathrm{d}t}\mathbf{W}_t \text{ is continuous on } [0,T] \\
\frac{\mathrm{D}}{\mathrm{d}t}\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}(0,0) &= \frac{\mathrm{D}}{\mathrm{d}t}\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}(T,0) = 0.
\end{aligned}$$

For subsequent theorems, we state without proof the following standard properties.

Fact III.1. *Let $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$ and $\mathbf{W}$ be vector fields, then the curvature tensor satisfies (see [15], page 53)*

$$\langle R(\mathbf{X},\mathbf{Y})\mathbf{Z},\mathbf{W}\rangle = \langle R(\mathbf{W},\mathbf{Z})\mathbf{Y},\mathbf{X}\rangle.$$

Fact III.2. *A one parameter variation $\mathbf{q}_\epsilon(t,u)$ satisfies (see [15], page 50)*

$$\frac{\mathrm{D}}{\partial \epsilon}\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial t} = \frac{\mathrm{D}}{\partial t}\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}.$$

Fact III.3. *Let $\mathbf{Y}$ be a vector field along $\mathbf{q}_\epsilon$, then (see [15], page 52)*

$$\frac{\mathrm{D}}{\partial \epsilon}\frac{\mathrm{D}}{\partial t}\mathbf{Y} - \frac{\mathrm{D}}{\partial t}\frac{\mathrm{D}}{\partial \epsilon}\mathbf{Y} = R\left(\frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial \epsilon}, \frac{\mathrm{D}\mathbf{q}_\epsilon}{\partial t}\right)\mathbf{Y}.$$

We also have the following lemma.

Lemma III.1. *Let $\mathbf{Z}$ be an arbitrary vector, $\mathbf{W}$ be a variational vector field and $\mathbf{v}$ be the velocity vector field. Then, we have*

$$\mathbf{Z}(\langle \mathbf{v},\mathbf{W}\rangle) = 0.$$

Proof In local coordinates, let $\mathbf{Z} = \sum_{k=1}^n \zeta_k \partial_k$, $\mathbf{v} = \sum_{i=1}^n v_i \partial_i$ and $\mathbf{W} = \sum_{j=1}^n w_j \partial_j$. Note that $w_j = \frac{\partial}{\partial \epsilon} q_j(t,\epsilon)$. Then we have:

$$\begin{aligned}
\mathbf{Z}(\langle \mathbf{v},\mathbf{W}\rangle) &= \sum_{k=1}^n \zeta_k \frac{\partial}{\partial q_k}\left(\sum_{i=1}^n v_i \frac{\partial q_i}{\partial \epsilon}\right) \\
&= \sum_{k=1}^n \zeta_k \sum_{i=1}^n v_i \frac{\partial^2 q_i}{\partial q_k \partial \epsilon} \\
&= 0,
\end{aligned}$$

where $\frac{\partial^2}{\partial q_k \partial \epsilon} q_i(t,\epsilon) = \frac{\partial^2}{\partial \epsilon \partial q_k} q_i(t,\epsilon) = \frac{\partial}{\partial \epsilon}\delta_{ik} = 0$ and $\frac{\partial v_i}{\partial q_k} = 0$ since the components of the velocity vector field $\mathbf{v}$ are independent of the local coordinates $q_1,\ldots,q_n$. ∎

Lemma III.2. *Let $\mathbf{Z}$ be an arbitrary vector, $\mathbf{W}$ be a variational vector field, $\mathbf{v}$ be the velocity vector field and $\mathbf{u} = \frac{\mathrm{D}\mathbf{v}}{\mathrm{d}t}$ be the control vector field. Then, we have*

$$\mathbf{Z}(\langle \mathbf{u},\mathbf{W}\rangle) = 0.$$

Proof of Lemma (III.2) is analogous to that of Lemma (III.1). Lemmas (III.1) and (III.2) and the identity

$$\mathbf{X}(\langle \mathbf{Y},\mathbf{Z}\rangle) = \langle \nabla_\mathbf{X}\mathbf{Y},\mathbf{Z}\rangle + \langle \mathbf{Y},\nabla_\mathbf{X}\mathbf{Z}\rangle$$

for arbitrary vector fields $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$, imply

$$\langle \mathbf{v},\nabla_\mathbf{Z}\mathbf{W}\rangle = -\langle \nabla_\mathbf{Z}\mathbf{v},\mathbf{W}\rangle \qquad (3.2)$$

and

$$\langle \mathbf{u},\nabla_\mathbf{Z}\mathbf{W}\rangle = -\langle \nabla_\mathbf{Z}\mathbf{u},\mathbf{W}\rangle. \qquad (3.3)$$

IV. Necessary Conditions for Optimality

In this section we first pursue a purely variational approach in deriving the necessary conditions for the problem (II.1) without the motion constraints (2.7). That is, Lagrange multipliers will not be introduced to the Lagrangian. This purely variational approach is used to derive necessary conditions in [2], [6] and [11]. However, to take into account the constraint that the control vector field $\mathbf{u}$ only spans a subspace of the tangent space to M at some point $\mathbf{q} \in M$, the following vector field will be introduced:

$$\mathbf{Z}_t = \sum_{k=m+1}^n \zeta_k(t)\mathbf{X}_k$$

such that

$$\left\langle \mathbf{X}_k, \frac{\mathrm{D}\mathbf{v}}{\mathrm{d}t}\right\rangle = \langle \mathbf{X}_k,\mathbf{u}\rangle = 0,\ k = m+1,\ldots,n \qquad (4.1)$$

and

$$\langle \mathbf{Z}_t,\mathbf{u}\rangle = 0. \qquad (4.2)$$

We will drop the subscript t in $\mathbf{Z}_t$ to become $\mathbf{Z}$.

Appending to the cost function the constraint (4.2), one obtains:

$$\mathcal{J}^\tau = \int_0^T \frac{1}{2}\langle \mathbf{u},\mathbf{u}\rangle + \frac{\tau^2}{2}\langle \mathbf{v},\mathbf{v}\rangle + \langle \mathbf{Z},\mathbf{u}\rangle\,\mathrm{d}t. \qquad (4.3)$$

A control vector $\mathbf{u}$ solves

$$\min_{\mathbf{u}} \mathcal{J}^\tau(\mathbf{q},\mathbf{v},\mathbf{u}) \qquad (4.4)$$

only if

$$\frac{\partial}{\partial \epsilon}\mathcal{J}^\tau(\mathbf{q}_\epsilon(t,\epsilon),\mathbf{v}_\epsilon(t,\epsilon),\mathbf{u}_\epsilon(t,\epsilon))\Big|_\epsilon = 0, \qquad (4.5)$$

where $\mathbf{v}_\epsilon(t,\epsilon)$ and $\mathbf{u}_\epsilon(t,\epsilon)$ are defined analogous to $\mathbf{q}_\epsilon(t,\epsilon)$ in the previous section. Replacing $\mathbf{u}$ with $\nabla_\mathbf{v}\mathbf{v}$ in Equation (4.3) and taking variations with respect to ϵ, one obtains:

$$\begin{aligned}
\frac{\partial \mathcal{J}^\tau}{\partial \epsilon}\Big|_{\epsilon=0} = \int_0^T &\left[\left\langle \frac{\mathrm{D}\mathbf{v}}{\partial t}, \frac{\mathrm{D}^2\mathbf{v}_\epsilon}{\partial \epsilon \partial t}\right\rangle + \tau^2\left\langle \mathbf{v}, \frac{\mathrm{D}\mathbf{v}_\epsilon}{\partial \epsilon}\right\rangle \right.\\
&\left. + \left\langle \frac{\mathrm{D}\mathbf{Z}}{\partial \epsilon}, \mathbf{u}_\epsilon\right\rangle + \left\langle \mathbf{Z}, \frac{\mathrm{D}^2\mathbf{v}_\epsilon}{\partial \epsilon \partial t}\right\rangle\right]_{\epsilon=0} \mathrm{d}t.
\end{aligned}$$

We set $\frac{\partial \mathbf{q}_\epsilon}{\partial \epsilon}\big|_{\epsilon=0} = \mathbf{W}$, use Facts (III.2) and (III.3) and the connection ∇ property $\frac{\mathrm{D}\mathbf{Z}}{\mathrm{d}\epsilon} = \nabla_\mathbf{W}\mathbf{Z} = \nabla_\mathbf{Z}\mathbf{W} + [\mathbf{W},\mathbf{Z}]$ so the right hand side of the above equation becomes

$$\begin{aligned}
&\int_0^T \left\langle \frac{\mathrm{D}\mathbf{v}}{\partial t} + \mathbf{Z}, R(\mathbf{W},\mathbf{v})\mathbf{v} + \frac{\mathrm{D}}{\partial t}\frac{\mathrm{D}\mathbf{v}}{\partial \epsilon}\right\rangle \\
&+\tau^2\left\langle \mathbf{v}, \frac{\mathrm{D}}{\partial t}\mathbf{W}\right\rangle + \langle \nabla_\mathbf{Z}\mathbf{W} + [\mathbf{W},\mathbf{Z}],\mathbf{u}\rangle\,\mathrm{d}t.
\end{aligned}$$

Using Fact (III.1) for the curvature in the first term, integrating the τ^2 term by parts and applying Lemma (III.2) to the last term in the integrand we obtain

$$\begin{aligned}
&\int_0^T \left\langle \mathbf{W}, R\left(\frac{\mathrm{D}\mathbf{v}}{\partial t} + \mathbf{Z}, \mathbf{v}\right)\mathbf{v}\right\rangle \\
&+\left\langle \frac{\mathrm{D}\mathbf{v}}{\partial t} + \mathbf{Z}, \frac{\mathrm{D}}{\partial t}\frac{\mathrm{D}}{\partial t}\mathbf{W}\right\rangle \\
&-\tau^2\left\langle \frac{\mathrm{D}\mathbf{v}}{\partial t}, \mathbf{W}\right\rangle + \mathbf{Z}\left(\left\langle \mathbf{W}, \frac{\mathrm{D}\mathbf{v}}{\partial t}\right\rangle\right) \\
&-\left\langle \mathbf{W}, \nabla_\mathbf{Z}\frac{\mathrm{D}\mathbf{v}}{\partial t}\right\rangle + \langle -[\mathbf{Z},\mathbf{W}],\mathbf{u}\rangle\,\mathrm{d}t.
\end{aligned}$$

Recall that the Lie bracket is skew-symmetric: $[\mathbf{W},\mathbf{Z}] = -[\mathbf{Z},\mathbf{W}]$ and that

$$\mathcal{L}_{\mathbf{Z}}\left\langle \mathbf{W}, \frac{\mathrm{D}}{\partial t}\mathbf{v}\right\rangle = \langle \mathcal{L}_{\mathbf{Z}}\mathbf{W},\mathbf{u}\rangle + \langle \mathbf{W},\mathcal{L}_{\mathbf{Z}}\mathbf{u}\rangle .$$

Observe, however, that

$$\mathcal{L}_{\mathbf{Z}}\left(\langle \mathbf{W},\mathbf{u}\rangle\right) = \mathbf{Z}\left(\langle \mathbf{W},\mathbf{u}\rangle\right) = 0$$

by Lemma (III.2). Hence, we have $\langle [\mathbf{Z},\mathbf{W}],\mathbf{u}\rangle = \langle \mathcal{L}_{\mathbf{Z}}\mathbf{W},\mathbf{u}\rangle = -\langle \mathbf{W},\mathcal{L}_{\mathbf{Z}}\mathbf{u}\rangle$. From this and by integrating the second inner product twice by parts, we obtain

$$\begin{aligned}&\int_0^T \langle \mathbf{W}, R(\mathbf{u}+\mathbf{Z},\mathbf{v})\mathbf{v}\rangle \\ &+\left\langle \nabla^3_{\mathbf{v}}\mathbf{v} + \nabla^2_{\mathbf{v}}\mathbf{Z},\mathbf{W}\right\rangle \\ &-\tau^2\left\langle \frac{\mathrm{D}\mathbf{v}}{\partial t},\mathbf{W}\right\rangle \\ &-\langle \mathbf{W},\nabla_{\mathbf{Z}}\mathbf{u}\rangle + \langle \mathbf{W},\mathcal{L}_{\mathbf{Z}}\mathbf{u}\rangle\,\mathrm{d}t.\end{aligned}$$

Collecting terms, we finally obtain

$$\begin{aligned}\left.\frac{\partial \mathcal{J}^\tau}{\partial \epsilon}\right|_{\epsilon=0} &= \int_0^T \Big\langle \mathbf{W}, R(\mathbf{u}+\mathbf{Z},\mathbf{v})\mathbf{v} + \nabla^3_{\mathbf{v}}\mathbf{v} + \nabla^2_{\mathbf{v}}\mathbf{Z} \\ &\quad -\tau^2\frac{\mathrm{D}\mathbf{v}}{\partial t} - \nabla_{\mathbf{Z}}\mathbf{u} + [\mathbf{Z},\mathbf{u}]\Big\rangle \mathrm{d}t. \end{aligned} \tag{4.6}$$

In obtaining Equation (4.6), repeated use has been made of the integration by parts identity, for example,

$$\int_0^T \left\langle \frac{\mathrm{D}}{\partial t}\mathbf{W},\frac{\mathrm{D}\mathbf{v}}{\partial t}\right\rangle = -\int_0^T\left\langle \mathbf{W},\frac{\mathrm{D}}{\partial t}\frac{\mathrm{D}\mathbf{v}}{\partial t}\right\rangle \mathrm{d}t,$$

and the fact that the variational vector field $\mathbf{W}$ is fixed at the boundary points 0 and T.

Since $\mathbf{W}$ is an arbitrary variational vector field, the condition (4.5) and Equation (4.6) immediately result in the main theorem of our paper.

Theorem IV.1. *A necessary condition for a control law $\mathbf{u}(t)$ to be an optimal solution for the problem (II.1) without the motion constraints (2.7), is that it satisfies the differential equation:*

$$\frac{\mathrm{D}^2\mathbf{u}}{\mathrm{d}t^2} + R(\mathbf{u}+\mathbf{Z},\mathbf{v})\mathbf{v} + \frac{\mathrm{D}^2\mathbf{Z}}{\mathrm{d}t^2} - \tau^2\mathbf{u} - \nabla_{\mathbf{Z}}\mathbf{u} + [\mathbf{Z},\mathbf{u}] = 0$$

and the condition (4.2).

It can be easily checked that this result reduces to previously published results in the literature. For example, if $m = n$, one can set $\mathbf{Z} = 0$ in the above theorem to obtain the necessary conditions for the fully actuated τ-elastic variational problem (see [2] and [11]).

We now derive the necessary conditions following the Lagrange multiplier approach and show that these are equivalent to those obtained in Theorem (IV.1).

First, we define a bilinear form $\mathbf{B}(\cdot,\cdot)$ that, for any vector field $\mathbf{Y}\sum_{i=1}^n y_i\mathbf{X}_i$, satisfies:

$$\begin{aligned}\nabla_{\mathbf{v}}\mathbf{Y} &= \dot{\mathbf{Y}} + \mathbf{B}(\mathbf{v},\mathbf{Y}) \\ \nabla_{\mathbf{W}}\mathbf{Y} &= \delta\mathbf{Y} + \mathbf{B}(\mathbf{W},\mathbf{Y}),\end{aligned}$$

where $\dot{\mathbf{Y}} = \sum_{i=1}^n \dot{y}_i\mathbf{X}_i$, $\delta\mathbf{Y} = \sum_{i=1}^n \frac{\partial y_i}{\partial\epsilon}\mathbf{X}_i$, $\mathbf{B}(\mathbf{v},\mathbf{Y}) = \sum_{i,j,k=1}^n v_iy_j\Gamma^k_{ij}\mathbf{X}_k$, $\mathbf{B}(\mathbf{W},\mathbf{Y}) = \sum_{i,j,k=1}^n w_iy_j\Gamma^k_{ij}\mathbf{X}_k$, and $\mathbf{W} = \sum_{i=1}^n w_i\mathbf{X}_i = \sum_{i=1}^n \frac{\partial q_{\epsilon_i}}{\partial\epsilon}\mathbf{X}_i$ is the variation vector field corresponding to the curve $\mathbf{q}(t)$. $\mathbf{B}(\mathbf{W},\cdot)$ is introduced in order to be able to separate variations in the components of a vector field, $\delta\mathbf{Y}$, from variations in the basis vector fields, which are contained in $\mathbf{B}(\mathbf{W},\mathbf{Y})$. It is important to separate these terms since the variations $\delta\mathbf{Y}$ are independent from $\mathbf{W}$. The notation $\mathbf{B}(\mathbf{v},\cdot)$ will be appreciated later when we treat the Lie group case.

We append the Lagrangian in Equation (2.5) with the dynamics (2.2) and the constraints (4.1) to obtain

$$\begin{aligned}\mathcal{J}^\tau &= \int_0^T \frac{1}{2}\langle\mathbf{u},\mathbf{u}\rangle + \frac{\tau^2}{2}\langle\mathbf{v},\mathbf{v}\rangle + \lambda_1\left(\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} - \mathbf{v}\right) \\ &\quad +\lambda_2\left(\frac{\mathrm{D}\mathbf{v}}{\mathrm{d}t} - \mathbf{u}\right) + \left\langle \mathbf{Z},\frac{\mathrm{D}\mathbf{v}}{\mathrm{d}t}\right\rangle \mathrm{d}t. \end{aligned}\tag{4.7}$$

Taking variations of this expression, one obtains

$$\begin{aligned}\frac{\partial\mathcal{J}^\tau}{\partial\epsilon} &= \langle \mathbf{u},\delta\mathbf{u} + \mathbf{B}(\mathbf{W},\mathbf{u})\rangle + \tau^2\langle\mathbf{v},\delta\mathbf{v} + \mathbf{B}(\mathbf{W},\mathbf{v})\rangle \\ &\quad +\lambda_1\left(\frac{\mathrm{D}\mathbf{W}}{\mathrm{d}t} - \delta\mathbf{v} - \mathbf{B}(\mathbf{W},\mathbf{v})\right) \\ &\quad +\lambda_2\left(\frac{\mathrm{D}^2\mathbf{v}}{\partial\epsilon\partial t} - \delta\mathbf{u} - \mathbf{B}(\mathbf{W},\mathbf{u})\right) \\ &\quad +\langle\nabla_{\mathbf{W}}\mathbf{Z},\mathbf{u}\rangle + \langle \mathbf{Z},\delta\mathbf{u} + \mathbf{B}(\mathbf{W},\mathbf{u})\rangle\,\mathrm{d}t. \end{aligned}\tag{4.8}$$

Now, note that

$$\begin{aligned}\int_0^T \lambda_1\left(\frac{\mathrm{D}\mathbf{W}}{\mathrm{d}t}\right)\mathrm{d}t &= -\int_0^T \frac{\mathrm{D}\lambda_1}{\mathrm{d}t}(\mathbf{W})\,\mathrm{d}t, \\ \int_0^T \lambda_2\left(\frac{\mathrm{D}^2\mathbf{v}}{\partial\epsilon\partial t}\right)\mathrm{d}t &= \int_0^T \lambda_2\left(\frac{\mathrm{D}^2\mathbf{v}}{\partial t\partial\epsilon} + R(\mathbf{W},\mathbf{v})\mathbf{v}\right) \\ &= \int_0^T -\frac{\mathrm{D}\lambda_2}{\mathrm{d}t}(\delta\mathbf{v} + \mathbf{B}(\mathbf{W},\mathbf{u})) \\ &\quad +\lambda_2(R(\mathbf{W},\mathbf{v})\mathbf{v})\,\mathrm{d}t \\ \langle\nabla_{\mathbf{W}}\mathbf{Z},\mathbf{u}\rangle &= \langle\nabla_{\mathbf{Z}}\mathbf{W} + [\mathbf{W},\mathbf{Z}],\mathbf{u}\rangle \\ &= -\langle\mathbf{W},\nabla_{\mathbf{Z}}\mathbf{u}\rangle + \langle\mathbf{W},[\mathbf{Z},\mathbf{u}]\rangle .\end{aligned}$$

Hence, we have

$$\begin{aligned}\frac{\partial\mathcal{J}^\tau}{\partial\epsilon} &= \int_0^T \langle\mathbf{u}+\mathbf{Z},\delta\mathbf{u}\rangle - \lambda_2(\delta\mathbf{u})\,\mathrm{d}t \\ &\quad +\int_0^T \tau^2\langle\mathbf{v},\delta\mathbf{v}\rangle - \lambda_1(\delta\mathbf{v}) - \frac{\mathrm{D}\lambda_2}{\mathrm{d}t}(\delta\mathbf{v}) \\ &\quad +\langle\mathbf{u},\delta\mathbf{u}\rangle + \tau^2\langle\mathbf{v},\mathbf{B}(\mathbf{W},\mathbf{v})\rangle - \lambda_1(\mathbf{B}(\mathbf{W},\mathbf{v})) \\ &\quad -\lambda_2(\mathbf{B}(\mathbf{W},\mathbf{u})) + \langle\mathbf{Z},\mathbf{B}(\mathbf{W},\mathbf{u})\rangle - \frac{\mathrm{D}\lambda_2}{\mathrm{d}t}(\delta\mathbf{v}) \\ &\quad +\langle\mathbf{W},[\mathbf{Z},\mathbf{u}]\rangle - \frac{\mathrm{D}\lambda_1}{\mathrm{d}t}(\mathbf{W}) \\ &\quad -\langle\mathbf{W},\nabla_{\mathbf{Z}}\mathbf{u} + R(\mathbf{u}+\mathbf{Z},\mathbf{v})\mathbf{v}\rangle\,\mathrm{d}t.\end{aligned}$$

By the independence of $\mathbf{W}$, $\delta\mathbf{v}$ and $\delta\mathbf{u}$, we have the following theorem.

Theorem IV.2. *A necessary condition for a control law $\mathbf{u}(t)$ to be an optimal solution for the problem (II.1) without the motion constraints (2.7), is that the first order differential equations:*

$$\begin{aligned}\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} &= \mathbf{v} \\ \frac{\mathrm{D}\mathbf{v}}{\mathrm{d}t} &= \mathbf{u} \\ \frac{\mathrm{D}\lambda_1}{\mathrm{d}t} &= \left([\mathbf{Z},\mathbf{u}] - \nabla_{\mathbf{Z}}\mathbf{u} + R(\mathbf{u}+\mathbf{Z},\mathbf{V})\mathbf{v}\right)^\flat \\ \frac{\mathrm{D}\lambda_2}{\mathrm{d}t} &= -\lambda_1 + \tau^2(\mathbf{v})^\flat\end{aligned}$$

$$\begin{aligned}\mathbf{u}+\mathbf{Z} &= \lambda_2^\sharp\\ \langle \mathbf{Z},\mathbf{u}\rangle &= 0\end{aligned}$$

be satisfied on $[0,T]$.

In Theorem (IV.2), $\sharp$ and $\flat$ denote the sharp and flat operators (see [9] for definitions). In the above, the Lagrange multipliers λ_i, $i=1,2$, are viewed as elements in the cotangent space T^*M.

We now show that the necessary conditions from Theorem (IV.2) are equivalent to those of theorem (IV.1). First note that for an arbitrary vector field $\mathbf{Y}$, Theorem (IV.2) implies that

$$\frac{\mathrm{D}^2\lambda_2}{\mathrm{d}t^2}(\mathbf{Y}) = \left\langle \frac{\mathrm{D}^2(\mathbf{u}+\mathbf{Z})}{\mathrm{d}t^2},\mathbf{Y}\right\rangle.$$

and that

$$\frac{\mathrm{D}^2\lambda_2}{\mathrm{d}t^2}(\mathbf{Y}) = -\frac{\mathrm{D}\lambda_1}{\mathrm{d}t}(\mathbf{Y}) + \tau^2\langle \mathbf{v},\mathbf{Y}\rangle$$
$$= \left\langle -[\mathbf{Z},\mathbf{u}] + \nabla_{\mathbf{Z}}\mathbf{u} - R(\mathbf{u}+\mathbf{Z},\mathbf{V})\mathbf{v} + \tau^2\mathbf{v},\mathbf{Y}\right\rangle.$$

Equating these two expressions immediately results in the fourth order necessary condition given in Theorem (IV.1). Hence, we have the following Lemma.

Lemma IV.1. *The necessary conditions of Theorem (IV.2) are equivalent to those of theorem (IV.1).*

V. Optimal Control on Compact Semi-Simple Lie Groups

In this section we derive the necessary conditions from Theorem (IV.1) where the manifold M is a compact semi-simple Lie group G. Let $\mathfrak{g}$ denote the Lie algebra of G and define the metric such that $\ll\cdot,\cdot\gg: -\frac{1}{2}\kappa(\cdot,\cdot)$, where κ denotes the Killing form on $\mathfrak{g}$. Recall that for semi-simple Lie groups, the Killing form is non-degenerate.

Let $J:\mathfrak{g}\to\mathfrak{g}$ be a positive definite linear mapping (the inertia tensor) that satisfies:

$$\begin{aligned}\ll J\mathbf{X},\mathbf{Y}\gg &= \ll \mathbf{X},J\mathbf{Y}\gg\\ \ll J\mathbf{X},\mathbf{X}\gg &= 0 \text{ if and only if } \mathbf{X}=0.\end{aligned}$$

Let $R_{\mathbf{g}}$ denote the right translation on G by $\mathbf{g}\in G$. If $\mathbf{X},\mathbf{Y}\in\mathfrak{g}$, then the corresponding right invariant vector fields are given by $\mathbf{X}^r_{\mathbf{g}} = \mathbf{X}^r(\mathbf{g}) = R_{\mathbf{g}^*}\mathbf{X}$ and $\mathbf{Y}^r_{\mathbf{g}} = \mathbf{Y}^r(\mathbf{g}) = R_{\mathbf{g}^*}\mathbf{Y}$, respectively. Hence a right invariant metric on G may be defined as

$$\langle \mathbf{X}^r(\mathbf{g}),\mathbf{Y}^r(\mathbf{g})\rangle := \ll \mathbf{X},J\mathbf{Y}\gg. \tag{5.1}$$

Corresponding to this metric there exists a unique Riemannian connection ∇, which, in turn, defines the bilinear form:

$$\begin{aligned}(\mathbf{X},\mathbf{Y})\to\mathbf{B}(\mathbf{X},\mathbf{Y}) &= \frac{1}{2}\{[\mathbf{X},\mathbf{Y}] + J^{-1}[\mathbf{X},J\mathbf{Y}]\\ &+ J^{-1}[\mathbf{Y},J\mathbf{X}]\},\end{aligned}\tag{5.2}$$

for any $\mathbf{X}$, $\mathbf{Y}\in\mathfrak{g}$. If J is the identity, then $\mathbf{B}(\mathbf{X},\mathbf{Y}) = \frac{1}{2}[\mathbf{X},\mathbf{Y}]$. In this paper we make the simplifying assumption $J=I$, the identity. The reason for doing this is that the derivation becomes very cumbersome and lengthy in the general case, which will be the focus of future work. We will drop the superscript r for right invariant vector fields in the rest of the paper.

With $\dot{\mathbf{q}} = \mathbf{v} = \sum_{i=1}^n v_i\mathbf{X}_i$, then we have:

$$\begin{aligned}\frac{\mathrm{D}\mathbf{v}}{\mathrm{d}t} &= \sum_{i=1}^n \dot{v}_i\mathbf{X}_i + \sum_{i,j=1}^n \frac{1}{2}v_iv_j[\mathbf{X}_i,\mathbf{X}_j] = \sum_{i=1}^m \dot{v}_i\mathbf{X}_i\\ \frac{\mathrm{D}^2\mathbf{v}}{\mathrm{d}t^2} &= \sum_{i=1}^n \ddot{v}_i\mathbf{X}_i + \frac{1}{2}\sum_{j,k=1}^n v_i\dot{v}_k[\mathbf{X}_i,\mathbf{X}_k]\\ \frac{\mathrm{D}^3\mathbf{v}}{\mathrm{d}t^3} &= \sum_{i=1}^n \dddot{v}_i\mathbf{X}_i + \sum_{i,j=1}^n \ddot{v}_iv_j[\mathbf{X}_j,\mathbf{X}_j]\\ &+\frac{1}{4}\sum_{i,j,k=1}^n v_i\dot{v}_jv_k[\mathbf{X}_k,[\mathbf{X}_i,\mathbf{X}_j]],\end{aligned}\tag{5.3}$$

where $\mathbf{B}(\mathbf{v},\mathbf{v}) = \nabla_{\mathbf{v}}\mathbf{v} = \sum_{i,j=1}^n \frac{1}{2}v_iv_j[\mathbf{X}_i,\mathbf{X}_j] = 0$ by the skew-symmetry of the Lie bracket. Note here that $\dot{v}_i = 0$ for $i=m+1,\ldots,n$. This is a standard result that can be found in [16]. We also need to compute $[\mathbf{Z},\mathbf{u}]$:

$$[\mathbf{Z},\mathbf{u}] = \sum_{i=m+1}^n\sum_{j=1}^m \zeta_iu_j[\mathbf{X}_i,\mathbf{X}_j]. \tag{5.4}$$

We now determine $R(\mathbf{u}+\mathbf{Z},\mathbf{v})\mathbf{v}$:

$$R(\mathbf{u}+\mathbf{Z},\mathbf{v})\mathbf{v} = \frac{1}{4}[[\mathbf{u}+\mathbf{Z},\mathbf{v}],\mathbf{v}] \tag{5.5}$$

$$= \frac{1}{4}\sum_{j,k=1}^n v_jv_k\left[\left[\sum_{i=1}^m u_i\mathbf{X}_i + \sum_{l=m+1}^n \zeta_l\mathbf{X}_l,\mathbf{X}_j\right],\mathbf{X}_k\right].$$

Finally, we need to compute the second-order time derivative of $\mathbf{Z}$. This is easily found to be:

$$\begin{aligned}\frac{\mathrm{D}^2\mathbf{Z}}{\mathrm{d}t^2} &= \sum_{i=m+1}^n \ddot{\zeta}_i\mathbf{X}_i + \sum_{i=m+1}^n\sum_{j=1}^n \dot{\zeta}_iv_j[\mathbf{X}_j,\mathbf{X}_i]\\ &+\frac{1}{2}\sum_{i=m+1}^n\sum_{j=1}^n \zeta_i\dot{v}_j[\mathbf{X}_j,\mathbf{X}_i]\\ &+\frac{1}{4}\sum_{i=m+1}^n\sum_{j,k=1}^n \zeta_iv_jv_k[\mathbf{X}_k,[\mathbf{X}_j,\mathbf{X}_i]].\end{aligned}\tag{5.6}$$

We are now in a position to state the necessary optimality conditions for the problem (II.1). By Theorem (IV.1), the necessary conditions are stated as:

$$\begin{aligned}&\dddot{v}_i\mathbf{X}_i - \tau^2u_k\mathbf{X}_k + \ddot{v}_iv_j[\mathbf{X}_j,\mathbf{X}_i] + \ddot{\zeta}_l\mathbf{X}_l\\ &+\dot{\zeta}_lv_i[\mathbf{X}_i,\mathbf{X}_l] + \zeta_lu_k[\mathbf{X}_l,\mathbf{X}_k]\\ &+\frac{1}{4}\Big\{v_hv_iv_j[\mathbf{X}_j,[\mathbf{X}_h,\mathbf{X}_i]]\\ &+v_iv_j[[u_k\mathbf{X}_k+\zeta_l\mathbf{X}_l,\mathbf{X}_i],\mathbf{X}_j]\\ &+\zeta_lv_iv_j[\mathbf{X}_j,[\mathbf{X}_i,\mathbf{X}_l]]\Big\} = 0,\end{aligned}\tag{5.7}$$

where we note that $\nabla_{\mathbf{Z}}\mathbf{u} = \mathbf{B}(\mathbf{Z},\mathbf{u}) = \frac{1}{2}[\mathbf{Z},\mathbf{u}]$ since $\mathbf{Z}$ and $\mathbf{u}$ are independent of the coordinate $\mathbf{q}$. We also used the fact that $\frac{\mathrm{D}^2\mathbf{u}}{\mathrm{d}t^2} = \frac{\mathrm{D}^3\mathbf{u}}{\mathrm{d}t^3}$. In Equation (5.7), we use the Einstein convention of summation over each (individual) term, where the indexes h,i,j are summed over $1,\ldots,n$, k over $1,\ldots,m$ and l over $m+1,\ldots,n$. Finally, note that the first term inside the curly brackets is zero, again, by the skew-symmetry of the Lie bracket. Hence, in final form, the necessary conditions are given by:

$$\begin{aligned}&\dddot{v}_i\mathbf{X}_i - \tau^2u_k\mathbf{X}_k + \ddot{v}_iv_j[\mathbf{X}_j,\mathbf{X}_i] + \ddot{\zeta}_l\mathbf{X}_l\\ &+\dot{\zeta}_lv_i[\mathbf{X}_i,\mathbf{X}_l] + \zeta_lu_k[\mathbf{X}_l,\mathbf{X}_k]\end{aligned}$$

$$+\frac{1}{4}\left\{v_i v_j\left[\left[u_k \mathbf{X}_k+\zeta_l \mathbf{X}_l, \mathbf{X}_i\right], \mathbf{X}_j\right]\right.$$
$$\left.+\zeta_l v_i v_j\left[\mathbf{X}_j,\left[\mathbf{X}_i, \mathbf{X}_l\right]\right]\right\}=0, \tag{5.8}$$

If we set $\tau=0$, this is the second order, dynamic version of the first order, kinematic problem found in [16] (Theorem 6) and [6]. Moreover, if we set $m=n$ (hence, $\mathbf{Z}=0$) and $\tau=0$, then the above equation reduces to:

$$\begin{aligned} 0 &= \dddot{v}_i \mathbf{X}_i+\ddot{v}_i v_j\left[\mathbf{X}_j, \mathbf{X}_i\right]+\frac{1}{4}\left\{v_i v_j v_k\left[\left[\mathbf{X}_k, \mathbf{X}_i\right], \mathbf{X}_j\right]\right\} \\ &= \dddot{v}_i \mathbf{X}_i+\ddot{v}_i v_j\left[\mathbf{X}_j, \mathbf{X}_i\right], \end{aligned}$$

where the term in curly brackets before the first equality sign is zero by the skew-symmetry of the Lie bracket. This is exactly what is found in [10] and Lemma 4 in [16].

Using the first equation in (5.3) as well as Equations (5.4) and (5.5), it is easy to derive the first order form of the necessary conditions as given in Theorem (IV.2). In the form of Theorem (IV.2), the necessary conditions are:

$$\sum_{i=1}^n \dot{q}_i \mathbf{X}_i=\sum_{i=1}^n v_i \mathbf{X}_i$$
$$\sum_{i=1}^m \dot{v}_i \mathbf{X}_i=\sum_{i=1}^m u_i \mathbf{X}_i$$
$$\left(\frac{\mathrm{D} \lambda_1}{\mathrm{dt}}\right)^{\sharp}=\frac{1}{2} \sum_{i=m+1}^n \sum_{j=1}^m \zeta_i u_j\left[\mathbf{X}_i, \mathbf{X}_j\right]$$
$$+\frac{1}{4} \sum_{j, k=1}^n v_j v_k\left[\left[\sum_{i=1}^m u_i \mathbf{X}_i+\sum_{l=m+1}^n \zeta_l \mathbf{X}_l, \mathbf{X}_j\right], \mathbf{X}_k\right]$$
$$\frac{\mathrm{D} \lambda_2}{\mathrm{dt}}=-\sum_i \lambda_1^i \boldsymbol{\Upsilon}_i+\tau^2 \sum_{i=1}^n v_i \boldsymbol{\Upsilon}_i,$$

where $\lambda_1=\sum_{i=1}^n \lambda_1^i \boldsymbol{\Upsilon}_i$ and $\lambda_2=\sum_{i=1}^n \lambda_2^i \boldsymbol{\Upsilon}_i$ and $\boldsymbol{\Upsilon}_i$, $i=1, \ldots, n$, is the co-frame for $T^* M$ such that $\Upsilon_i\left(\mathbf{X}_j\right)=\delta_{ij}$.

We now give an example on the three dimensional group of rigid body rotations $SO(3)$. In this case, we have $\left[\mathbf{X}_1, \mathbf{X}_2\right]=\mathbf{X}_3$, $\left[\mathbf{X}_2, \mathbf{X}_3\right]=\mathbf{X}_1$ and $\left[\mathbf{X}_3, \mathbf{X}_1\right]=\mathbf{X}_2$. We note, of course, that for the optimal control problem to be well defined, the system we consider must be controllable. For example, the under-actuated system:

$$\begin{aligned} \dot{\mathbf{q}} &= v_1 \mathbf{X}_1+v_2 \mathbf{X}_2+v_3 \mathbf{X}_3 \\ \frac{\mathrm{D}\mathbf{v}}{\mathrm{dt}} &= u_1 \mathbf{X}_1 \end{aligned} \tag{5.9}$$

is not controllable and, hence, the system can not be steered between any two arbitrary states. However, the system:

$$\begin{aligned} \dot{\mathbf{q}} &= v_1 \mathbf{X}_1+v_2 \mathbf{X}_2+v_3 \mathbf{X}_3 \\ \frac{\mathrm{D}\mathbf{v}}{\mathrm{dt}} &= u_1 \mathbf{X}_1+u_2 \mathbf{X}_2 \end{aligned} \tag{5.10}$$

is controllable. For this case we have $\mathbf{Z}=\zeta_3 \mathbf{X}_3$. After a long derivation, the fourth order necessary conditions in Equation (5.8) for $SO(3)$ can be shown to be:

$$\begin{aligned} \dddot{v}_1 &- \tau^2 u_1+\dot{\zeta}_3 v_2-\zeta_3 u_2 \\ &+ \frac{1}{4}\left[v_1 v_2 u_2+v_1 v_3 \zeta_3-u_1\left(v_2^2+v_3^2\right)\right]=0 \\ \dddot{v}_2 &- \tau^2 u_2-\dot{\zeta}_3 v_1+\zeta_3 u_1 \\ &+ \frac{1}{4}\left[v_1 v_2 u_1+v_2 v_3 \zeta_3-u_2\left(v_1^2+v_3^2\right)\right]=0 \\ \dddot{v}_3 &+ \ddot{\zeta}_3+\frac{1}{4}\left[v_1 v_3 u_3+v_2 v_3 u_2-\zeta_3\left(v_1^2+v_2^2\right)\right]=0. \end{aligned}$$

Remark V.1. Note that the dynamics (2.2) describe the motion of a rigid body with identity moments of inertia. Hence, the results presented in this paper pertain to symmetric bodies, but may easily be extended to systems with non-symmetric inertia properties.

VI. Conclusion

In this paper, we use geometric tools to derive coordinate-free necessary conditions for an optimal control problem, where the system is under-actuated and evolves on a Riemannian manifold. We apply the results to semi-simple compact Lie groups and relate our results to those appearing previously in the literature. Our current research focuses on the non-compact Lie group case, such as $SE(3)$. Future research will focus on adding more complexity to the model. Of particular interest are adding a drift (gravitation) term and the treatment of additional holonomic and/or nonholonomic constraints to the problem (II.1).

VII. Acknowledgements

The authors wish to thank Professor Peter Crouch for useful comments.

References

[1] M. Camarinha, "The geometry of cubic polynomials in Riemannian manifolds," Ph.D. dissertation, Universidade de Coimbra, 1996.

[2] I. I. Hussein and A. M. Bloch, "Dynamic interpolation on Riemannian manifolds: An application to interferometric imaging," *Proc. of the 2004 American Control Conference*, pp. 413–418, July 2004.

[3] S. Chakravorty, "Design and optimal control of multi-spacecraft interferometric imaging systems," Ph.D. dissertation, Aerospace Engineering, University of Michigan, 2004.

[4] I. I. Hussein, D. J. Scheeres, and D. C. Hyland, "Control of a satellite formation for imaging applications," *Proceedings of the 2003 American Control Conference*, pp. 308–313, June 2003.

[5] I. I. Hussein and A. M. Bloch, "Dynamic coverage optimal control for interferometric imaging spacecraft formations," *43rd IEEE Conference on Decision and Control*, December 2004, to appear.

[6] A. Bloch, P. Crouch, and T. Ratiu, "Sub-Riemannian optimal control problems," *Fields Institute Communications*, vol. 3, pp. 35–48, 1994.

[7] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*. NEw York, NY: Interscience Publishers, 1963.

[8] M. Spivak, *A Comprehensiive Introduction to Differential Geometry*. Oxford: Publish or Perish Inc., 1979, vol. 1.

[9] A. Bloch, J. Baillieul, P. E. Crouch, and J. E. Marsden, *Nonholonomic Mechanics and Control*. New York, NY: Springer-Verlag, 2003.

[10] L. Noakes, G. Heinzinger, and B. Paden, "Cubic splines on curved spaces," *IMA Journal of Mathematical Control and Information*, vol. 6, no. 4, pp. 465–473, 1989.

[11] P. Crouch and F. Silva-Leite, "The dynamic interpolation problem: On Riemannian manifolds, lie groups and symmetric spaces," *Journal of Dynamic and Control Systems*, vol. 1, no. 2, pp. 177–202, 1995.

[12] P. Crouch, F. Silva-Leite, and M. Camarinha, "Hamiltonian structure of generalized cubic polynomials," *Proceedings of the IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, pp. 13–18, 2000.

[13] I. I. Hussein and A. M. Bloch, "Optimal control on Riemannian manifolds with potential fields," *43rd IEEE Conference on Decision and Control*, December 2004, to appear.

[14] R. Brockett, "Control theory and singular Riemannian geometry," in *New Directions in Applied Mathematics*, P. J. Hilton and G. S. Young, Eds. Springer-Verlag, 1981, pp. 11–27.

[15] J. Milnor, *Morse Theory*. Princeton, NJ: Princeton University Press, 2002.

[16] P. Crouch and F. Silva-Leite, "Geometry and the dynamic interpolation problem," *Proceedigs of the 1991 American Control Conference*, pp. 1131–1136, 1991.