

**OPTIMAL CONTROL PROBLEMS ON PARALLELIZABLE RIEMANNIAN
MANIFOLDS: THEORY AND APPLICATIONS** *, **, ***RAM V. IYER¹, RAYMOND HOLSAPPLE¹ AND DAVID DOMAN²

Abstract. The motivation for this work is the real-time solution of a standard optimal control problem arising in robotics and aerospace applications. For example, the trajectory planning problem for air vehicles is naturally cast as an optimal control problem on the tangent bundle of the Lie Group $SE(3)$, which is also a parallelizable Riemannian manifold. For an optimal control problem on the tangent bundle of such a manifold, we use frame co-ordinates and obtain first-order necessary conditions employing calculus of variations. The use of frame co-ordinates means that intrinsic quantities like the Levi-Civita connection and Riemannian curvature tensor appear in the equations for the co-states. The resulting equations are singularity-free and considerably simpler (from a numerical perspective) than those obtained using a local co-ordinates representation, and are thus better from a computational point of view. The first order necessary conditions result in a two point boundary value problem which we successfully solve by means of a Modified Simple Shooting Method.

Mathematics Subject Classification. 49K99, 49M05, 49Q99, 90C99, 93C10.

Received January 6, 2004. Revised September 17, 2004.

1. INTRODUCTION

This paper studies regular solutions of optimal control problems for simple mechanical systems both from a theoretical and a computational point of view. The motivation for this paper is the real-time trajectory planning problem for hypersonic aircraft. The usual separation of time-scales assumption that is made for air-vehicles fails for such aircraft. Therefore the trajectory planning must be done simultaneously for both the attitude and position variables. An air-vehicle can be thought of as evolving on the tangent bundle of the Lie Group $SE(3)$ with a Riemannian metric that is obtained from the total kinetic energy. We assume that the center-of-mass and principal moments-of-inertias do not change during the flight, and consider the aerodynamic forces and moments to be input variables.

Jurdejevic [9] and Krishnaprasad [10] have considered optimal control problems for left-invariant systems on Lie Groups, with the input variables affecting the velocity vector field on the configuration space. Simple

Keywords and phrases. Regular optimal control, simple mechanical systems, calculus of variations, numerical solution, modified simple shooting method.

* *R.V.I. was supported by an NRC/USAF Summer Faculty Fellowship Grant during Summer 2002.*

** *R.H. was supported by an USAF Air-Vehicles Directorate Summer Research Program during Summer 2002.*

*** *Aerospace Engineer, Air Vehicles Directorate; Senior Member AIAA.*

¹ Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409, USA; ram.iyer@ttu.edu

² U.S. Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio 45433-7531, USA.

© EDP Sciences, SMAI 2006

mechanical systems that are the focus of this paper, with forces and moments as inputs, do not satisfy this framework. Sussmann [18] tackled the problem of generalizing the Pontryagin's Minimum Principle to manifolds (without any affine-connection structure), by developing the co-ordinate free maximum principle. For computational purposes, when this principle is applied to an air-vehicle problem, one employs local co-ordinates and the equations reduce to the necessary conditions for an optimal control problem in co-ordinates. Local co-ordinates might not be the best choice possible for the real-time computation of the optimal trajectory when one has an additional Lie Group structure. This is because a suitable choice of co-ordinates depends on the initial and final conditions on the optimal control problem, which makes it unsuitable for real-time computation of optimal trajectories. If the configuration space is a connected Lie Group G with Lie Algebra \mathcal{G} , then one can represent any point as a product of exponentials using the exponential map, $\exp : \mathcal{G} \rightarrow G$. In general, this map is not globally one-to-one or onto, but in the case of groups that semi-direct products of a connected and compact group and a connected Abelian group, it is globally one-to-one and onto (except on a set of measure zero). Such groups arise naturally in robotics and simple mechanical systems. The first order necessary conditions obtained in this paper using frame co-ordinates on the tangent bundle are especially suited for real-time computation in such applications. We obtain our results for a parallelizable Riemannian manifold, though our main interest lies in simple mechanical systems on Lie Groups. Thus the method is frame dependent rather than invariant. This approach can also be seen in the work of Crouch, Camarinha and Silva Leite [3–5]. Crouch *et al.* consider a special case of the problem investigated in this paper. Their work and ours differ in the nature of variations considered, and therefore the necessary conditions obtained in this paper are different from those of Crouch *et al.*

The main motivation for the choice of frame co-ordinates, is that the resulting first-order necessary conditions can be solved using a numerical method called the Modified Simple Shooting Method [8]. Though numerical methods for optimal control problems on local co-ordinates have a long history, there is a dearth of literature on numerical methods using frame co-ordinates. This paper, we believe, is the first to address this issue. Usually optimal control problems in local co-ordinates are tackled either by a direct method [1, 13] using non-linear programming methods, or by an indirect method that uses Pontryagin's Minimum Principle, resulting in a two-point boundary value problem [16, 17]. In this paper, we use the indirect approach to the solution of an optimal control problem. Among the indirect methods, the Modified Simple Shooting Method (MSSM) has been shown to be more accurate and resulting in faster computation times than other methods such as the Multiple Shooting (MSM), Finite Difference and Collocation methods [8]. In related research, it was found that the earlier MSM failed to converge numerically for an optimal control problem on $TSO(3)$ while using frame co-ordinates, while the MSSM converged successfully [7]. In this paper, we consider the more complex problem of a rigid body and show that the MSSM converges successfully in this case also.

2. OPTIMAL CONTROL ON PARALLELIZABLE RIEMANNIAN MANIFOLDS

In this section, we derive necessary conditions for regular solutions of optimal control problems on parallelizable Riemannian manifolds. Previous work in this area was done by P. Crouch, M. Camarinha and F. Silva Leite in a series of papers [3–5]. They considered a special case of the problem considered here and use a different approach to obtain first-order necessary conditions.

In related work, A. Lewis studied the first order necessary conditions arising from PMP for affine connection control systems that are also affine in the controls [11]. By studying the splitting of $T\star TM$ which arises from an Ehresmann connection associated with the affine connection, he directly derives the equation for the costates. This equation that he terms the adjoint Jacobi equation is compatible with the first-order necessary conditions derived in this paper, once suitable modifications of the cost function and system equations are carried out.

2.1. Parallelizable manifolds and Cartan formalism

The mathematical background in Riemannian geometry used in this paper can be found in standard sources such as Frankel [6] or Boothby [2]. If there exists a set of n smooth vector fields on a manifold M of dimension

n that are linearly independent at each point, then the manifold M is said to be parallelizable [2]. Such a set of n linearly independent vector fields is referred to as a *field of co-ordinate frames* or briefly as a *frame*. It is well known that the condition of being parallelizable is very special for a manifold. For example, among the spheres S^n , $n = 1, 2, \dots$, only S^1, S^3 and S^7 are parallelizable. However, it is well known that all Lie Groups are parallelizable [2]. In the case of air vehicles, the configuration space is the Lie Group $SE(3)$. However the manifold on which the air vehicle dynamics can be described is $TSE(3)$ – the tangent bundle of $SE(3)$. The following lemma states that $TSE(3)$ is parallelizable. The proof follows from the observation that locally TM is a product manifold $U \times \mathbb{R}^n$ where U is a co-ordinate neighborhood of M .

Lemma 2.1. *Let M be a parallelizable manifold. Then TM is parallelizable.*

Let M be a manifold with a symmetric, positive definite, and bilinear form (called a *Riemannian metric*) defined on T_qM where $q \in M$, denoted by $\langle \cdot, \cdot \rangle_q$. We assume it to be smooth for each $q \in M$ and C^1 as a function of q . A Riemannian metric defines a linear isomorphism: $\Sigma : T_qM \rightarrow T_q^*M$ (where $q \in M$ and T_q^*M is the dual of T_qM) by: $(\Sigma(v))(w) := \langle v, w \rangle_q$; $v, w \in T_qM$. As Σ is an isomorphism, we also have the inverse map $\Sigma^{-1} : T_q^*M \rightarrow T_qM$ where $q \in M$.

Assume M to be parallelizable and consider a frame $\{E_i\}$ and a dual co-frame $\{\theta^i\}$, $i = 1, 2, \dots, n$ with $\theta^i(E_j) = \delta_j^i$. Then the Levi-Civita connection is defined by the functions $\omega_{ij}^k(q)$ in the equation: $\nabla_{E_i}E_j = \omega_{ij}^k E_k$, $1 \leq i, j \leq n$. If we denote $\omega_j^k := \omega_{ij}^k \theta^i$, then we can define the $n \times n$ matrix ω of connection one-forms by $\omega = (\omega_j^k)$, where k indicates the row index and j the column index. The following proposition is useful in the computation of ω :

Proposition 2.1 [2]. *Let $\langle E_i, E_j \rangle = m_{ij}$. Define $d\theta = (d\theta^1, \dots, d\theta^n)$. Then there is a unique matrix of 1-forms $\omega = (\omega_j^i)$ such that: $d\theta = -\omega \wedge \theta$, that is, $d\theta^i = -\omega_j^i \wedge \theta^j$ and $\omega_j^i m_{rj} + \omega_r^i m_{ri} = 0$; $i, j = 1, \dots, n$.*

The computation of $d\theta^i$ is done using Theorem 4.25 of Frankel [6]. Cartan showed that the curvature tensor $R(\cdot, \cdot)\cdot$ can be computed from the connection matrix as follows:

Proposition 2.2 [6]. *The equations $\Omega = d\omega + \omega \wedge \omega$, that is, $\Omega_j^i = d\omega_j^i + \omega_k^i \wedge \omega_j^k$ define a skew-symmetric matrix of 2-forms that is related to the curvature tensor via $R(X, Y)E_j = \Omega_j^i(X, Y) \cdot E_i$.*

2.2. Derivation of first-order necessary conditions

Let M be a parallelizable Riemannian manifold. If $c : [0, 1] \rightarrow M$ is a differentiable curve on M , and $X : M \rightarrow TM$ is a differentiable vector field, then the co-variant derivative of X along $c(\cdot)$ is defined to be $\frac{DX}{dt} = \nabla_{\dot{c}(t)}X(t)$, $t \in (0, 1)$. Let $\{E_1, \dots, E_n\}$ be a frame of vector fields and let $\{\theta^1, \dots, \theta^n\}$ be a frame of co-vector fields on M , so that $\theta^i(E_j) = \delta_j^i$; $1 \leq i, j \leq n$. Let $q \in M$; $V \in T_qM$ and $u \in \mathbb{R}^m$ denote the control variables. We define a *control system* on TM by a second-order vector field $F : TM \times \mathbb{R}^m \rightarrow TTM$ defined as follows. If $\pi : TM \rightarrow M$ denotes the projection operator, then a second-order vector field is one that satisfies $d\pi \circ F_{((q,V),u)} = (q, V)$. Using the Levi-Civita connection on M , we can write the above system as one on TM described by the equations:

$$\dot{q} = V = V^i E_i, \quad \text{and} \quad \frac{DV}{dt} = f(q, V, u) = f^i(q, V, u) E_i. \quad (1)$$

The conditions on the function f will be set forth in our theorem on necessary conditions. Such a set of equations is useful in describing the equations of motion of an air vehicle that is subject to aerodynamic forces and moments (as well as gravity) that depend on its orientation with respect to its velocity vector, its altitude above sea level, current speed and the deflections of its control surfaces.

Now, let $\hat{q}_0, \hat{q}_f \in M$, $V_0 \in T_{\hat{q}_0}M$ and $V_f \in T_{\hat{q}_f}M$. Consider the space $C_2[t_0, t_f]$ of twice-differentiable maps $q : [t_0, t_f] \rightarrow M$ that satisfy equations (1), where $t_f > t_0$, $q(t_0) = \hat{q}_0$, $q(t_f) = \hat{q}_f$, $\dot{q}(t_0) = V_0$ and $\dot{q}(t_f) = V_f$. Then along one such map $q(\cdot)$ the control system takes the form: $\dot{q}(t) = V(t)$, and $\frac{DV}{dt} = f(q(t), V(t), u(t))$,

where $u(\cdot) \in C^m[t_0, t_f]$, the (m -vector valued) space of continuous functions. Suppose that one is required to find a function $u(\cdot)$ such that the above boundary conditions are satisfied by $q(\cdot)$ while minimizing:

$$J(u(\cdot)) = \int_{t_0}^{t_f} L(q(t), V(t), u(t)) dt. \quad (2)$$

We need the following standard construction to describe the notion of variations of a curve. Let $(t, \sigma) \rightarrow q(t, \sigma)$, $t \in [t_0, t_f]$ and $\sigma \in (-\epsilon, \epsilon)$, $\epsilon > 0$, be a parametrized family of curves satisfying

$$q(t, 0) = q(t); \quad q(t_0, \sigma) = \hat{q}_0; \quad q(t_f, \sigma) = \hat{q}_f; \quad \dot{q}(t_0, \sigma) = V_0; \quad \dot{q}(t_f, \sigma) = V_f. \quad (3)$$

For $V(t) \in T_{q(t)}M$ and $p_1(t), p_2(t) \in T_{q(t)}^*M$, we define the associated variations: $V(t, \sigma) = V^i(t, \sigma)E_i(q(t, \sigma)) \in T_{q(t, \sigma)}M$ and $p_1(t, \sigma) = p_{1_i}(t, \sigma)\theta^i(q(t, \sigma))$, $p_2(t, \sigma) = p_{2_i}(t, \sigma)\theta^i(q(t, \sigma)) \in T_{q(t, \sigma)}^*M$. Define the variational vector fields $W(t) = \delta q(t) := \frac{\partial q}{\partial \sigma}(t, 0) \in T_{q(t)}M$, $t \in [t_0, t_f]$, with $W(t_0) = W(t_f) = 0$; and $\delta V(t) := \frac{DV}{d\sigma}(t, 0) \in T_{q(t)}M$, $t \in [t_0, t_f]$. The variations in the input are denoted by $u(t, \sigma) \in \mathbb{R}^m$ with $\delta u(t) := \frac{\partial u}{\partial \sigma}(t, 0) \in \mathbb{R}^m$. In the following, any quantity that is described with the second variable σ suppressed, should be construed as having $\sigma = 0$ (so, $q(t) = q(t, 0)$).

Then we have the following lemma proved by Noakes, Heinzinger and Paden.

Lemma 2.2 [15]. $\frac{DV}{d\sigma}(t_0, \sigma) = 0$ and $\frac{DV}{d\sigma}(t_f, \sigma) = 0$ for all $\sigma \in (-\epsilon, \epsilon)$.

We need the following simple lemmata in the proof of the main theorem (Th. 2.1).

Lemma 2.3.

$$\int_{t_0}^{t_f} p_1(t) \left(\frac{D\dot{q}}{d\sigma}(t) \right) dt = - \int_{t_0}^{t_f} \frac{Dp_1}{dt}(\delta q(t)) dt. \quad (4)$$

Proof. We have $\frac{D\dot{q}}{d\sigma}(t) = \frac{DW}{dt}(t)$ because the Levi-Civita connection is symmetric and $[\frac{\partial}{\partial t}, \frac{\partial}{\partial \sigma}] = 0$. Therefore:

$$\begin{aligned} \int_{t_0}^{t_f} \frac{d}{dt} p_1(\delta q) dt &= \int_{t_0}^{t_f} \frac{d}{dt} \langle \Sigma^{-1} p_1, \delta q \rangle dt = \int_{t_0}^{t_f} \left(\left\langle \frac{d}{dt} \Sigma^{-1} p_1, \delta q \right\rangle + \left\langle \Sigma^{-1} p_1, \frac{D}{dt} \delta q \right\rangle \right) dt \\ p_1(\delta q) \Big|_{t_0}^{t_f} &= \int_{t_0}^{t_f} \left(\frac{Dp_1}{dt}(\delta q) + p_1 \left(\frac{D}{dt} \delta q \right) \right) dt. \end{aligned}$$

The result follows by noting that $W(t_0) = W(t_f) = 0$. □

Lemma 2.4.

$$\int_{t_0}^{t_f} p_2(t) \frac{D}{d\sigma} \frac{DV}{dt} dt = \int_{t_0}^{t_f} \left((\Sigma R(\Sigma^{-1} p_2, V)V) \delta q - \frac{Dp_2}{dt}(\delta V(t)) \right) dt. \quad (5)$$

Proof. First we note that: $\int_{t_0}^{t_f} p_2(t) \frac{D}{d\sigma} \frac{DV}{dt} dt = \int_{t_0}^{t_f} p_2(t) (R(W, V)V + \frac{D}{dt} \frac{DV}{d\sigma}) dt$, by the definition of the curvature tensor [2, 6], and the fact that $[\frac{\partial}{\partial t}, \frac{\partial}{\partial \sigma}] = 0$. By the definition of the linear isomorphism Σ , and standard properties of the curvature tensor [6] we have: $p_2(R(W, V)V) = \langle \Sigma^{-1} p_2, R(W, V)V \rangle = \langle R(\Sigma^{-1} p_2, V)V, W \rangle = (\Sigma R(\Sigma^{-1} p_2, V)V)(W)$. On integrating the second term by parts and using Lemma 2.2, we have the claim. □

For $p \in T^*M$, denote: $[\omega(f)]^* p := p_i \omega_j^i(f) \theta^j = p_i \omega_{kj}^i f^k \theta^j$. Let C_{kj}^i denote the structure constants for the Jacobi-Lie brackets of the coordinate vector fields. Also denote: $[C(f)]^* p := p_i C_j^i(f) \theta^j = p_i C_{kj}^i f^k \theta^j$. The following theorem is the main result of this paper. It establishes the first-order necessary conditions for the curve $(q_0, V_0, u_0)(t)$, $t \in [t_0, t_f]$ to be optimal.

Theorem 2.1. *Suppose that $(q_0, V_0, u_0)(t)$, $t \in [t_0, t_f]$ minimizes the cost function (2), while satisfying equations (1) and boundary conditions $q(t_0) = \hat{q}_0$, $q(t_f) = \hat{q}_f$, $\dot{q}(t_0) = V_0$ and $\dot{q}(t_f) = V_f$. Further suppose that f and L are differentiable functions of their arguments, and the linearized system is controllable at the origin. Then there exists one-forms $p_1(t), p_2(t)$ differentiable almost everywhere on $[t_0, t_f]$ such that:*

- $\frac{Dp_1}{dt} = L_q(q_0, V_0, u_0) + \Sigma R(\Sigma^{-1}p_2, V_0)V_0 - (f_q^*(q_0, V_0, u_0) + \omega^*(f) - C^*(f))p_2$
- $\frac{Dp_2}{dt} = -p_1 + L_V(q_0, V_0, u_0) - f_V^*(q_0, V_0, u_0)p_2;$
- $f_u^*(q_0, V_0, u_0)p_2 = L_u(q_0, V_0, u_0);$
- *the function $H(q_0, V_0, u_0, p_1, p_2)(t) = L(q_0, V_0, u_0)(t) - p_1(V_0)(t) - p_2(f(q_0, V_0, u_0))(t)$ is a constant for $t \in [t_0, t_f]$.*

The proof of the above theorem will be given shortly. The assumption of controllability of the linearized system is strong, and can be weakened along the lines of Pontryagin's maximum principle [18], once the symplectic 2-form on T^*TM is written in frame co-ordinates. Such a symplectic form is not intrinsic and depends on the Riemannian metric on TM as we show in a future publication. In any case, the results of this paper will serve as a check for the first order necessary conditions that can be obtained using more sophisticated tools of symplectic geometry. The proof here follows Luenberger [12], and uses the Lagrange Multiplier theorem to obtain the existence of the one-forms $p_1(\cdot), p_2(\cdot)$. Please refer to Luenberger to see how the controllability assumption enters the proof.

Proof of Theorem 2.1. For each $\sigma \in (-\epsilon, \epsilon)$, consider the augmented cost function:

$$\bar{J}(q(\cdot, \sigma), V(\cdot, \sigma), p_1(\cdot, \sigma), p_2(\cdot, \sigma)) = \int_{t_0}^{t_f} \left(L(q, V, u) + p_1(\dot{q} - V) + p_2 \left(\frac{DV}{dt} - f(q, V, u) \right) \right) dt, \quad (6)$$

where q, V, p_1, p_2 in the integral are functions of (t, σ) . By the Chain Rule:

$$\frac{\partial}{\partial \sigma} L(q, V, u)(t, 0) = L_q(q, V, u)(t, 0)(\delta q(t)) + L_V(q, V, u)(t, 0) \left(\frac{DV}{d\sigma}(t) \right) + L_u(q, V, u)(t, 0)(\delta u(t)). \quad (7)$$

$$\begin{aligned} \frac{\partial}{\partial \sigma} f(q, V, u)(t, 0) &= \left(f_q^k(q, V, u)(t, 0)(\delta q) + f_V^k(q, V, u)(t, 0) \left(\frac{DV}{d\sigma}(t) \right) \dots \right. \\ &\quad \left. + f_u^k(q, V, u)(t, 0) \left(\frac{\partial u}{\partial \sigma}(t) \right) \right) E_k(t, 0) + f^k(q, V, u)(t, 0) \frac{DE_k(t, 0)}{d\sigma}. \end{aligned} \quad (8)$$

Lets consider the last term in the above equation.

$$f^k(q, V, u)(t, 0) \frac{DE_k(t, 0)}{d\sigma} = f^k(q, V, u)(t) \omega_{jk}^i \delta q^j E_i = f^k(q, V, u)(t) (\omega_{kj}^i - C_{kj}^i) \delta q^j E_i. \quad (9)$$

Now by Lebesgue's Dominated Convergence Theorem, we have:

$$\frac{\partial \bar{J}}{\partial \sigma}(q, V, p_1, p_2)(\cdot, 0) = \int_{t_0}^{t_f} \left(\frac{\partial}{\partial \sigma} L(q, V, u) + p_1 \left(\frac{D\dot{q}}{d\sigma} \right) - p_1(\delta V) + p_2 \left(\frac{D}{d\sigma} \frac{DV}{dt} \right) - p_2 \frac{\partial}{\partial \sigma} f(q, V, u) \right) dt.$$

By Lemmas 2.3, 2.4 and equations (7)–(9), we get:

$$\begin{aligned} \frac{\partial \bar{J}}{\partial \sigma}(q, V, p_1, p_2)(\cdot, 0) &= \int_{t_0}^{t_f} \left(\left(L_q(q, V, u) - \frac{Dp_1}{dt} + \Sigma R(\Sigma^{-1}p_2, V)V - (f_q^* + [\omega(f)]^* - [C(f)]^*)p_2 \right) \delta q \right. \\ &\quad \left. + (L_V(q, V, u) - p_1 - \frac{Dp_2}{dt} - f_v^* p_2) \delta V + (L_u(q, V, u) - f_u^* p_2) \delta u \right) dt. \end{aligned}$$

As the variations δq , δV and δu are arbitrary, subject to $\delta q(t_0) = \delta q(t_f) = \delta V(t_0) = \delta V(t_f) = 0$ we have the first two statements of the theorem.

To prove the last statement of the theorem, consider

$$\begin{aligned}
\dot{H}(t) &= \frac{d}{dt}L(q_0, V_0, u_0)(t) - \frac{Dp_1}{dt}(V_0) - p_1 \left(\frac{DV}{dt} \right) (t) - \frac{Dp_2}{dt}(f(q_0, V_0, u_0))(t) - p_2 \left(\frac{D}{dt}f(q_0, V_0, u_0) \right) (t) \\
&= L_q(q_0, V_0, u_0)(V_0) + L_V(q_0, V_0, u_0) \left(\frac{DV}{dt} \right) + L_u(q_0, V_0, u_0)(\dot{u}) \\
&\quad - (L_q(q_0, V_0, u_0) + \Sigma R(\Sigma^{-1}p_2, V_0)V_0 - (f_q^*(q_0, V_0, u_0) + \omega^*(f) - C^*(f))p_2)(V_0)(t) \\
&\quad - p_1(f(q_0, V_0, u_0))(t) - (-p_1 + L_V(q_0, V_0, u_0) - f_V^*(q_0, V_0, u_0)p_2)(f(q_0, V_0, u_0))(t) \\
&\quad - p_2 \left(f_q(q_0, V_0, u_0)(V_0) + f_V(q_0, V_0, u_0) \left(\frac{DV}{dt} \right) + f_u(q_0, V_0, u_0)(\dot{u}) + f^i \frac{DE_i}{dt} \right) \\
&= 0,
\end{aligned}$$

due to the fact that $\langle R(\Sigma^{-1}p_2, V_0)V_0, V_0 \rangle = 0$ and $f^k(q, V, u)(t, 0) \frac{DE_k(t, 0)}{dt} = f^k(q, V, u)(t) \omega_{jk}^i \dot{q}^j E_i = f^k(q, V, u)(t) (\omega_{kj}^i - C_{kj}^i) V^j E_i$. \square

3. APPLICATIONS

3.1. Cubic splines on Riemannian manifolds

Here we specialize Theorem 2.1 and recover the Noakes, Heinzinger and Paden formula for cubic splines on Riemannian manifolds [15]. Let M be a parallelizable Riemannian manifold and let $q_0, q_1 \in M$, $V_0 \in T_{q_0}M$ and $V_1 \in T_{q_1}M$. Consider the problem: Minimize $J(u(\cdot)) = \int_{t_0}^{t_f} \|u(t)\|^2 dt$ subject to: $\dot{q}(t) = V(t)$, $\frac{DV}{dt} = u(t) = u^i(t)E_i(t)$, and boundary conditions $q(t_0) = q_0$; $q(t_f) = q_f$; $\dot{q}(t_0) = V_0$; $\dot{q}(t_f) = V_f$.

Thus we have $f(q, V, u) = u$ and Σ is the identity matrix. Therefore, there is an identification of vectors and co-vectors. Then, Theorem 2.1 asserts the existence of one-form sections $p_1(t), p_2(t)$ such that: $\frac{Dp_1}{dt} = R(p_2, V)V - (\omega^*(f) - C^*(f))p_2$, $\frac{Dp_2}{dt} = -p_1$ and $p_2 = u$, where we have used the identification of vectors and co-vectors in the last equation. Thus $\frac{D^2V}{dt^2} = \frac{Du}{dt} = \frac{Dp_2}{dt} = -p_1$ which implies $\frac{D^3V}{dt^3} = -\frac{Dp_1}{dt} = -R(p_2, V)V + (\omega^*(f) - C^*(f))p_2$. Now $((\omega^*(f) - C^*(f))p_2)w = p_{2_i}(\omega_{jk}^i(p_2) - C_{jk}^i(p_2))w^j = p_{2_i}\omega_{jk}^i p_2^k w^j = \omega_{kj}^i(w)p_2^k p_{2_i} = 0$ for all $w \in \Psi(M)$, because $\omega_{kj}^i = -\omega_{jk}^i$ by Proposition 2.1. Therefore, $\frac{D^3V}{dt^3} + R(\frac{DV}{dt}, V)V = 0$, which is the equation for a cubic spline that was first obtained by Noakes, Heinzinger and Paden [15].

3.2. Rigid body translation and rotation

In this subsection, we consider the problem of numerically solving the optimal control problem (1), (2) and (3) for a rigid body that is free to rotate and translate. Consider a rigid body with principal moment of inertia matrix \mathbb{I} and mass M (a scalar). The configuration space of a rigid body $SE(3)$ is the space of rotations given by the set of 3×3 matrices $SO(3) = \{Q \mid Q'Q = I; \det(Q) = 1\}$, and space of translations \mathbb{R}^3 . Q is the orientation of the rigid body with respect to an earth-fixed co-ordinate system, while b is the position of the center of mass of the rigid body with respect to the origin of the earth-fixed co-ordinate system. The angular velocity of the body in the principal axis system (called the body axis system) centered at the center of mass is defined as: $\Omega = Q'\dot{Q}$ where Q' denotes the transpose of Q , while the linear velocity of the center of mass expressed in the body axis system is denoted by: $v = Q'\dot{b}$. If we define the skew-symmetrization of

$\Omega' = [\Omega_1, \Omega_2, \Omega_3]$ to be $\hat{\Omega} = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix}$ then the Euler's equations for a rigid body can be written in a

compact form as:

$$\dot{Q} = Q\hat{\Omega} \quad (10)$$

$$\dot{b} = Qv \quad (11)$$

$$\dot{\Omega} - \mathbb{I}^{-1}(\mathbb{I}\Omega \times \Omega) = \mathbb{I}^{-1}T_e \quad (12)$$

$$\dot{v} + \Omega \times v = M^{-1}F_e, \quad (13)$$

where T_e, F_e are the moments and forces acting on the body (expressed in the principal body-axis system). In the notation of Theorem 2.1, $q = (Q, b)$, $V = (\Omega, v)$ and $u = (\mathbb{I}^{-1}T_e, M^{-1}F_e)$. Given $(Q, b, \Omega, v)(0)$, $(Q, b, \Omega, v)(1)$, and $L(q, V, u) = \frac{1}{2}\langle u, u \rangle$, the optimal control problem is to determine $u(t)$; $t \in [0, 1]$. We set up a two-point boundary value problem using Theorem 2.1, and then obtain a numerical solution using the Modified Simple Shooting method [8].

We will first find the Levi-Civita connection that corresponds to the Kinetic Energy metric: $K(Q, b, \dot{Q}, \dot{b}) = -\frac{1}{4}\text{trace}(\hat{\Omega}\mathbb{I}\hat{\Omega}) + \frac{1}{2}M\dot{b}^T\dot{b}$. Let $\{e_1, \dots, e_6\}$ form a basis for the Lie algebra $se(3) = so(3) \oplus \mathbb{R}^3$ of the Lie Group $SE(3)$, where e_1, e_2, e_3 form a basis for $so(3)$ and e_4, e_5, e_6 form a basis for \mathbb{R}^3 . Then one can form a parallel frame $E_i = Qe_i$, $i = 1, \dots, 6$ for $TSE(3)$ via the lift of the left translation action. For this parallel frame $\{E_1, \dots, E_6\}$ let the structure constants C_{ij}^k for the Jacobi-Lie bracket be given by:

$C_{12}^3 = 1$	$C_{23}^1 = 1$	$C_{31}^2 = 1$	$C_{15}^6 = 1$	$C_{16}^5 = -1$
$C_{24}^6 = -1$	$C_{26}^4 = 1$	$C_{34}^5 = 1$	$C_{35}^4 = -1$	$C_{ij}^k = 0$ for all other $1 \leq i, j, k \leq 6$.

The Riemannian metric is defined via the following table:

$\langle E_1, E_1 \rangle = \frac{1}{2}I_1$	$\langle E_2, E_2 \rangle = \frac{1}{2}I_2$	$\langle E_3, E_3 \rangle = \frac{1}{2}I_3$
$\langle E_4, E_4 \rangle = 1$	$\langle E_5, E_5 \rangle = 1$	$\langle E_6, E_6 \rangle = 1$
$\langle E_i, E_j \rangle = 0$ for all other $1 \leq i, j \leq 6$.		

We can compute the connection matrix for the rigid body using Proposition 2.1 and Theorem 4.25 of Frankel [6]. As $\theta^i(E_j) = \delta_j^i$, we have: $d\theta^i(E_j, E_k) = E_j(\theta^i(E_k)) - E_k(\theta^i(E_j)) - \theta^i([E_j, E_k]) = -\theta^i([E_j, E_k])$. Using Proposition 2.1 we compute the following connection matrix:

$$[\omega_i^k] = \begin{bmatrix} \alpha & 0_{3 \times 3} \\ 0_{3 \times 3} & \beta \end{bmatrix}, \text{ where } \alpha = \begin{bmatrix} 0 & \omega_{32}^1\theta^3 & \omega_{23}^1\theta^2 \\ \omega_{31}^2\theta^3 & 0 & \omega_{13}^2\theta^1 \\ \omega_{32}^3\theta^2 & \omega_{12}^3\theta^1 & 0 \end{bmatrix}, \beta = \begin{bmatrix} 0 & -\theta^3 & \theta^2 \\ \theta^3 & 0 & -\theta^1 \\ -\theta^2 & \theta^1 & 0 \end{bmatrix}$$

where $\omega_{32}^1 = \frac{1}{2} \left(\frac{-I_1 - I_2 + I_3}{I_1} \right)$, $\omega_{23}^1 = \frac{1}{2} \left(\frac{I_1 - I_2 + I_3}{I_1} \right)$, $\omega_{31}^2 = \frac{1}{2} \left(\frac{I_1 + I_2 - I_3}{I_2} \right)$, $\omega_{13}^2 = \frac{1}{2} \left(\frac{I_1 - I_2 - I_3}{I_2} \right)$, $\omega_{21}^3 = \frac{1}{2} \left(\frac{-I_1 + I_2 - I_3}{I_3} \right)$, and $\omega_{12}^3 = \frac{1}{2} \left(\frac{-I_1 + I_2 + I_3}{I_3} \right)$. Next, we compute the matrix $\nabla_E E$ [6]:

$$[\nabla_{E_j} E_i] = \begin{bmatrix} L & M \\ 0_{3 \times 3} & 0_{3 \times 3} \end{bmatrix}, \text{ where } L = \begin{bmatrix} 0 & \omega_{12}^3 E_3 & \omega_{13}^2 E_2 \\ \omega_{21}^3 E_3 & 0 & \omega_{23}^1 E_1 \\ \omega_{31}^2 E_2 & \omega_{32}^1 E_1 & 0 \end{bmatrix}, M = \begin{bmatrix} 0 & E_6 & -E_5 \\ -E_6 & 0 & E_4 \\ E_5 & -E_4 & 0 \end{bmatrix}.$$

Let $X = (\Omega, v), Y = (\xi, \bar{v}) \in \Psi(SE(3))$; $\Omega, v, \xi, \bar{v} \in \mathbb{R}^3$ and $q(t)$ be the curve obtained by solving the system $\dot{q}(t) = X$; $q(t_0) = \hat{q}_0$; $t \in [t_0, t_f]$. Then along the curve $q(\cdot)$, we compute:

$$\nabla_X Y = \begin{bmatrix} \dot{\xi} + \sum_{i,j,k=1}^3 \xi^i \omega_i^k(E_j) \Omega^j E_k \\ \dot{\bar{v}} - \bar{v} \times \Omega \end{bmatrix} = \begin{bmatrix} \bar{\nabla}_\Omega \xi \\ \dot{\bar{v}} - \bar{v} \times \Omega \end{bmatrix}, \quad (14)$$

where $\bar{\nabla}$ is the Levi-Civita connection on $SO(3)$ compatible with the Kinetic energy metric $\bar{K}(Q, \dot{Q}) = -\frac{1}{4}\text{trace}(\hat{\Omega}\mathbb{I}\hat{\Omega})$. In particular, $\frac{DX}{dt} = \nabla_X X = \begin{bmatrix} \dot{\Omega} - \mathbb{I}^{-1}(\mathbb{I}\Omega \times \Omega) \\ \dot{v} - v \times \Omega \end{bmatrix}$. We compute the curvature tensor $R(Y, X)X$ using Proposition 2.2. The curvature two-form Ω_i^k matrix turns out to be:

$$[\Omega_i^k] = \begin{bmatrix} P & 0_{3 \times 3} \\ 0_{3 \times 3} & R \end{bmatrix}, \quad \text{where } P = \begin{bmatrix} 0 & A\theta^1 \wedge \theta^2 & B\theta^3 \wedge \theta^1 \\ D\theta^1 \wedge \theta^2 & 0 & C\theta^2 \wedge \theta^3 \\ E\theta^3 \wedge \theta^1 & F\theta^2 \wedge \theta^3 & 0 \end{bmatrix},$$

and $R = \begin{bmatrix} 0 & H\theta^1 \wedge \theta^2 & I\theta^3 \wedge \theta^1 \\ K\theta^1 \wedge \theta^2 & 0 & J\theta^2 \wedge \theta^3 \\ L\theta^3 \wedge \theta^1 & M\theta^2 \wedge \theta^3 & 0 \end{bmatrix}$, where, $A = -\omega_{32}^1 - \omega_{23}^1 \omega_{12}^3$; $B = -\omega_{23}^1 + \omega_{32}^1 \omega_{13}^2$; $C = -\omega_{13}^2 - \omega_{31}^2 \omega_{23}^1$; $D = -\omega_{31}^2 + \omega_{13}^2 \omega_{21}^3$; $E = -\omega_{21}^3 - \omega_{31}^2 \omega_{12}^3$; $F = -\omega_{12}^3 + \omega_{21}^3 \omega_{32}^1$ and $H = I = J = K = L = M = 0$. Next, as $R(Y, X)E_j = \Omega_j^i((\xi, \bar{v}), (\Omega, v))E_i$, we have:

$$R((\xi, \bar{v}), (\Omega, v))(\Omega, v) = \begin{bmatrix} A\Omega^2(\xi^1\Omega^2 - \xi^2\Omega^1) + B\Omega^3(\xi^3\Omega^1 - \xi^1\Omega^3) \\ D\Omega^1(\xi^1\Omega^2 - \xi^2\Omega^1) + C\Omega^3(\xi^2\Omega^3 - \xi^3\Omega^2) \\ E\Omega^1(\xi^3\Omega^1 - \Omega^3\xi^1) + F\Omega^2(\xi^2\Omega^3 - \xi^3\Omega^2) \\ 0_{3 \times 1} \end{bmatrix}. \quad (15)$$

Finally, we compute the $(\omega^*(f) - C^*(f))p_2$ term that appears in the equation for $\frac{Dp_1}{dt}$.

$$(\omega^*(f) - C^*(f))p_2 = \begin{bmatrix} \frac{1}{I_1}(I_2\omega_{13}^2 + I_3\omega_{12}^3)\eta^2\eta^3 \\ \frac{1}{I_2}(I_1\omega_{23}^1 + I_3\omega_{21}^3)\eta^1\eta^3 \\ \frac{1}{I_3}(I_1\omega_{32}^1 + I_2\omega_{31}^2)\eta^2\eta^1 \\ 0_{3 \times 1} \end{bmatrix}. \quad (16)$$

If we define the vectors (ξ, \bar{v}_1) and (η, \bar{v}_2) via the identification $(\xi, \bar{v}_1)(t) = \Sigma^{-1}p_1(t)$ and $(\eta, \bar{v}_2)(t) = \Sigma^{-1}p_2(t)$ then, we can write the necessary conditions in terms of $(Q, b, \Omega, v, \xi, \bar{v}_1, \eta, \bar{v}_2)$. The full set of equations are (10)–(13) along with $u = (\eta, \bar{v}_2)$ and:

$$\begin{bmatrix} \dot{\xi} \\ \dot{\bar{v}}_1 \end{bmatrix} = - \begin{bmatrix} \sum_{i,j,k=1}^3 \xi^i \omega_{ji}^k \Omega^j E_k \\ -\bar{v}_1 \times \Omega \end{bmatrix} + R((\xi, \bar{v}), (\Omega, v))(\Omega, v) - (\omega^*(f) - C^*(f))p_2 \quad (17)$$

$$\begin{bmatrix} \dot{\eta} \\ \dot{\bar{v}}_2 \end{bmatrix} = - \begin{bmatrix} \sum_{i,j,k=1}^3 \eta^i \omega_{ji}^k \Omega^j E_k \\ -\bar{v}_2 \times \Omega \end{bmatrix} - \begin{bmatrix} \eta \\ \bar{v}_2 \end{bmatrix}, \quad (18)$$

with $(Q, \Omega, b, v)(0)$ and $(Q, \Omega, b, v)(1)$ specified.

4. NUMERICAL EXPERIMENTS

Equations (10)–(13), (17)–(18) along with $u = (\eta, \bar{v}_2)$ constitute a two-point boundary value problem. We used a modified shooting method technique [8] to numerically solve for the unknown ‘‘Lagrange multipliers’’

$(\xi, \bar{v}_1, \eta, \bar{v}_2)$ at initial time. The first equation (10) is a matrix equation that we integrated using the well-known Rodriguez's formula [14]:

$$Q(t+h) = Q(t) \left(I + \frac{\hat{\Omega}}{\|\Omega\|} \sin(\|\Omega\|h) + \frac{\hat{\Omega}^2}{\|\Omega\|^2} (1 - \cos(\|\Omega\|h)) \right), \quad (19)$$

where h is the time-step for integration. This results in a Q matrix that "stays" on the group $SO(3)$ at each time-step.

The moments of inertia constants for the numerical simulation were $I_1 = 10$; $I_2 = 7.5$; $I_3 = 5$. The initial time t_0 was set to 0 and the final time t_f was set to 1.

We chose the initial states $(Q, \Omega, b, v)(0)$ and the desired final states $(Q, \Omega, b, v)(1)$ using the MATLAB random number generator. Their values for a simulation run are listed in the following table (to 2 significant digits).

$Q(0)$			$\Omega(0)$	$b(0)$	$v(0)$
0.572	0.817	0.079	1.160	0.950	0.336
-0.783	0.572	-0.246	0.781	4.337	1.922
-0.246	0.079	0.966	3.693	7.092	4.714

$Q_{des}(1)$			$\Omega_{des}(1)$	$b_{des}(1)$	$v_{des}(1)$
-0.268	0.963	-0.011	2.319	1.901	0.673
-0.838	-0.227	0.497	1.562	8.673	3.843
0.476	0.142	0.868	7.385	4.185	9.427

The modified simple shooting method involves the choice of a continuous, time-parametrized reference path that connects the initial and final points. For $t \in [0, 1]$, we picked the reference path to be: $Q_{ref}(t) = Q(0) \exp(\hat{\phi}t)$, where $\hat{\phi} = \ln(Q(0)^{-1}Q_d)$, $\Omega_{ref}(t) = \Omega(0) + (\Omega_d - \Omega(0))t$, $b_{ref}(t) = b(0) + (b_d - b(0))t$, $v_{ref}(t) = v(0) + (v_d - v(0))t$.

The equations were integrated in the forward direction until at some time $t \in (0, 1]$, we had

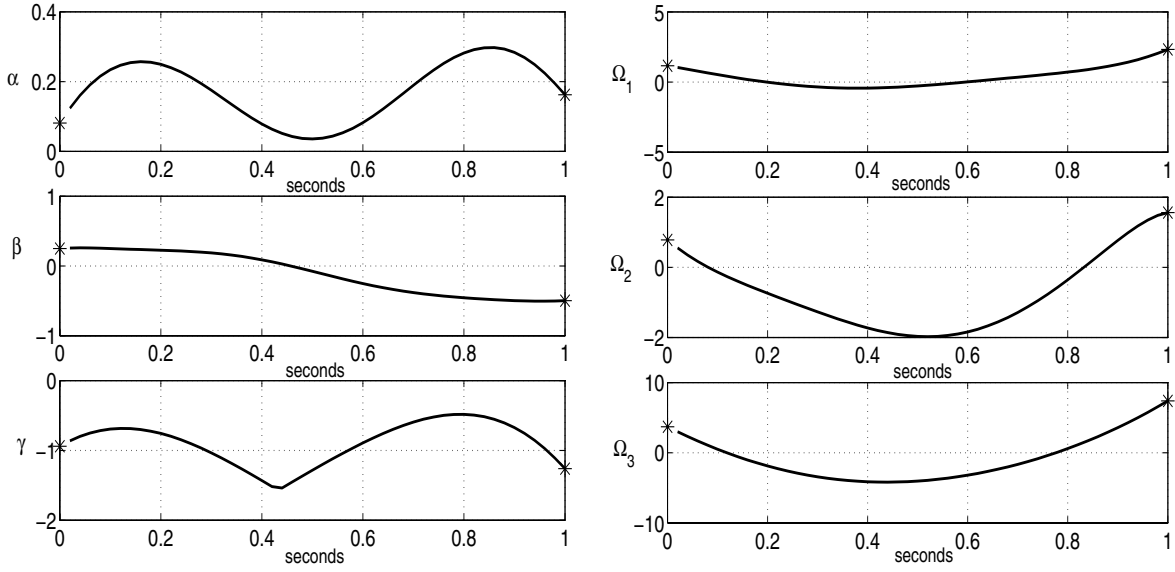
$$100 * \|\ln(Q(t)) - \ln(Q_{ref}(t))\| + \|\Omega(t) - \Omega_{ref}(t)\| + \|b(t) - b_{ref}(t)\| + \|v(t) - v_{ref}(t)\| \geq 60. \quad (20)$$

At this point, the initial guesses on the Lagrange multipliers were updated *via* the modified Newton's method until the sum on the left hand side is less than or equal to 0.5. Then we shoot forward again until the inequality in (20) was satisfied and the iteration was repeated. The norm on the orientation matrix was multiplied by 100 so that the final orientation is met accurately. The time step for the integration was 0.02 seconds. The CPU time taken for the computations run in a MATLAB environment running on a 1.8 GHz PC was 73 seconds. The solution of the two-point boundary value problem led to the following final states at $t = 1$:

$Q(1)$			$\Omega(1)$	$b(1)$	$v(1)$
-0.268	0.963	-0.011	2.319	1.901	0.672
-0.838	-0.227	0.497	1.562	8.674	3.843
0.476	0.142	0.868	7.384	4.185	9.426

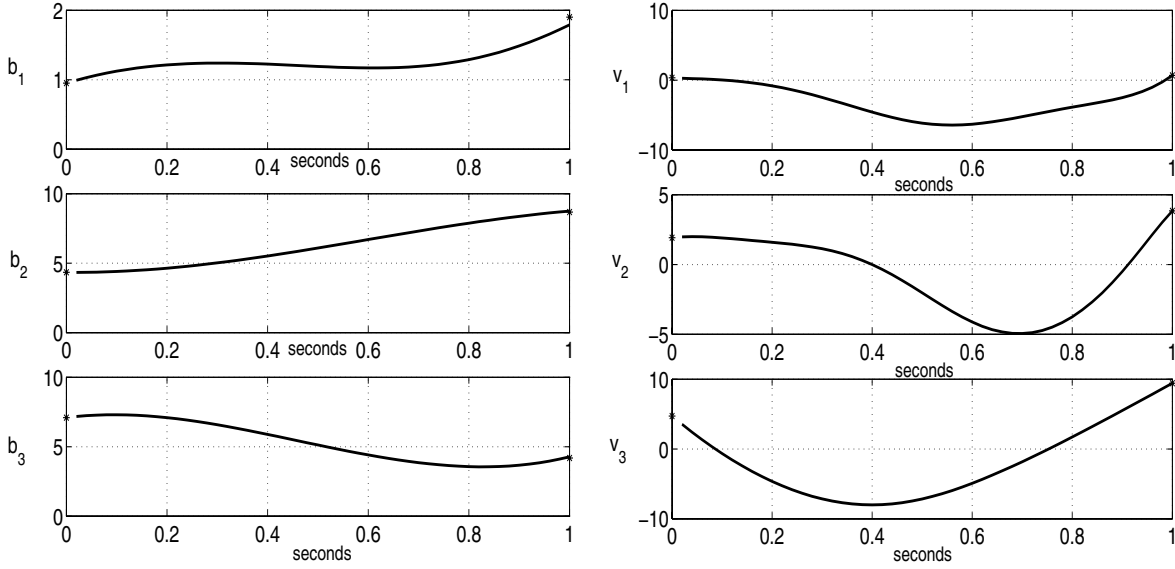
The initial value for the co-states $(\xi, \bar{v}_1, \eta, \bar{v}_2)(0)$ was chosen at random using the MATLAB random number generator (shown in the table in the left below) and converged to the table on the right below.

$\xi(0)$	$\bar{v}_1(0)$	$\eta(0)$	$\bar{v}(0)$	$\xi(0)$	$\bar{v}_1(0)$	$\eta(0)$	$\bar{v}(0)$
0.145	0.432	0.528	0.336	-12.839	-8.620	-6.766	-7.527
0.718	0.446	0.573	0.173	-18.404	-8.179	-9.251	-0.445
0.662	0.508	0.361	0.086	-80.533	-122.807	-36.136	-58.068



(a) Plot of ZYX Euler angles.

(b) Plot of angular velocities.



(c) Plot of position variables.

(d) Plot of linear velocities.

FIGURE 1. Rigid body rotation and translation problem.

Figures 1a–d show the results of the simulation. The ZYX Euler angles in Figure 1a was computed according to: $\beta = -\sin^{-1}(Q_{31})$; $\alpha = \sin^{-1}\left(\frac{Q_{32}}{\cos(\beta)}\right)$; $\gamma = \sin^{-1}\left(\frac{Q_{21}}{\cos(\beta)}\right)$.

5. CONCLUSION

In this paper, we have made three new contributions. Firstly, we have derived first order necessary conditions for an optimal control problem on a parallelizable Riemannian manifold, using frame co-ordinates. These equations specialize to those of cubic splines on Riemannian manifolds that were first discovered by Noakes, Heinzinger and Paden. Secondly, we have specialized the equations to a rigid body translation and rotation problem. Thirdly, we have presented the results of numerical experiments where we successfully computed the two point boundary value problem (TPBVP) resulting from the necessary conditions.

Acknowledgements. We wish to sincerely thank Prof. M. Toda, Department of Mathematics and Statistics, Texas Tech University for proof-reading this paper with great care.

REFERENCES

- [1] J.T. Betts, Survey of numerical methods for trajectory optimization. *Journal of Guidance, Control and Dynamics* **21** (1998) 193–207.
- [2] W.M. Boothby, *An introduction to Differential Geometry and Riemannian Manifolds*. Academic Press (1975).
- [3] P. Crouch, M. Camarinha and F. Silva Leite, Hamiltonian approach for a second order variational problem on a Riemannian manifold, in *Proc. of CONTROLO'98, 3rd Portuguese Conference on Automatic Control* (September 1998) 321–326.
- [4] P. Crouch, F. Silva Leite and M. Camarinha, Hamiltonian structure of generalized cubic polynomials, in *Proc. of the IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control* (2000) 13–18.
- [5] P. Crouch, F. Silva Leite and M. Camarinha, *A second order Riemannian variational problem from a Hamiltonian perspective*. Private Communication (2001).
- [6] T. Frankel, *The Geometry of Physics: An Introduction*. Cambridge University Press (1998).
- [7] R. Holsapple, R. Venkataraman and D. Doman, A modified simple shooting method for solving two point boundary value problems, in *Proc. of the IEEE Aerospace Conference, Big Sky, MT* (March 2003).
- [8] R. Holsapple, R. Venkataraman and D. Doman, A new, fast numerical method for solving two-point boundary value problems. *J. Guidance Control Dyn.* **27** (2004) 301–303.
- [9] V. Jurdejevic, *Geometric Control Theory*. Cambridge Studies in Advanced Mathematics (1997).
- [10] P.S. Krishnaprasad, *Optimal control and Poisson reduction*. TR 93–87, Institute for Systems Research, University of Maryland, (1993).
- [11] A. Lewis, *The geometry of the maximum principle for affine connection control systems*. Preprint, available online at <http://penelope.mast.queensu.ca/~andrew/cgibin/pslist.cgi?papers.db>, 2000.
- [12] D.G. Luenberger, *Optimization by Vector Space Methods*. John Wiley and Sons (1969).
- [13] M.B. Milam, K. Mushambi and R.M. Murray, A new computational approach to real-time trajectory generation for constrained mechanical systems, in *Proc. of 39th IEEE Conference on Decision and Control* **1** (2000) 845–851.
- [14] R.M. Murray, Z. Li and S.S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. CRC Press (1994).
- [15] L. Noakes, G. Heinzinger and B. Paden, Cubic splines on curved spaces. *IMA J. Math. Control Inform.* **6** (1989) 465–473.
- [16] H.J. Pesch, Real-time computation of feedback controls for constrained optimal control problems. Part 1: Neighbouring extremals. *Optim. Control Appl. Methods* **10** (1989) 129–145.
- [17] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, pp. 272–286; 502–535. Springer-Verlag, New York, second edition (1993).
- [18] H. Sussmann, An introduction to the coordinate-free maximum principle, in *Geometry of Feedback and Optimal Control*, B. Jakubczyk and W. Respondek Eds. Marcel Dekker, New York (1997) 463–557.